NONLINEAR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS WITH SINGULAR DIFFUSIVITY AND GRADIENT STRATONOVICH NOISE

IOANA CIOTIR AND JONAS M. TÖLLE

Abstract. We study existence and uniqueness of a variational solution in terms of stochastic variational inequalities (SVI) to stochastic nonlinear diffusion equations with a highly singular diffusivity term and multiplicative Stratonovich gradient-type noise. We derive a commutator relation for the unbounded noise coefficients in terms of a geometric Killing vector condition. The drift term is given by the total variation flow, respectively, by a singular p-Laplace-type operator. We impose nonlinear zero Neumann boundary conditions and precisely investigate their connection with the coefficient fields of the noise. This solves an open problem posed in [Barbu, Brzeźniak, Hausenblas, Tubaro; Stoch. Proc. Appl., 123 (2013)] and [Barbu, Röckner; J. Eur. Math. Soc., 17 (2015)].

1. Introduction

We consider existence and uniqueness of solutions to the following (multi-valued) nonlinear Stratonovich stochastic diffusion equation in $L^2(\mathcal{O})$,

$$\begin{cases} dX_t \in \operatorname{div}\left[\operatorname{sgn}(\nabla X_t)\right] dt + \displaystyle\sum_{i=1}^{N} \langle b_i, \nabla X_t \rangle \circ d\beta_t^i, & \text{in } (0,T)\times\mathcal{O}, \\ X_0 = x, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t}{\partial \nu} = 0, & \text{on } (0,T)\times\partial\mathcal{O}, \end{cases} \tag{1.1}$$

where $\mathcal{O}$ is an open, bounded domain in $\mathbb{R}^d$, $d \geq 2$, with (sufficiently) smooth boundary such that $\mathcal{O}$ or $\partial\mathcal{O}$ is convex. Here, for $N \geq 1$, $b_i : \overline{\mathcal{O}} \to \mathbb{R}^d$, $1 \leq i \leq N$ are "coefficient fields" and $\beta = (\beta^1, \dots, \beta^N)$ denotes an N-dimensional Brownian motion on a filtered (normal) probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$. The initial datum is chosen as $x \in L^2(\mathcal{O})$, or, more generally, as $x \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; L^2(\mathcal{O}))$.

Date: May 23, 2016.

2010 *Mathematics Subject Classification.* Primary: 35K55, 35K92, 60H15; Secondary: 49J40, 49J45, 60H25;

Key words and phrases. Stochastic variational inequality, singular diffusivity, nonlinear singular-degenerate stochastic partial differential equation, total variation flow, multiplicative gradient-type Stratonovich noise, singular p-Laplace, Killing vector field, binary tomography, stochastic methods in mathematical image processing.

I. C. was partially supported by the European Union with the European regional development fund (ERDF) and by the Haute-Normandie Regional Council (M2NUM Project). J. M. T. gratefully acknowledges funding granted by the CRC 701 "Spectral Structures and Topological Methods in Mathematics" (Bielefeld) of the German Research Foundation (DFG).

Here, ν denotes the outer unit normal on $\partial\mathcal{O}$. The multi-valued graph $\xi \mapsto \operatorname{sgn}(\xi)$ from $\mathbb{R}^d$ into $2^{\mathbb{R}^d}$ is defined by

$$\operatorname{sgn}(\xi) := \begin{cases} \dfrac{\xi}{|\xi|}, & \text{, if } \xi \neq 0, \\ \left\{\zeta \in \mathbb{R}^d \,|\, |\zeta| \leq 1\right\} & \text{, if } \xi = 0, \end{cases}$$

for all $\xi \in \mathbb{R}^d$. Because of the multi-valued diffusivity term, the equation becomes formally a *stochastic evolution inclusion*, as have been studied e.g. in [31,32,38]. We denote by $|\cdot|$ the Euclidean norm of $\mathbb{R}^d$, and by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product of $\mathbb{R}^d$.

Set

$$\mathbf{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} : \overline{\mathcal{O}} \to \mathbb{R}^{N \times d}, \tag{1.2}$$

and denote by $\mathbf{b}^*$ its transpose. We have that equation (1.1) is formally equivalent to the Itô stochastic partial differential equation,

$$\begin{cases} dX_t \in \operatorname{div}\left[\operatorname{sgn}\left(\nabla X_t\right)\right] dt + \dfrac{1}{2} \operatorname{div}[\mathbf{b}^* \mathbf{b} \nabla X_t] \, dt + \langle \mathbf{b} \nabla X_t, d\beta_t \rangle, & \text{in } (0,T) \times \mathcal{O}, \\ X_0 = x, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t}{\partial \nu} = 0. & \text{on } (0,T) \times \partial\mathcal{O}. \end{cases} \tag{1.3}$$

A similar equation was studied in [7] for the case of a dissipative drift, using the method of Brézis-Ekeland's variational principle[1]. On the other hand, equations with singular drift of the same form have been studied in [8–10,36] for additive and multiplicative bounded noise, respectively. See [37] for a multiplicative Stratonovich stochastic equation with a similar drift term. Those results do not apply to our case since the noise coefficient

$$u \mapsto \langle \mathbf{b} \nabla u, \cdot \rangle \tag{1.4}$$

is not bounded on the state space $L^2(\mathcal{O})$. In [30], existence and uniqueness as well as regularity have been investigated for the stochastic mean curvature flow with unbounded noise. The methods used are related to ours, however, the structure of the equation prevents a direct application to our situation.

Additionally, in our main Theorem 4.1, we will derive existence and uniqueness results also for the singular p-Laplace equations with $p \in (1,2)$,

$$\begin{cases} dX_t = \operatorname{div}\left[|\nabla X_t|^{p-2} \nabla X_t\right] dt + \displaystyle\sum_{i=1}^{N} \langle b_i, \nabla X_t \rangle \circ d\beta_t^i, & \text{in } (0,T) \times \mathcal{O}, \\ X_0 = x, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t}{\partial \nu} = 0. & \text{on } (0,T) \times \partial\mathcal{O}. \end{cases} \tag{1.5}$$

[1] Equation (1.3) with Dirichlet boundary conditions (instead of Neumann boundary conditions) is also being investigated in [42]. However, the preprint of [42] became publicly available after our revised work was submitted for publication. We point out that the method used in [42] is different from ours.

Due to the lack of strong coercivity of the drift operator, we shall employ so-called stochastic variational inequalities (SVI), with the aim to construct solutions to (1.5) in a weak variational sense. Even for bounded noise, singular equations of the above type are generally not known to satisfy an Itô integral equation — not even in the (analytically) weak sense. Compare with [8, 9, 30, 33] for related works employing SVI-frameworks. Using a rough path approach, equations with similar noise were studied in [17, 29]. A similar equation with linear drift is investigated in [13]. We would like to point out, that the solutions of the work at hand are strong solutions in the probabilistic sense, meaning, in particular, that the solutions are functions of the given Brownian motion.

The natural energy space for the (Neumann) total variation flow, the p-Laplace, respectively, would be $BV(\mathcal{O})$, the space of bounded variation functions, respectively, the Sobolev space $W^{1,p}(\mathcal{O})$. However, on the level of approximations, we shall work on the smaller space $H^1(\mathcal{O})$. One reason is, that we are using viscosity approximations, namely, we are adding a regularization term $\varepsilon\Delta$, and taking $\varepsilon \searrow 0$. In particular, this allows us to consider the gradient-type SPDE for the borderline case of a monotone drift operator ($p = 1$) which cannot be treated within the scope of *reflexive* Gelfand triples ($p > 1$), as e.g. has been done in [7, 10] for Dirichlet boundary conditions.

Another property, necessary for our arguments, is the mutual commutation behavior of the diffusion coefficients, as well as the question of commutation with the Neumann Laplace — in order to obtain these, we introduce a condition from differential geometry, similar to the notion of Killing vector fields, see Assumption 2.1 and Appendix A below. In this context, we prove that, under our assumptions, the first-order partial differential operator (1.4), which corresponds to an infinitesimal vector field action, preserves Neumann boundary conditions, see Lemma 2.8 below.

According to [46], the interest in studying this type of equation comes from its use for simulations in the tomographic reconstruction problem, which has several applications, for instance in medical imaging and general image processing.

More precisely, the binary tomography methods are proposed in [39] as a simpler inverse problem of reconstruction. Being still an ill-posed problem, it needs to be regularized, and this may be done for instance with the total variation (T.V.) regularization. In order to numerically solve the problem, a fast and efficient T.V./L^2 minimization algorithm based on the "Alternate Direction of Minimization Method" (A.D.M.M.) has been proposed in [1, 52]. Finally, a singular stochastic diffusion equation with gradient dependent noise is used to refine the solution obtained by the A.D.M.M. algorithm, see also the related Example 2.3 below. The time dependent (deterministic) T.V. image restoration problem has been studied e.g. in [15]. We refer to [34] and the references therein for a stationary stochastic approach.

Therefore, the present work gives rigorous theory to support the use of this kind of equation for numerical results such that those in [46]. However, the authors of [46] are posing the problem for an Itô-equation instead of a Stratonovich one, see also [51].

Another possible interest of studying stochastic differential equations perturbed by this type of noise comes from the applications in modes of turbulence (see [40]).

Discussion of an approach via transformation. Following the classical works [26, 48], we can also think of an alternative access to our equation, which, however, must fail even on a heuristic level. Here, we shall briefly discuss this approach and point out the difficulties.

Let $y \in L^2(\mathcal{O})$ and consider the following deterministic PDE

$$dY_t(\xi) \in \operatorname{div}[\operatorname{sgn}(\nabla Y_t(\xi))]\,dt, \quad Y_0(\xi) = y(\xi), \quad t \in (0,T],\ \xi \in \mathcal{O},$$

where we impose Neumann boundary conditions. For initial datum $y \in H^1(\mathcal{O})$, a unique weak solution in the Gelfand triple $H^1 \subset L^2 \subset H^{-1}$ was constructed in [32, Theorem 2.6]. For initial conditions $y \in L^2(\mathcal{O})$, see [2, 3, 23].

Let $\mathbf{b}$ be as in (1.2), and assume merely that $b_i \in C^1(\overline{\mathcal{O}};\mathbb{R}^d)$ for $1 \le i \le N$. For $t \in [0,T]$, $\xi \in \mathcal{O}$, $\omega \in \Omega$, define

$$X_t(\xi)(\omega) := Y_t(\xi + \mathbf{b}^*(\xi)\beta_t(\omega)), \quad X_0(\xi) = y(\xi). \tag{1.6}$$

A similar transformation approach can be found in [13] for linear equations and in [29] for the case of conservation laws. See also [14] for other nonlinear SPDEs treated by this transformation.

Assume for a while, that we have a pathwise Itô formula available (that is, for $\omega \in \Omega$, fixed), ignoring the lack of regularity of $(x,t) \mapsto Y_t(\xi + \mathbf{b}^*(\xi)x) =: F(x,t)$ for a moment:

$$F(\beta_t,t) = F(0,0) + \sum_{i=1}^N \int_0^t \partial_{x_i} F(\beta_s,s) \circ d\beta_s^i + \int_0^t \partial_t F(\beta_s,s)\,ds,$$

see [11, 12, 27, 28]. By the chain rule, we would obtain that for dt-a.e. $t \in [0,T]$, possibly outside an exceptional subset of $\mathcal{O}$,

$$X_t \in y + \int_0^t \operatorname{div}[\operatorname{sgn}(\nabla X_s)]\,ds + \sum_{i=1}^N \sum_{j=1}^d \int_0^t b_i^j \partial_{\xi_j} X_s \circ d\beta_s^i,$$

which is a pathwise representation of equation (1.1). The Stratonovich correction term is formally given by

$$\frac{1}{2}[\nabla F(\beta,\cdot),\beta]_t = \frac{1}{2}\sum_{k=1}^d \sum_{i=1}^N \int_0^t \partial_{\xi_k}(\langle b_i, \nabla X_s\rangle)\,ds = \frac{1}{2}\int_0^t \operatorname{div}[\mathbf{b}^*\mathbf{b}\nabla X_s]\,ds,$$

where $t \mapsto [\cdot,\cdot]_t$ denotes the quadratic covariation process, compare with [27, 28].

Even if one finds a way to deal with the measurability issues, the direct application of this approach must fail due to the lack of regularity, since, according to [18, 19, 25], good Hölder estimates for the solution (and for the gradient of the solution) to the parabolic p-Laplace equation usually hold only if $p > \frac{2d}{d+2}$, thus sorting out the total variation flow.

Organization of the paper. After a brief part on notational conventions of this work, we shall give our assumptions and discuss the resulting properties of the noise coefficient operators (1.4) in Section 2 — in particular, we establish the commutation relations which we shall need subsequently. Our notion of SVI-solutions (to equations with gradient-type multiplicative Stratonovich noise) is provided in Section 3. In Section 4, we shall first derive a useful a priori estimate in $H^2(\mathcal{O})$ and after that go through several approximation steps necessary for proving the existence of a solution. The uniqueness of SVI solutions is proved in Subsection 4.2. For the reader's convenience, we shall provide some results on Killing vector fields in the appendix.

Notation. We shall recall a few standard definitions and fix notation which will be used later.

We set $V := H^1(\mathcal{O}) = W^{1,2}(\mathcal{O})$, the standard first order square integrable Sobolev space and $H := L^2(\mathcal{O})$, the Hilbert space of (classes of) square integrable functions with respect to the Lebesgue measure. We also consider the second order square integrable Sobolev space $H^2(\mathcal{O}) = W^{2,2}(\mathcal{O})$. We shall write H^2, H^1, L^2, and so on, if the context is clear. We denote the inner product in H by $(\cdot,\cdot)_H$. V^* denotes the topological dual of V with dualization denoted by $\langle\cdot,\cdot\rangle$. Let $W^{1,p}(\mathcal{O})$ be the usual first order p-integrable Sobolev space. For $u \in L^1(\mathcal{O})$ we define the total variation semi-norm by

$$\|u\|_{TV} := \sup\left\{\int_{\mathcal{O}} u \operatorname{div}\eta \, d\xi \,\middle|\, \eta \in C_0^\infty(\mathcal{O};\mathbb{R}^d),\ \|\eta\|_{L^\infty(\mathcal{O};\mathbb{R}^d)} \le 1\right\}$$

and let BV be the space of functions of bounded variation, that is,

$$BV(\mathcal{O}) := \{u \in L^1(\mathcal{O}) \mid \|u\|_{TV} < \infty\}.$$

For a proper, convex, lower semi-continuous (l.s.c.) function $\Phi : H \to [0,+\infty]$, we denote the subdifferential by $\partial\Phi$. The graph of $\partial\Phi$ consists precisely of the pairs of elements $(x,y) \in \partial\Phi \subseteq H \times H$ that satisfy $(y, z-x)_H \le \Phi(z) - \Phi(x)$ for all $z \in H$. In this context, we may also write $y \in \partial\Phi(x)$, where we identify the subdifferential as a multi-valued map $\partial\Phi : H \to 2^H$.

We say that a function $X \in L^1([0,T]\times\Omega;H)$ is $\{\mathcal{F}_t\}$-progressively measurable if $X1_{[0,t]}$ is $\mathcal{B}([0,t])\otimes\mathcal{F}_t$-measurable for all $t \in [0,T]$. The domain of (unbounded) linear operators A is denoted by $\operatorname{dom}(A)$, and by the same notation, we denote the effective domain of convex functionals or multi-valued graphs. By C, we denote a positive constant that may change its value from line to line.

2. Hypotheses and commutation relation

Suppose that $\mathcal{O} \subset \mathbb{R}^d$ is a sufficiently smooth, open, bounded domain. Denote the surface element on $\partial\mathcal{O}$ by S^{d-1}. Denote by ν the outer unit normal on $\partial\mathcal{O}$.

Below, we collect our assumptions on the "diffusion matrix" $\mathbf{b}$ and prove some essential properties of the associated partial differential operators. Briefly summarized, we are assuming conditions to ensure that

- the first-order partial differential operators associated to the rows of $\mathbf{b}$ are well-defined unbounded skew-symmetric linear operators on $L^2(\mathcal{O})$, see Assumption 2.1 (i) and Lemma 2.6 below;
- the groups of diffeomorphisms generated by the rows of $\mathbf{b}$ mutually commute, see Assumption 2.1 (ii) and Lemma 2.7 below;
- the partial differential operators associated to the rows of $\mathbf{b}$ leave the domain of the Neumann Laplace invariant and commute with its resolvent, see Assumption 2.1 (i), (iii), (iv) and Lemmas 2.8 and 2.10 below.

We note that the commutation assumptions are typical for gradient-type noise, even for linear stochastic equations, see [20, 21] and [22, Section 6.5].

Assumption 2.1. *Suppose that the diffusion coefficients $b_i \in C^2(\overline{\mathcal{O}};\mathbb{R}^d)$, $1 \le i \le N$, satisfy the following conditions:*

(i) *$\langle b_i, \nu\rangle = 0$ on $\partial\mathcal{O}$ for all $1 \le i \le N$.*

(ii) *Either $N = 1$, or $b_i^k \partial_k b_l^j = b_l^k \partial_k b_i^j$ on $\overline{\mathcal{O}}$ for all $1 \le k, j \le d$, $1 \le i, l \le N$, $i \ne l$.*

(iii) $\operatorname{div} b_i = 0$ *and* $\langle \Delta b_i, b_i\rangle = 0$ *on* $\overline{\mathcal{O}}$ *for all* $1 \le i \le N$ *(where the Laplace operator acts componentwise).*

(iv) $\langle\langle \nabla b_i, \nu\rangle, b_i\rangle + \langle\langle \nabla b_i, b_i\rangle, \nu\rangle = 0$ *on* $\partial\mathcal{O}$ *for all* $1 \le i \le N$ *(where the gradient acts componentwise).*

By Lemma A.2 in the appendix, sufficiently smooth vector fields b_i that satisfy (iii) and (iv) above, are precisely the so-called *Killing vector fields*, see (A.1) in the appendix for the definition.

Remark 2.2. Condition (i) and (iii) in the above Assumption sort out any examples with stochastic perturbation for the case $d = 1$. Indeed, let $\mathcal{O} = I$ be a bounded interval, so that clearly $\nu = \pm 1$ at the endpoints of I. One the one hand, condition (i) implies that $b = 0$ on ∂I. On the other hand, $\operatorname{div} b = b' = 0$ implies that b must be constant on $\overline{I}$. Hence $b \equiv 0$.

Altogether, condition (i) ensures that the noise coefficients respect Neumann boundary conditions[2], see Lemma 2.8 below.

Example 2.3. Let $N = 1$, $d = 2$. Let $\mathcal{O} = \{\zeta \in \mathbb{R}^2 \mid |\zeta| < R\}$, $R > 0$. Let $b(\xi) := (\xi_2, -\xi_1)$. Then (1.1) becomes

$$\begin{cases} dX_t \in \operatorname{div}\left[\operatorname{sgn}(\nabla X_t)\right] dt + (\xi_2\partial_1 X_t - \xi_1\partial_2 X_t) \circ d\beta_t, & \text{in } (0,T)\times\mathcal{O}, \\ X_0 = x, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t}{\partial \nu} = 0, & \text{on } (0,T)\times\partial\mathcal{O}. \end{cases}$$

Example 2.4. Let $N = 1$, $d = 3$. Let $\mathcal{O} = \{\zeta \in \mathbb{R}^3 \mid |\zeta| < R\}$, $R > 0$. Let $b(\xi) := (\xi_3 - \xi_2, \xi_1 - \xi_3, \xi_2 - \xi_1)$ and denote $\mathbf{1} := (1,1,1)$ (clearly, $b(\xi) = \xi \times \mathbf{1}$). Then b is a Killing vector field and (1.1) becomes

$$\begin{cases} dX_t \in \operatorname{div}\left[\operatorname{sgn}(\nabla X_t)\right] dt + \langle \xi \times \nabla X_t, \mathbf{1}\rangle \circ d\beta_t, & \text{in } (0,T)\times\mathcal{O}, \\ X_0 = x, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t}{\partial \nu} = 0, & \text{on } (0,T)\times\partial\mathcal{O}. \end{cases}$$

One can replace $\mathbf{1}$ by any constant vector $\zeta_0 \in \mathbb{R}^3 \setminus \{0\}$ and get that $\tilde{b}(\xi) := \xi \times \zeta_0$ still satisfies Assumption 2.1.

Remark 2.5. Note that:

(i) The above vector fields $\xi \mapsto (\xi_2, -\xi_1)$ and $\xi \mapsto \xi \times \zeta_0$ resp. are the infinitesimal generators of the rotation groups $SO(2)$ and $SO(3)$ resp., see e.g. [35]. They generate groups of rotations around the origin, leaving balls centered at the origin invariant, which explains why the respective domains are chosen as above (ζ_0 spans the axis of rotation).

(ii) Let $d = N$. The example of constant vector fields $b_i^j = \delta_{i,j}$ are precisely the infinitesimal generators of groups of translations (violating Assumption 2.1 (i) on balls). The d-torus $\mathbb{T}^d$ leaves the translation groups invariant, and is still a bounded, convex domain, leading either to periodic boundary conditions or to a setting for compact manifolds without boundary.

[2] If $b \equiv 1$ on $\mathcal{O} = I = (0, 2\pi)$, then $\xi \mapsto \cos\xi$ has Neumann boundary conditions, however, $b \cdot (\cos\xi)'$ does not.

Recall that the domain $\mathrm{dom}(-\Delta)$ of the Neumann Laplace in the weak sense is given by all elements $u \in H^1(\mathcal{O})$ such that $\Delta u \in L^2(\mathcal{O})$ and such that

$$\int_{\mathcal{O}} v \Delta u \, d\xi = -\int_{\mathcal{O}} \langle \nabla v, \nabla u \rangle \, d\xi \quad \forall v \in H^1(\mathcal{O}).$$

For $u \in \mathrm{dom}(-\Delta)$, the normal derivative $\frac{\partial u}{\partial \nu}$ belongs to $H^{-1/2}(\partial\mathcal{O})$ (being the dual of the space of traces $H^{1/2}(\partial\mathcal{O})$) and is zero, see e.g. [24, p. 250] for details. As we assume smooth boundary, the normal derivative is given by $\frac{\partial u}{\partial \nu} = \langle \nabla u, \nu \rangle$ S^{d-1}-a.e., whenever $u \in C^2(\overline{\mathcal{O}})$. Hence,

$$\mathcal{C} := \{u \in C^2(\overline{\mathcal{O}}) \mid \langle \nabla u, \nu \rangle = 0 \ S^{d-1}\text{-a.e.}\} \tag{2.1}$$

is a *core* for the Neumann Laplace, that is, dense in $\mathrm{dom}(-\Delta)$ w.r.t. to the graph norm

$$\|u\|^2_{\mathrm{dom}(-\Delta)} := \int_{\mathcal{O}} (|\Delta u|^2 + |u|^2) \, d\xi.$$

On the domain $H^1(\mathcal{O})$, we define the linear operators B_i, $1 \le i \le N$ as

$$B_i : H^1(\mathcal{O}) \to L^2(\mathcal{O})$$

$$\begin{aligned} B_i(u)(\xi) &:= \langle b_i(\xi), \nabla u(\xi) \rangle \\ &= \mathrm{div}[b_i(\xi) u(\xi)], \quad \forall u \in H^1(\mathcal{O}), \end{aligned}$$

where b_i satisfies Assumption 2.1.

Lemma 2.6. *Assume Assumption 2.1. Let us collect the following properties:*

(i) *The space $H^1(\mathcal{O})$ is the domain of skew-adjointness of B_i, $1 \le i \le N$, that is*

$$B_i u = -B_i^* u, \quad 1 \le i \le N, \quad u \in H^1(\mathcal{O}),$$

where B_i^* denotes the adjoint operator in $L^2(\mathcal{O})$.

(ii) *For all $u \in H^1(\mathcal{O})$, it holds that*

$$\int_{\mathcal{O}} u B_i u \, d\xi = 0, \quad 1 \le i \le N. \tag{2.2}$$

Proof. Let $u \in H^1(\mathcal{O})$ and fix $1 \le i \le N$.

(i): By the Gauss-Green theorem, for $v \in H^1(\mathcal{O})$, taking Assumption 2.1 (i) into account,

$$\begin{aligned} (B_i^* u, v)_{L^2(\mathcal{O})} &= (u, B_i v)_{L^2(\mathcal{O})} \\ &= \int_{\mathcal{O}} u \langle b_i, \nabla v \rangle \, d\xi \\ &= -\int_{\mathcal{O}} \mathrm{div}(b_i u) v \, d\xi + \int_{\partial\mathcal{O}} uv \langle b_i, \nu \rangle \, dS^{d-1} \\ &= -\int_{\mathcal{O}} \langle b_i, \nabla u \rangle v \, d\xi \\ &= -(B_i u, v)_{L^2(\mathcal{O})}. \end{aligned}$$

The density of $H^1(\mathcal{O}) \subset L^2(\mathcal{O})$ yields (i).

(ii): This follows directly from (i).

□

For $1 \leq i \leq N$ fixed, let $e^{tB_i} : L^2(\mathcal{O}) \to L^2(\mathcal{O})$, $t \in \mathbb{R}$, denote the C_0-group of linear operators associated to B_i, such that, in particular,

$$\frac{d}{dt} e^{tB_i} u\Big|_{t=0} = B_i u, \quad u \in H^1(\mathcal{O}).$$

Lemma 2.7. *The groups e^{tB_i}, $1 \leq i \leq N$, $t \in \mathbb{R}$ mutually commute, whenever Assumption 2.1 holds.*

Proof. For $N = 1$, there is nothing to prove. Let b_i, b_l, $i \neq l$ be as above. Let $f \in C^2(\overline{\mathcal{O}})$. Define the commutator $[b_i, b_l]f := B_i B_l f - B_l B_i f$. By Leibniz's rule,

$$\begin{aligned}
[b_i, b_l]\, f &= \sum_{1 \leq k,j \leq d} b_i^k \partial_k (b_l^j \partial_j f) - b_l^k \partial_k (b_i^j \partial_j f) \\
&= \sum_{1 \leq k,j \leq d} b_i^k b_l^j \partial_k \partial_j f + b_i^k \partial_k b_l^j \partial_j f - b_l^k b_i^j \partial_k \partial_j f - b_l^k \partial_k b_i^j \partial_j f \\
&= \sum_{1 \leq k,j \leq d} (b_i^k \partial_k b_l^j - b_l^k \partial_k b_i^j) \partial_j f \\
&= 0,
\end{aligned}$$

where we have used Assumption 2.1 (ii) in the last step.

Now, for any $1 \leq i \leq N$, denote by $Z_t^i : \overline{\mathcal{O}} \to \overline{\mathcal{O}}$, $t \in [0, \infty)$ the flow of diffeomorphisms on $\overline{\mathcal{O}}$ corresponding to the vector field action of b_i, that is,

$$\frac{d}{dt} Z_t^i = b_i(Z_t^i), \quad t \geq 0, \quad Z_0^i = \xi \in \overline{\mathcal{O}}.$$

Then, since we have proved above that $[b_i, b_l] = 0$ for any $i \neq l$, and from the property of the vector fields to be divergence free, we get that the Z^i, $1 \leq i \leq N$ mutually commute (in the sense of composition of maps), see [43, Ch. I.2, Exercise 3]. However, it holds that $(e^{tB_i} u)(\xi) = u(Z_t^i(\xi))$ and one easily deduces that the groups of operators commute on L^2. □

Lemma 2.8. *Suppose that Assumption 2.1 holds. Let $u \in C^3(\overline{\mathcal{O}})$ be a scalar function with $\langle \nabla u, \nu \rangle = 0$ on $\partial \mathcal{O}$. Then it holds that $\langle \nabla(\langle b_i, \nabla u \rangle), \nu \rangle = 0$ on $\partial \mathcal{O}$ for every $1 \leq i \leq N$.*

Proof. Fix $1 \leq i \leq N$ and set $b := b_i$. Let $\eta \in C^2(\overline{\mathcal{O}})$ be a test-function. We claim that for any $u \in C^3(\overline{\mathcal{O}})$ with $\langle \nabla u, \nu \rangle = 0$ on $\partial \mathcal{O}$, it holds that

$$\int_{\partial \mathcal{O}} \eta \langle \nabla(\langle b, \nabla u \rangle), \nu \rangle \, dS^{d-1} = 0 \quad \forall \eta \in C^2(\overline{\mathcal{O}}). \tag{2.3}$$

In order to prove (2.3), we first apply Gauss's divergence theorem to the vector field $F := \eta \nabla(\langle b, \nabla u \rangle)$, and get that

$$\int_{\partial \mathcal{O}} \langle \eta \nabla(\langle b, \nabla u \rangle), \nu \rangle \, dS^{d-1} = \int_{\mathcal{O}} \operatorname{div} F \, d\xi.$$

However, div $F = \langle \nabla \eta, \nabla(\langle b, \nabla u \rangle) \rangle + \eta \Delta(\langle b, \nabla u \rangle)$. Let us begin with investigating the second term. By the Killing assumption, we have the commutation on sufficiently smooth functions (cf. Theorem A.3 in the appendix), thus, $\eta \Delta(\langle b, \nabla u \rangle) = \eta \langle b, \nabla \Delta u \rangle = \langle \eta b, \nabla \Delta u \rangle$. Integrating by parts, we get that

$$\int_{\mathcal{O}} \langle \eta b, \nabla \Delta u \rangle \, d\xi = -\int_{\mathcal{O}} \operatorname{div}(\eta b) \Delta u \, d\xi + \int_{\partial \mathcal{O}} \eta \langle b, \nu \rangle \Delta u \, dS^{d-1}.$$

The latter term is zero by $\langle b, \nu\rangle = 0$. Also, since both η and b are smooth up to the boundary, $\operatorname{div}(\eta b) \in H^1(\mathcal{O})$. So we can use the Neumann boundary condition for u to get that

$$-\int_{\mathcal{O}} \operatorname{div}(\eta b)\Delta u \, d\xi = \int_{\mathcal{O}} \langle \nabla \operatorname{div}(\eta b), \nabla u\rangle \, d\xi.$$

Clearly, as $\operatorname{div} b = 0$, we get that $\nabla \operatorname{div}(\eta b) = \nabla(\langle b, \nabla\eta\rangle)$. Hence

$$\int_{\mathcal{O}} \operatorname{div} F \, d\xi = \int_{\mathcal{O}} [\langle \nabla\eta, \nabla(\langle b, \nabla u\rangle)\rangle + \langle \nabla(\langle b, \nabla\eta\rangle), \nabla u\rangle] \, d\xi,$$

differentiating out this term yields

$$\int_{\mathcal{O}} \operatorname{div} F \, d\xi = \int_{\mathcal{O}} [\langle (Db)\cdot\nabla u, \nabla\eta\rangle + \langle (D^2 u)\cdot b, \nabla\eta\rangle + \langle (Db)\cdot\nabla\eta, \nabla u\rangle + \langle (D^2\eta)\cdot b, \nabla u\rangle] \, d\xi,$$

where, Db denotes the Jacobian of b and D^2 denotes the Hessian of a scalar function, "$\cdot$" denotes matrix multiplication.

However, Db is skew-symmetric with respect to the Euclidean scalar product due to the Killing assumption, see (A.1) in the appendix. Hence

$$\langle (Db)\cdot\nabla u, \nabla\eta\rangle = -\langle (Db)\cdot\nabla\eta, \nabla u\rangle$$

and the above term becomes

$$\int_{\mathcal{O}} \operatorname{div} F \, d\xi = \int_{\mathcal{O}} [\langle (D^2 u)\cdot b, \nabla\eta\rangle + \langle (D^2\eta)\cdot b, \nabla u\rangle] \, d\xi.$$

With Einstein's summation convention, interchanging the order of differentiation,

$$\int_{\mathcal{O}} [\partial_i\partial_j u b^j \partial_i\eta + \partial_i\partial_j\eta b^j \partial_i u] \, d\xi = \int_{\mathcal{O}} [\partial_j\partial_i u b^j \partial_i\eta + \partial_i\partial_j\eta b^j \partial_i u] \, d\xi.$$

Integrating by parts in the first term yields

$$\int_{\mathcal{O}} \partial_j\partial_i u b^j \partial_i\eta \, d\xi = -\int_{\mathcal{O}} \partial_i u \partial_j(b^j\partial_i\eta) \, d\xi + \int_{\partial\mathcal{O}} \partial_i u b^j \partial_i\eta\nu_j \, dS^{d-1}.$$

Now, in the boundary integral term, we can separate the sums over j and i resp. and get that this term becomes zero by $\langle b, \nu\rangle = 0$. Furthermore,

$$-\partial_i u\partial_j(b^j\partial_i\eta) = -\partial_i u\partial_j b^j\partial_i\eta - \partial_i u b^j\partial_j\partial_i\eta = -\partial_i u b^j\partial_j\partial_i\eta,$$

as we have that $\operatorname{div}\, b = 0$. Finally, the remaining terms cancel, and we get that

$$0 = \int_{\mathcal{O}} \operatorname{div} F \, d\xi = \int_{\partial\mathcal{O}} \eta\langle \nabla(\langle b, \nabla u\rangle), \nu\rangle \, dS^{d-1} \quad \forall \eta \in C^2(\overline{\mathcal{O}}).$$

□

We shall need the following commutation result. Denote the resolvent of the Neumann Laplace by $J_\delta := (\operatorname{Id} - \delta\Delta)^{-1}$, $\delta > 0$.

Theorem 2.9 (Shigekawa). *Fix $1 \le i \le N$. Suppose that there exists a linear subspace $\mathcal{D} \subset \operatorname{dom}(-\Delta)$ such that the following conditions hold:*

(i) $\Delta(\mathcal{D}) \subseteq \operatorname{dom}(B_i)$,
(ii) $B_i(\mathcal{D}) \subseteq \operatorname{dom}(-\Delta)$,
(iii) *$\mathcal{D}$ is a core (see* (2.1) *for the terminology) for* $(-\Delta, \operatorname{dom}(-\Delta))$,
(iv) $\operatorname{dom}(-\Delta) \subseteq \operatorname{dom}(B_i)$ *and* $\operatorname{dom}(-\Delta) \subseteq \operatorname{dom}(B_i^*)$,
(v) *For any $u \in \mathcal{D}$, it holds that*

$$B_i\Delta u = \Delta B_i u.$$

Then for all $\delta > 0$, and every $u \in \operatorname{dom}(B_i)$, it holds that

$$B_i J_\delta u = J_\delta B_i u.$$

Proof. See [44, Theorem 3.1 and Proposition 3.2]. □

Lemma 2.10. *Assumption 2.1 implies that all of the conditions of Theorem 2.9 are satisfied for all $1 \le i \le N$.*

Proof. Assume the conditions of Assumption 2.1. Fix $1 \le i \le N$. First note that Lemma 2.6 implies that $\operatorname{dom}(B_i) = \operatorname{dom}(B_i^*) = H^1(\mathcal{O})$. Let $\mathcal{D} := C^\infty(\overline{\mathcal{O}}) \cap \mathcal{C}$, where $\mathcal{C} := \{u \in C^2(\overline{\mathcal{O}}) \mid \langle \nabla u, \nu \rangle = 0 \ S^{d-1}\text{-a.e.}\}$. Obviously, $\mathcal{D}$ is a core for $\operatorname{dom}(-\Delta)$. Hence (i), (iii), (iv) are clearly satisfied. (ii) follows from Lemma 2.8. The commutation on smooth functions (v) follows from Theorem A.3, since b_i is a Killing field by Assumption 2.1 and Lemma A.2. □

Let us also define $B_i^2 : H^1(\mathcal{O}) \to (H^1(\mathcal{O}))^*$ by

$$B_i^2 u := -B_i^* B_i u, \quad u \in H^1(\mathcal{O}), \quad 1 \le i \le N.$$

In the sense of Schwartz distributions, it holds that

$$\sum_{i=1}^N B_i^2 u = \operatorname{div}[\mathbf{b}^* \mathbf{b} \nabla u].$$

Set $S := H^1(\mathcal{O})$. We thus have a Gelfand triple

$$S \subset H \subset S^*.$$

3. Stochastic variational inequalities (SVI)

Let $\beta = (\beta^1, \ldots, \beta^N)$ be a N-dimensional Brownian motion on a filtered (normal) probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P}\right)$ (with expected value $\mathbb{E}[Y] = \int_\Omega Y \, d\mathbb{P}$, $Y \in L^1(\Omega)$). We consider the following SPDE on $H = L^2(\mathcal{O})$, where $\mathcal{O} \subset \mathbb{R}^d$ is a smooth, open, bounded domain such that $\mathcal{O}$ or $\partial\mathcal{O}$ is convex,

(3.1)
$$\begin{cases} dX_t \in \operatorname{div}[\Psi(\nabla X_t)]\, dt + \dfrac{1}{2}\displaystyle\sum_{i=1}^N B_i^2 X_t \, dt + \sum_{i=1}^N B_i X_t \, d\beta_t^i, & \text{in } (0,T) \times \mathcal{O}, \\ X_0 = x, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t}{\partial \nu} = 0, & \text{on } (0,T) \times \partial\mathcal{O}, \end{cases}$$

here, $\Psi := \partial\varphi \subseteq \mathbb{R}^d \times \mathbb{R}^d$ is the subdifferential of $\varphi := \frac{1}{p}|\cdot|^p$ for $p \in [1,2)$, which is multi-valued for $p = 1$, i.e. $\partial(\xi \mapsto |\xi|)(\cdot) = \operatorname{sgn}(\cdot)$. More precisely, after fixing $p \in [1,2)$, let

$$\tilde{\Phi}(u) := \begin{cases} \int_{\mathcal{O}} \varphi(\nabla u(\xi))\, d\xi, & \text{if } u \in H^1, \\ +\infty, & \text{if } u \in L^2 \setminus H^1. \end{cases}$$

$\tilde{\Phi}$ is a proper convex functional on L^2 but might fail to be lower semi-continuous. Let us define

$$\Phi(u) := \operatorname{cl}\tilde{\Phi}(u) := \inf\left\{ \liminf_{n\to\infty} \ \tilde{\Phi}(u_n) \;\middle|\; u_n \to u \in L^2(\mathcal{O}) \text{ strongly} \right\},$$

the so-called *lower semi-continuous envelope* of $\tilde{\Phi}$, cf. [5, Proposition 11.1.1]. The l.s.c. envelope is given by, for $p \in (1,2)$,

$$\Phi(u) := \begin{cases} \int_{\mathcal{O}} \varphi(\nabla u)\, d\xi & \text{if } u \in W^{1,p}(\mathcal{O}) \cap L^2(\mathcal{O}), \\ +\infty & \text{if } u \in L^2(\mathcal{O}) \setminus W^{1,p}(\mathcal{O}), \end{cases}$$

and for $p = 1$,

$$\Phi(u) := \begin{cases} \|u\|_{TV} & \text{if } u \in BV(\mathcal{O}) \cap L^2(\mathcal{O}), \\ +\infty & \text{if } u \in L^2(\mathcal{O}) \setminus BV(\mathcal{O})., \end{cases}$$

where we suppress the dependence on p in the notation. Obviously, Φ is convex and it is easy to see that Φ is lower semi-continuous on H. Moreover, $\tilde{\Phi}$ is Gâteaux-differentiable in u with derivative given by

$$D\tilde{\Phi}(u)(v) = \int_{\mathcal{O}} \langle \eta, \nabla v \rangle \, d\xi,$$

with $\eta(\xi) \in \Psi(\nabla u(\xi))$ for a.e. $\xi \in \mathcal{O}$. In fact, Φ coincides with the lower semi-continuous hull of $\tilde{\Phi}$ on H, and we have for $u \in H^1$ that

$$\{-\operatorname{div} \eta \mid \eta \in H^1(\mathcal{O};\mathbb{R}^d),\, \eta \in \Psi(\nabla u),\, d\xi\text{-a.e.}\} \subseteq \partial\Phi(u).$$

However, the full characterization of $\partial\Phi$ (already in the space $L^1(\mathcal{O})$) is involved. We shall omit its precise characterization and instead refer to [5].

Equation (3.1) is then written in relaxed form as

$$\begin{cases} dX_t \in -\partial\Phi(X_t)\, dt + \frac{1}{2}\sum_{i=1}^N B_i^2 X_t\, dt + \sum_{i=1}^N B_i X_t \, d\beta_t^i, \quad t \in (0,T), \\ X_0 = x. \end{cases} \tag{3.2}$$

Motivated by [8,9], let us define our notion of a solution to (3.2).

Definition 3.1. Let $x \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; L^2(\mathcal{O}))$, $T > 0$. An $\{\mathcal{F}_t\}$-progressively measurable process $X \in L^2([0,T] \times \Omega; L^2(\mathcal{O}))$ is called an *SVI-solution* to (3.2) if

(i) *(Regularity)*

$$\Phi(X) \in L^1([0,T] \times \Omega). \tag{3.3}$$

(ii) *(Variational inequality)* For every $Z \in L^2([0,T] \times \Omega; H^1(\mathcal{O}))$ such that there exist $Z_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H^1(\mathcal{O}))$, $G \in L^2([0,T] \times \Omega; L^2(\mathcal{O}))$, $\{\mathcal{F}_t\}$-progressively measurable, such that the following equality holds $L^2(\mathcal{O})$, that is,

$$Z_t = Z_0 + \int_0^t G_s\, ds + \frac{1}{2}\sum_{i=1}^N \int_0^t B_i^2 Z_s\, ds + \sum_{i=1}^N \int_0^t B_i Z_s\, d\beta_s^i, \tag{3.4}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$, we have that the following variational inequality holds true

$$\begin{aligned} &\frac{1}{2}\mathbb{E}\|X_t - Z_t\|_{L^2(\mathcal{O})}^2 + \mathbb{E}\int_0^t \Phi(X_s)\, ds \\ &\qquad \le \frac{1}{2}\mathbb{E}\|x - Z_0\|_{L^2(\mathcal{O})}^2 + \mathbb{E}\int_0^t \Phi(Z_s)\, ds \\ &\qquad\quad - \mathbb{E}\int_0^t (G_s, X_s - Z_s)_{L^2(\mathcal{O})}\, ds, \end{aligned} \tag{3.5}$$

for almost all $t \in [0,T]$.

Moreover, if $X \in L^2(\Omega; C([0,T]; L^2(\mathcal{O})))$, we say that X is a *(time-) continuous SVI solution* to (3.2).

Remark 3.2. Practically, the test-process Z needs to satisfy $Z \in L^2([0,T]\times\Omega; H^2(\mathcal{O}))$, we shall provide in (4.5) below that a process Z of the form (3.4) in fact exists (see also (4.4) below).

Inequality (3.5) is obtained by formally applying the Itô formula for the square of the H-norm to the process

$$d(X-Z) = (-\partial\Phi(X) - G)\,dt + \frac{1}{2}\sum_{i=1}^{N} B_i^2(X-Z)\,dt + \sum_{i=1}^{N} B_i(X-Z)\,d\beta^i,$$

taking expectation and using the subdifferential property.

4. Existence and uniqueness

4.1. Existence.

Theorem 4.1. *Let $x \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Then there is a unique continuous SVI solution $X \in L^2(\Omega; C([0,T]; H))$ to (3.2) in the sense of Definition 3.1. For two SVI solutions X, Y with initial conditions $x, y \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, resp., we have*

$$\operatorname*{ess\,sup}_{t\in[0,T]} \mathbb{E}\|X_t - Y_t\|_H^2 \le \mathbb{E}\|x-y\|_H^2.$$

Proof. Recall the notation $H = L^2$, $S = H^1$. We first assume an initial condition $x_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; S)$ and, in the last part of the proof, we shall generalize to $x_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$.

In order to prove the existence of the solution, we need to take a threefold approximation for equation (3.1). Therefore, we consider the following regularized equation,

(4.1)
$$\begin{cases} dX_t^{\varepsilon,\lambda,\delta} = J_\delta \operatorname{div}\Psi^\lambda\left(\nabla J_\delta X_t^{\varepsilon,\lambda,\delta}\right)dt + \varepsilon\Delta X_t^{\varepsilon,\lambda,\delta}dt & \text{in } (0,T)\times\mathcal{O}, \\ \qquad +\dfrac{1}{2}\sum\limits_{i=1}^{N}(B_i^\delta)^2(X_t^{\varepsilon,\lambda,\delta})dt + \sum\limits_{i=1}^{N} B_i^\delta\left(X_t^{\varepsilon,\lambda,\delta}\right)d\beta_t^i, & \\ X_0^{\varepsilon,\lambda,\delta} = x_0, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t^{\varepsilon,\lambda,\delta}}{\partial\nu} = 0, & \text{on } (0,T)\times\partial\mathcal{O}, \end{cases}$$

where Ψ^λ, $\lambda > 0$, is the Yosida approximation of Ψ (cf. [6, p. 37]), J_δ, $\delta > 0$, is the resolvent of the Neumann Laplacian $L := -\Delta$, i.e., $J_\delta = (\operatorname{Id} - \delta\Delta)^{-1}$ and $B_i^\delta(\cdot) = B_i(J_\delta(\cdot))$.

By [7, Theorem 2.4], we have that the there exists a unique $\{\mathcal{F}_t\}$-adapted solution with $X \in C([0,T]; H) \cap L^2([0,T]; S)$ $\mathbb{P}$-a.s. such that (4.1) holds $\mathbb{P}$-a.s. in $L^2([0,T]; S^*)$. We note that our hypotheses guarantee that the conditions needed for [7, Theorem 2.4] are satisfied.

Step I (the estimate in $H^2(\mathcal{O})$):

Considering J_α, $\alpha > 0$, the resolvent of the Neumann Laplace operator $-\Delta$, we define the sequence of semi-inner products on H

$$(u,v)_\alpha := ((-\Delta)_\alpha u, v)_H, \quad u,v \in H,$$

where $(-\Delta)_\alpha$ is the Yosida approximation of the operator $-\Delta$, i.e., $(-\Delta)_\alpha = \frac{1}{\alpha}(\mathrm{Id} - J_\alpha) = -\Delta J_\alpha$ and the induced semi-norms

$$\|u\|_\alpha := \left\|(-\Delta)_\alpha^{\frac{1}{2}} u\right\|_H, \quad u \in H,$$

where $(-\Delta)_\alpha^{\frac{1}{2}}$ denotes the operator square root.

Since they are continuous on H and for all $u \in S$ we have that

$$\|u\|_\alpha \longrightarrow \|\nabla u\|_{L^2(\mathcal{O};\mathbb{R}^d)} \quad \text{as } \alpha \to 0.$$

We shall apply the Itô formula [41, Theorem 4.2.5] to (4.1) with the functional $u \mapsto \|u\|_\alpha^2$, for ε, λ and δ fixed, and we get that for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned}
\left\|X_t^{\varepsilon,\lambda,\delta}\right\|_\alpha^2 &= \|x_0\|_\alpha^2 + 2\int_0^t \left((-\Delta)_\alpha X_s^{\varepsilon,\lambda,\delta}, J_\delta \operatorname{div} \Psi^\lambda\left(\nabla J_\delta X_s^{\varepsilon,\lambda,\delta}\right)\right)_H ds \\
&\quad + 2\varepsilon \int_0^t {}_S\langle(-\Delta)_\alpha X_s^{\varepsilon,\lambda,\delta}, \Delta X_s^{\varepsilon,\lambda,\delta}\rangle_{S^*} ds \\
&\quad + \sum_{i=1}^N \int_0^t \left((-\Delta)_\alpha X_s^{\varepsilon,\lambda,\delta}, (B_i^\delta)^2 X_s^{\varepsilon,\lambda,\delta}\right)_H ds \\
&\quad + 2\sum_{i=1}^N \int_0^t \left((-\Delta)_\alpha X_s^{\varepsilon,\lambda,\delta}, B_i^\delta X_s^{\varepsilon,\lambda,\delta} d\beta_s^i\right)_H \\
&\quad + \sum_{i=1}^N \int_0^t \left\|(-\Delta)_\alpha^{\frac{1}{2}} B_i^\delta X_s^{\varepsilon,\lambda,\delta}\right\|_H^2 ds.
\end{aligned} \tag{4.2}$$

By well-known properties of the resolvent (as symmetry in L^2, commutation with the Yosida approximation) and keeping in mind that the the operators B_i commute with the resolvent of the Neumann Laplace by Theorem 2.9 and Lemma 2.10, we can easily see that, by setting,

$$\Phi_\lambda(u) := \int_{\mathcal{O}} \varphi^\lambda(\nabla u)\, d\xi, \quad u \in S,$$

and setting $v = X_s^{\varepsilon,\lambda,\delta}$, that

$$\left((-\Delta)_\alpha v, J_\delta \operatorname{div} \Psi^\lambda\left(\nabla J_\delta v\right)\right)_H = -\frac{1}{\alpha}(v - J_\alpha v, \partial(\Phi_\lambda \circ J_\delta)v)_H,$$

cf. [45, Proposition II.7.8] for the chain rule for subdifferentials. By using the argument of [33, Equation (3.7)][3] (here, the convexity assumption on the boundary is needed, see also [32, Example 7.11], where the heat kernel estimates of [49, 50] are applied), we see that

$$\begin{aligned}
&\frac{1}{\alpha}(J_\alpha v - v, \partial(\Phi_\lambda \circ J_\delta)v)_H \\
\le &\frac{1}{\alpha}\left(\Phi_\lambda(J_\delta J_\alpha v) - \Phi_\lambda(J_\delta v)\right) = \frac{1}{\alpha}\left(\Phi_\lambda(J_\alpha J_\delta v) - \Phi_\lambda(J_\delta v)\right) \le 0.
\end{aligned}$$

[3]For Dirichlet boundary conditions on piecewise convex domains, this result has been proved directly without the use of heat kernel estimates in in [9, Appendix].

Note that, since $v \in H^1(\mathcal{O})$, we have that

$$_S\langle (-\Delta)_\alpha v, \Delta v\rangle_{S^*} \le - \|(-\Delta)_\alpha v\|_H^2 .$$

To see this, just take into account that

$$0 \le \frac{1}{\alpha}(\nabla v - \nabla J_\alpha v, \nabla v - \nabla J_\alpha v)_{L^2(\mathcal{O};\mathbb{R}^d)} = ((-\Delta)_\alpha v, \Delta J_\alpha v)_H - {}_S\langle (-\Delta)_\alpha v, \Delta v\rangle_{S^*}.$$

Furthermore, by commutation (see Theorem 2.9 and Lemma 2.10),

$$\begin{aligned}
&\left((-\Delta)_\alpha v, (B_i^\delta)^2 v\right)_H \\
&= - \left((-\Delta)_\alpha v, J_\delta B_i^* B_i^\delta v\right)_H = - \left(B_i^\delta (-\Delta)_\alpha v, B_i^\delta v\right)_H \\
&= - \left(B_i^\delta (-\Delta)_\alpha^{\frac{1}{2}} v, B_i^\delta (-\Delta)_\alpha^{\frac{1}{2}} v\right)_H = - \left\|(-\Delta)_\alpha^{\frac{1}{2}} B_i^\delta v\right\|_H^2 .
\end{aligned}$$

By going back and replacing in (4.2) we get that $\mathbb{P}\otimes ds$-a.s.,

$$\begin{aligned}
\left\|X_t^{\varepsilon,\lambda,\delta}\right\|_\alpha^2 \quad \le \quad & \|x_0\|_\alpha^2 - 2\varepsilon \int_0^t \left\|(-\Delta)_\alpha X_s^{\varepsilon,\lambda,\delta}\right\|_H^2 ds \\
& - \sum_{i=1}^N \int_0^t \left\|(-\Delta)_\alpha^{\frac{1}{2}} B_i^\delta \left(X_s^{\varepsilon,\lambda,\delta}\right)\right\|_H^2 ds \\
& + 2\sum_{i=1}^N \int_0^t \left((-\Delta)_\alpha X_s^{\varepsilon,\lambda,\delta}, B_i^\delta X_s^{\varepsilon,\lambda,\delta} d\beta_s^i\right)_H \\
& + \sum_{i=1}^N \int_0^t \left\|(-\Delta)_\alpha^{\frac{1}{2}} B_i^\delta \left(X_s^{\varepsilon,\lambda,\delta}\right)\right\|_H^2 ds.
\end{aligned} \tag{4.3}$$

Taking the expectation and letting $\alpha \to 0$ yields

$$\mathbb{E}\left\|\nabla X_t^{\varepsilon,\lambda,\delta}\right\|_{L^2(\mathcal{O};\mathbb{R}^d)}^2 + 2\varepsilon\mathbb{E}\int_0^t \left\|\Delta X_s^{\varepsilon,\lambda,\delta}\right\|_H^2 ds \le \mathbb{E}\left\|\nabla x_0\right\|_{L^2(\mathcal{O};\mathbb{R}^d)}^2 . \tag{4.4}$$

Step II ($\delta \to 0$):

We shall pass to the limit in (4.1) for $\delta \to 0$ by using Theorem 2.2 from [7]. Note that Fatou's lemma (after passing on to an a.e. convergent subsequence) and $\Phi_\lambda(J_\delta \cdot) \le \Phi_\lambda(\cdot)$ (which holds e.g. by [32, Example 7.11]) imply that $\Phi_\lambda \circ J_\delta \longrightarrow \Phi_\lambda$ in Mosco sense as $\delta \to 0$ (for the terminology, see [4]). Therefore, we have that

$$J_\delta \operatorname{div} \Psi^\lambda\left(\nabla J_\delta(\cdot)\right) + \varepsilon\Delta(\cdot) \xrightarrow{G} \operatorname{div}\Psi^\lambda\left(\nabla(\cdot)\right) + \varepsilon\Delta(\cdot),$$

as $\delta \to 0$ and so for the corresponding inverse subdifferential operators[4]. Also, it is clear that for $v \in S$,

$$\nabla(J_\delta v) \longrightarrow \nabla v, \quad \text{strongly in } L^2(\mathcal{O};\mathbb{R}^d),$$

as $\delta \to 0$, which is sufficient for the strong convergence of the C_0-groups of linear operators associated to B_i^δ to the C_0-group associated to B_i, see e.g. [16]. Therefore, we can apply [7, Theorems 2.2 and 2.3] (note that we do not need that the

[4]Note that, $\Phi^n \to \Phi$ in Mosco sense implies that $\partial\Phi^n \to \partial\Phi$ in G-sense and $(\partial\Phi^n)^{-1} \to (\partial\Phi)^{-1}$ in G-sense, see [4, 6].

semigroups converge in $C^1([0,T];H)$, as we do not assume any time dependence for our noise coefficients) and obtain that $\mathbb{P}$-a.s. as $\delta \to 0$,

$$X^{\varepsilon,\lambda,\delta} \quad \longrightarrow \quad X^{\varepsilon,\lambda}, \quad \text{weakly in } L^2\left([0,T];S\right) \text{ and} \\ \text{weakly}^*\text{in } L^\infty\left([0,T];H\right).$$

Combining with (4.4), we get by weak lower semicontinuity of the norm that

$$\operatorname*{ess\,sup}_{t\in[0,T]} \mathbb{E}\left\|\nabla X_t^{\varepsilon,\lambda}\right\|^2_{L^2(\mathcal{O};\mathbb{R}^d)} + 2\varepsilon\mathbb{E}\int_0^T \left\|\Delta X_s^{\varepsilon,\lambda}\right\|_H^2 ds \le \mathbb{E}\left\|\nabla x_0\right\|^2_{L^2(\mathcal{O};\mathbb{R}^d)}.$$

We have proved that there exists a strong solution (in the sense of [7]) to

$$\tag{4.5} \begin{cases} dX_t^{\varepsilon,\lambda} = \operatorname{div}\Psi^\lambda\left(\nabla X_t^{\varepsilon,\lambda}\right)dt + \varepsilon\Delta X_t^{\varepsilon,\lambda}\,dt & \text{in } (0,T)\times\mathcal{O}, \\ \qquad +\frac{1}{2}\sum_{i=1}^N B_i^2(X_t^{\varepsilon,\lambda})dt + \sum_{i=1}^N B_i(X_t^{\varepsilon,\lambda})\,d\beta_t^i, & \\ X_0^{\varepsilon,\lambda} = x_0, & \text{in } \mathcal{O}, \\ \dfrac{\partial X_t^{\varepsilon,\lambda}}{\partial\nu} = 0, & \text{on } (0,T)\times\partial\mathcal{O}, \end{cases}$$

for initial datum $x_0 \in L^2(\Omega;S)$ which is of the particular form (3.4) as claimed in Remark 3.2.

Step III ($\lambda \to 0$):

By applying the Itô formula with $u \mapsto \frac{1}{2}\|u\|_H^2$ and the expectation to the difference

$$\begin{aligned} d\left(X_t^{\varepsilon,\lambda} - Z_t\right) &= \left(\operatorname{div}\left(\Psi^\lambda\left(\nabla X_t^{\varepsilon,\lambda}\right)\right) + \varepsilon\Delta X_t^{\varepsilon,\lambda} - G_t\right)dt \\ &\quad +\frac{1}{2}\sum_{i=1}^N\left(B_i^2(X_t^{\varepsilon,\lambda}) - B_i^2(Z_t)\right)dt + \sum_{i=1}^N B_i\left(X_t^{\varepsilon,\lambda} - Z_t\right)d\beta_t^i, \end{aligned}$$

for Z and G considered as in Definition 3.1, we see that $X^{\varepsilon,\lambda}$ is also a SVI solution to (4.1), i.e.

$$\tag{4.6} \begin{aligned} &\frac{1}{2}\mathbb{E}\left\|X_t^{\varepsilon,\lambda} - Z_t\right\|_H^2 + \mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi^\lambda\left(\nabla X_s^{\varepsilon,\lambda}\right)d\xi ds \\ &+\varepsilon\mathbb{E}\int_0^t\int_{\mathcal{O}}\left\langle\nabla X_s^{\varepsilon,\lambda}, \nabla X_s^{\varepsilon,\lambda} - \nabla Z_s\right\rangle d\xi ds \\ \le\ &\frac{1}{2}\mathbb{E}\left\|x_0 - Z_0\right\|_H^2 + \mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi^\lambda\left(\nabla Z_s\right)d\xi ds \\ &-\mathbb{E}\int_0^t\int_{\mathcal{O}} G_s\left(X_s^{\varepsilon,\lambda} - Z_s\right)d\xi ds. \end{aligned}$$

In order to pass to the limit we shall need the following a-priori estimates.

First we apply the Itô formula for the functional $u \mapsto \frac{1}{2}\|u\|_H^2$ to the equation

$$\text{(4.7)}\qquad \begin{cases} dX_t^{\varepsilon,\lambda} = \operatorname{div}\Psi^\lambda\left(\nabla X_t^{\varepsilon,\lambda}\right)dt + \varepsilon\Delta X_t^{\varepsilon,\lambda}dt & \text{in } (0,T)\times\mathcal{O} \\ \qquad\qquad + \frac{1}{2}\sum_{i=1}^N B_i^2(X_t^{\varepsilon,\lambda})dt + \sum_{i=1}^N B_i\left(X_t^{\varepsilon,\lambda}\right)d\beta_t^i, & \\ X_0^{\varepsilon,\lambda} = x_0, & \text{in } \mathcal{O} \\ \dfrac{\partial X_t^{\varepsilon,\lambda}}{\partial\nu} = 0, & \text{on } (0,T)\times\partial\mathcal{O} \end{cases}$$

in order to get that

$$\text{(4.8)}\qquad \begin{aligned} &\frac{1}{2}\mathbb{E}\left\|X_t^{\varepsilon,\lambda}\right\|_H^2 + \mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi^\lambda\left(\nabla X_s^{\varepsilon,\lambda}\right)d\xi ds + \varepsilon\mathbb{E}\int_0^t\int_{\mathcal{O}}\left|\nabla X_s^{\varepsilon,\lambda}\right|^2 d\xi ds \\ \le\ &\frac{1}{2}\mathbb{E}\|x_0\|_H^2, \quad \forall\lambda,\ t\in[0,T]. \end{aligned}$$

Moreover, in order to verify (3.3), we see that by the Mosco convergence (see e.g. [33, Proposition 6.2])

$$\Phi_\lambda = \int_{\mathcal{O}}\varphi^\lambda(\nabla\cdot)\,d\xi \longrightarrow \Phi \quad \text{in Mosco sense as } \lambda\to 0,$$

and Fatou's lemma (after passing to an a.e. convergent subsequence — strong L^2-convergence is justified below), we get that

$$\mathbb{E}\int_0^t\Phi\left(X_s^\varepsilon\right)ds \le \liminf_{\lambda\to 0}\mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi^\lambda\left(\nabla X_s^{\varepsilon,\lambda}\right)d\xi ds < \infty.$$

On the other hand, also by the Itô formula and Lemma 2.6 (i), we get that $\mathbb{P}$-a.s., $t\in[0,T]$,

$$\begin{aligned} &\frac{1}{2}\left\|X_t^{\varepsilon,\lambda_1} - X_t^{\varepsilon,\lambda_2}\right\|_H^2 \\ &+\int_0^t\left(\Psi^{\lambda_1}\left(\nabla X_s^{\varepsilon,\lambda_1}\right) - \Psi^{\lambda_2}\left(\nabla X_s^{\varepsilon,\lambda_2}\right), \nabla X_s^{\varepsilon,\lambda_1} - \nabla X_s^{\varepsilon,\lambda_2}\right)_{L^2(\mathcal{O};\mathbb{R}^d)}ds \\ &+\varepsilon\int_0^t\left\|\nabla X_s^{\varepsilon,\lambda_1} - \nabla X_s^{\varepsilon,\lambda_2}\right\|_{L^2(\mathcal{O};\mathbb{R}^d)}^2 ds \\ =\ &\sum_{i=1}^N\int_0^t\left(B_i\left(X_s^{\varepsilon,\lambda_1} - X_s^{\varepsilon,\lambda_2}\right)d\beta_s^i\ , X_s^{\varepsilon,\lambda_1} - X_s^{\varepsilon,\lambda_2}\right)_H. \end{aligned}$$

By using [33, eq. (A.6) in Appendix A], we have for all $\xi,\zeta\in\mathbb{R}^d$ and some positive constant $C>0$ that

$$\langle\Psi^{\lambda_1}(\xi) - \Psi^{\lambda_2}(\zeta), \xi-\zeta\rangle \ge -C\left(\lambda_1+\lambda_2\right)\left(1+|\xi|^2+|\zeta|^2\right).$$

We obtain $\mathbb{P}\otimes ds$-a.s. that

$$\begin{aligned} &\left(\Psi^{\lambda_1}\left(\nabla X^{\varepsilon,\lambda_1}\right) - \Psi^{\lambda_2}\left(\nabla X^{\varepsilon,\lambda_2}\right), \nabla X^{\varepsilon,\lambda_1} - \nabla X^{\varepsilon,\lambda_2}\right)_{L^2(\mathcal{O};\mathbb{R}^d)} \\ \ge\ & -C\left(\lambda_1+\lambda_2\right)\int_{\mathcal{O}}\left(1+\left|\nabla X^{\varepsilon,\lambda_1}\right|^2 + \left|\nabla X^{\varepsilon,\lambda_2}\right|^2\right)d\xi \\ \ge\ & -C\left(\lambda_1+\lambda_2\right)\left(1+\left|X^{\varepsilon,\lambda_1}\right|_S^2 + \left|X^{\varepsilon,\lambda_2}\right|_S^2\right), \end{aligned}$$

and then, by (4.4) and (4.8), we get for the expectation, that

$$\mathbb{E}\int_0^t \left(\Psi^{\lambda_1}\left(\nabla X_s^{\varepsilon,\lambda_1}\right) - \Psi^{\lambda_2}\left(\nabla X_s^{\varepsilon,\lambda_2}\right), \nabla X_s^{\varepsilon,\lambda_1} - \nabla X_s^{\varepsilon,\lambda_2}\right)_{L^2(\mathcal{O};\mathbb{R}^d)} ds \geq -C\left(\lambda_1+\lambda_2\right).$$

Now, by the Burkholder-Davis-Gundy inequality, taking (2.2) into account, and by the above computation concerning Ψ^λ, we obtain that for all $t \in [0,T]$,

$$\frac{1}{2}\mathbb{E}\sup_{0\leq s\leq t}\left\|X_s^{\varepsilon,\lambda_1} - X_s^{\varepsilon,\lambda_2}\right\|_H^2 + \varepsilon\mathbb{E}\int_0^t \left\|\nabla X_s^{\varepsilon,\lambda_1} - \nabla X_s^{\varepsilon,\lambda_2}\right\|_H^2 ds \leq C(\lambda_1+\lambda_2).$$

Consequently, we have that

$$\lim_{\lambda\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}\left\|X_t^{\varepsilon,\lambda} - X_t^{\varepsilon}\right\|_H^2\right] = 0. \tag{4.9}$$

We can now pass to the limit for $\lambda \to 0$ in (4.6) in order to obtain (recall that $\int_{\mathcal{O}} \varphi^\lambda(\cdot)\,d\xi \leq \Phi$)

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\left\|X_t^\varepsilon - Z_t\right\|_H^2 + \mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi\left(\nabla X_s^\varepsilon\right)d\xi ds \\
&+\varepsilon\mathbb{E}\int_0^t\int_{\mathcal{O}}\left\langle \nabla X_s^\varepsilon, \nabla X_s^\varepsilon - \nabla Z_s\right\rangle d\xi ds \\
\leq\; &\frac{1}{2}\mathbb{E}\left\|x_0 - Z_0\right\|_H^2 + \mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi\left(\nabla Z_s\right)d\xi ds \\
&-\mathbb{E}\int_0^t\int_{\mathcal{O}} G_s\left(X_s^\varepsilon - Z_s\right)d\xi ds.
\end{aligned} \tag{4.10}$$

Step IV ($\varepsilon \to 0$):

Arguing as in the previous step, we get that

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\left\|X_t^\varepsilon\right\|_H^2 + \mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi\left(\nabla X_s^\varepsilon\right)d\xi ds + \varepsilon\mathbb{E}\int_0^t\int_{\mathcal{O}}\left|\nabla X_s^\varepsilon\right|^2 d\xi ds \\
\leq\; &\frac{1}{2}\mathbb{E}\left\|x_0\right\|_H^2, \quad t\in[0,T].
\end{aligned} \tag{4.11}$$

By Itô's formula, this time, considering the process for fixed $\lambda > 0$ and not $\varepsilon > 0$ fixed, as previously, and by monotonicity, we get that

$$\begin{aligned}
&\frac{1}{2}\left\|X_t^{\varepsilon_1,\lambda} - X_t^{\varepsilon_2,\lambda}\right\|_H^2 e^{-t} \\
&+\int_0^t e^{-s}\left(\varepsilon_1\nabla X_s^{\varepsilon_1,\lambda} - \varepsilon_2\nabla X_s^{\varepsilon_2,\lambda}, \nabla X_s^{\varepsilon_1,\lambda} - \nabla X_s^{\varepsilon_2,\lambda}\right)_{L^2(\mathcal{O};\mathbb{R}^d)} ds \\
&+\frac{1}{2}\int_0^t e^{-s}\|X_s^{\varepsilon_1,\lambda} - X_s^{\varepsilon_2,\lambda}\|_H^2\, ds \\
\leq\; &\sum_{i=1}^N\int_0^t e^{-s}\left(B_i\left(X_s^{\varepsilon_1,\lambda} - X_s^{\varepsilon_2,\lambda}\right)d\beta_s^i\;, X_s^{\varepsilon_1,\lambda} - X_s^{\varepsilon_2,\lambda}\right)_H
\end{aligned}$$

Since $\mathbb{P}\otimes ds$-a.s.,

$$\begin{aligned}&\left(\varepsilon_1\nabla X^{\varepsilon_1,\lambda}-\varepsilon_2\nabla X^{\varepsilon_2,\lambda},\nabla X^{\varepsilon_1,\lambda}-\nabla X^{\varepsilon_2,\lambda}\right)_{L^2(\mathcal{O};\mathbb{R}^d)}\\ = \;& -\left(\varepsilon_1\Delta X^{\varepsilon_1,\lambda}-\varepsilon_2\Delta X^{\varepsilon_2,\lambda},X^{\varepsilon_1,\lambda}-X^{\varepsilon_2,\lambda}\right)_H\\ \geq \;& -\frac{1}{2}\left(\varepsilon_1^2\left\|\Delta X^{\varepsilon_1,\lambda}\right\|_H^2+\varepsilon_2^2\left\|\Delta X^{\varepsilon_2,\lambda}\right\|_H^2\right)-\frac{1}{2}\left\|X^{\varepsilon_1,\lambda}-X^{\varepsilon_2,\lambda}\right\|_H^2\end{aligned}$$

and by using again the Burkholder-Davis-Gundy inequality and (2.2), we get that for all $t\in[0,T]$,

$$\begin{aligned}&\mathbb{E}\sup_{0\leq s\leq t}e^{-s}\left\|X_s^{\varepsilon_1,\lambda}-X_s^{\varepsilon_2,\lambda}\right\|_H^2\\ \leq \;& \varepsilon_1^2\mathbb{E}\int_0^t e^{-s}\left\|\Delta X_s^{\varepsilon_1,\lambda}\right\|_H^2 ds+\varepsilon_2^2\mathbb{E}\int_0^t e^{-s}\left\|\Delta X_s^{\varepsilon_2,\lambda}\right\|_H^2 ds.\end{aligned}$$

Keeping in mind that for initial data in $L^2(\Omega;S)$, by *Step II* above, in particular, by (4.4),

$$\varepsilon\mathbb{E}\int_0^T\left\|\Delta X_s^{\varepsilon,\lambda}\right\|_H^2 ds\leq C, \tag{4.12}$$

uniformly in $\lambda>0$, we obtain by (4.9), for letting first $\lambda\to 0$,

$$\lim_{\varepsilon\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}\|X_t^\varepsilon-X_t\|_H^2\right]=0$$

for some limiting process $X\in C([0,T];L^2(\Omega;H))$. Note that by weak convergence in $L^2([0,T];S)$, we get that $\mathbb{P}\otimes dt$-a.s. $X\in H^1(\mathcal{O})$.

Finally, by computing $\mathbb{P}\otimes ds$-a.e.

$$\begin{aligned}\varepsilon\int_{\mathcal{O}}\langle\nabla X^\varepsilon,\nabla X^\varepsilon-\nabla Z\rangle\, d\xi &= -\varepsilon\int_{\mathcal{O}}\Delta X^\varepsilon\left(X^\varepsilon-Z\right)d\xi\\ &\geq -\frac{1}{2}\varepsilon^{\frac{4}{3}}\|\Delta X^\varepsilon\|_H^2-\frac{1}{2}\varepsilon^{\frac{2}{3}}\|X^\varepsilon-Z\|_H^2\end{aligned}$$

and using again (4.12) we can pass to the limit in (4.10) and get that X is a continuous SVI solution (which satisfies (3.3) by passing to the limit in (4.11))

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\|X_t-Z_t\|_H^2+\mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi\left(\nabla X_s\right)d\xi ds\\ \leq\;& \frac{1}{2}\mathbb{E}\|x_0-Z_0\|_H^2+\mathbb{E}\int_0^t\int_{\mathcal{O}}\varphi\left(\nabla Z_s\right)d\xi ds\\ &-\mathbb{E}\int_0^t\int_{\mathcal{O}}G_s\left(X_s-Z_s\right)d\xi ds.\end{aligned}$$

Step V (general initial conditions):

In order to conclude the proof of existence we only need to extend the solution for arbitrary $x\in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$. Let X,X^* be continuous SVI solutions starting in x,x^*, resp. Note that S is dense in H, so this follows directly from

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|X_t-X_t^*\|_H^2\right]\leq\mathbb{E}\|x_0-x_0^*\|_H^2,$$

which can easily be obtained by arguments similar to those from the previous steps (using monotonicity, Burkholder-Davis-Gundy and (2.2)).

□

4.2. **Uniqueness.**

Proof of Theorem 4.1 (continued). The existence of a continuous SVI solution is proved in the section above. We follow an argument from [33]. Let X be any continuous SVI solution to (1.1) with initial condition $x \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ and let $Y^{\varepsilon,\lambda,n}$ be the strong approximating solution to (4.5) with initial condition $y^n \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; S)$ such that $y^n \to y$ in $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, where $y \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Then the following variational inequality holds (with $Z_0 = y^n$, $Z = Y^{\varepsilon,\lambda,n}$, $G = \operatorname{div} \Psi^\lambda\left(\nabla Y^{\varepsilon,\lambda,n}\right) + \varepsilon\Delta Y^{\varepsilon,\lambda,n}$),

$$\begin{aligned}\frac{1}{2}\mathbb{E}\|X_t - Y_t^{\varepsilon,\lambda,n}\|_H^2 &+ \mathbb{E}\int_0^t \Phi(X_s)\,ds \\ &\leq \frac{1}{2}\mathbb{E}\|x-y^n\|_H^2 + \mathbb{E}\int_0^t \Phi(Y_s^{\varepsilon,\lambda,n})\,ds \\ &\quad - \mathbb{E}\int_0^t (\varepsilon\Delta Y_s^{\varepsilon,\lambda,n} + \operatorname{div}\Psi^\lambda(\nabla Y_s^{\varepsilon,\lambda,n}), X_s - Y_s^{\varepsilon,\lambda,n})_H\,ds,\end{aligned}$$

for a.e. $t \in [0,T]$.

By [33, Appendix A] for all $z \in H^1$ we have

$$-(\operatorname{div}\Psi^\lambda(\nabla Y^{\varepsilon,\lambda,n}), z - Y^{\varepsilon,\lambda,n})_H + \Phi(Y^{\varepsilon,\lambda,n}) \leq \Phi(z) + C\lambda(1+\Phi(Y^{\varepsilon,\lambda,n})) \quad ds\otimes\mathbb{P}\text{-a.e.}$$

Since Φ is the lower-semicontinuous envelope of $\tilde{\Phi} = \Phi|_{H^1}$ (i.e., Φ restricted to H^1), for $ds \otimes \mathbb{P}$-a.e. $(s,\omega) \in [0,T]\times\Omega$, we can choose a sequence $z^m \in H^1$ such that $z^m \to X_s(\omega)$ in H and $\Phi(z^m) \to \Phi(X_s(\omega))$.

Hence,

$$-(\operatorname{div}\Psi^\lambda(\nabla Y^{\varepsilon,\lambda,n}), X - Y^{\varepsilon,\lambda,n})_H + \Phi(Y^{\varepsilon,\lambda,n}) \leq \Phi(X) + C\lambda(1+\Phi(Y^{\varepsilon,\lambda,n})) \quad ds\otimes\mathbb{P}\text{-a.e.}$$

Thus,

$$\begin{aligned}\frac{1}{2}\mathbb{E}\|X_t - Y_t^{\varepsilon,\lambda,n}\|_H^2 &\leq \frac{1}{2}\mathbb{E}\|x-y^n\|_H^2 + C\lambda\mathbb{E}\int_0^t \left(1+\Phi(Y_s^{\varepsilon,\lambda,n})\right)ds \\ &\quad + \frac{1}{2}\mathbb{E}\int_0^t \left(\varepsilon^{\frac{4}{3}}\|\Delta Y_s^{\varepsilon,\lambda,n}\|_H^2 + \varepsilon^{\frac{2}{3}}\|X_s - Y_s^{\varepsilon,\lambda,n}\|_H^2\right)ds.\end{aligned}$$

Taking first $\lambda \to 0$ then $\varepsilon \to 0$ (using the H^2 bound (4.12), which is uniform in λ) and then $n \to \infty$ yields

$$\mathbb{E}\|X_t - Y_t\|_H^2 \leq \mathbb{E}\|x-y\|_H^2,$$

for a.e. $t \in [0,T]$.

□

Appendix A. Vector fields of Killing

In this section, suppose that $\mathcal{O} \subset \mathbb{R}^d$, open, bounded, with smooth boundary $\partial\mathcal{O}$.

Definition A.1. A C^1-vector field $b : \overline{\mathcal{O}} \to \mathbb{R}^d$, is called *Killing vector field*, if the following condition is satisfied on $\overline{\mathcal{O}}$

$$\partial_j b^i + \partial_i b^j = 0 \quad \forall 1 \leq i,j \leq d. \tag{A.1}$$

The following lemma is based on ideas from [54], see also [53].

Lemma A.2. *A (sufficiently smooth) vector field $b : \overline{\mathcal{O}} \to \mathbb{R}^d$ is a Killing vector field if and only if*

$$\langle \Delta b, b\rangle = 0, \quad \operatorname{div} b = 0 \quad on\ \overline{\mathcal{O}}. \tag{A.2}$$

$$\sum_{1\le i,j\le d} (\partial_j b^i + \partial_i b^j)\nu_j b^i = 0 \quad on\ \partial\mathcal{O}, \tag{A.3}$$

where the Laplace acts componentwise and where $\nu : \partial\mathcal{O} \to \mathbb{R}^d$ denotes the outer unit normal of $\overline{\mathcal{O}}$.

Proof. One easily sees that,

$$\begin{aligned}
&\partial_j[(\partial_j b^i + \partial_i b^j)b^i - b^j(\partial_i b^i)]\\
=&(\partial_j\partial_j b^i)b^i + (\partial_j b^i)(\partial_j b^i)\\
&+ (\partial_j\partial_i b^j)b^i + (\partial_i b^j)(\partial_j b^i)\\
&- (\partial_j b^j)(\partial_i b^i) - b^j(\partial_j\partial_i b^i).
\end{aligned}$$

Assuming the above conditions, interchanging the order of differentiation, and summing over $1 \le i, j \le d$, we obtain that

$$\begin{aligned}
&\sum_{1\le i,j\le d} \partial_j[(\partial_j b^i + \partial_i b^j)b^i - b^j(\partial_i b^i)]\\
=&\langle \Delta b, b\rangle + \sum_{1\le i,j\le d} [(\partial_j b^i)^2 + (\partial_i b^j)(\partial_j b^i)] - (\operatorname{div} b)^2\\
&+ \sum_{1\le i,j\le d} (\partial_j\partial_i b^j)b^i - (\partial_i\partial_j b^i)b^j\\
=&\frac{1}{2}\sum_{1\le i,j\le d} (\partial_j b^i)^2 + 2(\partial_i b^j)(\partial_j b^i) + (\partial_i b^j)^2.\\
=&\frac{1}{2}\sum_{1\le i,j\le d} (\partial_j b^i + \partial_i b^j)^2.
\end{aligned}$$

Denote by S^{d-1} the surface element on $\partial\mathcal{O}$. By Gauss's divergence theorem, we get that,

$$\begin{aligned}
&\frac{1}{2}\sum_{1\le i,j\le d} \int_{\overline{\mathcal{O}}} (\partial_j b^i + \partial_i b^j)^2\, d\xi\\
=&\sum_{1\le i,j\le d} \int_{\partial\mathcal{O}} [(\partial_j b^i + \partial_i b^j)b^i - b^j(\partial_i b^i)]\nu_j\, dS^{d-1}\\
=&\sum_{1\le i,j\le d} \int_{\partial\mathcal{O}} [(\partial_j b^i + \partial_i b^j)b^i]\nu_j\, dS^{d-1}\\
=&0,
\end{aligned}$$

and hence $\partial_j b^i + \partial_i b^j = 0$ on $\overline{\mathcal{O}}$.

Suppose conversely, that b is a Killing vector field. Then (A.3) is automatically satisfied. Also, $\operatorname{div} b = 0$ by choosing $i = j$. Clearly, also

$$0 = b^i\partial_j\partial_j b^i + b^i\partial_j\partial_i b^j = b^i\partial_j\partial_j b^i + b^i\partial_i\partial_j b^j,$$

and summing over $1 \leq i, j \leq d$ yields,

$$\langle \Delta b, b \rangle = 0,$$

and hence (A.2) is satisfied, too. □

Theorem A.3. *Let $b : \overline{\mathcal{O}} \to \mathbb{R}^d$ be a C^1-vector field. In order that the first order differential operator $u \mapsto \langle b, \nabla u \rangle$ commutes with the Laplace operator $u \mapsto -\Delta u$ on the space of smooth functions on $\overline{\mathcal{O}}$ it is necessary and sufficient that b is a Killing vector field.*

Proof. See [47, Theorem 2.1]. □

Acknowledgements

The first author would like to thank Viorel Barbu and Witold Respondek for helpful comments. The second author would like to thank the colleagues at the Laboratoire de Mathématiques de l'INSA de Rouen for a pleasant stay at their department where part of this work was done. He also would like to express his gratitude to Benjamin Gess, who made some important remarks on a preliminary version of this work. Both authors are grateful for some clarifying remarks by the referee.

References

[1] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.*, 20(3):681–695, 2011.

[2] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential and Integral Equations. An International Journal for Theory & Applications*, 14(3):321–360, 2001.

[3] F. Andreu-Vaillo, V. Caselles, and J. M. Mazón. *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*. Progress in Mathematics. Birkhäuser Basel, 2012.

[4] H. Attouch. *Variational convergence for functions and operators*. Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[5] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Applications to PDEs and optimization.

[6] V. Barbu. *Nonlinear differential equations of monotone types in Banach spaces*. Springer Monographs in Mathematics. Springer, 2010.

[7] V. Barbu, Z. Brzeźniak, E. Hausenblas, and L. Tubaro. Existence and convergence results for infinite dimensional nonlinear stochastic equations with multiplicative noise. *Stochastic Process. Appl.*, 123(3):934–951, 2013.

[8] V. Barbu, G. Da Prato, and M. Röckner. Stochastic nonlinear diffusion equations with singular diffusivity. *SIAM J. Math. Anal.*, 41(3):1106–1120, 2009.

[9] V. Barbu and M. Röckner. Stochastic variational inequalities and applications to the total variation flow perturbed by linear multiplicative noise. *Arch. Rational Mech. Anal.*, 209(3):797–834, 2013.

[10] V. Barbu and M. Röckner. An operatorial approach to stochastic partial differential equations driven by linear multiplicative noise. *J. Eur. Math. Soc.*, pages 1789–1815, 2015.

[11] X. Bardina and M. Jolis. An extension of Ito's formula for elliptic diffusion processes. *Stochastic Processes and their Applications*, 69(1):83–109, 1997.

[12] X. Bardina and C. Rovira. On Itô's formula for elliptic diffusion processes. *Bernoulli*, 13(3):820–830, 2007.

[13] L. Beck, F. Flandoli, M. Gubinelli, and M. Maurelli. Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Preprint*, pages 1–59, 2014. arXiv:1401.1530, `http://arxiv.org/abs/1401.1530`.

[14] F. E. Benth, T. Deck, J. Potthoff, and L. Streit. Nonlinear evolution equations with gradient coupled noise. *Letters in Mathematical Physics*, 43:267–278, 1998.

[15] V. Bögelein, F. Duzaar, and P. Marcellini. A time dependent variational approach to image restoration. *SIAM Journal on Imaging Sciences*, 8(2):968–1006, 2015.

[16] H. Brézis and A. Pazy. Convergence and approximation of semigroups of nonlinear operators in Banach spaces. *J. Functional Analysis*, 9:63–74, 1972.

[17] M. Caruana, P. K. Friz, and H. Oberhauser. A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(1):27–46, 2011.

[18] Y.-Z. Chen. Hölder continuity of the gradient of solutions of nonlinear degenerate parabolic systems. *Acta Mathematica Sinica*, 2(4):309–331, 1986.

[19] Y.-Z. Chen and E. DiBenedetto. Hölder estimates of solutions of singular parabolic equations with measurable coefficients. *Archive for Rational Mechanics and Analysis*, 118(3):257–271, 1992.

[20] G. Da Prato, M. Iannelli, and L. Tubaro. An existence result for a linear abstract stochastic equation in Hilbert spaces. *Rendiconti del Seminario Matematico dell'Università di Padova*, 67:171–180, 1982.

[21] G. Da Prato, M. Iannelli, and L. Tubaro. Some results on linear stochastic differential equations in Hilbert spaces. *Stochastics*, 6(2):105–116, 1982.

[22] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[23] S. S. de León and C. M. Webler. Global existence and uniqueness for the inhomogeneous 1-Laplace evolution equation. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1213–1246, 2015.

[24] F. Demengel and G. Demengel. Functional Spaces for the Theory of Elliptic Partial Differential Equations. Springer London, London, 2012.

[25] E. Di Benedetto. *Degenerate parabolic equations*. Universitext. Springer, Berlin–Heidelberg–New York, 1993.

[26] H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *C. R. Acad. Sci. Paris Sér. A-B*, 283(13):Ai, A939–A942, 1976.

[27] N. Eisenbaum. On Itô's formula of Föllmer and Protter. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 390–395. Springer, Berlin, 2001.

[28] H. Föllmer, P. Protter, and A. N. Shiryayev. Quadratic Covariation and an Extension of Itô's Formula. *Bernoulli*, 1(1/2):149, 1995.

[29] P. K. Friz and B. Gess. Stochastic scalar conservation laws driven by rough paths. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis, in press*, 2015. http://dx.doi.org/10.1016/j.anihpc.2015.01.009.

[30] B. Gess and M. Röckner. Stochastic variational inequalities and regularity for degenerate stochastic partial differential equations. *Preprint*, pages 1–26, 2014. arXiv:1405.5866, http://arxiv.org/abs/1405.5866.

[31] B. Gess and M. Röckner. Singular-Degenerate Multivalued Stochastic Fast Diffusion Equations. *SIAM Journal on Mathematical Analysis*, 47(5):4058–4090, 2015.

[32] B. Gess and J. M. Tölle. Multi-valued, singular stochastic evolution inclusions. *J. Math. Pures Appl.*, 101(6):789–827, 2014.

[33] B. Gess and J. M. Tölle. Stability of solutions to stochastic partial differential equations. *J. Differential Equations*, 260(6):4973–5025, 2016.

[34] H. Hakula, N. Hyvönen, and M. Leinonen. Reconstruction algorithm based on stochastic Galerkin finite element method for electrical impedance tomography. *Inverse Problems*, 30(6):065006, 2014. http://stacks.iop.org/0266-5611/30/i=6/a=065006.

[35] B. C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. Springer, 2015.

[36] J. U. Kim. On a stochastic singular diffusion equation in R^d. *J. Funct. Anal.*, 262(6):2928–2977, 2012.

[37] A. Krause, M. Lewis, and B. Wang. Dynamics of the non-autonomous stochastic p-Laplace equation driven by multiplicative noise. *Applied Mathematics and Computation*, 246(C):365–376, 2014.

[38] P. Krée. Diffusion equation for multivalued stochastic differential equations. *J. Funct. Anal.*, 49(1):73–90, 1982.
[39] A. Kuba and G. T. Herman. Introduction. In *Advances in discrete tomography and its applications*, Appl. Numer. Harmon. Anal., pages 1–16. Birkhäuser Boston, Boston, MA, 2007.
[40] R. Mikulevicius and B. L. Rozovskii. Global L_2-solutions of stochastic Navier-Stokes equations. *Ann. Probab.*, 33(1):137–176, 2005.
[41] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin–Heidelberg–New York, 2007.
[42] M. Röckner and I. Munteanu. The total variation flow perturbed by gradient linear multiplicative noise. *Preprint*, pages 1–22, 2016.
[43] T. Sakai. *Riemannian Geometry*. Fields Institute Communications. American Mathematical Soc., 1996.
[44] I. Shigekawa. Semigroup domination on a Riemannian manifold with boundary. *Acta Appl. Math.*, 63(1-3):385–410, 2000.
[45] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*. Mathematical surveys and monographs, Amer. Math. Soc., 1997.
[46] B. Sixou, L. Wang, and F. Peyrin. Stochastic diffusion equation with singular diffusivity and gradient-dependent noise in binary tomography. *J. Phys.: Conf. Ser.*, 542. conference 1:1–6, 2014.
[47] T. Sumitomo. On the commutator of differential operators. *Hokkaido Mathematical Journal*, 1(1):30–42, 1972.
[48] H. J. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probability*, 6(1):19–41, 1978.
[49] F.-Y. Wang. Second fundamental form and gradient of Neumann semigroups. *J. Funct. Anal.*, 256(10):3461–3469, 2009.
[50] F.-Y. Wang and L. Yan. Gradient estimate on convex domains and applications. *Proc. Amer. Math. Soc.*, 141(3):1067–1081, 2013.
[51] L. Wang, B. Sixou, and F. Peyrin. Binary tomography reconstructions of bone microstructure from few projections with stochastic level-set methods. *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1778–1782, 2014.
[52] Y. Wang, J. Yang, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.*, 1(3):248–272, 2008.
[53] K. Yano. Harmonic and Killing vector fields in compact orientable Riemannian spaces with boundary. *Annals of Mathematics. Second Series*, 69(3):588–597, 1959.
[54] K. Yano and M. Ako. Vector fields in Riemannian and Hermitian manifolds with boundary. *Kodai Mathematical Seminar Reports*, 17(3):129–157, 1965.

Normandie Univ, INSA Rouen, LMI, 76000 Rouen, France
E-mail address: ioana.ciotir@insa-rouen.fr

Department of Mathematics and Systems Analysis, School of Science, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland
E-mail address: jonas.tolle@aalto.fi