MARTINGALE SOLUTIONS OF STOCHASTIC NONLOCAL CROSS-DIFFUSION SYSTEMS

M. BENDAHMANE AND K. H. KARLSEN

ABSTRACT. We establish the existence of solutions for a class of stochastic reaction-diffusion systems with cross-diffusion terms modeling interspecific competition between two populations. More precisely, we prove the existence of weak martingale solutions employing appropriate Faedo-Galerkin approximations and the stochastic compactness method. The nonnegativity of solutions is proved by a stochastic adaptation of the well-known Stampacchia approach.

1. INTRODUCTION

This work is devoted to the mathematical analysis of a stochastic reaction-diffusion system with cross-diffusion modeling the interaction between two populations. Cross-diffusion expresses that the population flux of a given subpopulation is affected by the presence of other subpopulations. The (deterministic) dynamics of interacting species with cross-diffusion were investigated by many authors, including Levin [25], Levin and Segel [24], Okubo and Levin [31], Mimura and Murray [27], Mimura and Kawasaki [26], Mimura and Yamaguti [28], Galiano *et al.* [17, 18], Bendahmane *et al.* [1, 6] (see also [2, 3, 5, 7]) to name a few. We consider a spatially distributed population wherein $u = u(t, x)$ and $v = v(t, x)$ are the respective densities of two subpopulations at time t and location $x \in \Omega$. The variables u and v may represent predator and prey densities. In the context of dispersal of an epidemic disease, the two variables u and v may represent predator and prey densities for susceptible (those who can catch the disease) and infectious individuals (those who are infected and can transmit the disease). Let $p = u + v$ be the total population density. The population in each subclass is given by

$$U(t) = \int_\Omega u(t, x)\, dx, \qquad V(t) = \int_\Omega v(t, x)\, dx,$$

whereas the total population is

$$P(t) = \int_\Omega (u + v)\,(t, x)\, dx = \int_\Omega p(t, x)\, dx,$$

where Ω is a bounded open domain of $\mathbb{R}^d$ ($d = 3$), with C^3 boundary $\partial\Omega$ and outward unit normal ν. In this work, we assume that the diffusion of individuals follows a Fick law modified by various other processes such as searching for food, escaping high infection risks, or avoiding large concentrations of individuals. This means that the mobility in each subclass is influenced by the spatial gradient of the other subclass (cf. e.g. [29, 30, 31]).

Date: April 29, 2022.

2010 *Mathematics Subject Classification.* Primary: 60H15, 35K57; Secondary: 35M10, 35A05.

Key words and phrases. Stochastic partial differential equation, cross-diffusion system, existence, martingale solution, population dynamics.

The work of Bendahmane is supported by the Fincome project. The work of Karlsen is supported by the Research Council of Norway (“Stochastic Conservation Laws”, 250674/F20).

A prototype of stochastic reaction-diffusion systems with nonlocal diffusion and cross-diffusion terms is

$$
\begin{aligned}
&du - \nabla\cdot\left(D_u\Big(\int_\Omega u(t,x)\,dx\Big)\nabla u + \mathcal{A}_{11}(u,v)\nabla u + \mathcal{A}_{12}(u,v)\nabla v\right)\,dt \\
&\qquad\qquad = F(u,v)\,dt + \sigma_u(u)\,dW_u(t), \\
&dv - \nabla\cdot\left(D_v\left(\int_\Omega v(t,x)\,dx\Big)\nabla v + \mathcal{A}_{21}(u,v)\nabla u + \mathcal{A}_{22}(u,v)\nabla v\right)\,dt \\
&\qquad\qquad = G(u,v)\,dt + \sigma_v(v)\,dW_v(t),
\end{aligned}
\tag{1.1}
$$

which is posed in the time-space cylinder $\Omega_T := (0,T)\times\Omega$. This system is supplemented with nonnegative initial data,

$$
u(0,x) = u_0(x) \ge 0, \quad v(0,x) = v_0(x) \ge 0, \qquad x\in\Omega, \tag{1.2}
$$

and zero-flux boundary conditions on $\Sigma_T := (0,T)\times\partial\Omega$:

$$
\begin{aligned}
&\left(D_u\Big(\int_\Omega u(t,x)\,dx\Big)\nabla u + \mathcal{A}_{11}(u,v)\nabla u + \mathcal{A}_{12}(u,v)\nabla v\right)\cdot\nu = 0, \\
&\left(D_v\Big(\int_\Omega v(t,x)\,dx\Big)\nabla v + \mathcal{A}_{21}(u,v)\nabla u + \mathcal{A}_{22}(u,v)\nabla v\right)\cdot\nu = 0.
\end{aligned}
\tag{1.3}
$$

In the system (1.1), W_w is a cylindrical Wiener process, with noise function σ_w for $w = u, v$. Formally, we can think of $\sigma_w(w)\,dW_w$ as $\sum_{k\ge1}\sigma_{w,k}(w)\,dW_{k,w}(t)$, where $\{W_{w,k}\}_{k\ge1}$ is a sequence of independent 1D Brownian motions and $\{\sigma_{w,k}\}_{k\ge1}$ a sequence of noise coefficients. The processes W_u and W_v are independent, and the terms $\sigma_u(u)\,dW_u$ and $\sigma_v(v)\,dW_v$ model environmental noise.

In (1.1),

$$
F(u,v) := -\theta(u,v) - \mu u \qquad G(u,v) := \theta(u,v) - \gamma v - \mu v \tag{1.4}
$$

are the reaction terms. In the dispersal of an epidemic disease, the constants $\mu,\gamma>0$ are the biological parameters of the system (think of $1/\gamma$ as the duration of the infectious stage and μ as the mortality rate). The incidence function θ takes a proportionate mixing form: for some constant $\alpha>0$,

$$
\theta(u,v) = \alpha\frac{uv}{u+v}, \qquad u,v\ge 0. \tag{1.5}
$$

For later use, note that

$$
0 \le \theta(u,v) \le \alpha\min(u,v), \qquad u,v\ge 0. \tag{1.6}
$$

The diffusion rates (given by $D_u(\cdot)$ and $D_v(\cdot) > 0$) are assumed to be "nonlocal", depending on the whole of each population rather than on the local density; in other words, the diffusion of individuals is guided by the global state of the population in the medium. For example, if we want to model species tending to leave crowded zones, a natural assumption is that $D_u(\cdot), D_v(\cdot)$ are increasing functions. Otherwise, for species attracted by a growing population, one may assume that the nonlocal diffusion coefficients $D_u(\cdot), D_v(\cdot)$ are decreasing functions. We assume that $D_u, D_v : \mathbb{R}\to\mathbb{R}$ are continuous functions satisfying the following conditions: $\exists C_m,\, C_M > 0$ such that for $w = u, v$,

$$
D_w(I) \ge C_m, \qquad |D_w(I_1) - D_w(I_2)| \le C_M\,|I_1 - I_2|\,, \quad \forall I, I_1, I_2 \in\mathbb{R}. \tag{1.7}
$$

In (1.1), $\mathcal{A}(u,v) = \{\mathcal{A}_{ij}(u,v)\}_{i,j=1}^2$ is the cross-diffusion matrix. For simplicity of presentation, we introduce the short-hand notation

$$\mathcal{A}(u,v)\begin{pmatrix}\nabla u\\ \nabla v\end{pmatrix} = \begin{pmatrix}\mathcal{A}_{11}(u,v)\nabla u + \mathcal{A}_{12}(u,v)\nabla v\\ \mathcal{A}_{21}(u,v)\nabla u + \mathcal{A}_{22}(u,v)\nabla v\end{pmatrix}.$$

We assume that the matrix $\mathcal{A}$ has as C^2 entries and satisfies the following conditions:

$$\begin{aligned}
&\forall u,v \geq 0, \quad \mathcal{A}_{12}(0,v) = 0, \quad \mathcal{A}_{21}(u,0) = 0,\\
&\forall u,v \geq 0,\ \forall \mathbf{w} := \begin{pmatrix}\mathbf{w}_1\\ \mathbf{w}_2\end{pmatrix} \in \mathbb{R}^{2d}, \quad \big(\mathcal{A}(u,v)\mathbf{w},\mathbf{w}\big) \geq \frac{1}{C}\left|\mathcal{A}(u,v)\right|\,\left|\mathbf{w}\right|^2,\\
&\forall u_1,u_2,v_1,v_2 \geq 0, \quad |\mathcal{A}(u_1,v_1) - \mathcal{A}(u_2,v_2)| \leq C\big(|u_1-u_2| + |v_1-v_2|\big),
\end{aligned} \tag{1.8}$$

where $(\cdot,\cdot)$ is the usual scalar product on $\mathbb{R}^2$, with corresponding norm $|\cdot|$. Moreover, $|\mathcal{A}(\cdot,\cdot)| = \max_{i,j=1,2} |\mathcal{A}_{ij}(\cdot,\cdot)|$ and C is a positive constant. Notice that (1.8) implies

$$\mathcal{A}_{11}(u,v) \geq 0, \quad \mathcal{A}_{22}(u,v) \geq 0, \quad \forall u,v \geq 0.$$

A typical example of a cross-diffusion matrix is

$$\mathcal{A}(u,v) = \begin{pmatrix} a_{11}u + a_{12}v & a_{13}u\\ a_{21}v & a_{22}u + a_{23}v\end{pmatrix},$$

where the coefficients $a_{ij} > 0$ are known as self-diffusion rates. This matrix is nonnegative if $8a_{11}a_{21} \geq a_{12}^2$ and $8a_{22}a_{12} \geq a_{21}^2$, cf. [4] for more details.

Remark 1.1. *For the upcoming analysis we need to extend the definitions of $\mathcal{A}$, cf.* (1.8)*, F and G to all $u,v \in \mathbb{R}$. We do this by assuming the following (for $i,j = 1,2$):*

if $u,v \geq 0$, then $\mathcal{A}_{ij}(u,v) \geq 0$, otherwise $\mathcal{A}_{ij}(u,v) = 0$ ($i \neq j$) and $\mathcal{A}_{ii}(u,v) \geq 0$,

$$F(u,v) = \begin{cases} -\theta(u,v) - \mu u, & \text{if } u,v \geq 0,\\ -\mu u, & \text{if } u \geq 0 \text{ and } v < 0,\\ 0, & \text{if } u < 0 \text{ and } v \geq 0,\end{cases}$$

$$G(u,v) = \begin{cases} \theta(u,v) - \gamma v - \mu v, & \text{if } u,v \geq 0,\\ 0, & \text{if } u \geq 0 \text{ and } v < 0,\\ -\gamma v - \mu v, & \text{if } u < 0 \text{ and } v \geq 0.\end{cases}$$

Our analysis is restricted to positive cross-diffusion matrices $\mathcal{A}$. Positive matrices are motivated by their applications in population dynamics. In a forthcoming work, we decipher stability and instability conditions for the spatially constant stationary state. Moreover, we define and prove the existence of suitably defined solutions satisfying these conditions. The “natural” solutions are determined when the nonlinearities and cross diffusivities obey certain constraints. In the deterministic case [3], these constraints are not fully satisfied for realistic parameters, yielding instabilities. The interesting open question is, which type of solution experiences instabilities? Degenerate cross-diffusion systems and numerical methods will be the subject of another forthcoming work.

Historically, cross-diffusion models are deterministic,, meaning that the input data determine the solution at each moment in time. In deterministic models, non-predictable environmental factors are not considered, although it is well-known that a combination of random perturbations and nonlinearities can strongly influence solutions. Multiple factors may influence the population’s growth in the environment, such as food, water, temperature, etc., each element easily being thought of as stochastic. It is natural to employ noise to model these environmental fluctuations by adding a stochastic forcing term to the deterministic system, resulting in (1.1).

Let us now put the mathematical contributions of this paper into perspective. First, note that the standard theory for parabolic systems does not apply naturally to the cross-diffusion model because of the strong coupling in the highest derivatives. As a result, no traditional maximum principle applies. A stochastic forcing term further complicates the maximum principle approach. The existence result for (1.1) is based on martingale solutions and the introduction of suitable approximate (Faedo-Galerkin) solutions. We derive a series of system-specific a priori estimates in $L^2_{\omega,t}H^1 \cap L^2_\omega L^\infty_t L^2_x \cap L^1_\omega C_t(W^{1,4}_x)^\star$ for the Faedo-Galerkin approximations and use a compactness method to conclude convergence. The system's nonlinear structure requires strong convergence of the approximate solutions in suitable norms. However, one cannot directly deduce strong convergence in the probability variable. To handle this issue, we establish weak compactness of the probability laws of the approximate solutions, which follows from tightness and Prokhorov's theorem. We then construct a.s. convergent versions of the approximations using Skorokhod's representation theorem, which makes it possible to show that the limit constitutes a martingale solution of (1.1). We demonstrate that the constructed solutions are nonnegative by adapting the Stampacchia approach to the stochastic setting, following Chekroun, Park, and Temam [10]. Finally, we mention that the pathwise uniqueness of the solution for the deterministic and stochastic cross-diffusion systems remains an open problem.

In [14], the authors prove the existence of solutions for a related stochastic cross-diffusion system (with $F, G, D_u, D_v \equiv 0$) using the entropy method, assuming that the cross-diffusion matrix exhibits a quadratic entropy structure. A critical difference between our work and [14] is that the cross-diffusion term in the predator-prey system (1.1) does not have an entropy structure. Besides, the system (1.1) contains nonlocal diffusion terms, which further breaks the entropy structure in [14].

For the existence of martingale solutions for other classes of SPDEs, we refer [8, 9, 11, 12, 15, 16, 20, 21, 22, 34, 35], to mention a few inspirational examples.

The paper is organized as follows: In Section 2, we present the stochastic framework and state the noise coefficients' hypotheses. Section 3 supplies the definition of a weak martingale solution and declares the main result. We construct approximate solutions by the Faedo-Galerkin method in Section 4. Uniform estimates for these approximations are established in Sections 5 and 6. Section 7 proves the tightness of the probability laws generated by the Faedo-Galerkin approximations. The tightness and Skorokhod's representation theorem is used to show that a weakly convergent sequence of the probability laws has a limit that can be represented as the law of an almost surely convergent sequence of random variables defined on a common probability space. The limit of this sequence is proved to be a weak martingale solution of the stochastic reaction-diffusion system in Section 8, while its nonnegativity is deferred to Section 9.

Throughout this paper, we will frequently use the letters C, K, etc., to denote a generic constant independent of n, that may take different values at different occurrences.

2. STOCHASTIC FRAMEWORK

This section recalls basic concepts and results from stochastic analysis (see e.g. [11, 33] for more details). We consider a complete probability space $(D, \mathcal{F}, P)$, along with a complete right-continuous filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$.

In passing, note that the letter Ω is reserved for the physical domain in this paper. In contrast, we use D for the probability domain (in the stochastic literature, Ω denotes the probability domain).

Given a separable Banach space $\mathbb{B}$, which is equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{B})$, a $\mathbb{B}$-valued random variable X is a measurable mapping from $(D, \mathcal{F}, P)$ to $(\mathbb{B}, \mathcal{B}(\mathbb{B}))$, $D \ni \omega \mapsto X(\omega) \in \mathbb{B}$. The expectation of a random variable X is $\mathbb{E}[X] := \int_D X \, dP$. For $p \geq 1$, the Banach space $L^p(D; \mathbb{B}) = L^p(D, \mathcal{F}, P; \mathbb{B})$ is the collection of all $\mathbb{B}$-valued random variables, equipped with the following norm

$$
\begin{aligned}
\|X\|_{L^p(D;\mathbb{B})} &= \|X\|_{L^p(D,\mathcal{F},P;\mathbb{B})} := \left(\mathbb{E}\big[\|X\|_{\mathbb{B}}^p\big]\right)^{\frac{1}{p}} \quad (p < \infty), \\
\|X\|_{L^p(D;\mathbb{B})} &= \|X\|_{L^\infty(D,\mathcal{F},P;\mathbb{B})} := \sup_{\omega \in D} \|X(\omega)\|_{\mathbb{B}} .
\end{aligned}
$$

We use the abbreviation a.s. (or almost surely) for "P-almost every $\omega \in D$". A stochastic process $X = \{X(t)\}_{t\in[0,T]}$ is a collection of $\mathbb{B}$-valued random variables $X(t)$. We assume that X is *measurable*, which means that the map $X : D \times [0, T] \to \mathbb{B}$ is measurable from $\mathcal{F} \times \mathcal{B}([0, T])$ to $\mathcal{B}(\mathbb{B})$. The paths $t \to X(\omega, t)$ are then automatically Borel measurable.

We refer to

$$
\mathcal{S} = \Big(D, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, P, \{W_k\}_{k=1}^\infty\Big) \tag{2.1}
$$

as a (Brownian) *stochastic basis*, where $\{W_k\}_{k=1}^\infty$ is a sequence of independent one-dimensional Wiener processes adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$.

A stochastic process X is *adapted* if $X(t)$ is $\mathcal{F}_t$ measurable for all $t \in [0, T]$. When a filtration is involved there are additional notions of measurability (predictable, optional and progressive) that occasionally are more convenient to work with. Herein we use the (stronger) notion of a predictable process. A *predictable* process is a $\mathcal{P}_T \times \mathcal{B}([0, T])$ measurable map $D \times [0, T] \to \mathbb{B}$, $(\omega, t) \mapsto X(\omega, t)$, where $\mathcal{P}_T$ is the predictable σ-algebra on $D \times [0, T]$ associated with $\{\mathcal{F}_t\}_{t\in[0,T]}$, i.e., the σ-algebra generated by all left-continuous adapted processes.

Consider a Hilbert space $\mathbb{U}$ equipped with a complete orthonormal basis $\{\psi_k\}_{k\geq 1}$. A cylindrical Wiener process W on $\mathbb{U}$ is defined by $W := \sum_{k\geq 1} W_k \psi_k$. The vector space of all bounded linear operators from $\mathbb{U}$ to $L^2(\Omega)$ is denoted $L(\mathbb{U}, L^2(\Omega))$. Denote by $L_2(\mathbb{U}, L^2(\Omega))$ the space of Hilbert-Schmidt operators from $\mathbb{U}$ to $L^2(\Omega)$, i.e., $R \in L_2(\mathbb{U}, L^2(\Omega)) \Leftrightarrow R \in L(\mathbb{U}, L^2(\Omega))$, $\|R\|^2_{L_2(\mathbb{U},L^2(\Omega))} := \sum_{k\geq 1} \|R\psi_k\|_{L^2(\Omega)} < \infty$. We recall that $L_2(\mathbb{U}, L^2(\Omega))$ is a Hilbert space. As is well-known, there is an auxiliary Hilbert space $\mathbb{U}_0 \supset \mathbb{U}$, with a Hilbert-Schmidt embedding $J : \mathbb{U} \to \mathbb{U}_0$, on which the infinite series $\sum_{k\geq 1} W_k \psi_k$ converges.

For a given cylindrical Wiener process W_w, the $L^2(\Omega)$-valued Itô stochastic integral $\int \sigma \, dW_w$ is defined as follows (see for, e.g., [11, 33]):

$$
\int_0^t \sigma_w \, dW_w = \sum_{k=1}^\infty \int_0^t \sigma_{w,k} \, dW_{w,k}, \qquad \sigma_{w,k} := \sigma_w \psi_k, \tag{2.2}
$$

for any $L^2(\Omega)$-valued predictable integrand

$$
\sigma \in L^2\left(D, \mathcal{F}, P; L^2\big(0, T; L_2(\mathbb{U}, L^2(\Omega))\big)\right).
$$

Throughout the paper, we assume several conditions on the noise coefficients σ_u, σ_v appearing in (1.1). For each $z \in L^2(\Omega)$, we assume that $\sigma_w(z) : \mathbb{U} \to L^2(\Omega)$, for $w = u, v$, is defined by

$$
\sigma_w(z)\psi_k = \sigma_{w,k}(z(\cdot)), \quad k \geq 1,
$$

for some real-valued functions $\sigma_{w,k}(\cdot):\mathbb{R}\to\mathbb{R}$ that satisfy

$$
\begin{aligned}
&\sum_{k\geq 1}|\sigma_{w,k}(z)|^2 \leq C_\sigma\left(1+|z|^2\right), \qquad \forall z\in\mathbb{R},\\
&\sum_{k\geq 1}|\sigma_{w,k}(z_1)-\sigma_{w,k}(z_2)|^2 \leq C_\sigma\,|z_1-z_2|^2, \qquad \forall z_1,z_2\in\mathbb{R},
\end{aligned}
\tag{2.3}
$$

for a constant $C_\sigma>0$. A consequence of (2.3) is

$$
\begin{aligned}
&\|\sigma_w(z)\|^2_{L_2(\mathbb{U},L^2(\Omega))} \leq C_\sigma\left(1+\|z\|^2_{L^2(\Omega)}\right), \quad z\in L^2(\Omega),\\
&\|\sigma_w(z_1)-\sigma_w(z_2)\|^2_{L_2(\mathbb{U},L^2(\Omega))} \leq C_\sigma\,\|z_1-z_2\|^2_{L^2(\Omega)}, \quad z_1,z_2\in L^2(\Omega).
\end{aligned}
\tag{2.4}
$$

Under these conditions (2.4), the stochastic integral (2.2) is an $L^2(\Omega)$-valued square integrable martingale, satisfying the Burkholder-Davis-Gundy (BDG) inequality

$$
\mathbb{E}\left[\sup_{t\in[0,T]}\left\|\int_0^t \sigma_w\,dW_w\right\|^p_{L^2(\Omega)}\right] \leq C\,\mathbb{E}\left[\left(\int_0^T \|\sigma_w\|^2_{L_2(\mathbb{U},L^2(\Omega))}\,dt\right)^{\frac{p}{2}}\right],
\tag{2.5}
$$

where C is a constant depending on $p\geq 1$.

We need the following convergence result for stochastic integrals [12, Lemma 2.1].

Lemma 2.1 (convergence of stochastic integrals)**.** *For each $n\in\mathbb{N}$, consider a stochastic basis $\mathcal{S}_n=\big(D,\mathcal{F},\{\mathcal{F}^n_t\},P,W^n\big)$ and a $\{\mathcal{F}^n_t\}$–predictable process G_n, which belongs to $L^2\big(0,T;L_2(\mathbb{U},L^2(\Omega))\big)$, almost surely. Furthermore, suppose there exist a stochastic basis $\mathcal{S}=\big(D,\mathcal{F},\{\mathcal{F}_t\},P,W\big)$ and a $\{\mathcal{F}_t\}$–predictable process G, which belongs to $L^2\big(0,T;L_2(\mathbb{U},L^2(\Omega))\big)$ a.s., such that*

$$
W^n \overset{n\uparrow\infty}{\longrightarrow} W \quad \text{in } C([0,T];\mathbb{U}_0), \text{ in probability}
$$

$$
G^n \overset{n\uparrow\infty}{\longrightarrow} G \quad \text{in } L^2\big(0,T;L_2(\mathbb{U};L^2(\Omega))\big), \text{ in probability.}
$$

Then

$$
\int_0^t G^n\,dW^n \overset{n\uparrow\infty}{\longrightarrow} \int_0^t G\,dW \quad \text{in } L^2(0,T;L^2(\Omega)), \text{ in probability.}
$$

Let $\mathbb{S}$ be a Polish space. We denote by $\mathcal{B}(\mathbb{S})$ the collection Borel subsets of $\mathbb{S}$ and by $\mathcal{P}(\mathbb{S})$ the family of all Borel probability measures on $\mathbb{S}$. A sequence of probability measures $\{\mu_n\}_{n\geq 1}$ on $(\mathbb{S},\mathcal{B}(\mathbb{S}))$ is tight [11] if for every $\epsilon>0$ there is a compact set $\mathbb{K}_\epsilon\subset\mathbb{S}$ such that $\mu_n(\mathbb{K}_\epsilon)>1-\epsilon$ for all $n\geq 1$. According to Prokhorov's theorem (see e.g. [11, Theorem 2.3]), tightness is a criterion for weak compactness: If $\{\mu_n\}_{n\geq 1}$ is tight, then there exists a subsequence $\{\mu_{n_j}\}_{j\geq 1}$ that converges weakly to a probability measure μ, where weak convergence means that $\int_{\mathbb{S}}\phi(w)\,d\mu_{n_j}(w)\to\int_{\mathbb{S}}\phi(w)\,d\mu(w)$, for any continuous bounded function $\phi:\mathbb{S}\to\mathbb{R}$.

Any random variable $X:D\to\mathbb{S}$ induces a probability measure $\mathcal{L}$ on $(\mathbb{S},\mathcal{B}(\mathbb{S}))$ via the pushforward of P through X, often $\mathcal{L}=P\circ X^{-1}$ is referred to as the law of X. Let $\{X_k\}_{k\geq 1}$ be a sequence of random variables whose laws $\mathcal{L}_k$ converge weakly to $\mathcal{L}$. Then a well-known result of Skorokhod (see e.g. [11, Theorem 2.4]) says that there exist a probability space $(\tilde{D},\tilde{\mathcal{F}},\tilde{P})$ and random variables $\tilde{X}_k,\tilde{X}:\tilde{\mathcal{D}}\to\mathbb{S}$ such that the law of $\tilde{X}_k$ is $\mathcal{L}_k$, the law of $\tilde{X}$ is $\mathcal{L}$, and $\tilde{X}_k\to\tilde{X}$ $\tilde{P}$-almost surely as $k\to\infty$.

3. Notion of solution and main result

We will utilize the following notion of solution for the stochastic cross-diffusion system.

Definition 3.1 (weak martingale solution). *Let μ_{u_0}, μ_{v_0} be probability measures on $L^2(\Omega)$. A weak martingale solution of the stochastic cross-diffusion system (1.1), with initial-boundary data (1.2) and (1.3), is a triplet $\big(\mathcal{S}, u, v\big)$ satisfying the following conditions:*

(1) $\mathcal{S} = \big(D, \mathcal{F}, \{\mathcal{F}_t\}, P, \{W_{u,k}\}_{k=1}^{\infty}, \{W_{v,k}\}_{k=1}^{\infty}\big)$ *is a stochastic basis;*
(2) $W_u := \sum_{k\geq 1} W_{k,u}\psi_k$ *and* $W_v := \sum_{k\geq 1} W_{k,v}\psi_k$ *are two independent cylindrical Wiener processes, adapted to the filtration* $\{\mathcal{F}_t\}$*;*
(3) *The elements* u *and* v *are nonnegative, belong to*

$$L^2\big(D, \mathcal{F}, P; L^2(0,T;H^1(\Omega))\big) \bigcap L^2\big(D, \mathcal{F}, P; L^\infty(0,T;L^2(\Omega))\big),$$

and satisfy

$$\sqrt{|\mathcal{A}_{ij}(u,v)|}\nabla u \in L^2\big(D, \mathcal{F}, P; L^2(0,T;L^2(\Omega))\big), \quad i,j = 1,2.$$

Finally, $u, v \in C\left([0,T]; \big(H^1(\Omega)\big)^\star\right)$ *a.s., and* u, v *are predictable in* $\big(H^1(\Omega)\big)^\star$.
(4) *The laws of* $u_0 := u(0)$ *and* $v_0 := v(0)$ *are respectively* μ_{u_0} *and* μ_{v_0}*;*
(5) *The following equations hold* P*-almost surely, for any* $t \in [0,T]$*:*

$$\begin{aligned}
&\int_\Omega u(t)\varphi_u \,dx - \int_\Omega u_0\varphi_u\,dx \\
&\quad + \int_0^t \int_\Omega \left(D_u\Big(\int_\Omega u(t,x)\,dx\Big)\nabla u + \mathcal{A}_{11}(u,v)\nabla u + \mathcal{A}_{12}(u,v)\nabla v\right)\cdot\nabla\varphi_u\,dx\,ds \\
&\qquad = \int_0^t\int_\Omega F(u,v)\varphi_u\,dx\,ds + \int_0^t\int_\Omega \sigma_u(u)\varphi_u\,dx\,dW_u(s),
\end{aligned}$$

(3.1)

$$\begin{aligned}
&\int_\Omega v(t)\varphi_v \,dx - \int_\Omega v_0\varphi_v\,dx \\
&\quad + \int_0^t \int_\Omega \left(D_v\Big(\int_\Omega v(t,x)\,dx\Big)\nabla v + \mathcal{A}_{21}(u,v)\nabla u + \mathcal{A}_{22}(u,v)\nabla v\right)\cdot\nabla\varphi_v\,dx\,ds \\
&\qquad = \int_0^t\int_\Omega G(u,v)\varphi_v\,dx\,ds + \int_0^t\int_\Omega \sigma_v(v)\varphi_v\,dx\,dW_v(s),
\end{aligned}$$

for all $\varphi_u, \varphi_v \in W^{1,4}(\Omega)$.

Remark 3.2. *In Definition 3.1, we use the standard Sobolev spaces*

$$\begin{aligned}
&H^1(\Omega) = W^{1,2}(\Omega), \text{ and for } p \in (1,\infty), \\
&W^{1,p}(\Omega) = \left\{u \in L^p(\Omega) : \nabla u \in L^p(\Omega;\mathbb{R}^d)\right\},
\end{aligned}$$

along with the corresponding dual spaces $\big(H^1(\Omega)\big)^*$ *and* $\big(W^{1,p}(\Omega)\big)^*$*. Later we also use the space* $H^2(\Omega)$ *consisting of all functions* $u \in L^2(\Omega)$ *for which* $\nabla u \in L^2(\Omega;\mathbb{R}^d)$ *and* $\nabla^2 u \in L^2(\Omega;\mathbb{R}^{d\times d})$*. Throughout the paper we use* $\big(W^{1,p}(\Omega)\big)^\star$ *to denote the dual of* $W^{1,p}(\Omega)$*, which is a Banach space with norm*

$$\|L\|_{(W^{1,p}(\Omega))^\star} = \sup\left\{\big|\langle L, \phi\rangle\big| \,:\, \phi \in W^{1,p}(\Omega),\ \|\phi\|_{W^{1,p}(\Omega)} \leq 1\right\},$$

where $\langle\cdot,\cdot\rangle$ *is the duality pairing between* $\big(W^{1,p}(\Omega)\big)^\star$ *and* $W^{1,p}(\Omega)$.

Recall that $L \in \left(W^{1,p}(\Omega)\right)^\star$ if and only if there exist functions $f_0, f_1, \dots, f_d \in L^{p'}(\Omega)$, $p' = \frac{p}{p-1}$, such that

$$\langle L, \phi \rangle = \int_\Omega f_0 \phi + \sum_{i=1}^d f_i \partial_{x_i} \phi \, dx, \quad \forall \phi \in W^{1,p}(\Omega),$$

and $\|L\|_{(W^{1,p}(\Omega))^\star} = \left(\sum_{i=0}^d \|f_i\|_{L^{p'}(\Omega)}\right)^{1/p'}$ [23, Theorem 10.41]. *Note that bounded linear functionals over $W^{1,p}(\Omega)$ are not distributions.*

Remark 3.3. *1. Given the regularity conditions imposed in Definition 3.1, one can show that the deterministic and the stochastic integrals in* (3.1) *are all well-defined. Regarding the stochastic terms $\int_0^t \int_\Omega \sigma_w(w)\varphi_w \, dx \, dW_w(s)$, $w = u, v$, they are interpreted as in* (2.2).

2. For martingale solutions, one prescribes the initial data in terms of probability measures μ_{u_0}, μ_{v_0} on $L^2(\Omega)$, For probabilistic strong solutions (not considered here), one prescribes the initial data in terms of random variables $u_0, v_0 \in L^2_{\omega,x} := L^2\left(D; L^2(\Omega)\right)$.

3. Part (3) *of Definition 3.1 implies that u, v belong to the space $L^\infty(0,T; L^2(\Omega)) \cap C\big([0,T]; \left(H^1(\Omega)\right)^\star\big)$, almost surely. Hence, $u, v \in C_w([0,T]; L^2(\Omega))$ a.s., i.e., for any $\phi \in L^2(\Omega)$, $[0,T] \ni t \mapsto \int_\Omega w(t)\phi \, dx$ is continuous a.s., for $w = u, v$. We do not have $u, v \in C([0,T]; L^2(\Omega))$ (strong time-continuity in L^2). As $W^{1,4}(\Omega) \subset H^1(\Omega)$ with continuous embedding (recall that $\Omega \subset \mathbb{R}^3$ bounded), $\left(H^1(\Omega)\right)^\star \subset \left(W^{1,4}(\Omega)\right)^\star$ with continuous embedding, and therefore $u, v \in C\left([0,T]; \left(W^{1,4}(\Omega)\right)^\star\right)$ a.s., which is consistent with requiring the equations* (3.1) *to hold for all $\varphi_u, \varphi_v \in W^{1,4}(\Omega)$.*

Remark 3.4. *A significant difficulty for the analysis of* (1.1) *is the strong coupling in the highest derivatives. However, since these terms are zero on the boundary, cf.* (1.3)*, the nonlinear boundary conditions will "disappear" in the weak martingale formulation.*

Our main result is

Theorem 3.5 (existence)**.** *Suppose conditions* (1.4), (1.5), (1.6), (1.8), (1.7), *and* (2.3) *hold, and that the initial data u_0, v_0 are random variables with laws μ_{u_0}, μ_{v_0} satisfying*

$$\int_{L^2(\Omega)} \|w\|_{L^2(\Omega)}^{q_0} \, d\mu_{w_0}(w) < \infty, \quad \text{for some } q_0 > 3, \quad w := u, v. \tag{3.2}$$

Then the stochastic cross-diffusion system (1.1)*, with initial-boundary data* (1.2) *and* (1.3)*, possesses a weak martingale solution in the sense of Definition 3.1. Moreover, assuming $\sigma_u(0) = \sigma_v(0) = 0$, this martingale solution is nonnegative.*

The proof of Theorem 3.5 is organized into several sections. First, in Section 4, we construct the Faedo-Galerkin solutions. Energy-type estimates are derived in Section 5. Convergence of the approximate solutions (along a subsequence) to a limit follows from these estimates, a temporal translation estimate, cf. 6, and the tightness of the probability laws generated by the Faedo-Galerkin solutions, cf. Section 7. In Section 8, we show that the limit is a weak martingale solution. Finally, we prove the nonnegativity of the constructed martingale solution, cf. Section 9.

4. Construction of approximate solutions

In this section, we define precisely the Faedo-Galerkin equations and prove that there exists a solution to these equations. We begin by fixing a stochastic basis $\mathcal{S}$, cf. (2.1), and $\mathcal{F}_0$-measurable initial data $u_0, v_0 \in L^2(D; L^2(\Omega))$, with respective laws μ_{u_0}, μ_{v_0} on $L^2(\Omega)$. We look for approximate solutions obtained from the projection of (1.1), (1.2) and (1.3) onto a finite dimensional space $\mathbb{X}_n := \mathrm{Span}\{e_1, \ldots, e_n\}$.

Let us make precise the basis functions $e_1, \ldots, e_n$. The following discussion is well-known but is included for the sake of readability. First, we introduce the spaces

$$\begin{aligned} L_0^2 &:= \left\{ u \in L^2(\Omega) \,:\, \overline{u} := \frac{1}{|\Omega|} \int_\Omega u \, dx = 0 \right\}, \\ H_N^2 &:= \left\{ u \in H^2(\Omega) \,:\, \frac{\partial u}{\partial \nu} = 0 \text{ on } \partial\Omega \right\}, \\ (H^1)_0^\star &:= \left\{ u \in \left(H^1(\Omega)\right)^\star \,:\, \overline{u} := \frac{1}{|\Omega|} \langle u, 1 \rangle_{(H^1)^\star, H^1} = 0 \right\}. \end{aligned}$$

The embeddings $H_N^2 \subset H^1 \subset L^2 \cong \left(L^2\right)^\star \subset \left(H^1\right)^\star \subset \left(H_N^2\right)^\star$ are continuous, dense and compact. We have $\langle u, v \rangle_{(H^1)^\star, H^1} = (u, v) := \int_\Omega uv \, dx$ for $u \in L^2(\Omega)$, $v \in H^1(\Omega)$. Similarly, $\langle u, v \rangle_{(H_N^2)^\star, H_N^2} = (u, v)$ for $u \in L^2(\Omega)$, $v \in H_N^2$.

The Neumann-Laplace operator $-\Delta_N : H^1(\Omega) \cap L_0^2(\Omega) \to (H^1)_0^\star$ is defined by

$$\langle -\Delta_N u, v \rangle_{(H^1)^\star, H^1} = \int_\Omega \nabla u \cdot \nabla v \, dx, \quad u, v \in H^1(\Omega).$$

The Neumann-Laplace operator is positive and self-adjoint. By the Lax-Milgram theorem and the Poincaré inequality, the inverse operator $(-\Delta_N)^{-1} : (H^1)_0^\star \to H^1(\Omega) \cap L_0^2$ is compact, positive and symmetric. By the spectral theorem, $(-\Delta_N)^{-1}$ admits a sequence of eigenfunctions $\{w_l\}_{l=1}^\infty$ that forms a complete orthonormal basis in L_0^2. The eigenfunctions of $-\Delta_N$ is $e_1 := 1/|\Omega|^{\frac{1}{2}}$ and $e_l := w_{l-1}$ for $l \geq 2$. The sequence $\{e_l\}_{l=1}^\infty$ is an orthonormal basis of $L^2(\Omega)$. The L^2 orthogonal projection is denoted by

$$\Pi_n : L^2(\Omega) \to \mathbb{X}_n = \mathrm{Span}\{e_1, \ldots, e_n\}, \quad \Pi_n u := \sum_{l=1}^n (u, e_l) \, e_l. \tag{4.1}$$

Then $\Pi_n u \to u$ in $L^2(\Omega)$ as $n \to \infty$ and $\|\Pi_n u\|_{L^2(\Omega)} \leq \|u\|_{L^2(\Omega)}$.

Denoting the corresponding eigenvalues by $\{\lambda_l\}_{l=1}^\infty$, we have

$$-\Delta e_l = \lambda_l e_l \text{ in } \Omega, \qquad \frac{\partial e_l}{\partial \nu} = 0 \text{ on } \partial\Omega, \tag{4.2}$$

for each $l \in \mathbb{N}$. The eigenvalues form a nondecreasing sequence with $\lambda_1 = 0$ and $\lambda_l \to \infty$ as $l \to \infty$. By elliptic regularity theory, each eigenfunction e_l belongs to $H_N^2 \subset L^\infty(\Omega)$, $e_l \in C^\infty(\Omega)$, and e_l is as smooth in $\overline{\Omega}$ as $\partial\Omega$ deems possible (e.g. $e_l \in C^\infty$ if $\partial\Omega$ is C^∞). By [19, Lemma 3.1], the space H_N^2 is dense in $H^1(\Omega)$ and in $W^{1,5}(\Omega)$. The same proof applies to $W^{1,p}(\Omega)$ for any $p \in [1, 6]$. It is further known that $\{e_l\}_{l=1}^\infty$ forms a basis of H_N^2. Indeed, for any $u \in H_N^2$,

$$\begin{aligned} \Delta \Pi_n u &= \sum_{l=1}^n (u, e_l) \, \Delta e_l = \sum_{l=1}^n (u, -\lambda e_l) \, e_l \\ &= \sum_{l=1}^n (u, \Delta e_l) \, e_l = \sum_{l=1}^n (\Delta u, e_l) \, e_l = \Pi_n \Delta u. \end{aligned}$$

As a result, $\Delta \Pi_n u$ converges in L^2 to Δu as $n \to \infty$. We can therefore conclude that $\Pi_n u \to u$ in H_N^2. Hence, the sequence $\{e_l\}_{l=1}^\infty$ forms a basis of H_N^2. Later we will make use of the estimate

$$\|\Pi_n u\|_{H^2_N} \le C \|u\|_{H^2_N},$$

for a constant C that is independent of n.

From the weak form of (4.2) with test function $v = e_m$,

$$\int_\Omega \nabla e_l \cdot \nabla e_m \, dx = \lambda_l \int_\Omega e_l e_m \, dx = \lambda_l \delta_{lm}, \quad \forall l, m \in \mathbb{N};$$

thus $(u,v)_{H^1(\Omega)} = (1+\lambda_l)\delta_{lm}$ and $\|e_l\|_{H^1(\Omega)} = \sqrt{(u,u)_{H^1(\Omega)}} = (1+\lambda_l)^{\frac{1}{2}}$, i.e., $\{e_l\}_{l=1}^\infty$ is an orthonormal basis of $L^2(\Omega)$ that is orthogonal in $H^1(\Omega)$. Set $\tilde{e}_l := e_l/(1+\lambda_l)^{\frac{1}{2}}$. Then $\{\tilde{e}_l\}_{l=1}^\infty$ forms an orthonormal basis of $H^1(\Omega)$. To see this, note that $\{\tilde{e}_l\}_{l=1}^\infty$ is clearly an orthonormal sequence in $H^1(\Omega)$. To prove that it is a basis, it suffices to establish that $(u, \tilde{e}_l)_{H^1(\Omega)} = 0 \; \forall l$ implies $u = 0$, for any $u \in H^1(\Omega)$. Suppose $(u, \tilde{e}_l)_{H^1(\Omega)} = 0 \; \forall l$. From integration by parts and (4.2),

$$0 = \int_\Omega \nabla u \cdot \nabla \tilde{e}_l \, dx + \int_\Omega u \tilde{e}_l \, dx = (1+\lambda_l)^{\frac{1}{2}} \int_\Omega u e_l \, dx,$$

so that $(u, e_l) = 0 \; \forall l$. Since $\{e_l\}_{l=1}^\infty$ is a basis of $L^2(\Omega)$, this implies that $u = 0$.

Let us note that the restriction of Π_n to $H^1(\Omega)$ coincides with $\tilde{\Pi}_n$, the H^1 orthogonal projection onto the space $\mathrm{Span}\{\tilde{e}_1, \dots, \tilde{e}_n\}$: for any $u \in H^1(\Omega)$,

$$\tilde{\Pi}_n u = \sum_{l=1}^n (u, \tilde{e}_l)_{H^1(\Omega)} \tilde{e}_l = \sum_{l=1}^n (1+\lambda_l)^{\frac{1}{2}} (u, e_l) \tilde{e}_l = \sum_{l=1}^n (u, e_l) e_l = \Pi_n u.$$

Consequently,

$$\Pi_n u \overset{n\uparrow\infty}{\longrightarrow} u \text{ in } H^1(\Omega), \quad \|\Pi_n u\|_{H^1(\Omega)} \le \|u\|_{H^1(\Omega)}.$$

Finally, we will continue to use the symbol Π_n for the operator

$$\Pi_n : X^\star \to \mathrm{Span}\{e_1, \dots, e_n\}, \quad \Pi_n u := \sum_{l=1}^n \langle u, e_l \rangle_{X^\star, X} e_l,$$

where $X = H^1(\Omega)$ or $X = H^2_N$. The restriction of this operator to $L^2(\Omega)$ coincides with the L^2 orthogonal projection defined before (4.1). It is easy to verify that

$$(\Pi_n u, v) = \langle u, \Pi_n v \rangle_{X^\star, X}, \quad u \in X^\star, \; v \in X,$$

as $\left(\sum_{l=1}^n \langle u, e_l \rangle_{X^\star,X} e_l, v\right) = \sum_{l=1}^n \langle u, e_l \rangle_{X^\star,X} (e_l, v) = \left\langle u, \sum_{l=1}^n (v, e_l) e_l \right\rangle_{X^\star,X}$.

We can now define our Faedo-Galerkin approximations

$$u^n, v^n : [0,T] \to \mathbb{X}_n, \quad u^n(t) = \sum_{l=1}^n c_l^n(t) e_l, \quad v^n(t) = \sum_{l=1}^n d_l^n(t) e_l, \tag{4.3}$$

where the coefficients $c^n = \{c_l^n(t)\}_{l=1}^n$ and $d^n = \{d_l^n\}_{l=1}^n$ are determined such that the following equations hold (for $l = 1, \dots, n$):

$$
\begin{aligned}
&(du^n, e_l) + D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\,(\nabla u^n, \nabla e_l)\;dt \\
&\qquad\qquad + \big(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(u^n,v^n)\nabla v^n, \nabla e_l\big)\,dt \\
&\qquad = (F(u^n,v^n), e_l)\;dt + \sum_{k=1}^n \big(\sigma_{u,k}^n(u^n), e_l\big)\;dW_{u,k}(t), \\
&(dv^n, e_l) + D_v\Big(\int_\Omega v^n(t,x)\,dx\Big)\,(\nabla v^n, \nabla e_l)\;dt \\
&\qquad\qquad + \big(\mathcal{A}_{21}(u^n,v^n)\nabla u^n + \mathcal{A}_{22}(u^n,v^n)\nabla v^n, \nabla e_l\big)\,dt \\
&\qquad = (G(u^n,v^n), e_l)\;dt + \sum_{k=1}^n \big(\sigma_{v,k}^n(v^n), e_l\big)\;dW_{v,k}(t),
\end{aligned}
\tag{4.4}
$$

and, with reference to the initial data,

$$
\begin{aligned}
u^n(0) = u_0^n &:= \sum_{l=1}^n c_l^n(0) e_l, \quad c_l^n(0) := (u_0^n, e_l)_{L^2(\Omega)}\,, \\
v^n(0) = v_0^n &:= \sum_{l=1}^n d_l^n(0) e_l, \quad d_l^n(0) := (v_0, e_l)_{L^2(\Omega)}\,.
\end{aligned}
\tag{4.5}
$$

In (4.4) we have used the following approximations of the noise coefficients:

$$
\begin{aligned}
&\sigma_{w,k}^n(w^n) := \sum_{l=1}^n \sigma_{w,k,l}(w^n) e_l, \quad \text{where} \\
&\sigma_{w,k,l}(w^n) := (\sigma_{w,k}(w^n), e_l)_{L^2(\Omega)}\,, \quad w = u, v.
\end{aligned}
\tag{4.6}
$$

Using the Faedo-Galerkin equations (4.4), the regularity $u^n(t), v^n(t) \in H_N^2 \subset L^\infty$, and basic properties of the projection operator Π_n, we obtain

$$
\begin{aligned}
&u^n(t) - u_0^n \\
&\quad - \int_0^t \Pi_n\left[\nabla\cdot\left(D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\nabla u^n\right)\right]\,ds \\
&\qquad - \int_0^t \Pi_n\left[\nabla\cdot\left(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(u^n,v^n)\nabla v^n\right)\right]\,ds \\
&\qquad\quad = \int_0^t \Pi_n\left[F(u^n,v^n)\right]\,ds + \int_0^t \sigma_u^n(u^n)\,dW_u^n(s) \quad \text{in } L^2(\Omega), \\
&v^n(t) - v_0^n \\
&\quad - \int_0^t \Pi_n\left[\nabla\cdot\left(D_v\Big(\int_\Omega v^n(t,x)\,dx\Big)\nabla v^n\right)\right]\,ds \\
&\qquad - \int_0^t \Pi_n\left[\nabla\cdot\left(\mathcal{A}_{21}(u^n,v^n)\nabla u^n + \mathcal{A}_{22}(u^n,v^n)\nabla v^n\right)\right]\,ds \\
&\qquad\quad = \int_0^t \Pi_n\left[G(u^n,v^n)\right]\,ds + \int_0^t \sigma_v^n(v^n)\,dW_v^n(s) \quad \text{in } L^2(\Omega),
\end{aligned}
\tag{4.7}
$$

with initial data $u_0^n = \Pi_n u_0$ and $v_0^n = \Pi_n v_0$, where $\sigma_w^n(w^n)\,dW_w^n$ is short-hand notation for $\sum_{k=1}^n \sigma_{w,k}^n(w^n)\,dW_{w,k}$, $w = u, v$. The formulation (4.7) allows us to treat u^n, v^n as stochastic processes in $\mathbb{R}^n$, so that one can apply the finite dimensional Itô formula to the Faedo-Galerkin equations.

Remark 4.1. *Our construction of approximate solutions makes use of Neumann boundary conditions, which are encoded in the space H^2_N. The zero-flux boundary conditions (1.3) are recovered when we pass to the limit to identify the weak martingale solution.*

The existence of pathwise solutions to the finite-dimensional problem (4.4), (4.5) is guaranteed by the next lemma.

Lemma 4.2. *For each $n \in \mathbb{N}$, the Faedo-Galerkin equations (4.3), (4.4), (4.5) possess a unique adapted solution $(u^n(t), v^n(t))$ on $[0,T]$. Furthermore, $u^n, v^n \in C([0,T]; \mathbb{X}_n)$ a.s., where $\mathbb{X}_n$ is defined in (4.1), and $\mathbb{E}\big[\|w^n(t)\|^2_{L^2(\Omega)}\big] \lesssim_{T,n} 1$, $\forall t \in [0,T]$, $w = u, v$.*

Proof. We look for a stochastic process C^n taking values in $\mathbb{X}_n \times \mathbb{X}_n$ that is a solution to the following system of stochastic differential equations:

$$dC^n = M(C^n)\,dt + \Gamma(C^n)\,dW^n, \tag{4.8}$$

where $C^n = \begin{pmatrix} u^n \\ v^n \end{pmatrix}$, $M(C^n) = \begin{pmatrix} A_u\,(C^n) \\ A_v\,(C^n) \end{pmatrix}$, and

$$\begin{aligned} A_u\,(C^n) &= -\Pi_n \nabla \cdot \left(D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\nabla u^n \right) \\ &\quad - \Pi_n \nabla \cdot \Big(\mathcal{A}_{11}(u^n, v^n)\nabla u^n + \mathcal{A}_{12}(u^n, v^n)\nabla v^n\Big) + \Pi_n F(u^n, v^n), \\ A_v\,(C^n) &= -\Pi_n \nabla \cdot \left(D_v\Big(\int_\Omega v^n(t,x)\,dx\Big)\nabla v^n \right) \\ &\quad - \Pi_n \nabla \cdot \Big(\mathcal{A}_{21}(u^n, v^n)\nabla u^n + \mathcal{A}_{22}(u^n, v^n)\nabla v^n\Big) + \Pi_n G(u^n, v^n). \end{aligned}$$

Moreover, $\Gamma(C^n)\,dW^n$ is short-hand notation for $\begin{pmatrix} \sigma^n_u\,(u^n)\; dW^n_u \\ \sigma^n_v\,(v^n)\; dW^n_v \end{pmatrix}$. We complete (4.8) with initial data $C^n(0) = C^n_0$, where C^n_0 is the vector defined by (4.5).

To prove the existence and uniqueness of a pathwise solution to (4.8), we will use [33, Theorem 3.1.1] (see also Theorem 5.1.3 in [33]), which asks that M and Γ satisfy the following conditions:

(i) — *local weak monotonicity*. For all $C_1 = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$ and $C_2 = \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$ with $u_i, v_i \in \mathbb{X}_n$ such that $\|u^n_i\|_{L^2(\Omega)}, \|v^n_i\|_{L^2(\Omega)} \le r$, for any $r > 0$ and $i = 1, 2$, we have

$$\begin{aligned} &2\big(M(C_1) - M(C_2), C_1 - C_2\big) + \|\Gamma(C_1) - \Gamma(C_2)\|^2_{L^2(\Omega)} \\ &\qquad \le K(r)\,\|C_1 - C_2\|^2_{L^2(\Omega)}, \end{aligned} \tag{4.9}$$

for a constant $K(r)$ that may depend on r, where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ inner product.

(ii) — *weak coercivity*. For all $C = \begin{pmatrix} u \\ v \end{pmatrix}$ with $u, v \in \mathbb{X}_n$,

$$2\big(M(C), C\big) + \|\Gamma(C)\|^2_{L^2(\Omega)} \le K\left(1 + \|C\|^2_{L^2(\Omega)}\right), \tag{4.10}$$

for some constant $K > 0$.

The weak coercivity condition (4.10) is easily verified using the assumption (1.8) and the global Lipschitz continuity of F, G, Γ.

Let us verify the weak monotonicity condition (4.9) in some detail. Fix a real number $r > 0$ and set $\overline{u} := u_1 - u_2$ and $\overline{v} := v_1 - v_2$, where u_i, v_i are arbitrary functions in $\mathbb{X}_n$ for which $\|u_i\|_{L^2(\Omega)}, \|v_i\|_{L^2(\Omega)} \le r$ for $i = 1, 2$. In view of (1.8) and Young's inequality,

$$\big(M(C_1) - M(C_2), C_1 - C_2\big) + \|\Gamma(C_1) - \Gamma(C_2)\|_{L^2(\Omega)}^2 = \sum_{i=0}^{6} I_i, \tag{4.11}$$

where $I_0 = \|\Gamma(C_1) - \Gamma(C_2)\|_{L^2(\Omega)}^2 \overset{(2.4)}{\lesssim} \|C_1 - C_2\|_{L^2(\Omega)}^2$ and

$$\begin{aligned}
I_1 &= -\sum_{w=u,v} D_w\Big(\int_\Omega w_1\,dx\Big)\big(\nabla\overline{w}, \nabla\overline{w}\big),\\
I_2 &= -\sum_{w=u,v} \left(D_w\Big(\int_\Omega w_1\,dx\Big) - D_w\Big(\int_\Omega w_2\,dx\Big)\right)\big(\nabla w_2, \nabla\overline{w}\big),\\
I_3 &= -\left(\mathcal{A}(u_1, v_1)\begin{pmatrix}\nabla\overline{u}\\ \nabla\overline{v}\end{pmatrix}, \begin{pmatrix}\nabla\overline{u}\\ \nabla\overline{v}\end{pmatrix}\right),\\
I_4 &= -\left(\Big(\mathcal{A}(u_1, v_1) - \mathcal{A}(u_2, v_2)\Big)\begin{pmatrix}\nabla u_2\\ \nabla v_2\end{pmatrix}, \begin{pmatrix}\nabla\overline{u}\\ \nabla\overline{v}\end{pmatrix}\right),\\
I_5 &= \big(F(u_1, v_1) - F(u_2, v_2), \overline{u}\big), \quad I_6 = \big(G(u_1, v_1) - G(u_2, v_2), \overline{v}\big).
\end{aligned}$$

Recall that the basis functions e_l belong to H_N^2 and $H_N^2 \subset W^{1,p}(\Omega) \cap L^\infty(\Omega)$, for any $p \in [1, 6]$ (as $\Omega \subset \mathbb{R}^3$ is bounded). Hence, the assumption $\|w_i\|_{L^2(\Omega)} \le r$ implies $\|w_i\|_{H_N^2} \lesssim_{r,n} 1$, for $w = u, v$ and $i = 1, 2$. In view of (1.7),

$$|I_2| \lesssim \sum_{w=u,v} \|w_1 - w_2\|_{L^1(\Omega)} \|\nabla w_2\|_{L^2(\Omega)} \|\nabla\overline{w}\|_{L^2(\Omega)},$$

and so $|I_2| \lesssim_{r,n} \sum_{w=u,v} \|w_1 - w_2\|_{L^2(\Omega)}$. Similarly, given the assumption (1.8),

$$\begin{aligned}
|I_4| &\lesssim \sum_{w=u,v} \|w_1 - w_2\|_{L^2(\Omega)} \sum_{w=u,v} \|\nabla w_2\|_{L^4(\Omega)} \sum_{w^n=u,v} \|\nabla\overline{w}\|_{L^4(\Omega)}\\
&\lesssim \sum_{w=u,v} \|w_1 - w_2\|_{L^2(\Omega)} \sum_{w=u,v} \|\nabla w_2^n\|_{H_N^2} \sum_{w=u,v} \|\nabla\overline{w}\|_{H_N^2},
\end{aligned}$$

and so $|I_4| \lesssim_{r,n} \sum_{w=u,v} \|w_1 - w_2\|_{L^2(\Omega)}$. In view of the global Lipschitz continuity of the reaction functions F and G, cf. (1.4), it follows that

$$|I_5| + |I_6| \lesssim \sum_{w=u,v} \|w_1 - w_2\|_{L^2(\Omega)} \sum_{w=u,v} \|w^n\|_{L^2(\Omega)},$$

so that $|I_5| + |I_6| \lesssim_r \sum_{w=u,v} \|w_1 - w_2\|_{L^2(\Omega)}$. Finally, by (1.7) and (1.8), $I_1, I_3 \le 0$.

Referring to (4.11), this implies $\sum_{i=0}^{6} I_i \lesssim_{r,n} \|C_1^n - C_2^n\|_{L^2(\Omega)}^2$, and (4.9) thus holds. □

5. Basic a priori estimates

We start with a series of basic energy-type estimates.

Lemma 5.1. *Let $u^n(t), v^n(t)$, $t \in [0, T]$, satisfy (4.4), (4.5). There is a constant $C > 0$, independent of n, such that*

$$\mathbb{E}\Big[\|u^n(t)\|_{L^2(\Omega)}^2\Big] + \mathbb{E}\Big[\|v^n(t)\|_{L^2(\Omega)}^2\Big] \le C, \qquad \forall t \in [0, T]; \tag{5.1}$$

$$\mathbb{E}\left[\int_0^T \int_\Omega |\nabla u^n|^2 \, dx\, dt\right] + \mathbb{E}\left[\int_0^T \int_\Omega |\nabla v^n|^2 \, dx\, dt\right] \le C; \tag{5.2}$$

$$\mathbb{E}\left[\int_0^T \int_\Omega |\mathcal{A}_{ij}(u^n, v^n)| \left(|\nabla u^n|^2 + |\nabla v^n|^2\right) dx\, dt\right] \le C, \quad i, j = 1, 2; \tag{5.3}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|u^n(t)\|^2_{L^2(\Omega)}\right] + \mathbb{E}\left[\sup_{t\in[0,T]} \|v^n(t)\|^2_{L^2(\Omega)}\right] \le C. \tag{5.4}$$

Proof. By Itô's formula, $dS(w^n) = S'(w^n)\, dw^n + \frac{1}{2} S''(w^n) \sum_{k=1}^n (\sigma_{w,k}(w^n))^2 \, dt$, $w = u, v$, for any C^2 function $S : \mathbb{R} \to \mathbb{R}$. Hence, with $S(w) = \frac{1}{2} |w|^2$,

$$\begin{aligned}
&\frac{1}{2} \sum_{w=u,v} \|w^n(t)\|^2_{L^2(\Omega)} + \sum_{w=u,v} \int_0^t D_w\Big(\int_\Omega w^n(t,x)\, dx\Big) \int_\Omega |\nabla w^n|^2 \, dx\, ds \\
&\qquad + \int_0^t \Big(\mathcal{A}_{11}(u^n, v^n)\nabla u^n + \mathcal{A}_{12}(u^n, v^n)\nabla v^n, \nabla u^n\Big)_{L^2(\Omega)} ds \\
&\qquad + \int_0^t \Big(\mathcal{A}_{21}(u^n, v^n)\nabla u^n + \mathcal{A}_{22}(u^n, v^n)\nabla v^n, \nabla v^n\Big)_{L^2(\Omega)} ds \\
&\quad = \frac{1}{2} \sum_{w=u,v} \|w^n(0)\|^2_{L^2(\Omega)} + \int_0^t \Big(F(u^n, v^n), u^n\Big)_{L^2(\Omega)} ds \\
&\qquad + \int_0^t \Big(G(u^n, v^n), v^n\Big)_{L^2(\Omega)} ds \\
&\qquad + \sum_{w=u,v} \sum_{k=1}^n \int_0^t \int_\Omega w^n \sigma^n_{w,k}(w^n)\, dx\, dW_{w,k} \\
&\qquad + \frac{1}{2} \sum_{w=u,v} \sum_{k=1}^n \int_0^t \int_\Omega \left(\sigma^n_{w,k}(w^n)\right)^2 \, dx\, ds \\
&\quad \le \frac{1}{2} \sum_{w=u,v} \|w^n(0)\|^2_{L^2(\Omega)} + C \int_0^t \Big(1 + \|u^n(t)\|^2_{L^2(\Omega)} + \|v^n(t)\|^2_{L^2(\Omega)}\Big) ds \\
&\qquad + \sum_{w=u,v} \sum_{k=1}^n \int_0^t \int_\Omega w^n \sigma^n_{w,k}(w^n)\, dx\, dW_{w,k}(s),
\end{aligned} \tag{5.5}$$

where have put to good use (1.4), (1.5), and also (1.6). By the fundamental assumption (1.8), the sum of the $\mathcal{A}_{ij}$ terms is lower bounded by $|\mathcal{A}(u^n,v^n)|\left(|\nabla u^n|^2+|\nabla v^n|^2\right)$, so

$$
\begin{aligned}
&\sum_{w=u,v}\|w^n(t)\|_{L^2(\Omega)}^2+\sum_{w=u,v}C_m\int_0^t\int_\Omega|\nabla w^n|^2\,dx\,ds\\
&\qquad+\sum_{w=u,v}\int_0^t\int_\Omega|\mathcal{A}(u^n,v^n)|\,|\nabla w^n|^2\,dx\,ds\\
&\qquad\quad\le\sum_{w=u,v}\|w^n(0)\|_{L^2(\Omega)}^2+C\int_0^t\left(1+\sum_{w=u,v}\|w^n(t)\|_{L^2(\Omega)}^2\right)ds\\
&\qquad\qquad+\sum_{w=u,v}\sum_{k=1}^n\int_0^t\int_\Omega w^n\sigma_{w,k}^n(w^n)\,dx\,dW_{w,k}(s).
\end{aligned}\tag{5.6}
$$

where we have also used (1.7). Applying $\mathbb{E}[\cdot]$ to (5.6) and using the Gronwall inequality, we arrive at (5.1), (5.2), and (5.3), recalling that the initial data u_0,v_0 belong to L^2.

To prove the final estimate (5.4), we take $\sup_{t\in[0,T]}$ and then $\mathbb{E}[\cdot]$ in (5.5). Using (5.1) and the L^2 boundedness of the initial data, we end up with the estimate

$$
\sum_{w=u,v}\mathbb{E}\left[\sup_{t\in[0,T]}\|w^n(t)\|_{L^2(\Omega)}^2\right]\le C\left(1+\sum_{w=u,v}I_w\right),\tag{5.7}
$$

where $I_w:=\mathbb{E}\left[\sup_{t\in[0,T]}\left|\sum_{k=1}^n\int_0^t\int_\Omega w^n\sigma_{w,k}^n(w^n)\,dx\,dW_{w,k}(s)\right|\right]$. Using the BDG inequality (2.5), the Cauchy-Schwarz inequality, (2.3), Cauchy's inequality, and (5.1), we proceed as follows for $w=u,v$:

$$
\begin{aligned}
|I_w|&\le C\mathbb{E}\left[\left(\int_0^T\sum_{k=1}^n\left|\int_\Omega w^n\sigma_{w,k}^n(w^n)\,dx\right|^2dt\right)^{\frac12}\right]\\
&\le C\mathbb{E}\left[\left(\int_0^T\left(\int_\Omega|w^n|^2\,dx\right)\left(\sum_{k=1}^n\int_\Omega\left|\sigma_{w,k}^n(w^n)\right|^2dx\right)dt\right)^{\frac12}\right]\\
&\le C\mathbb{E}\left[\left(\sup_{t\in[0,T]}\int_\Omega|w^n|^2\,dx\right)^{\frac12}\left(\int_0^T\sum_{k=1}^n\int_\Omega\left|\sigma_{w,k}^n(w^n)\right|^2dx\,dt\right)^{\frac12}\right]\\
&\le\alpha\mathbb{E}\left[\sup_{t\in[0,T]}\int_\Omega|w^n|^2\,dx\right]+C(\alpha)\mathbb{E}\left[\int_0^T\sum_{k=1}^n\int_\Omega\left|\sigma_{w,k}^n(w^n)\right|^2dx\,dt\right]\\
&\le\alpha\mathbb{E}\left[\sup_{t\in[0,T]}\|w^n(t)\|_{L^2(\Omega)}^2\right]+C,
\end{aligned}\tag{5.8}
$$

for any number $\alpha>0$. Combining the inequalities (5.7) and (5.8), and choosing $\alpha>0$ small, we arrive at the estimate (5.4). □

Later we will need to convert a.s. convergence into L^2 convergence. To this end, the next lemma—containing improved integrability estimates—is useful.

Corollary 5.2. *Let $u^n(t),v^n(t)$, $t\in[0,T]$, satisfy (4.4), (4.5). Suppose u_0,v_0 belong to $L^q\left(D,\mathcal{F},P;L^2(\Omega)\right)$ with $q\in(2,q_0]$, cf. (3.2). Then there exists a constant $C>0$,* *independent of n, such that*

$$\mathbb{E}\left[\sup_{0\le t\le T}\|w^n(t)\|_{L^2(\Omega)}^q\right]\le C,\quad \mathbb{E}\left[\|\nabla w^n\|_{L^2((0,T)\times\Omega)}^q\right]\le C,\quad w=u,v, \tag{5.9}$$

and

$$\mathbb{E}\left[\left|\int_0^T\int_\Omega |\mathcal{A}_{ij}(u^n,v^n)|\left(|\nabla u^n|^2+|\nabla v^n|^2\right)dx\,dt\right|^{\frac{q}{2}}\right]\le C,\quad i,j=1,2. \tag{5.10}$$

Proof. Starting off from (5.5), the following estimate holds for any $(\omega,t)\in D\times[0,T]$:

$$\begin{aligned}\sum_{w=u,v}\sup_{0\le\tau\le t}\|w^n(\tau)\|_{L^2(\Omega)}^2 &\le \sum_{w=u,v}\|w^n(0)\|_{L^2(\Omega)}^2 + C\sum_{w=u,v}\int_0^t\|w^n(s)\|_{L^2(\Omega)}^2\,ds\\ &\quad + C\sum_{w=u,v}\sup_{0\le\tau\le t}\left|\sum_{k=1}^n\int_0^\tau\int_\Omega w^n\sigma_{w,k}^n(w^n)\,dx\,dW_{w,k}(s)\right|,\end{aligned}$$

for some constant C independent of n. Next, we raise both sides of this inequality to power $q/2$ and take the expectation, eventually obtaining

$$\begin{aligned}\sum_{w=u,v}\mathbb{E}\left[\sup_{0\le\tau\le t}\|u^n(\tau)\|_{L^2(\Omega)}^q\right] &\le C\sum_{w=u,v}\mathbb{E}\left[\|w^n(0)\|_{L^2(\Omega)}^q\right]+C\,(1+t)^{\frac{q}{2}}\\ &\quad + C\sum_{w=u,v}\int_0^t\|w^n(s)\|_{L^2(\Omega)}^q\,ds+\sum_{w=u,v}I_w,\end{aligned} \tag{5.11}$$

where $I_w=\mathbb{E}\left[\sup_{0\le\tau\le t}\left|\sum_{k=1}^n\int_0^\tau\int_\Omega w^n\sigma_{w,k}^n(w^n)\,dx\,dW_{w,k}(s)\right|^{\frac{q}{2}}\right]$. Relying on the martingale inequality (2.5), we proceed as in (5.8):

(5.12)

$$\begin{aligned}I_w &\le C\mathbb{E}\left[\left(\int_0^t\sum_{k=1}^n\left|\int_\Omega w^n\sigma_{w,k}^n(w^n)\,dx\right|^2ds\right)^{\frac{q}{4}}\right]\\ &\le C\mathbb{E}\left[\left(\int_0^t\left(\int_\Omega|w^n|^2\,dx\right)\left(\sum_{k=1}^n\int_\Omega\left|\sigma_{w,k}^n(w^n)\right|^2dx\right)ds\right)^{\frac{q}{4}}\right]\\ &\le C\mathbb{E}\left[\left(\sup_{\tau\in[0,t]}\int_\Omega|w^n|^2\,dx\right)^{\frac{q}{4}}\left(\int_0^t\sum_{k=1}^n\int_\Omega\left|\sigma_{k,w}^n(w^n)\right|^2dx\,ds\right)^{\frac{q}{4}}\right]\\ &\le \alpha\mathbb{E}\left[\left(\sup_{\tau\in[0,t]}\int_\Omega|w^n|^2\,dx\right)^{\frac{q}{2}}\right]+C(\alpha)\mathbb{E}\left[\left(\int_0^t\sum_{k=1}^n\int_\Omega\left|\sigma_{k,u}^n(w^n)\right|^2dx\,ds\right)^{\frac{q}{2}}\right]\\ &\le \alpha\mathbb{E}\left[\sup_{\tau\in[0,t]}\|w^n(\tau)\|_{L^2(\Omega)}^q\right]+C\mathbb{E}\left[\int_0^t\|w^n(s)\|_{L^2(\Omega)}^q\,ds\right]+C,\end{aligned}$$

for any number $\alpha>0$. Choosing α small, we conclude from (5.11), (5.12) that

$$\sum_{w=u,v}\mathbb{E}\left[\sup_{0\le\tau\le t}\|w^n(\tau)\|_{L^2(\Omega)}^q\right]$$

$$\leq C \sum_{w=u,v} \mathbb{E}\left[\|w^n(0)\|_{L^2(\Omega)}^q\right] + C \sum_{w=u,v} \int_0^t \mathbb{E}\left[\|w^n(s)\|_{L^2(\Omega)}^q \, ds\right] + C,$$

for some constant $C > 0$ independent of n. An application of Grönwall's inequality now yields the sought-after estimate (5.9).

Finally, we use (5.6), the first part of (5.12), and (5.9) to conclude that there is a constant $C > 0$, independent of n, such that

$$\sum_{w=u,v} \mathbb{E}\left[\left|\int_0^t \int_\Omega |\nabla w^n|^2 \, dx \, ds\right|^{\frac{q}{2}}\right] \leq C, \quad w = u, v,$$

and the second part of (5.9) follows. Similarly, we derive (5.10). □

6. Temporal translation estimates

Given Lemma 5.1, it is easy to see that $\mathcal{A}_{i1}(u^n, v^n)\nabla u^n$ and $\mathcal{A}_{2j}(u^n, v^n)\nabla v^n$ are uniformly bounded in L^q for some $q < 2$, for $i, j = 1, 2$. As a result, we cannot control the time translation of the approximate solution in the space $\big(H^1(\Omega)\big)^\star$. Although we expect the exact solution to be continuous in time with values in $\big(W^{1,4}(\Omega)\big)^\star$ (evident by inspecting the proof below), the fact that the sequence $\{e_l\}_{l=1}^\infty$ is not a basis of $W^{1,4}(\Omega)$—but it is for $H_N^2 \subset W^{1,4}$—we cannot control the projection operator in $\|\cdot\|_{W^{1,4}(\Omega)}$—but we can in $\|\cdot\|_{H_N^2}$. To ensure strong $L^2_{t,x}$ compactness of a sequence of Faedo-Galerkin solutions, we will therefore establish a temporal translation estimate in the larger space $\big(H_N^2\big)^\star \supset \big(W^{1,4}(\Omega)\big)^\star \supset \big(H^1(\Omega)\big)^\star$, which is enough to work out the required $L^2_{t,x}$ compactness (and tightness).

Lemma 6.1. *Extend the Faedo-Galerkin functions $u^n(t), v^n(t)$, $t \in [0, T]$, which satisfy* (4.4) *and* (4.5)*, by zero outside of $[0, T]$. There exists a constant $C = C(T, \Omega) > 0$, independent of n, such that*

$$\mathbb{E}\left[\sup_{|\tau| \in (0,\delta)} \|w^n(t+\tau) - w^n(t)\|_{(H_N^2)^\star}\right] \leq C\delta^{1/4}, \quad \forall t \in [0, T], \tag{6.1}$$

for any sufficiently small $\delta > 0$, $w = u, v$.

Proof. In what follows, we write $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_{(H_N^2)^\star, H_N^2}$. We will estimate the expected value of

$$\begin{aligned} I(t,\tau) &:= \|u^n(t+\tau,\cdot) - u^n(t,\cdot)\|_{(H_N^2)^\star} \\ &= \sup\left\{\left|\langle u^n(t+\tau,\cdot) - u^n(t,\cdot), \phi\rangle\right| \,:\, \phi \in H_N^2,\, \|\phi\|_{H_N^2} \leq 1\right\} \\ &= \sup\left\{\int_\Omega \big(u^n(t+\tau,x) - u^n(t,x)\big)\phi(x)\,dx \,:\, \phi \in H_N^2,\, \|\phi\|_{H_N^2} \leq 1\right\}, \end{aligned}$$

for $\tau \in (0, \delta)$, $\delta > 0$. The same estimate can derived for $\tau \in (-\delta, 0)$.

By (4.3),

$$I(t,\tau) := \|u^n(t+\tau,\cdot) - u^n(t,\cdot)\|_{(H_N^2)^\star} \leq \sum_{i=1}^4 I_i(t,\tau),$$

where

$$I_1(t,\tau) = \left\|\int_t^{t+\tau} \Pi_n\left[\nabla \cdot \left(D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\nabla u^n\right)\right] ds\right\|_{(H_N^2)^\star},$$

$$I_2(t,\tau) = \left\| \int_t^{t+\tau} \Pi_n \left[\nabla \cdot \left(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(u^n,v^n)\nabla v^n \right) \right] \, ds \right\|_{\left(H^2_N\right)^\star},$$
$$I_3(t,\tau) = \left\| \int_t^{t+\tau} \Pi_n \left[F(u^n,v^n) \right] \, ds \right\|_{\left(H^2_N\right)^\star},$$
$$I_4(t,\tau) = \left\| \sum_{k=1}^n \int_t^{t+\tau} \sigma^n_{u,k}(u^n) \, dW_{u,k}(s) \right\|_{\left(H^2_N\right)^\star}.$$

Estimate of I_2. Setting $L^n_{2,u} := \Pi_n \left[\nabla \cdot \left(\mathcal{A}_{11}(u^n,v^n)\nabla u^n \right) \right]$, let us estimate

$$\begin{aligned}
&\left\| \int_t^{t+\tau} L^n_{2,u} \, ds \right\|_{\left(H^2_N\right)^\star} \\
&\quad = \sup \left\{ \left| \left\langle \int_t^{t+\tau} L^n_{2,u} \, ds, \phi \right\rangle \right| \, : \, \phi \in H^2_N, \, \|\phi\|_{H^2_N} \le 1 \right\} \\
&\quad = \sup \left\{ \left| \int_t^{t+\tau} \int_\Omega L^n_{2,u} \phi \, dx \, ds \right| \, : \, \phi \in H^2_N, \, \|\phi\|_{H^2_N} \le 1 \right\} \\
&\quad = \sup \left\{ \left| \int_t^{t+\tau} \int_\Omega \mathcal{A}_{11}(u^n,v^n)\nabla u^n \cdot \nabla \Pi_n \phi \, dx \, ds \right| \, : \, \phi \in H^2_N, \, \|\phi\|_{H^2_N} \le 1 \right\}
\end{aligned}$$

by bounding the term

$$I := \left| \int_t^{t+\tau} \int_\Omega \mathcal{A}_{11}(u^n,v^n)\nabla u^n \cdot \nabla \Pi_n \phi \, dx \, ds \right|.$$

By the generalised Hölder inequality,

$$\begin{aligned}
I \le \tau^{1/4} &\left\| \sqrt{|\mathcal{A}_{11}(u^n,v^n)|} \right\|_{L^4((0,T)\times\Omega)} \\
&\times \left\| \sqrt{|\mathcal{A}_{11}(u^n,v^n)|} \, |\nabla u^n| \right\|_{L^2((0,T)\times\Omega)} \|\nabla \Pi_n \phi\|_{L^4(\Omega)}.
\end{aligned}$$

Now we use that H^2_N is continuously embedded in $W^{1,p}(\Omega)$ $\forall p \in [1,6]$ (recalling that $\Omega \subset \mathbb{R}^3$ bounded), so

$$\|\nabla \Pi_n \phi\|_{L^4(\Omega)} \le \|\Pi_n \phi\|_{W^{1,4}(\Omega)} \lesssim \|\Pi_n \phi\|_{H^2_N}.$$

As $\{e_l\}_{l=1}^\infty$ is a basis of H^2_N, $\|\Pi_n \phi\|_{H^2_N} \lesssim \|\phi\|_{H^2_N}$ and thus $\|\nabla \Pi_n \phi\|_{L^4(\Omega)} \lesssim \|\phi\|_{H^2_N}$. Using this bound and Young's product inequality,

$$\begin{aligned}
I \lesssim \tau^{1/4} \Bigg(&\left\| \sqrt{|\mathcal{A}_{11}(u^n,v^n)|} \right\|^2_{L^4((0,T)\times\Omega)} \\
&+ \left\| \sqrt{|\mathcal{A}_{11}(u^n,v^n)|} \, |\nabla u^n| \right\|^2_{L^2((0,T)\times\Omega)} \Bigg) \|\phi\|_{H^2_N},
\end{aligned}$$

Note that

$$\begin{aligned}
&\left\| \sqrt{|\mathcal{A}_{11}(u^n,v^n)|} \right\|^2_{L^4((0,T)\times\Omega)} \\
&\quad = \|\mathcal{A}_{11}(u^n,v^n)\|_{L^2((0,T)\times\Omega)} \lesssim \|1 + u^n + v^n\|_{L^2((0,T)\times\Omega)} \\
&\quad \lesssim_{T,\Omega} 1 + \|u^n\|_{L^\infty(0,T;L^2(\Omega))} + \|v^n\|_{L^\infty(0,T;L^2(\Omega))}.
\end{aligned}$$

Consequently, after taking the expectation and using (5.3) and (5.4),

$$\mathbb{E}\left[I\right] \lesssim_{T,\Omega} \tau^{1/4} \left\|\phi\right\|_{H^2_N}.$$

Summarising,

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)} \sup\left\{\left|\left\langle\int_t^{t+\tau} L^n_{2,u}\,ds,\phi\right\rangle\right| \,:\, \phi\in H^2_N,\, \|\phi\|_{H^2_N}\le 1\right\}\right] \lesssim \delta^{\frac14},$$

i.e.,

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)} \left\|\int_t^{t+\tau} L^n_{u,2}\,ds\right\|_{\left(H^2_N\right)^\star}\right] \lesssim \delta^{\frac14}.$$

A similar estimate holds for $L^n_{2,v} := \Pi_n\left[\nabla\cdot\left(\mathcal{A}_{12}(u^n,v^n)\nabla v^n\right)\right]$, and therefore

$$\mathbb{E}\left[\sup_{0\le\tau\le\delta} I_2(t,\tau)\right] \lesssim \delta^{1/4}, \quad \text{uniformly in } t\in[0,T].$$

Estimate of I_1. Set $L^n_1 := \Pi_n\left[\nabla\cdot\left(D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\nabla u^n\right)\right]$. Given (1.7),

$$\left|D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\right|^2 \lesssim 1 + \left(\int_\Omega |u^n(t,x)|\,dx\right)^2 \lesssim_\Omega 1 + \|u^n\|^2_{L^\infty(0,T;L^2(\Omega))}.$$

Using this, we bound

$$\left|\left\langle\int_t^{t+\tau} L^n_1\,ds,\phi\right\rangle\right| = \left|\int_t^{t+\tau}\int_\Omega D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\nabla u^n\cdot\nabla\Pi_n\phi\,dx\,ds\right|$$

by a constant times

$$\tau^{1/2}\left(\int_0^T\int_\Omega\left(1+\|u^n\|^2_{L^\infty(0,T;L^2(\Omega))}\right)|\nabla u^n|^2\,dx\,ds\right)^{\frac12}\|\nabla\Pi_n\phi\|_{L^2(\Omega)}$$
$$\lesssim \tau^{1/2}\left(1+\|u^n\|_{L^\infty(0,T;L^2(\Omega))}\right)\|\nabla u^n\|_{L^2((0,T)\times\Omega)}\|\Pi_n\phi\|_{H^1(\Omega)}.$$

Recalling that the sequence $\{e_l\}_{l=1}^\infty$ is an orthogonal basis of $H^1(\Omega)$, we have

$$\|\Pi_n\phi\|_{H^1(\Omega)} \le \|\phi\|_{H^1(\Omega} \le \|\phi\|_{H^2_N}.$$

Taking the expectation and using Young's inequality,

$$\mathbb{E}\left[\left(1+\|u^n\|_{L^\infty(0,T;L^2(\Omega))}\right)\|\nabla u^n\|_{L^2(\Omega)}\right]$$
$$\lesssim 1+\mathbb{E}\left[\|u^n\|^2_{L^\infty(0,T;L^2(\Omega))}\right]+\mathbb{E}\left[\|\nabla u^n\|^2_{L^\infty(0,T;L^2(\Omega))}\right] \overset{(5.2),(5.4)}{\lesssim} 1,$$

and thus we conclude that

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)} \sup\left\{\left|\left\langle\int_t^{t+\tau} L^n_1\,ds,\phi\right\rangle\right| \,:\, \phi\in H^2_N,\, \|\phi\|_{H^2_N}\le 1\right\}\right] \lesssim \delta^{1/2},$$

i.e.,

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)} I_1(t,\tau)\right] \lesssim \delta^{1/2}, \quad \text{uniformly in } t\in[0,T].$$

Estimate of I_3. Set $L_3^n := \Pi_n\left[F(u^n, v^n)\right]$. The function F is linearly growing in both its arguments, which follows from (1.4), (1.5) and (1.6). Using this, we bound

$$\left|\left\langle \int_t^{t+\tau} L_3^n\, ds, \phi\right\rangle\right| = \left|\int_t^{t+\tau}\int_\Omega F(u^n, v^n)\Pi_n\phi\, dx\, ds\right|$$

by a constant times

$$\begin{aligned} &\tau^{1/2}\left\|1+u^n+v^n\right\|_{L^2((0,T)\times\Omega)}\left\|\Pi_n\phi\right\|_{L^2(\Omega)} \\ &\quad\lesssim \tau^{1/2}\left(1+\left\|u^n\right\|_{L^2((0,T)\times\Omega)}^2+\left\|u^n\right\|_{L^2((0,T)\times\Omega)}^2\right)\left\|\phi\right\|_{H_N^2}, \end{aligned}$$

where we have used Young's inequality and that the sequence $\{e_l\}_{l=1}^\infty$ is an orthonormal basis of $L^2(\Omega)$, so that $\left\|\Pi_n\phi\right\|_{L^2(\Omega)} \le \left\|\phi\right\|_{L^2(\Omega)} \le \left\|\phi\right\|_{H_N^2}$. Hence

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)}\sup\left\{\left|\left\langle\int_t^{t+\tau} L_3^n\, ds, \phi\right\rangle\right| \,:\, \phi\in H_N^2,\ \left\|\phi\right\|_{H_N^2}\le 1\right\}\right] \lesssim \delta^{1/2},$$

i.e.,

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)} I_3(t,\tau)\right] \lesssim \delta^{1/2}, \quad \text{uniformly in } t\in[0,T].$$

Estimate of I_4. Set $L_4^n := \sum_{k=1}^n\int_t^{t+\tau}\sigma_{u,k}^n(u^n)\, dW_{u,k}(s)$. We bound

$$\left|\left\langle\int_t^{t+\tau} L_3^n\, ds, \phi\right\rangle\right| = \left|\int_\Omega\sum_{k=1}^n\int_t^{t+\tau}\sigma_{u,k}^n(u^n)\, dW_{u,k}(s)\,\phi\, dx\right|$$

by a constant times

$$\left\|\sum_{k=1}^n\int_t^{t+\tau}\sigma_{u,k}^n(u^n)\, dW_{u,k}(s)\right\|_{L^2(\Omega)}\left\|\phi\right\|_{L^2(\Omega)}$$

where $\left\|\phi\right\|_{L^2(\Omega)} \le \left\|\phi\right\|_{H_N^2}$. By the Burkholder-Davis-Gundy inequality (2.5),

$$\begin{aligned} &\mathbb{E}\left[\sup_{\tau\in(0,\delta)}\left\|\sum_{k=1}^n\int_t^{t+\tau}\sigma_{u,k}^n(u^n)\, dW_{u,k}(s)\right\|_{L^2(\Omega)}\right] \\ &\qquad\lesssim \mathbb{E}\left[\sum_{k=1}^n\int_t^{t+\delta}\int_\Omega\left(\sigma_{u,k}^n(u^n)\right)^2\, dx\, ds\right]^{\frac12} \\ &\qquad\overset{(2.3)}{\lesssim_\Omega} \delta^{1/2}\left(1+\mathbb{E}\left[\left\|u^n\right\|_{L^\infty(0,T;L^2(\Omega))}\right]\right), \end{aligned}$$

where $\mathbb{E}\left[\left\|u^n\right\|_{L^\infty(0,T;L^2(\Omega))}\right] \overset{(5.4)}{\lesssim} 1$. As a result,

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)}\sup\left\{\left|\left\langle\int_t^{t+\tau} L_4^n\, ds, \phi\right\rangle\right| \,:\, \phi\in H_N^2,\ \left\|\phi\right\|_{H_N^2}\le 1\right\}\right] \lesssim \delta^{1/2},$$

i.e.,

$$\mathbb{E}\left[\sup_{\tau\in(0,\delta)} I_4(t,\tau)\right] \lesssim \delta^{1/2}, \quad \text{uniformly in } t\in[0,T].$$

Summarising our estimates of $I_1,\dots,I_4$ concludes the proof of (6.1) for $w=u$. The proof for $w=v$ is the same. □

7. TIGHTNESS AND SKOROKHOD A.S. REPRESENTATIONS

In this section we establish the tightness of the probability measures (laws) generated by the Faedo-Galerkin solutions $\{(u^n, v^n, W_u^n, W_v^n, u_0^n, v_0^n)\}_{n\geq 1}$. Note that the strong convergence of u^n, v^n in $L^2_{t,x}$ is a consequence of the spatial H^1 bound (5.2) and the time translation estimate (6.1), recalling that $H^1 \subset L^2 \subset \left(H^2_N\right)^\star$. To secure the strong (almost sure) convergence in the probability variable $\omega \in D$, we need to use some results of Skorokhod linked to tightness (weak compactness) of probability measures and almost sure representations of random variables.

We choose the following phase space for the probability laws of the Faedo-Galerkin approximations:

$$\mathcal{H} := \mathcal{H}_u \times \mathcal{H}_v \times \mathcal{H}_{W_u} \times \mathcal{H}_{W_v} \times \mathcal{H}_{u_0} \times \mathcal{H}_{v_0},$$

where

$$\mathcal{H}_u,\, \mathcal{H}_v = L^2(0,T;L^2(\Omega)) \bigcap C\big(0,T;(H^1(\Omega))^\star\big)$$

and ($\mathbb{U}_0$ is defined in Section 2)

$$\mathcal{H}_{W_u},\, \mathcal{H}_{W_v} = C([0,T];\mathbb{U}_0), \quad \mathcal{H}_{u_0} = \mathcal{H}_{v_0} = L^2(\Omega).$$

As $\mathcal{X}_1 = L^2(0,T;L^2(\Omega))$, $\mathcal{X}_2 = C\big(0,T;(H^1(\Omega))^\star\big)$ are Polish spaces, the intersection space $\mathcal{X}_1 \cap \mathcal{X}_2$ is Polish. It is also a fact that products of Polish spaces are Polish. Therefore, since $C([0,T];\mathbb{U}_0)$ and $L^2(\Omega)$ are Polish, $\mathcal{H}$ is a Polish space. We denote by $\mathcal{B}(\mathcal{H})$ the σ-algebra of Borel subsets of $\mathcal{H}$, and introduce the measurable mapping

$$\Psi_n : (D,\mathcal{F},P) \to (\mathcal{H},\mathcal{B}(\mathcal{H})),$$
$$\Psi_n(\omega) = \big(u^n(\omega), v^n(\omega), W_u^n(\omega), W_v^n(\omega), u_0^n(\omega), v_0^n(\omega)\big).$$

We define a probability measure $\mathcal{L}_n$ on $(\mathcal{H},\mathcal{B}(\mathcal{H}))$ by

$$\mathcal{L}_n(\mathcal{A}) = \left(P \circ \Psi^{-1}\right)(A) = P\left(\Psi_n^{-1}(\mathcal{A})\right), \quad \mathcal{A} \in \mathcal{B}(\mathcal{H}). \tag{7.1}$$

Denote by $\mathcal{L}_{u^n}$, $\mathcal{L}_{v^n}$, $\mathcal{L}_{W_u^n}$, $\mathcal{L}_{W_v^n}$, $\mathcal{L}_{u_0^n}$, $\mathcal{L}_{v_0^n}$ the respective laws of u^n, v^n, W_u^n, W_v^n, u_0^n and v_0^n, which are defined respectively on $(\mathcal{H}_u,\mathcal{B}(\mathcal{H}_u))$, $(\mathcal{H}_v,\mathcal{B}(\mathcal{H}_v))$, $(\mathcal{H}_{W_u},\mathcal{B}(\mathcal{H}_{W_u}))$, $(\mathcal{H}_{W_v},\mathcal{B}(\mathcal{H}_{W_v}))$ $(\mathcal{H}_{u_0},\mathcal{B}(\mathcal{H}_{u_0}))$ and $(\mathcal{H}_{v_0},\mathcal{B}(\mathcal{H}_{v_0}))$. Thus

$$\mathcal{L}_n = \mathcal{L}_{u^n} \times \mathcal{L}_{v^n} \times \mathcal{L}_{W_u^n} \times \mathcal{L}_{W_v^n} \times \mathcal{L}_{u_0^n} \times \mathcal{L}_{v_0^n}.$$

Remark 7.1. *As a cartesian product of topological spaces, $\mathcal{H}$ is always equipped with the product topology and, thus, the Borel σ-algebra $\mathcal{B}(\mathcal{H})$ generated by the product topology. Of course, on $\mathcal{H}$ there are two natural σ-algebras: the product of the Borel σ-algebras and the already introduced $\mathcal{B}(\mathcal{H})$ for the product topology. For Polish (and separable metric) spaces, these two coincide. This implies that coordinatewise measurability and tightness is the same as joint measurability and tightness, which is important since we use the product of the Borel σ-algebras in the computations below leading up to the joint tightness and weak convergence in the product space $(\mathcal{H},\mathcal{B}(\mathcal{H}))$.*

Given sequences $\{r_m\}_{m\geq 1}, \{\nu_m\}_{m\geq 1}$ of positive numbers tending to zero as $m \to \infty$ (to be specified below), introduce the set

$$\mathcal{Z}_{r_m,\nu_m} := \Bigg\{ z \in L^\infty\left(0,T;L^2(\Omega)\right) \cap L^2\left(0,T;H^1(\Omega)\right) : \sup_{m\geq 1} \frac{1}{\nu_m} \sup_{\tau\in(0,r_m)} \|z(\cdot+\tau) - z\|_{L^\infty\left(0,T-\tau;(H^2_N)^\star\right)} < \infty \Bigg\}.$$

It is easy to see that $\mathcal{Z}_{r_m,\nu_m}$ is a Banach space under the norm

$$\|z\|_{\mathcal{Z}_{r_m,\nu_m}} := \|z\|_{L^\infty(0,T;L^2(\Omega))} + \|w\|_{L^2(0,T;H^1(\Omega))} + \sup_{m\geq 1} \frac{1}{\nu_m} \sup_{0\leq\tau\leq r_m} \|z(\cdot+\tau) - z\|_{L^\infty\left(0,T-\tau;\left(H^2_N\right)^\star\right)}.$$

In view of [36], we have

$$\mathcal{Z}_{r_m,\nu_m} \subset\subset L^2(0,T;L^2(\Omega)) \cap C\big([0,T];(H^1(\Omega))^\star\big),$$

where $X \subset\subset Y$ means that X is compactly embedded in Y. Indeed, to conclude this we need Theorem 5 in [36] on the compactness of functions with values in an intermediate space. Let X_1, X_0, X_{-1} be Banach spaces with continuous embeddings $X_1 \subset X_0 \subset X_{-1}$ and X_1 compactly embedded in X_0. Then [36, Theorem 5] ensures that $\mathcal{Z}$ is relatively compact in $L^p(0,T;X_0)$, with $p \in [1,\infty)$, if $\mathcal{Z}$ is bounded in $L^p(0,T;X_1)$ and, as $\tau \to 0$, there holds that $\|u(\cdot+\tau) - u\|_{L^p(0,T-\tau;X_{-1})} \to 0$, uniformly for $u \in \mathcal{Z}$, if p is finite. If $p = \infty$, then the relative compactness is in $C([0,T];X_0)$. First, we will apply this result with $X_1 = H^1(\Omega)$, $X_0 = L^2(\Omega)$, $X_{-1} = \left(H^2_N\right)^\star$ and $p = 2$, which implies relative compactness in $L^2(0,T;L^2(\Omega))$. Second, we will apply it with $X_1 = L^2(\Omega)$, $X_0 = (H^1(\Omega))^\star$, $X_{-1} = \left(H^2_N\right)^\star$ and $p = \infty$, to conclude relative compactness in the space $C\big([0,T];(H^1(\Omega))^\star\big)$.

Now we verify that the laws $\mathcal{L}_n$, cf. (7.1), of the Faedo-Galerkin solutions are tight.

Lemma 7.2. *The sequence $\{\mathcal{L}_n\}_{n\geq 1}$ of probability measures is (uniformly) tight, and therefore weakly compact, on the phase space $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$.*

Proof. For each $\delta > 0$, we need to produce compact sets

$$\mathbf{C}_{1,\delta} \subset L^2(0,T;L^2(\Omega)) \bigcap C\big(0,T;(H^1(\Omega))^\star\big),$$
$$\text{and} \quad \mathbf{C}_{2,\delta} \subset C([0,T];\mathbb{U}_0), \quad \mathbf{C}_{3,\delta} \subset L^2(\Omega),$$

such that $\mathcal{L}_n(\mathbf{C}_\delta) = P(\{\Phi_n \in \mathbf{C}_\delta\}) > 1 - \delta$, where $\mathbf{C}_\delta$ is short-hand notation for $(\mathbf{C}_{1,\delta})^2 \times (\mathbf{C}_{2,\delta})^2 \times (\mathbf{C}_{3,\delta})^2$. This follows if we show that $\mathcal{L}_n\left(\mathbf{C}^c_{i,\delta}\right) \leq \delta/6$ for $i = 1,2,3$.

To this end, pick the sequences $\{r_m\}_{m=1}^\infty$, $\{\nu_m\}_{m=1}^\infty$ such that

$$\sum_{m=1}^\infty \frac{r_m^{1/4}}{\nu_m} < \infty, \tag{7.2}$$

and take

$$\mathbf{C}_{1,\delta} := \left\{ z \in \mathcal{Z}_{r_m,\nu_m} : \|z\|_{\mathcal{Z}_{r_m,\nu_m}} \leq R_{1,\delta} \right\},$$

where $R_{1,\delta} > 0$ is a number to be determined later. In view of [36, Theorem 5], $\mathbf{C}^{1,\delta}$ is a compact subset of $L^2(0,T;L^2(\Omega))$. For $w = u, v$, we have

$$\begin{aligned}
&P\left(\left\{\omega \in D : w^n(\omega) \notin \mathbf{C}^{1,\delta}\right\}\right) \\
&\quad \leq P\left(\left\{\omega \in D : \|w^n(\omega)\|_{L^\infty(0,T;L^2(\Omega))} > R_{1,\delta}\right\}\right) \\
&\qquad + P\left(\left\{\omega \in D : \|w^n(\omega)\|_{L^2(0,T;H^1(\Omega))} > R_{1,\delta}\right\}\right) \\
&\qquad + P\left(\left\{\omega \in D : \sup_{\tau\in(0,r_m)} \|w^n(\cdot+\tau) - w^n\|_{L^\infty\left(0,T-\tau;\left(H^2_N\right)^\star\right)} > R_{1,\delta}\,\nu_m\right\}\right) \\
&\quad =: P_{1,1} + P_{1,2} + P_{1,3} \quad \text{(for any } m \geq 1).
\end{aligned}$$

Repeated applications of the Chebyshev inequality supply

$$
\begin{aligned}
P_{1,1} &\le \frac{1}{R_{1,\delta}}\mathbb{E}\left[\|w^n(\omega)\|_{L^\infty(0,T;L^2(\Omega))}\right] \le \frac{C}{R_{1,\delta}},\\
P_{1,2} &\le \frac{1}{R_{1,\delta}}\mathbb{E}\left[\|w^n(\omega)\|_{L^2(0,T;H^1(\Omega))}\right] \le \frac{C}{R_{1,\delta}},\\
P_{1,3} &\le \sum_{m=1}^{\infty}\frac{1}{R_{1,\delta}\,\nu_m}\mathbb{E}\left[\sup_{0\le\tau\le r_m}\|w^n(\cdot+\tau)-w^n\|_{L^\infty\left(0,T-\tau;\left(H^2_N\right)^\star\right)}\right]\\
&\le \frac{C}{R_{1,\delta}}\sum_{m=1}^{\infty}\frac{r_m^{1/4}}{\nu_m} \overset{(7.2)}{\le} \frac{C}{R_{1,\delta}},
\end{aligned}
$$

where we have used (5.2), (5.4), and (6.1). From this, we can choose $R_{1,\delta}$ such that

$$
\mathcal{L}_{w^n}\left(\mathbf{C}^c_{1,\delta}\right) = P\left(\{\omega\in D : w^n(\omega)\notin \mathbf{C}_{1,\delta}\}\right) \le \frac{\delta}{6}, \quad w=u,v.
$$

Regarding the finite-dimensional approximations of the Wiener processes, we know that the finite series W^n_u, W^n_v are P-a.s. convergent in $C([0,T];\mathbb{U}_0)$ as $n\to\infty$. This implies that the laws $\mathcal{L}_{W^n_u}, \mathcal{L}_{W^n_v}$ converge weakly. Now we use Prokhorov's weak compactness characterization (see e.g. [11, Theorem 2.3])) to conclude the tightness of $\left\{\mathcal{L}_{W^n_u}\right\}_{n\ge1}$ and $\left\{\mathcal{L}_{W^n_u}\right\}_{n\ge1}$; thus, for any $\delta>0$, there exists a compact set $\mathbf{C}_{2,\delta}$ in $C([0,T];\mathbb{U}_0)$ such that

$$
\mathcal{L}_{W^n_w}\left(\mathbf{C}^c_{2,\delta}\right) = P\left(\{\omega\in D : W^n_w(\omega)\notin \mathbf{C}_{2,\delta}\}\right) \le \frac{\delta}{6}, \quad w=u,v.
$$

Similarly, the initial data approximations u^n_0, v^n_0 are P-a.s. convergent in $L^2(\Omega)$ as $n\to\infty$, and so the laws $\mathcal{L}_{u^n_0}, \mathcal{L}_{v^n_0}$ converge weakly (with $\mathcal{L}_{u^n_0}\rightharpoonup \mu_{u_0}$, $\mathcal{L}_{v^n_0}\rightharpoonup \mu_{v_0}$). As a result, these laws are tight and thus

$$
\mathcal{L}_{w^n_0}\left(\mathbf{C}_{3,\delta}\right) = P\left(\{\omega\in D : w^n_0(\omega)\notin \mathbf{C}_{3,\delta}\}\right) \le \frac{\delta}{6}, \quad w=u,v.
$$

Summarising, $\{\mathcal{L}_n\}_{n\ge1}$ is a tight sequence of probability measures. By Prokhorov's theorem [11, Theorem 2.3], this implies the weak compactness of $\{\mathcal{L}_n\}_{n\ge1}$. □

As the probability measures $\mathcal{L}_n$ linked to the Faedo-Galerkin approximations form a sequence that is weakly compact on $(\mathcal{H},\mathcal{B}(\mathcal{H}))$, we deduce that $\mathcal{L}_n$ converges weakly to a probability measure $\mathcal{L}$ on $\mathcal{H}$, up to a subsequence that we do not relabel. We can then apply the Skorokhod representation theorem (see e.g. [11, Theorem 2.4]) to deduce the existence of a new (complete) probability space $(\tilde{D},\tilde{\mathcal{F}},\tilde{P})$ and new random variables

$$
\tilde{\Psi}_n = \left(\tilde{u}^n, \tilde{v}^n, \tilde{W}^n_u, \tilde{W}^n_v, \tilde{u}^n_0, \tilde{v}^n_0\right), \quad \tilde{\Psi} = \left(\tilde{u}, \tilde{v}, \tilde{W}_u, \tilde{W}_v, \tilde{u}_0, \tilde{v}_0\right), \tag{7.3}
$$

with respective joint laws $\tilde{\mathcal{L}}_n = \mathcal{L}_n$ and $\tilde{\mathcal{L}} = \mathcal{L}$, such that $\tilde{\Psi}_n\to\tilde{\Psi}$ almost surely in the topology of $\mathcal{X}$, i.e., the following convergences hold $\tilde{P}$-almost surely as $n\to\infty$:

$$
\begin{aligned}
&\tilde{u}^n\to\tilde{u}, \quad \tilde{v}^n\to\tilde{v} \quad \text{in } L^2(0,T;L^2(\Omega)),\\
&\tilde{u}^n\to\tilde{u}, \quad \tilde{v}^n\to\tilde{v} \quad \text{in } C\big([0,T];\left(H^1(\Omega)\right)^\star\big),\\
&\tilde{W}^n_u\to\tilde{W}_u, \quad \tilde{W}^n_v\to\tilde{W}_v \quad \text{in } C([0,T];\mathbb{U}_0),\\
&\tilde{u}^n_0\to\tilde{u}_0, \quad \tilde{v}^n_0\to\tilde{v}_0 \quad \text{in } L^2(\Omega).
\end{aligned}
\tag{7.4}
$$

By equality of the laws, the estimates in Lemma 5.1 and Corollary 5.2 continue to hold for the new random variables $\tilde{w}^n$ $(w=u,v)$. In fact, all statistical estimates for the

Faedo-Galerkin approximations w^n are valid for the "tilde" approximations $\tilde{w}^n$ defined on the new probability space $(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P})$. Recall $w^n \in C([0,T];\mathbb{X}_n)$ P-a.s., where $\mathbb{X}_n = \mathrm{Span}\{e_1,\dots,e_n\}$ and each e_l belongs to $H^2_N \subset L^\infty$ Besides, by elliptic regularity, e_l is smooth in Ω. Since w^n and $\tilde{w}^n$ have the same laws $\mathcal{L}_{w^n}$ and $C([0,T];\mathbb{X}_n)$ is a Borel subset of $L^2(0,T;L^2(\Omega))\bigcap C\big(0,T;(H^1(\Omega))^\star\big)$, it follows that $\mathcal{L}_{w^n}\big(C([0,T];\mathbb{X}_n)\big) = 1$ and $\tilde{w}^n \in C([0,T];\mathbb{X}_n)$ $\tilde{P}$-a.s., $w = u, v$. Moreover, we have

Lemma 7.3. *Let $\tilde{u}^n(t)$, $\tilde{v}^n(t)$, $\tilde{W}^n_u(t)$, $\tilde{W}^n_v(t)$, $\tilde{u}^n_0$, $\tilde{v}^n_0$ be the Skorokhod representations of the Faedo-Galerkin approximations, cf.* (7.3). *There exists a constant $C > 0$, independent of n, such that*

$$\tilde{\mathbb{E}}\left[\|\tilde{w}^n\|^q_{L^\infty(0,T;L^2(\Omega))}\right] \le C, \quad \tilde{\mathbb{E}}\left[\|\nabla\tilde{w}^n\|^q_{L^2((0,T)\times\Omega)}\right] \le C, \quad w = u, v, \tag{7.5}$$

for any $q \in [2, q_0]$, see (3.2) *and Corollary 5.2 for the appearance of q_0.*

Proof. We prove the first estimate in (7.5), as the other ones can be proved in the same way. Let $f : \mathcal{X}_1 \to \mathcal{X}_2$ be a continuous injection between Polish spaces. According to the Lusin-Suslin theorem, $f(\mathcal{X}_1)$ is a Borel set in $\mathcal{X}_2$. Since $\mathcal{X}_1 := L^\infty(0,T;L^2(\Omega))$ is continuously embedded in $\mathcal{X}_2 := L^2(0,T;L^2(\Omega))$, we can apply the Lusin-Suslin theorem to ensure that $\mathcal{X}_1$ is a Borel set in $\mathcal{X}_2$. Hence, as $\mu := \mathcal{L}_{w^n}$ is a measure on $\mathcal{X}_2$ and $|\cdot|^q : \mathcal{X}_2 \to \mathbb{R}$ is continuous ($\Rightarrow$ Borel measurable), the integration $\int_{\mathcal{X}_1} |w|^q \, d\mu(w)$ makes sense. Consequently, $\tilde{\mathbb{E}}\left[\|\tilde{w}^n\|^q_{L^\infty_t L^2_x}\right] = \int_{\mathcal{X}_1} |w|^q \, d\mu(w) = \mathbb{E}\left[\|w^n\|^q_{L^\infty_t L^2_x}\right] \le C$. □

Recalling (7.3), consider the associated stochastic basis

$$\tilde{\mathcal{S}}_n = \big(\tilde{D}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}^n_t\}_{t\in[0,T]}, \tilde{P}, \tilde{W}^n_u, \tilde{W}^n_v\big), \tag{7.6}$$

where $\tilde{\mathcal{F}}^n_t = \sigma\big(\sigma\big(\tilde{\Psi}_n\big|_{[0,t]}\big)\bigcup\{N \in \tilde{\mathcal{F}} : \tilde{P}(N) = 0\}\big)$. The filtration $\{\tilde{\mathcal{F}}^n_t\}_{n\ge1}$ is the smallest one making all the "tilde processes" $\tilde{u}^n$, $\tilde{v}^n$, $\tilde{W}^n_u$, $\tilde{W}^n_v$, $\tilde{u}^n_0$, and $\tilde{v}^n_0$ adapted.

A cylindrical Wiener process is fully determined by its law. By equality of the laws and Lévy's martingale characterization of a Wiener process, see [11, Theorem 4.6], we conclude that $\tilde{W}^n_u$ and $\tilde{W}^n_v$ are cylindrical Wiener processes with respect to their canonical filtrations. Furthermore, we claim that $\tilde{W}^n_u$, $\tilde{W}^n_v$ are cylindrical Wiener processes relative to the filtration $\{\tilde{\mathcal{F}}^n_t\}_{n\ge1}$ defined in (7.6). To prove this, we must verify that $\tilde{W}^n_w(t)$ is $\tilde{\mathcal{F}}^n_t$ measurable and $\tilde{W}^n_w(t) - \tilde{W}^n_w(s)$ is independent of $\tilde{\mathcal{F}}^n_s$, for all $0 \le s < t \le T$, $w = u, v$. These properties are simple consequences of the fact that $\tilde{W}^n_w$ and W^n_w have the same laws and that $W^n_w(t)$ is $\mathcal{F}_t$ measurable and $W^n_w(t) - W^n_w(s)$ is independent of $\mathcal{F}_s$.

Hence, there exist sequences $\{\tilde{W}^n_{u,k}\}_{k\ge1}$, $\{\tilde{W}^n_{v,k}\}_{k\ge1}$ of mutually independent real-valued Wiener processes adapted to $\{\tilde{\mathcal{F}}^n_t\}_{t\in[0,T]}$ such that

$$\tilde{W}^n_w = \sum_{k\ge1} \tilde{W}^n_{w,k}\psi_k, \quad \text{for } w = u, v, \tag{7.7}$$

recalling that $\{\psi_k\}_{k\ge1}$ is the basis of $\mathbb{U}$ and the series converge in $\mathbb{U}_0 \supset \mathbb{U}$ (cf. Sect. 2).

In what follows, we will use the following n-truncated sums

$$\tilde{W}^{(n)}_w = \sum_{k=1}^n \tilde{W}^n_{w,k}\psi_k, \quad w = u, v,$$

which converges to $\tilde{W}_w$ in $C([0,T];\mathbb{U}_0)$, $\tilde{P}$-almost surely; the convergence claim follows from (7.4) and standard arguments (see e.g. [11, Section 4.2.2]).

Arguing as in [8], using (4.7) and equality of the laws, the following equations hold $\tilde{P}$-almost surely on the new probability space $\big(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}\big)$:

$$
\begin{aligned}
&\tilde{u}^n(t) - \int_0^t \Pi_n \left[\nabla \cdot \left(D_u\Big(\int_\Omega \tilde{u}^n(t,x)\,dx\Big)\nabla \tilde{u}^n\right)\right] ds \\
&\qquad - \int_0^t \Pi_n \left[\nabla \cdot \left(\mathcal{A}_{11}(\tilde{u}^n,\tilde{v}^n)\nabla \tilde{u}^n + \mathcal{A}_{12}(\tilde{u}^n,\tilde{v}^n)\nabla \tilde{v}^n\right)\right] ds \\
&\quad = \tilde{u}_0^n + \int_0^t \Pi_n \left[F(\tilde{u}^n,\tilde{v}^n)\right] ds + \int_0^t \sigma_u^n(\tilde{u}^n)\, d\tilde{W}_u^{(n)}(s) \quad \text{in } L^2(\Omega), \\
&\tilde{v}^n(t) - \int_0^t \Pi_n \left[\nabla \cdot \left(D_v\Big(\int_\Omega \tilde{v}^n(t,x)\,dx\Big)\nabla \tilde{v}^n\right)\right] ds \\
&\qquad - \int_0^t \Pi_n \left[\nabla \cdot \left(\mathcal{A}_{21}(\tilde{u}^n,\tilde{v}^n)\nabla \tilde{u}^n + \mathcal{A}_{22}(\tilde{u},\tilde{v})\nabla \tilde{v}_n\right)\right] ds \\
&\quad = \tilde{v}_0^n + \int_0^t \Pi_n \left[G(\tilde{u}^n,\tilde{v}^n)\right] ds + \int_0^t \sigma_v^n(\tilde{v}^n)\, d\tilde{W}_v^{(n)}(s) \quad \text{in } L^2(\Omega),
\end{aligned}
\tag{7.8}
$$

for any $t \in [0,T]$, where $\sigma_w^n(\tilde{w}^n)\, d\tilde{W}_w^{(n)} = \sum_{k=1}^n \sigma_{w,k}^n(\tilde{w}^n)\, d\tilde{W}_{w,k}^n$, $w = u, v$. Let us sketch the proof of the first equation in (7.8), with the second one following in the same way. Consider the first equation in (4.7) and introduce the $L^2(\Omega)$-valued stochastic process

$$
\begin{aligned}
I_n(\omega,t,x) := u^n(t) - u_0^n - \int_0^t \Pi_n &\left[\nabla \cdot \left(D_u\Big(\int_\Omega u^n(t,x)\,dx\Big)\nabla u^n\right)\right] ds \\
- \int_0^t \Pi_n &\left[\nabla \cdot \left(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(u^n,v^n)\nabla v^n\right)\right] ds \\
&- \int_0^t \Pi_n \left[F(u^n,v^n)\right] ds - \int_0^t \sigma_u^n(u^n)\, dW_u^n(s).
\end{aligned}
$$

Replacing u^n, v^n, u_0^n, W_u^n by $\tilde{u}^n, \tilde{v}^n, \tilde{u}_0^n, \tilde{W}_u^n$, we denote the resulting process by $\tilde{I}_n$. Let us also introduce the random variables

$$
\mathcal{I}_n(\omega) = \left\| I_n(\omega,\cdot,\cdot)\right\|_{L^2(0,T;L^2(\Omega))}^2, \quad \tilde{\mathcal{I}}_n(\omega) = \left\| \tilde{I}_n(\omega,\cdot,\cdot)\right\|_{L^2(0,T;L^2(\Omega))}^2 .
$$

By (4.7), $\mathcal{I}_n = 0$ a.s. and thus $\mathbb{E}\left[\frac{\mathcal{I}_n}{1+\mathcal{I}_n}\right] = 0$. Recalling that $\int_0^t \sigma_u^n(u^n)\, dW_u^n(s) = \sum_{k=1}^n \int_0^t \sigma_{u,k}^n(u^n)\, dW_{u,k}^n$, let us replace the integrand $\sigma_{u,k}^n(u^n)$ by the time-regularised function $\sigma_{u,k}^\delta(t) = \frac{1}{\delta}\int_0^t e^{-\frac{t-s}{\delta}} \sigma_{u,k}^n(u^n(s))\, ds$, for $\delta > 0$, in which case the stochastic integral can be viewed as a continuous function of the Wiener process $W_{u,k}^n$ (after an integration by parts). Denote by $\mathcal{I}_n^\delta$ the analog of $\mathcal{I}_n$ with $\sigma_{u,k}^n(u^n)$ replaced by $\sigma_{u,k}^\delta$. We use a similar definition of $\tilde{\mathcal{I}}_n^\delta$. It is now possible to write $\frac{\mathcal{I}_n^\delta}{1+\mathcal{I}_n^\delta} = L(\Psi_n)$, $\frac{\tilde{\mathcal{I}}_n^\delta}{1+\tilde{\mathcal{I}}_n^\delta} = L(\tilde{\Psi}_n)$, for some continuous function $L : \mathcal{X} \to \mathbb{R}$. By equality of the laws,

$$
\tilde{\mathbb{E}}\left[\frac{\tilde{\mathcal{I}}_n^\delta}{1+\tilde{\mathcal{I}}_n^\delta}\right] = \int_{\mathcal{X}} L(\Psi)\, d\mathcal{L}_n(\Psi) = \mathbb{E}\left[\frac{\mathcal{I}_n^\delta}{1+\mathcal{I}_n^\delta}\right].
$$

Sending $\delta \downarrow 0$ yields $\tilde{\mathbb{E}}\left[\frac{\tilde{\mathcal{I}}_n}{1+\tilde{\mathcal{I}}_n}\right] = 0$, implying that the first equation in (7.8) holds.

The next estimate was not stated in Lemma 7.3, but it can be derived from the "tilde" equations in (7.8), following the proofs of (5.3) and (5.10). For any $q \in [2, q_0]$,

$$\mathbb{E}\left[\left|\int_0^T \int_\Omega |\mathcal{A}_{ij}(\tilde{u}^n, \tilde{v}^n)| \left(|\nabla \tilde{u}^n|^2 + |\nabla \tilde{v}^n|^2\right) dx\, dt\right|^{\frac{q}{2}}\right] \le C, \quad i,j = 1,2, \tag{7.9}$$

where the constant C is independent of n.

8. Passing to the limit in the Faedo-Galerkin equations

A stochastic basis is needed for the limit of the Skorokhod representations, i.e., for the variables $\tilde{\Psi} := \left(\tilde{u}, \tilde{v}, \tilde{W}_u, \tilde{W}_v, \tilde{u}_0, \tilde{v}_0\right)$, cf. (7.3): namely,

$$\tilde{\mathcal{S}} = \left(\tilde{D}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\in[0,T]}, \tilde{P}, \tilde{W}_u, \tilde{W}_v\right), \tag{8.1}$$

where $\tilde{\mathcal{F}}_t = \sigma\left(\sigma\left(\tilde{\Psi}\big|_{[0,t]}\right) \bigcup \left\{N \in \tilde{\mathcal{F}} : \tilde{P}(N) = 0\right\}\right)$. Recall that $\tilde{W}_u^n$, $\tilde{W}_v^n$ are cylindrical Wiener processes with respect to $\tilde{\mathcal{S}}_n$, see (7.6) and (7.7). Since $\tilde{W}_u^n \to \tilde{W}_u$, $\tilde{W}_v^n \to \tilde{W}_v$ in the sense of (7.4), it is more or less obvious that also the limits $\tilde{W}_u$, $\tilde{W}_v$ are cylindrical Wiener processes with respect to $\tilde{\mathcal{S}}$, see for example [32, Lemma 9.9] or [13, Proposition 4.8]. As a result, there exist sequences $\left\{\tilde{W}_{u,k}\right\}_{k\ge1}$, $\left\{\tilde{W}_{v,k}\right\}_{k\ge1}$ of real-valued Wiener processes adapted to the filtration $\left\{\tilde{\mathcal{F}}_t\right\}_{t\in[0,T]}$, cf. (8.1), such that $\tilde{W}_u = \sum_{k\ge1} \tilde{W}_{u,k}\psi_k$ and $\tilde{W}_v = \sum_{k\ge1} \tilde{W}_{v,k}\psi_k$.

Given the n-independent estimates in Lemma 7.3 and the almost sure convergences in (7.4), we deduce the following result:

Lemma 8.1 (convergence)**.** *The limits $\tilde{u}$, $\tilde{v}$, $\tilde{W}_u$, $\tilde{W}_v$, $\tilde{u}_0$ and $\tilde{v}_0$, see* (7.3) *and also* (7.4)*, satisfy*

$$\tilde{u}, \tilde{v} \in L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; H^1(\Omega))\right) \bigcap L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^\infty(0,T; L^2(\Omega))\right)$$
$$\bigcap L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; C\left([0,T]; \left(H^1(\Omega)\right)^\star\right)\right),$$

and $\sqrt{|\mathcal{A}_{ij}(\tilde{u}, \tilde{v})|}\nabla\tilde{u} \in L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; L^2(\Omega))\right)$, for $i,j = 1,2$.

Let $\tilde{u}^n(t)$, $\tilde{v}^n(t)$, $\tilde{W}_u^n(t)$, $\tilde{W}_v^n(t)$, $\tilde{u}_0^n$, $\tilde{v}_0^n$ be the Skorokhod representations of the Faedo-Galerkin approximations, cf. (7.3)*. Then, passing if necessary to subsequence as $n \to \infty$,*

$$\begin{aligned}
&\text{(i)} && \tilde{u}^n \to \tilde{u}, \quad \tilde{v}^n \to \tilde{v} \quad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; L^2(\Omega))\right),\\
&\text{(ii)} && \tilde{u}^n \rightharpoonup \tilde{u}, \quad \tilde{v}^n \rightharpoonup \tilde{v} \quad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; H^1(\Omega))\right),\\
&\text{(iii)} && \tilde{u}^n \overset{\star}{\rightharpoonup} \tilde{u}, \quad \tilde{v}^n \overset{\star}{\rightharpoonup} \tilde{v} \quad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^\infty(0,T; L^2(\Omega))\right),\\
&\text{(iv)} && \tilde{u}^n \to \tilde{u}, \quad \tilde{v}^n \to \tilde{v} \quad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; C\left([0,T]; \left(H^1(\Omega)\right)^\star\right)\right),\\
&\text{(v)} && \sqrt{|\mathcal{A}_{i1}(\tilde{u}^n, \tilde{v}^n)|}\nabla\tilde{u}^n \rightharpoonup \sqrt{|\mathcal{A}_{i1}(\tilde{u}, \tilde{v})|}\nabla\tilde{u}\\
&&& \qquad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; L^2(\Omega))\right), \quad i = 1,2,\\
&\text{(vi)} && \sqrt{|\mathcal{A}_{i2}(\tilde{u}^n, \tilde{v}^n)|}\nabla\tilde{v}^n \rightharpoonup \sqrt{|\mathcal{A}_{i2}(\tilde{u}, \tilde{v})|}\nabla\tilde{v}\\
&&& \qquad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; L^2(\Omega))\right), \quad i = 1,2,\\
&\text{(vii)} && \tilde{W}_u^n \to \tilde{W}_u, \quad \tilde{W}_v^n \to \tilde{W}_v \quad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; C([0,T]; \mathbb{U}_0)\right),\\
&\text{(viii)} && \tilde{u}_0^n \to \tilde{u}_{i,0}, \quad \tilde{v}_0^n \to \tilde{v}_0 \quad \textit{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(\Omega)\right).
\end{aligned} \tag{8.2}$$

Proof. The strong convergences (i) follow from (7.4), the moment estimate (7.5) with $q > 2$, and Vitali's convergence theorem. The strong convergences (vii) and (viii) follow in a similar way. The weak convergences (ii), (iii) are consequences of the n-uniform bounds on $\tilde{u}^n, \tilde{v}^n$ in $L^2_\omega L^2_t H^1_x$ and in $L^2_\omega L^\infty_t L^2_x$, cf. (7.5), passing if necessary to a subsequence.

Part (iv) is a consequence of (7.4) and Vitali's convergence theorem, given the moment bounds (with some $q > 2$)

$$\tilde{\mathbb{E}} \left\| w \right\|^q_{C\left([0,T];(H^1(\Omega))^\star\right)} \lesssim \tilde{\mathbb{E}} \left[\left\| w \right\|^{q-1/2}_{L^\infty(0,T;L^2(\Omega))} \left\| w \right\|^{1/2}_{C\left([0,T];(H^1(\Omega))^\star\right)} \right] \lesssim 1,$$

for $w = \tilde{u}^n, \tilde{v}^n, \tilde{u}, \tilde{v}$, where we have used that w is bounded in $L^{2q-1}_\omega L^\infty_t L^2_x$, see (7.5).

Let us verify part (v). Set $a_n := \sqrt{|\mathcal{A}_{i2}(\tilde{u}^n, \tilde{v}^n)|}$, $b_n := \nabla \tilde{u}^n$, $c_n := a_n b_n$ and $a := \sqrt{|\mathcal{A}_{i2}(\tilde{u}, \tilde{v})|}$, $b := \nabla \tilde{u}$, $c = ab$. By (ii), $b_n \rightharpoonup b$ in $L^2_{\omega,t,x}$. By (i), passing to a subsequence (not relabelled), we may as well assume that $\tilde{u}^n \to \tilde{u}$, $\tilde{v}^n \to \tilde{v}$ almost everywhere in (ω, t, x). By the global Lipschitz continuity of $\mathcal{A}_{i2}(\cdot,\cdot)$, this transfers to $a_n \to a$ almost everywhere in (ω, t, x). Besides, since $\tilde{u}^n$ and $\tilde{v}^n$ are uniformly bounded in $L^2_{\omega,t,x}$, a_n is uniformly bounded in $L^4_{\omega,t,x}$. Vitali's convergence theorem then implies that $a_n \to a$ in $L^2_{\omega,t,x}$. Next, given the bound (7.9) (with $q = 2$), c_n converges weakly to some limit c in $L^2_{\omega,t,x}$, passing if necessary to a subsequence (not relabelled). At the same time, $a_n \to a$ and $b_n \rightharpoonup b$ in $L^2_{\omega,t,x}$, and so the strong-weak product $a_n b_n$ converges weakly to ab in $L^1_{\omega,t,x}$, which allows us to identify the weak limit $c \in L^2_{\omega,t,x}$ as ab, i.e., $c_n = a_n b_n \rightharpoonup c = ab$ in $L^2_{\omega,t,x}$. This proves (v). The verification of (vi) is similar. □

Our final step is to pass to the limit in the Faedo-Galerkin equations (7.8).

Lemma 8.2 (limit equations)**.** *The limits $\tilde{u}$, $\tilde{v}$, $\tilde{W}_u$, $\tilde{W}_v$, $\tilde{u}_0$, $\tilde{v}_0$ of the Skorokhod a.s. representations of the Faedo-Galerkin approximations—constructed in* (7.3)*,* (7.4)*—satisfy the following equations $\tilde{P}$-a.s., for all $t \in [0,T]$:*

$$\begin{aligned}
&\int_\Omega \tilde{u}(t)\varphi_u \, dx - \int_\Omega \tilde{u}_0 \, \varphi_u \, dx \\
&\qquad + \int_0^t \int_\Omega \left(D_u\Big(\int_\Omega \tilde{u}(t,x)\, dx\Big) \nabla \tilde{u} + \mathcal{A}_{11}(\tilde{u},\tilde{v}) \nabla \tilde{u} + \mathcal{A}_{12}(\tilde{u},\tilde{v}) \nabla \tilde{v} \right) \cdot \nabla \varphi_u \, dx \, ds \\
&\qquad = \int_0^t \int_\Omega F(\tilde{u},\tilde{v}) \varphi_u \, dx\, ds + \int_0^t \int_\Omega \sigma_u(\tilde{u}) \varphi_u \, dx \, d\tilde{W}_u(s),
\end{aligned} \tag{8.3}$$

$$\begin{aligned}
&\int_\Omega \tilde{v}(t)\varphi_v \, dx - \int_\Omega \tilde{v}_0 \varphi_v \, dx \\
&\qquad + \int_0^t \int_\Omega \left(D_v\Big(\int_\Omega \tilde{v}(t,x)\, dx\Big) \nabla \tilde{v} + \mathcal{A}_{21}(\tilde{u},\tilde{v}) \nabla u + \mathcal{A}_{22}(\tilde{u},\tilde{v}) \nabla v \right) \cdot \nabla \varphi_v \, dx \, ds \\
&\qquad = \int_0^t \int_\Omega G(\tilde{u},\tilde{v}) \varphi_v \, dx\, ds + \int_0^t \int_\Omega \sigma_v(\tilde{v}) \varphi_v \, dx \, d\tilde{W}_v(s),
\end{aligned} \tag{8.4}$$

for all $\varphi_u, \varphi_v \in W^{1,4}(\Omega)$, where the laws of $\tilde{u}_0$ and $\tilde{v}_0$ are μ_{u_0} and μ_{v_0}, respectively.

Proof. We will focus on (8.3). The second equation (8.4) can be treated similarly. First, recall that the space H^2_N is dense in $W^{1,4}(\Omega)$. Therefore, it is sufficient to establish (8.3) under the assumption that $\varphi_u \in H^2_N \subset L^\infty$. Indeed, given the bounds in Lemma 8.1, all terms in (8.3)—except for cross-diffusion and the stochastic integral—are bounded by a

(ω-dependent) constant times the $L^2(\Omega)$ or $H^1(\Omega)$ norm of φ_u. Via the BDG inequality (2.5), the stochastic integral is bounded in expectation by a constant times the $L^2(\Omega)$ norm of φ_u. Finally, as $\mathcal{A}_{11}(\tilde{u},\tilde{v})\nabla\tilde{u}$ and $\mathcal{A}_{12}(\tilde{u},\tilde{v})\nabla\tilde{v}$ can be bounded in $L^{4/3}((0,T)\times\Omega)$, the cross-diffusion terms are bounded by a (ω-dependent) constant times $\|\varphi\|_{W^{1,4}(\Omega)}$.

Fix $\varphi_u \in H^2_N$, and write (8.3) symbolically as $I_u(\omega,t)=0$, for $(\omega,t)\in\tilde{D}\times(0,T)$. As in [12], the strategy of the proof is to demonstrate that

$$\|I_u\|^2_{L^2(\tilde{D}\times(0,T))} = \tilde{\mathbb{E}}\int_0^T \left(I_u(\omega,t)\right)^2\,dt = 0,$$

which would imply that $I_u = 0$ for $d\tilde{P}\times dt$-a.e. $(\omega,t)\in\tilde{D}\times(0,T)$ and thus, by the Fubini theorem, $I_u=0$ $\tilde{P}$-a.s., for a.e. $t\in(0,T)$. Since the simple functions are dense in L^2, it enough to prove that

$$\mathbb{E}\left[\int_0^T \mathbf{1}_Z(\omega,t)I_u(\omega,t)\right]\,dt = 0, \tag{8.5}$$

for a measurable set $Z\subset\tilde{D}\times(0,T)$, where $\mathbf{1}_Z(\omega,t)\in L^\infty\left(\tilde{D}\times(0,T);\tilde{d}P\times dt\right)$ denotes the characteristic function of Z.

The Faedo-Galerkin equations (7.8) holds in $L^2(\Omega)$, and hence pointwise in x. Multiplying the first (pointwise) equation with $\varphi_u\in H^2_N$ and then doing spatial integration by parts, using the fact that $\tilde{u}^n,\tilde{v}^n\in H^2_N$—and thus $\frac{\partial\tilde{u}^n}{\partial\nu}=\frac{\partial\tilde{v}^n}{\partial\nu}=0$ on $\partial\Omega$—and basic properties of the projection operator Π_n, we eventually arrive at

$$\begin{aligned}
&\int_\Omega \tilde{u}^n(t)\varphi_u\,dx + \int_0^t\int_\Omega D_u\Big(\int_\Omega \tilde{u}^n(t,x)\,dx\Big)\nabla\tilde{u}^n\cdot\nabla\Pi_n\varphi_u\,dx\,ds\\
&\qquad + \int_0^t\int_\Omega\Big(\mathcal{A}_{11}(\tilde{u}^n,\tilde{v}^n)\nabla\tilde{u}^n + \mathcal{A}_{12}(\tilde{u}^n,\tilde{v}^n)\nabla\tilde{v}^n\Big)\cdot\nabla\Pi_n\varphi_u\,dx\,ds\\
&\quad = \int_\Omega \tilde{u}^n_0\varphi_u\,dx + \int_0^t\int_\Omega F(\tilde{u}^n,\tilde{v}^n)\Pi_n\varphi_u\,dx\,ds\\
&\qquad + \int_0^t\int_\Omega \sigma_u^n(\tilde{u}^n)\Pi_n\varphi_u\,dx\,d\tilde{W}_u^{(n)}(s).
\end{aligned}\tag{8.6}$$

We multiply (8.6) with $\mathbf{1}_Z(\omega,t)$, integrate the result over (ω,t), and then we pass to the limit $n\to\infty$ in each term separately.

By part (viii) of (8.2), we obtain $\tilde{\mathbb{E}}\int_0^T\int_\Omega \mathbf{1}_Z\tilde{u}^n_0\varphi_u\,dx \overset{n\uparrow\infty}{\longrightarrow} \tilde{\mathbb{E}}\int_0^T\int_\Omega \mathbf{1}_Z\tilde{u}_0\varphi_u\,dx$. Recall that $u^n_0=\Pi_n u_0\to u_0$ in $L^2(\Omega)$ and $u_0\sim\mu_{u_0}$ (cf. Theorem 3.5 for the appearance of μ_{u_0}). Hence, as the laws of u^n_0 and $\tilde{u}^n_0$ are the same, we conclude that $\tilde{u}_0\sim\mu_{u_0}$.

In what follows, we will make repeated use of the following simple fact: If $X_n\rightharpoonup X$ in $L^p(\tilde{D}\times(0,T))$, $p\in[1,\infty)$, then $\int_0^t X_n\,ds\rightharpoonup\int_0^t X\,ds$ in $L^p(\tilde{D}\times(0,T))$ as well. Furthermore, we will use that

$$\begin{aligned}
\mathbf{1}_Z(\omega,t)\varphi_u(x) &\in L^\infty\big(\tilde{D}\times(0,T)\times\Omega\big) =: L^\infty_{\omega,t,x},\\
\mathbf{1}_Z(\omega,t)\nabla\varphi_u(x) &\in L^2\big(\tilde{D}\times(0,T)\times\Omega\big) =: L^2_{\omega,t,x}.
\end{aligned}$$

The weak convergence in $L^2_{\omega,t,x}$ of $\tilde{\nabla}u^n$, cf. (8.2)–(ii), implies that

$$\tilde{\mathbb{E}}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t\int_\Omega D_u\Big(\int_\Omega \tilde{u}^n(t,x)\,dx\Big)\nabla\tilde{u}^n\cdot\nabla\Pi_n\varphi_u\,dx\,ds\right)\,dt\right]$$

$$\overset{n\uparrow\infty}{\longrightarrow} \tilde{\mathbb{E}}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t \int_\Omega D_u\Big(\int_\Omega \tilde{u}(t,x)\,dx\Big)\nabla\tilde{u}\cdot\nabla\varphi_u\,dx\,ds\right)\,dt\right],$$

where we have used that $D_u\Big(\int_\Omega \tilde{u}^n(\cdot,x)\,dx\Big)\nabla\Pi_n\varphi_u \overset{n\uparrow\infty}{\longrightarrow} D_u\Big(\int_\Omega \tilde{u}(\cdot,x)\,dx\Big)\nabla\varphi_u$ strongly in $L^2_{\omega,t,x}$, recalling $\nabla\Pi_n\varphi_u \to \nabla\varphi_u$ in L^2_x and noting that the strong $L^2_{\omega,t,x}$ convergence of $\tilde{u}^n$, cf. (8.2)–(i), and (1.7) imply the strong $L^2_{\omega,t}$ convergence of $D_u\Big(\int_\Omega \tilde{u}^n(\cdot,x)\,dx\Big)$.

Regarding the cross-diffusion terms, set (for $w=u,v$ and $i=1,2$)

$$\begin{aligned}
&a_n := \sqrt{|\mathcal{A}_{1i}(\tilde{u}^n,\tilde{v}^n)|}, \quad b_n := \nabla\Pi_n\varphi_u, \quad c_n := a_n b_n,\\
&a := \sqrt{|\mathcal{A}_{1i}(\tilde{u},\tilde{v})|}, \quad b := \nabla\varphi_u, \quad c := ab,\\
&d_n = \sqrt{|\mathcal{A}_{1i}(\tilde{u}^n,\tilde{v}^n)|}\nabla\tilde{w}^n, \quad d = \sqrt{|\mathcal{A}_{1i}(\tilde{u},\tilde{v})|}\nabla\tilde{w}
\end{aligned}$$

and write

$$\mathcal{A}_{1i}(\tilde{u}^n,\tilde{v}^n)\nabla\tilde{w}^n\cdot\nabla\Pi_n\varphi_u = c_n\cdot d_n.$$

Recalling that d_n is weakly convergent to d in $L^2_{\omega,t,x}$, cf. (8.2)–(v), we need to prove that $c_n = a_n b_n$ is strongly convergent to $c = ab$ in $L^2_{\omega,t,x}$, in order to conclude that $c_n\cdot d_n \rightharpoonup c\cdot d$ in $L^1_{\omega,t,x}$. First, $b_n \to b$ in L^4:

$$\|b-b_n\|_{L^4(\Omega)} \le \|\varphi_u - \Pi_n\varphi_u\|_{W^{1,4}(\Omega)} \lesssim \|\varphi_u - \Pi_n\varphi_u\|_{H^2_N} \overset{n\uparrow\infty}{\longrightarrow} 0,$$

where we have used that $H^2_N \subset W^{1,4}(\Omega)$ and $\{e_l\}_{l=1}^\infty$ is a basis of H^2_N. We also claim that $a_n \to a$ in $L^4_{\omega,t,x}$. To see this, note that (8.2)–(i) and (1.8) imply

$$a_n^2 = |\mathcal{A}_{1i}(\tilde{u}^n,\tilde{v}^n)| \to |\mathcal{A}_{1i}(\tilde{u},\tilde{v})| = a^2 \quad \text{in } L^2_{\omega,t,x}.$$

Thus, by the Brezis-Lieb lemma,

$$\left\|a_n^4\right\|^4_{L^4_{\omega,t,x}} \overset{n\uparrow\infty}{\longrightarrow} \left\|a^4\right\|^4_{L^4_{\omega,t,x}}.$$

Passing to a subsequence if necessary, we may as well assume that $a_n \to a$ a.e., and further note that a_n is uniformly bounded in $L^4_{\omega,t,x}$, because $\tilde{u}^n$, $\tilde{v}^n$ are uniformly bounded in $L^2_{\omega,t,x}$ and $\mathcal{A}_{1i}(\cdot,\cdot)$ is globally Lipschitz continuous, cf. (1.8). Another application of the Brezis-Lieb lemma then guarantees that $a_n \to a$ in $L^4_{\omega,t,x}$ as $n\to\infty$. Summarising, $c_n = a_n b_n \to c = ab$ in $L^2_{\omega,t,x}$, $d_n \rightharpoonup d$ in $L^2_{\omega,t,x}$, and thus $c_n\cdot d_n \rightharpoonup c\cdot d$ in $L^1_{\omega,t,x}$

As a result of $\int_\Omega c_n\cdot d_n\,dx \rightharpoonup \int_\Omega c\cdot d\,dx$ in $L^1_{\omega,t}$, we obtain ($w=u,v$, $i=1,2$)

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t\int_\Omega \mathcal{A}_{1i}(\tilde{u}^n,\tilde{v}^n)\nabla\tilde{w}^n\cdot\nabla\Pi_n\varphi_u\,dx\,ds\right)\,dt\right]\\
&\qquad \overset{n\uparrow\infty}{\longrightarrow} \mathbb{E}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t\int_\Omega \mathcal{A}_{1i}(\tilde{u},\tilde{v})\nabla\tilde{w}\cdot\nabla\varphi_u\,dx\,ds\right)\,dt\right].
\end{aligned}$$

Using that F is globally Lipschitz, cf. (1.4) and (1.5), and the strong convergences $\tilde{u}^n \to \tilde{u}$, $\tilde{v}^n \to \tilde{v}$ in $L^2_{\omega,t,x}$, cf. (8.2)–(i), and recalling $\Pi_n\varphi_u \to \varphi_u$ in $L^2(\Omega)$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t\int_\Omega F(\tilde{u}^n,\tilde{v}^n)\Pi_n\varphi_u\,dx\,ds\right)\,dt\right]\\
&\qquad \overset{n\uparrow\infty}{\longrightarrow} \mathbb{E}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t\int_\Omega F(\tilde{u},\tilde{v})\varphi_u\,dx\,ds\right)\,dt\right].
\end{aligned}$$

For the stochastic integral, we will use Lemma 2.1 to prove that

$$\int_0^t \sigma_u^n(\tilde{u}^n)\, d\tilde{W}_u^{(n)}(s) \overset{n\uparrow\infty}{\longrightarrow} \int_0^t \sigma_u(\tilde{u})\, d\tilde{W}_u(s) \quad \text{in } L^2\left(0,T;L^2(\Omega)\right), \tag{8.7}$$

in probability (with respect to $\tilde{P}$). Since $\tilde{W}_u^{(n)} \to \tilde{W}_u$ in $C([0,T];\mathbb{U}_0)$, $\tilde{P}$-a.s. and thus in probability, cf. (7.4), it remains to prove that

$$\sigma_u^n(\tilde{u}^n) \to \sigma_u(\tilde{u}) \quad \text{in } L^2\big(0,T;L_2(\mathbb{U};L^2(\Omega))\big),\ \tilde{P}\text{-almost surely.} \tag{8.8}$$

Clearly,

$$\begin{aligned} &\int_0^T \|\sigma_u(\tilde{u}) - \sigma_u^n(\tilde{u}^n)\|_{L_2(\mathbb{U};L^2(\Omega))}^2 \, dt \\ &\quad \le \int_0^T \|\sigma_u(\tilde{u}) - \sigma_u(\tilde{u}^n)\|_{L_2(\mathbb{U};L^2(\Omega))}^2 \, dt \\ &\qquad + \int_0^T \|\sigma_u(\tilde{u}) - \sigma_u^n(\tilde{u})\|_{L_2(\mathbb{U};L^2(\Omega))}^2 \, dt =: I_1 + I_2. \end{aligned} \tag{8.9}$$

By (2.4) and (7.4), we easily obtain

$$I_1 \overset{n\uparrow\infty}{\longrightarrow} 0, \quad \tilde{P}\text{-almost surely.} \tag{8.10}$$

For the I_2-term, we proceed as follows:

$$\begin{aligned} I_2 &= \int_0^T \sum_{k\ge1} \left\|\sigma_{u,k}(\tilde{u}) - \sigma_{u,k}^n(\tilde{u})\right\|_{L^2(\Omega)}^2 \, dt \\ &= \int_0^T \sum_{k\ge1} \left\|\sigma_{u,k}(\tilde{u}) - \sum_{l=1}^n \sigma_{u,k,l}(\tilde{u}) e_l\right\|_{L^2(\Omega)}^2 \, dt \\ &= \int_0^T \sum_{k\ge1} \left\|\sigma_{u,k}(\tilde{u}) - \Pi_n\big(\sigma_{u,k}(\tilde{u})\big)\right\|_{L^2(\Omega)}^2 \, dt =: \int_0^T \Sigma_n(t)\, dt, \end{aligned}$$

where $\sigma_{u,k}$, $\sigma_{u,k,l}$ are defined respectively in (2.2), (4.6).

The integrand can be dominated by an $L_t^1 := L^1(0,T)$ function ($\tilde{P}$-a.s.):

$$\begin{aligned} 0 \le \Sigma_n(t) &\le 4 \sum_{k\ge1} \|\sigma_{u,k}(\tilde{u}(t))\|_{L^2(\Omega)}^2 = 4\, \|\sigma_u(\tilde{u}(t))\|_{L_2(\mathbb{U};L^2(\Omega))}^2 \\ &\overset{(2.4)}{\le} C\left(1 + \|\tilde{u}(t)\|_{L^2(\Omega)}^2\right) \in L_t^1, \end{aligned}$$

recalling that $\tilde{u} \in L_\omega^2 L_t^\infty L_x^2$ and thus $t \mapsto \|\tilde{u}(t)\|_{L^2(\Omega)}^2 \in L^1(0,T)$ (a.s.). This calculation also shows that $\|\sigma_u(\tilde{u})\|_{L_2(\mathbb{U};L^2(\Omega))}^2 \in L_t^1$ a.s. and $\sum_{k\ge1} |\sigma_{u,k}(\tilde{u})|^2 \in L_{t,x}^1$ a.s., so that

$$\Pi_n\left(\sum_{k\ge1} \sigma_{u,k}(\tilde{u})\right) \overset{n\uparrow\infty}{\longrightarrow} \sum_{k\ge1} \sigma_{u,k}(\tilde{u}) \quad \text{in } L^2(\Omega),$$

for a.e. t and almost surely. In view of these facts and

$$\Sigma_n(t) \overset{n\uparrow\infty}{\longrightarrow} 0, \quad \text{a.e. on } [0,T] \text{ (and a.s),}$$

an application of Lebesgue's dominated convergence theorem supplies

$$I_2 \overset{n\uparrow\infty}{\longrightarrow} 0, \quad \tilde{P}\text{-almost surely.} \tag{8.11}$$

Combining (8.9), (8.10) and (8.11), we arrive at (8.8). By Lemma 2.1, this implies (8.7).

Passing to a subsequence (not relabeled), we may replace "in probability" by "$\tilde{P}$-almost surely" in (8.7). Fixing any number $q \in (2, q_0]$, cf. (3.2), we use the Burkholder-Davis-Gundy inequality (2.5) and (2.3), (7.5) to work out the following estimate:

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left\|\int_0^t \sigma_u^n(\tilde{u}^n)\, d\tilde{W}_u^{(n)}\right\|_{L^2((0,T);L^2(\Omega))}^q\right] \\
&\quad = \tilde{\mathbb{E}}\left[\left(\int_0^T \left\|\sum_{k=1}^n \int_0^t \sigma_{u,k}^n(\tilde{v}^n)\, d\tilde{W}_{u,k}^n\right\|_{L^2(\Omega)}^2 dt\right)^{\frac{q}{2}}\right] \\
&\quad \le \bar{C}_T \tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left\|\sum_{k=1}^n \int_0^t \sigma_{u,k}^n(\tilde{u}^n)\, d\tilde{W}_{u,k}^n\right\|_{L^2(\Omega)}^q\right] \\
&\quad \le C_T \tilde{\mathbb{E}}\left[\left(\int_0^T \sum_{k=1}^n \left\|\sigma_{u,k}^n(\tilde{u}^n)\right\|_{L^2(\Omega)}^2 dt\right)^{\frac{q}{2}}\right] \le C_{\sigma,T}.
\end{aligned}$$

Hence, by Vitali's convergence theorem, (8.7) implies

$$\int_0^t \sigma_u^n(\tilde{u}^n)\, d\tilde{W}_u^{(n)}(s) \to \int_0^t \sigma_u(\tilde{u})\, d\tilde{W}_u(s) \quad \text{in } L^2\left(\tilde{D}, \tilde{\mathcal{F}}, \tilde{P}; L^2(0,T; L^2(\Omega))\right).$$

Using this and the fact that $\Pi_n \varphi_u \to \varphi_u$ in $L^2(\Omega)$, we arrive at

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t \int_\Omega \sigma_u^n(\tilde{u}^n)\Pi_n\varphi_u\, dx\, d\tilde{W}_u^n(s)\right) dt\right] \\
&\quad = \tilde{\mathbb{E}}\left[\int_0^T \int_\Omega \left(\int_0^t \sigma_u^n(\tilde{u}^n)\, d\tilde{W}_u^{(n)}(s)\right)\big(\mathbf{1}_Z(\omega,t)\Pi_n\varphi_u(x)\big)\, dx\, dt\right] \\
&\quad \overset{n\uparrow\infty}{\longrightarrow} \tilde{\mathbb{E}}\left[\int_0^T \mathbf{1}_Z(\omega,t)\left(\int_0^t \int_\Omega \sigma_u(\tilde{u})\varphi_u\, dx\, d\tilde{W}_u(s)\right) dt\right].
\end{aligned}$$

This concludes the proof of (8.5), which implies that the desired (8.3) holds. □

Remark 8.3. *We have proved that the Skorokhod representations* (7.3), (7.4) *satisfy the weak formulation* (8.3), (8.4) *for a.e.* $t \in [0,T]$. *As* $\tilde{u}, \tilde{v} \in C\big([0,T]; \big(H^1(\Omega)\big)^\star\big)$ *a.s., the weak form* (8.3), (8.4) *actually holds for every* $t \in [0,T]$. *This weak continuity property also ensures that* $\tilde{u}, \tilde{v}$ *are predictable in* $\big(H^1(\Omega)\big)^\star$.

9. Nonnegativity of solutions

This section proves that the martingale solution (u, v) constructed as the limit of the Faedo-Galerkin approximations (u^n, v^n) is non-negative, thereby ending the proof of Theorem 3.5. The proof is based on the Stampacchia method, which was properly adapted to the stochastic setting in [10]. It uses Itô's formula to derive the SDEs satisfied by the negative parts $(u^{n,-}, v^{n,-})$ of the Faedo-Galerkin solutions, an energy estimate, and a limiting process with $n \to \infty$, arriving eventually at $\mathbb{E}\,\|(u^-(t), v^-(t))\|_{L^2(\Omega)} = 0$, if the initial data are nonnegative. We write a^- for the negative part, $\max(-a, 0)$, of $a \in \mathbb{R}$. Below we work with a smooth approximation $S_\varepsilon(\cdot)$ of $(\cdot)^-$.

The nonnegativity result is contained in

Lemma 9.1. *The solution (u,v) constructed in Theorem 3.5 is non-negative.*

Proof. In this proof we drop the tildes on the relevant functions, writing for example u^n, u instead of $\tilde{u}^n, \tilde{u}$. For $\varepsilon > 0$, denote by $S_\varepsilon(w)$ the C^2 approximation of $(w^-)^2$ defined by

$$S_\varepsilon(w) = \begin{cases} w^2 - \frac{\varepsilon^2}{6} & \text{if } w < -\varepsilon, \\ -\frac{w^4}{2\varepsilon^2} - \frac{4w^3}{3\varepsilon} & \text{if } -\varepsilon \le w < 0, \\ 0 & \text{if } w \ge 0. \end{cases}$$

Observe that

$$S_\varepsilon'(w) = \begin{cases} 2w & w < -\varepsilon, \\ -\frac{2w^3}{\varepsilon^2} - \frac{4w^2}{\varepsilon} & w \in [-\varepsilon, 0), \\ 0 & w \ge 0 \end{cases} \qquad S_\varepsilon''(w) = \begin{cases} 2 & w < -\varepsilon, \\ -\frac{6w^2}{\varepsilon^2} - \frac{8w}{\varepsilon} & w \in [-\varepsilon, 0), \\ 0 & w \ge 0. \end{cases}$$

It is easy to see that $S_\varepsilon(w) \ge 0$, $S_\varepsilon'(w) \le 0$, and $S_\varepsilon''(w) \ge 0$ for all $w \in \mathbb{R}$. Besides, as $\varepsilon \to 0$, the following convergences hold, uniformly in $w \in \mathbb{R}$: $S_\varepsilon(w) \to (w^-)^2$, $S_\varepsilon'(w) \to -2w^-$, and $S_\varepsilon''(w) \to \begin{cases} 2 & \text{if } w < 0 \\ 0 & \text{if } w \ge 0 \end{cases}$. An application of Itô formula to $S_\varepsilon(u^n)$, where u^n solves (4.7), gives

$$\begin{aligned} &\int_\Omega S_\varepsilon(u^n(t))\,dx - \int_\Omega S_\varepsilon(u^n(0))\,dx \\ &\quad = -\int_0^t \int_\Omega S_\varepsilon''(u^n(s)) D_u\Big(\int_\Omega u^n(s,x)\,dx\Big) \left|\nabla u^n\right|^2\,dx\,ds \\ &\qquad - \int_0^t \int_\Omega S_\varepsilon''(u^n(s)) \Big(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(\tilde{u}^n,\tilde{v}^n)\nabla v^n\Big) \cdot \nabla u^n\,dx\,ds \\ &\qquad + \int_0^t \int_\Omega S_\varepsilon'(u^n(s)) F(u^n,v^n)\,dx\,ds \\ &\qquad + \sum_{k=1}^n \int_0^t \int_\Omega S_\varepsilon'(u^n(s)) \sigma_{u,k}^n(u^n)\,dx\,dW_{u,k}^n \\ &\qquad + \frac{1}{2}\sum_{k=1}^n \int_0^t \int_\Omega S_\varepsilon''(u^n(s)) \left(\sigma_{u,k}^n(u^n)\right)^2 dx\,ds =: \sum_{i=1}^5 I_i. \end{aligned} \tag{9.1}$$

It is easy to see that $I_1 \le 0$. In view of (1.8) and Remark 1.1,

$$\begin{aligned} &S_\varepsilon''(w) = 0 \quad \text{for } w \ge 0, \quad \text{and} \quad S_\varepsilon''(w) \ge 0 \quad \text{for } w \in \mathbb{R}, \\ &\mathcal{A}_{11}(w,\cdot) \ge 0 \quad \text{and} \quad \mathcal{A}_{12}(w,\cdot) = 0, \quad \text{for } w \le 0. \end{aligned} \tag{9.2}$$

As a result,

$$\begin{aligned} I_2 &:= -\int_0^t \int_\Omega S_\varepsilon''(u^n(t)) \\ &\qquad\qquad \times \Big(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(\tilde{u}^n,\tilde{v}^n)\nabla v^n\Big) \cdot \nabla u^n\,dx\,ds \\ &= -\iint_{\{u^n(t,x)\ge 0\}} S_\varepsilon''(u^n(t)) \\ &\qquad\qquad \times \Big(\mathcal{A}_{11}(u^n,v^n)\nabla u^n + \mathcal{A}_{12}(u^n,v^n)\nabla v^n\Big) \cdot \nabla u^n\,dx\,ds \end{aligned}$$

$$
\begin{aligned}
&- \iint_{\{u^n(t,x)<0\}} S''_\varepsilon(u^n(t)) \\
&\qquad\qquad \times \Big(\mathcal{A}_{11}(u^n, v^n)\nabla u^n + \mathcal{A}_{12}(u^n, v^n)\nabla v^n \Big) \cdot \nabla u^n \, dx\, ds \\
&= - \iint_{\{u^n(t,x)<0\}} S''_\varepsilon(u^n(t)) \\
&\qquad\qquad \times \Big(\mathcal{A}_{11}(u^n, v^n)\nabla u^n + \mathcal{A}_{12}(u^n, v^n)\nabla v^n \Big) \cdot \nabla u^n \, dx\, ds \overset{(9.2)}{\leq} 0.
\end{aligned}
$$

Similarly, from the definition of the function F, cf. (1.4) and (1.5), it follows that $I_3 = 0$.

Keeping in mind the convergences in (8.2) (see also [10, Section 3.2]), we send $n \to \infty$ in (9.1) to arrive at the inequality:

$$
\begin{aligned}
&\mathbb{E}\left[\|S_\varepsilon(u(t))\|^2_{L^2(\Omega)} \right] - \mathbb{E}\left[\|S_\varepsilon(u(0))\|^2_{L^2(\Omega)} \right] \\
&\qquad \leq \mathbb{E}\left[\sum_{k=1}^{\infty} \int_0^t \int_\Omega S''_\varepsilon(u(t)) \left(\sigma^n_{k,u}(u)\right)^2 \, dx\, ds \right], \qquad t \in [0,T].
\end{aligned} \tag{9.3}
$$

Sending $\varepsilon \to 0$ in (9.3), and proceeding exactly as in [10, Section 3.4], we arrive at

$$
\mathbb{E}\left[\left\|u^-(t)\right\|^2_{L^2(\Omega)} \right] - \mathbb{E}\left[\left\|u^-(0)\right\|^2_{L^2(\Omega)} \right] \leq C\, \mathbb{E}\left[\int_0^t \left\|u^-(s)\right\|^2_{L^2(\Omega)} \, ds \right], \tag{9.4}
$$

for a.e. $t \in [0,T]$ where $C > 0$ is a constant. Finally, by the nonnegativity of $u(0)$ and applying Gronwall's inequality in (9.4), we conclude that $u^- = 0$ a.e. in $(0,T) \times \Omega$, almost surely. Along the same lines, it follows that $v \geq 0$ a.e. in $(0.T) \times \Omega$, almost surely. □

REFERENCES

[1] B. Ainseba, M. Bendahmane, and A. Noussair. A reaction-diffusion system modeling predator-prey with prey-taxis. *Nonlinear Anal. Real World Appl.*, 9(5):2086–2105, 2008. 1

[2] V. Anaya, M. Bendahmane, M. Langlais, and M. Sepúlveda. A convergent finite volume method for a model of indirectly transmitted diseases with nonlocal cross-diffusion. *Comput. Math. Appl.*, 70(2):132–157, 2015. 1

[3] V. Anaya, M. Bendahmane, M. Langlais, and M. Sepúlveda. Pattern formation for a reaction diffusion system with constant and cross diffusion. In *Numerical mathematics and advanced applications—ENUMATH 2013*, volume 103 of *Lect. Notes Comput. Sci. Eng.*, pages 153–161. Springer, Cham, 2015. 1, 3

[4] V. Anaya, M. Bendahmane, M. Langlais, and M. Sepúlveda. Remarks about spatially structured SI model systems with cross diffusion. In *Contributions to partial differential equations and applications*, volume 47 of *Comput. Methods Appl. Sci.*, pages 43–64. Springer, Cham, 2019. 3

[5] M. Bendahmane. Analysis of a reaction-diffusion system modeling predator-prey with prey-taxis. *Netw. Heterog. Media*, 3(4):863–879, 2008. 1

[6] M. Bendahmane, K. H. Karlsen, and J. M. Urbano. On a two-sidedly degenerate chemotaxis model with volume-filling effect. *Math. Methods Appl. Sci.*, 17(5):783—804, 2007. 1

[7] M. Bendahmane, T. Lepoutre, A. Marrocco, and B. Perthame. Conservative cross diffusions and pattern formation through relaxation. *J. Math. Pures Appl. (9)*, 92(6):651–667, 2009. 1

[8] A. Bensoussan. Stochastic navier-stokes equations. *Acta Applicandae Mathematica*, 38(3):267–304, 1995. 4, 25

[9] H. Bessaih. Martingale solutions for stochastic Euler equations. *Stochastic Anal. Appl.*, 17(5):713–725, 1999. 4

[10] M. D. Chekroun, E. Park, and R. Temam. The Stampacchia maximum principle for stochastic partial differential equations and applications. *J. Differential Equations*, 260(3):2926–2972, 2016. 4, 31, 33

[11] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014. 4, 5, 6, 23, 24

[12] A. Debussche, N. Glatt-Holtz, and R. Temam. Local martingale and pathwise solutions for an abstract fluids model. *Phys. D*, 240(14-15):1123–1144, 2011. 4, 6, 28

[13] A. Debussche, M. Hofmanová, and J. Vovelle. Degenerate parabolic stochastic partial differential equations: Quasilinear case. *Ann. Probab.*, 44(3):1916–1955, 2016. 26

[14] G. Dhariwal, A. Jüngel, and N. Zamponi. Global martingale solutions for a stochastic population cross-diffusion system. *Stochastic Process. Appl.*, 129(10):3792–3820, 2019. 4

[15] F. Flandoli. An introduction to 3D stochastic fluid dynamics. In *SPDE in hydrodynamic: recent progress and prospects*, volume 1942 of *Lecture Notes in Math.*, pages 51–150. Springer, Berlin, 2008. 4

[16] F. Flandoli and D. Gatarek. Martingale and stationary solutions for stochastic Navier-Stokes equations. *Probab. Theory Related Fields*, 102(3):367–391, 1995. 4

[17] G. Galiano, M. L. Garzón, and A. Jüngel. Analysis and numerical solution of a nonlinear cross-diffusion system arising in population dynamics. *RACSAM. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. AMat.*, 95(2):281–295, 2001. 1

[18] G. Galiano, M. L. Garzón, and A. Jüngel. Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numer. Math.*, 93(4):655–673, 2003. 1

[19] H. Garcke and K. Lam. Global weak solutions and asymptotic limits of a cahn–hilliard–darcy system modelling tumour growth. *AIMS Math.*, 1(3):318–360, 2016. 9

[20] N. Glatt-Holtz, R. Temam, and C. Wang. Martingale and pathwise solutions to the stochastic Zakharov-Kuznetsov equation with multiplicative noise. *Discrete Contin. Dyn. Syst. Ser. B*, 19(4):1047–1085, 2014. 4

[21] E. Hausenblas, P. A. Razafimandimby, and M. Sango. Martingale solution to equations for differential type fluids of grade two driven by random force of Lévy type. *Potential Anal.*, 38(4):1291–1331, 2013. 4

[22] M. Hofmanová. Degenerate parabolic stochastic partial differential equations. *Stochastic Process. Appl.*, 123(12):4294–4336, 2013. 4

[23] G. Leoni. *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2017. 8

[24] S. A. Levin. A more functional response to predator-prey stability. *The American Naturalist*, 111(978):381–383, 1977. 1

[25] S. A. Levin and L. A. Segel. Hypothesis for origin of planktonic patchiness. *Nature*, 259(5545):659–659, 1976. 1

[26] M. Mimura and K. Kawasaki. Spatial segregation in competitive interaction-diffusion equations. *J. Math. Biol.*, 9(1):49–64, 1980. 1

[27] M. Mimura and J. D. Murray. On a diffusive prey-predator model which exhibits patchiness. *J. Theoret. Biol.*, 75(3):249–262, 1978. 1

[28] M. Mimura and M. Yamaguti. Pattern formation in interacting and diffusing systems in population biology. *Advances in Biophysics*, 15:19 – 65, 1982. 1

[29] J. D. Murray. *Mathematical biology. I*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2002. An introduction. 1

[30] J. D. Murray. *Mathematical biology. II*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2003. Spatial models and biomedical applications. 1

[31] A. Okubo and S. A. Levin. *Diffusion and ecological problems: modern perspectives*, volume 14 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, second edition, 2001. 1

[32] M. Ondreját. Stochastic nonlinear wave equations in local Sobolev spaces. *Electron. J. Probab.*, 15:no. 33, 1041–1091, 2010. 26

[33] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. 4, 5, 12

[34] P. A. Razafimandimby and M. Sango. Existence and large time behavior for a stochastic model of modified magnetohydrodynamic equations. *Z. Angew. Math. Phys.*, 66(5):2197–2235, 2015. 4

[35] M. Sango. Density dependent stochastic Navier-Stokes equations with non-Lipschitz random forcing. *Rev. Math. Phys.*, 22(6):669–697, 2010. 4

[36] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987. 22

(Mostafa Bendahmane)
INSTITUT DE MATHÉMATIQUES DE BORDEAUX UMR CNRS 525
UNIVERSITÉ DE BORDEAUX
F-33076 BORDEAUX CEDEX, FRANCE
E-mail address: mostafa.bendahmane@u-bordeaux.fr

(Kenneth H. Karlsen)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OSLO
P.O. BOX 1053, BLINDERN, N–0316 OSLO, NORWAY
E-mail address: kennethk@math.uio.no