SINGULAR-DEGENERATE MULTIVALUED STOCHASTIC FAST DIFFUSION EQUATIONS

BENJAMIN GESS

Department of Mathematics
University of Chicago
USA

MICHAEL RÖCKNER

Faculty of Mathematics
University of Bielefeld
Germany

Abstract. We consider singular-degenerate, multivalued stochastic fast diffusion equations with multiplicative Lipschitz continuous noise. In particular, this includes the stochastic sign fast diffusion equation arising from the Bak-Tang-Wiesenfeld model for self-organized criticality. A well-posedness framework based on stochastic variational inequalities (SVI) is developed, characterizing solutions to the stochastic sign fast diffusion equation, previously obtained in a limiting sense only. Aside from generalizing the SVI approach to stochastic fast diffusion equations we develop a new proof of well-posedness, applicable to general diffusion coefficients. In case of linear multiplicative noise, we prove the existence of (generalized) strong solutions, which entails higher regularity properties of solutions than previously known.

1. Introduction

We consider singular-degenerate, multivalued stochastic fast diffusion equations (SFDE) of the type

$$
\begin{aligned}
dX_t &\in \Delta(|X_t|^{m-1}X_t)dt + B(t,X_t)dW_t, \\
X_0 &= x_0
\end{aligned}
\tag{1.1}
$$

with $m \in [0,1]$, on bounded, smooth domains $\mathcal{O} \subseteq \mathbb{R}^d$ with zero Dirichlet boundary conditions, where $|r|^{-1}r := \mathrm{Sgn}(r)$ denotes the maximal monotone extension of the sign function. In particular, we include the multivalued case $m = 0$ and general diffusion coefficients B. In the following W is a cylindrical Wiener process on some separable Hilbert space U and the diffusion coefficients $B : [0,T] \times H^{-1} \times$

E-mail addresses: gess@uchicago.edu, roeckner@mathematik.uni-bielefeld.de.

Date: December 24, 2014.

2010 *Mathematics Subject Classification.* Primary: 60H15; Secondary: 76S05.

Key words and phrases. singular degenerate SPDE, multivalued SPDE, self-organized criticality, stochastic fast diffusion, sign fast diffusion, regularity, stochastic variational inequalities.

Acknowledgements: We thank the DFG for support through the research project “Random dynamical systems and regularization by noise for stochastic partial differential equations” and through CRC 701. B.G. thanks Jonas Tölle for stimulating discussion.

$\Omega \to L_2(U, H^{-1})$ take values in the space of Hilbert-Schmidt operators $L_2(U, H^{-1})$, where H^{-1} is the dual of $H_0^1(\mathcal{O})$.

Our main results are twofold: First, in the case of general diffusion coefficients and initial data we introduce a notion of stochastic variational inequalities (SVI) to (1.1) and establish a new method of proof of well-posedness. In particular, this new methods allows treatment of general diffusion coefficients, whereas previously the approach of SVI solutions was restricted to additive or linear multiplicative noise (cf. (1.3), (1.4) below). In this sense, our results generalize those of [3, 7]. The second main result yields regularity properties of solutions in the case of linear multiplicative noise (cf. (1.5) below). In particular, we prove the existence of strong solutions, which extends the results from [8] from the degenerate case $m > 1$ to the singular case $m \in [0, 1]$.

In the case $m > 0$ a variational approach to (1.1) has been developed in [14] for $x_0 \in L^2(\Omega; H^{-1})$ based on the coercivity property

$$_{V^*}\langle \Delta(|v|^{m-1}v), v\rangle_V \geq c\|v\|_V^{m+1} \quad \forall v \in V = L^{m+1}(\mathcal{O}).$$

In the multivalued limiting case $m = 0$ two complications appear: First, the reflexivity of the energy space $L^{m+1}(\mathcal{O})$ is lost, making the variational methods from [14] inapplicable in this case. Second, the operator $\Delta(|v|^{m-1}v) = \Delta\mathrm{Sgn}(v)$ becomes multivalued. Recently, for regular initial data $x_0 \in L^2(\Omega; L^2(\mathcal{O}))$ an alternative variational approach to (1.1) has been developed in [11], proving well posedness for $m \in [0, 1]$. However, for general initial conditions $x_0 \in L^2(\Omega; H^{-1})$ solutions could be constructed in a limiting sense only. That is, it has been shown that for each approximating sequence $x_0^n \in L^2(\Omega; L^2(\mathcal{O}))$ with $x_0^n \to x$ in $L^2(\Omega; H^{-1})$ the corresponding variational solutions X^n converge to a limit X independent of the chosen approximating sequence x_0^n. The characterization of X in terms of a generalized notion of solution to (1.1) remained open. This problem is solved in this paper by introducing a notion of stochastic variational inequalities for (1.1) and proving well-posedness in this framework. The limiting solution X is thus characterized as an SVI solution to (1.1).

The difficulties for (1.1) as described above are similar to the ones for the stochastic total variation flow

$$\begin{aligned} dX_t &\in \mathrm{div}\left(\frac{\nabla X_t}{|\nabla X_t|}\right) dt + B(t, X_t) dW_t \\ X_0 &= x_0. \end{aligned} \tag{1.2}$$

As for (1.1), in case of regular initial data $x_0 \in L^2(\Omega; H_0^1(\mathcal{O}))$ variational solutions have been constructed in [11]. For general initial data $x_0 \in L^2(\Omega; L^2(\mathcal{O}))$ solutions to (1.2) could be constructed in a limiting sense only. In the special case of additive noise, i.e.

$$dX_t \in \mathrm{div}\left(\frac{\nabla X_t}{|\nabla X_t|}\right) dt + dW_t \tag{1.3}$$

and assuming $d = 1, 2$, a notion of SVI solution to (1.3) has been introduced in [3]. Only recently, well-posedness of SVI solutions and thus characterization of limiting solutions in the linear multiplicative case

$$dX_t \in \mathrm{div}\left(\frac{\nabla X_t}{|\nabla X_t|}\right) dt + \sum_{k=1}^{\infty} f_k X_t d\beta_t^k \tag{1.4}$$

has been shown in [7]. In this sense, our results on well-posedness of SVI solutions to (1.1) parallel those of [3, 7] in the case of stochastic fast diffusion equations. In both cases (1.3) and (1.4), the SPDE may be transformed into a random PDE, i.e.

a PDE with random coefficients. This technique is a crucial ingredient in the proofs given in [3, 7] and requires the restriction to either additive or linear multiplicative noise. In contrast to this, in the first part of this paper (see Section 2) we consider (1.1) for general multiplicative noise and introduce an alternative method to prove well-posedness of SVI solutions that does not rely on a transformation into a random PDE. This allows to treat general noise, while significantly simplifying the proof as compared to [7]. Moreover, in contrast to [3] no restrictions on the dimension d will be required.

In the second part of this paper (see Section 3) we prove regularity properties for solutions to (1.1) in the case of linear multiplicative noise, i.e. for

$$\begin{aligned} dX_t &\in \Delta(|X_t|^{m-1}X_t)dt + \sum_{k=1}^{\infty} f_k X_t d\beta_t^k, \\ X_0 &= x_0, \end{aligned} \tag{1.5}$$

with $m \in [0,1]$, $d \in \mathbb{N}$ and $x_0 \in L^2(\Omega; H^{-1})$. More precisely, we prove the existence of (generalized) strong solutions (cf. Definition A.1 below), in particular implying that X takes values in the domain of $\Delta(|\cdot|^{m-1}\cdot)$, $dt \otimes \mathbb{P}$-almost everywhere. This extends regularity results obtained in [8] where the degenerate case $m \geq 1$ was considered by entirely different methods. The case of singular diffusions ($m \in [0,1)$) could not be handled in [8] due to the singularity of the non-linearity $\phi(r) = |r|^{m-1}r$ at zero. Roughly speaking, this singularity has to be compensated by sufficient decay of the diffusion coefficients at zero; a problem not appearing in degenerate, non-singular cases treated in [10]. This requires a careful choice of approximating problems and leads to entirely different methods than those developed in [8]. In particular, it turns out that different approximations of the nonlinearity ϕ need to be considered in the proof of regularity for (1.5) and in the proof of well-posedness of SVI solutions for (1.1). We underline that also the methods developed in the second part of this paper do not rely on a transformation of (1.5) into a random PDE and therefore depend only loosely on the linear structure of the noise in (1.5). In fact, as pointed out above, the crucial structural condition in (1.5) is *not* the linearity of the diffusion coefficients, but their decay behavior at zero.

Stochastic fast diffusion equations of the type (1.1) have been intensively investigated in recent years. For the single-valued case $m > 0$ we refer to [12, 14] and the references therein. As the multivalued, limiting case $m = 0$ is concerned, mostly the case of linear multiplicative noise (1.5) has been considered in the literature. Well-posedness for regular initial data $x_0 \in L^4(\mathcal{O})$ and $d = 1, 2, 3$ was first proven in [4]. Finite time extinction for (1.5) has been investigated in [2, 4–6, 9, 17]. For bounded initial data $x_0 \in L^\infty(\mathcal{O})$ and finite driving noise, that is $f_k \equiv 0$ for all k large enough, the existence of strong solutions (cf. Definition A.1 below) to (1.5) has been proven in [9] by entirely different methods, relying on a transformation of (1.5) into a random PDE. Well-posedness for (1.1) with $m = 1$ and with general multiplicative noise has been obtained in [11] for the first time, proving well-posedness in terms of variational solutions for regular initial data $x_0 \in L^2(\mathcal{O})$. For background on the deterministic fast diffusion equation we refer to [18, 19] and the references therein.

1.1. **Notation.** In the following let $\mathcal{O} \subseteq \mathbb{R}^d$ be a bounded set with smooth boundary. $L^p := L^p(\mathcal{O})$ denotes the usual Lebesgue space with norm $\|\cdot\|_{L^p}$ and inner product $(\cdot,\cdot)_2$ if $p = 2$. Further, $H_0^1 := H_0^1(\mathcal{O})$ denotes the Sobolev space of order one in L^2 equipped with the inner product $(v,w)_{H_0^1} = (\nabla v, \nabla w)_2$ and norm $\|\cdot\|_{H_0^1}$. Let $(H^{-1}, (\cdot,\cdot)_{H^{-1}})$ with norm $\|\cdot\|_{H^{-1}}$ be the dual of H_0^1. Moreover, we let $C_0(\bar{\mathcal{O}})$

denote the set of all continuous functions on $\mathcal{O}$ vanishing at the boundary. In the proofs, as usual, constants may change from line to line.

2. Stochastic variational inequalities

In this section we consider stochastic singular fast diffusion equations of the type

$$dX_t \in \Delta\left(|X_t|^{m-1}X_t\right)dt + B(t,X_t)dW_t, \\ X_0 = x_0 \tag{2.1}$$

for $m \in [0,1]$, on a bounded, smooth domain $\mathcal{O} \subseteq \mathbb{R}^d$ with Dirichlet boundary conditions and general diffusion coefficients B, in particular including additive and linear multiplicative noise. The precise definition of the nonlinear part on the right hand side of (2.1) including its domain, as well as the definition of a solution to (2.1) will be given below. We emphasize that the multivalued, limiting case $m = 0$ is included.

Here W is a cylindrical Wiener process in some separable Hilbert space U defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t\geq 0}$ and the diffusion coefficients B take values in the space of Hilbert-Schmidt operators $L_2(U,H)$. As compared to the regularity results obtained in Section 3 below, for this general choice of diffusion coefficients, we cannot expect (generalized) strong solutions to exist for arbitrary initial conditions $x_0 \in H^{-1}$. Instead we introduce a notion of stochastic variational inequalities for (2.1) which we prove to uniquely characterize solutions.

We suppose that $B : [0,T] \times H^{-1} \times \Omega$ is progressively measurable and satisfies

$$\|B(t,v) - B(t,w)\|^2_{L_2(U,H^{-1})} \leq C\|v-w\|^2_{H^{-1}} \quad \forall v,w \in H^{-1} \\ \|B(t,v)\|^2_{L_2(U,L^2)} \leq C(1+\|v\|^2_{L^2}) \quad \forall v \in L^2, \tag{2.2}$$

for some constant $C > 0$ and all $(t,\omega) \in [0,T] \times \Omega$.

Let $\mathcal{M}$ be the space of all signed Radon measures on $\mathcal{O}$ with finite total variation. For $\mu \in \mathcal{M}$ we let $|\mu|$ be its variation with total variation

$$\mathrm{TV}(\mu;\mathcal{O}) = |\mu|(\mathcal{O}).$$

We define (cf. Appendix D)

$$L^{m+1} \cap H^{-1} := \left\{ v \in L^{m+1} | \int vh dx \leq C\|h\|_{H_0^1},\ \forall h \in C_c^1(\mathcal{O}) \text{ for some } C \geq 0 \right\}$$

$$\mathcal{M} \cap H^{-1} := \left\{ \mu \in \mathcal{M} | \int h(x) d\mu(x) \leq C\|h\|_{H_0^1},\ \forall h \in C_c^1(\mathcal{O}) \text{ for some } C \geq 0 \right\}.$$

Note that $\mathcal{M} \cap H^{-1}$ is known as the space of finite measures of bounded energy (cf. e.g. [16]). Clearly, we have $L^{m+1} \cap H^{-1} \subseteq H^{-1}$ and $\mathcal{M} \cap H^{-1} \subseteq H^{-1}$. For $m > 0$ and $v \in H^{-1}$ we define $\varphi : H^{-1} \to [0,\infty]$ by

$$\varphi(v) := \begin{cases} \frac{1}{m+1}\|v\|^{m+1}_{m+1} & , \quad v \in L^{m+1} \cap H^{-1} \\ +\infty & , \quad \textit{otherwise.} \end{cases}$$

By Lemma D.1 in Appendix D below, φ defines a convex, lower-semicontinuous function on H^{-1}. Moreover, by Lemma D.2

$$\varphi(\mu) := \begin{cases} \mathrm{TV}(\mu;\mathcal{O}) & , \quad \mu \in \mathcal{M} \cap H^{-1} \\ +\infty & , \quad \textit{otherwise.} \end{cases}$$

is the lower-semicontinuous hull on H^{-1} of $\|\cdot\|_1$ defined on $L^1 \cap H^{-1}$. For $m \geq 0$ we set

$$\psi(r) = \frac{1}{m+1}|r|^{m+1}, \quad r \in \mathbb{R},$$

and note that

$$\partial\psi(r) = |r|^m \mathrm{Sgn}(r) =: \phi(r), \quad r \in \mathbb{R}$$

where Sgn is the maximal monotone, multivalued extension of the sign function.

Concerning the subgradient $\partial\varphi$ of φ we have (see Lemma D.3 below)

$$\partial\varphi(v) \supseteq \{-\Delta w | w \in H_0^1, w \in \phi(v) \text{ a.e.}\},$$

for all $v \in L^{m+1} \cap H^{-1}$. Hence, we may rewrite (2.1) in the *relaxed* form

$$\begin{aligned} dX_t &\in -\partial\varphi(X_t)dt + B(t, X_t)dW_t, \\ X_0 &= x_0 \end{aligned} \tag{2.3}$$

and define (generalized) strong solutions according to Definition A.1 in Appendix A below.

Definition 2.1. Let $x_0 \in L^2(\Omega; H^{-1})$. An $\mathcal{F}_t$-adapted process $X \in L^2(\Omega; C([0,T]; H^{-1}))$ is said to be an SVI solution to (2.1) if

i. [Regularity]

$$\varphi(X) \in L^1([0,T] \times \Omega).$$

ii. [Variational inequality] For each $\mathcal{F}_t$-progressively measurable process $G \in L^2([0,T]\times\Omega; H^{-1})$ and each $\mathcal{F}_t$-adapted process $Z \in L^2(\Omega; C([0,T]; H^{-1})) \cap L^2([0,T] \times \Omega; L^2)$ solving the equation

$$Z_t - Z_0 = \int_0^t G_s ds + \int_0^t B(s, Z_s) dW_s, \quad \forall t \in [0,T],$$

we have

$$\begin{aligned} &\mathbb{E}\|X_t - Z_t\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi(X_r)dr \\ &\le \mathbb{E}\|x_0 - Z_0\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi(Z_r)dr \\ &- 2\mathbb{E}\int_0^t (G_r, X_r - Z_r)_{H^{-1}} dr + C\mathbb{E}\int_0^t \|X_r - Z_r\|_{H^{-1}}^2 dr \quad \forall t \ge 0, \end{aligned} \tag{2.4}$$

for some $C > 0$.

Remark 2.2. If (X, η) is a strong solution to (2.1) satisfying $\varphi(X) \in L^1([0,T] \times \Omega)$ then X is an SVI solution to (2.1).

Proof. Definition 2.1 (i) is satisfied by assumption. For (ii): Let $Z \in L^2(\Omega; C([0,T]; H^{-1})) \cap L^2([0,T] \times \Omega; L^2)$ be a solution to

$$dZ_t = G_t dt + B(t, Z_t) dW_t$$

for some $\mathcal{F}_t$-progressively measurable $G \in L^2([0,T] \times \Omega; H^{-1})$. Then Itô's formula implies:

$$\begin{aligned} \mathbb{E}\|X_t - Z_t\|_{H^{-1}}^2 =& \mathbb{E}\|x_0 - Z_0\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t (\eta_r - G_r, X_r - Z_r)_{H^{-1}} dr \\ &+ \mathbb{E}\int_0^t \|B(r, X_r) - B(r, Z_r)\|_{L_2(U, H^{-1})}^2 dr \quad \forall t \in [0,T]. \end{aligned}$$

Since $\eta_r \in -\partial\varphi(X_r)$ we have

$$(\eta_r, X_r - Z_r)_{H^{-1}} \le \varphi(Z_r) - \varphi(X_r), \quad dt \otimes d\mathbb{P} - \text{a.e.}$$

which, using (2.2), implies (2.4). $\square$

The main result of the current section is the proof of well-posedness of (2.1) in the sense of Definition 2.1:

Theorem 2.3. *Let $x_0 \in L^2(\Omega; H^{-1})$. Then there is a unique SVI solution X to* (2.1) *in the sense of Definition 2.1. For two SVI solutions X, Y with initial conditions $x_0, y_0 \in L^2(\Omega; H^{-1})$ we have*

$$\sup_{t\in[0,T]} \mathbb{E}\|X_t - Y_t\|^2_{H^{-1}} \leq C\mathbb{E}\|x_0 - y_0\|^2_{H^{-1}}.$$

The unique SVI solution X coincides with the limiting solution to (2.1) *constructed in [11].*

Proof. We construct SVI solutions to (2.1) by considering appropriate approximations by strong solutions. The specific form of the construction will also be a crucial ingredient in the proof of uniqueness.

Step 1: Existence

We consider approximating SPDE of the form

$$\begin{aligned} dX^\varepsilon_t &= \varepsilon\Delta X^\varepsilon_t dt + \Delta\phi^\varepsilon(X^\varepsilon_t)dt + B(t, X^\varepsilon_t)dW_t, \\ X^\varepsilon_0 &= x_0, \end{aligned} \tag{2.5}$$

with $\varepsilon > 0$, $x_0 \in L^2(\Omega; L^2)$ and $\psi^\varepsilon, \phi^\varepsilon$ as in Appendix C. By Lemma B.1 there is a unique strong solution X^ε to (2.5) satisfying

$$\mathbb{E} \sup_{t\in[0,T]} \|X^\varepsilon_t\|^2_2 + 2\varepsilon\mathbb{E}\int_0^T \|X^\varepsilon_r\|^2_{H^1_0} dr \leq C(\mathbb{E}\|x_0\|^2_2 + 1), \tag{2.6}$$

for some $C > 0$ independent of $\varepsilon > 0$. For two solutions $X^{\varepsilon_1}, X^{\varepsilon_2}$ to (2.5) with initial conditions $x^1_0, x^2_0 \in L^2(\Omega; L^2)$ we have

$$\begin{aligned} e^{-Kt}\|X^{\varepsilon_1}_t - X^{\varepsilon_2}_t\|^2_{H^{-1}} =& \|x^1_0 - x^2_0\|^2_{H^{-1}} \\ &+ 2\int_0^t e^{-Kr}(\varepsilon_1\Delta X^{\varepsilon_1}_r - \varepsilon_2\Delta X^{\varepsilon_2}_r, X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} dr \\ &+ 2\int_0^t e^{-Kr}(\Delta\phi^{\varepsilon_1}(X^{\varepsilon_1}_r) - \Delta\phi^{\varepsilon_2}(X^{\varepsilon_2}_r), X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} dr \\ &+ 2\int_0^t e^{-Kr}(X^{\varepsilon_1}_r - X^{\varepsilon_2}_r, B(r, X^{\varepsilon_1}_r) - B(r, X^{\varepsilon_2}_r))_{H^{-1}} dW_r \\ &+ \int_0^t e^{-Kr}\|B(r, X^{\varepsilon_1}_r) - B(r, X^{\varepsilon_1}_r)\|^2_{L_2} dr \\ &- K\int_0^t e^{-Kr}\|X^{\varepsilon_1}_r - X^{\varepsilon_1}_r\|^2_{H^{-1}} dr. \end{aligned}$$

Using (C.5) we note that

$$\begin{aligned} (\Delta\phi^{\varepsilon_1}(X^{\varepsilon_1}_r) - \Delta\phi^{\varepsilon_2}(X^{\varepsilon_2}_r), X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} &= -\int_{\mathcal{O}} (\phi^{\varepsilon_1}(X^{\varepsilon_1}_r) - \phi^{\varepsilon_2}(X^{\varepsilon_2}_r))(X^{\varepsilon_1}_r - X^{\varepsilon_2}_r) dx \\ &\leq C(\varepsilon_1 + \varepsilon_2)(1 + \|X^{\varepsilon_1}_r\|^2_2 + \|X^{\varepsilon_2}_r\|^2_2) \end{aligned}$$

and

$$\begin{aligned} (\varepsilon_1\Delta X^{\varepsilon_1}_r - \varepsilon_2\Delta X^{\varepsilon_2}_r, X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} &= \int_{\mathcal{O}} (\varepsilon_1 X^{\varepsilon_1}_r - \varepsilon_2 X^{\varepsilon_2}_r)(X^{\varepsilon_1}_r - X^{\varepsilon_2}_r) dx \\ &\leq C(\varepsilon_1 + \varepsilon_2)(\|X^{\varepsilon_1}_r\|^2_2 + \|X^{\varepsilon_2}_r\|^2_2), \end{aligned}$$

$dt \otimes d\mathbb{P}$-a.e.. Thus,

$$
\begin{aligned}
e^{-Kt}\|X_t^{\varepsilon_1} - X_t^{\varepsilon_2}\|_{H^{-1}}^2 \leq & \|x_0^1 - x_0^2\|_{H^{-1}}^2 \\
& + C(\varepsilon_1+\varepsilon_2)\int_0^t (\|X_r^{\varepsilon_1}\|_2^2 + \|X_r^{\varepsilon_1}\|_2^2 + 1)dr \\
& + 2\int_0^t e^{-Kr}(X_r^{\varepsilon_1} - X_r^{\varepsilon_2}, B(r, X_r^{\varepsilon_1}) - B(r, X_r^{\varepsilon_2}))_{H^{-1}} dW_r \\
& + C\int_0^t e^{-Kr}\|X_r^{\varepsilon_1} - X_r^{\varepsilon_1}\|_{H^{-1}}^2 dr \\
& - K\int_0^t e^{-Kr}\|X_r^{\varepsilon_1} - X_r^{\varepsilon_1}\|_{H^{-1}}^2 dr.
\end{aligned}
$$

Using the Burkholder-Davis-Gundy inequality and (2.6) we obtain

$$
\begin{aligned}
\mathbb{E}\sup_{t\in[0,T]} e^{-Kt}\|X_t^{\varepsilon_1} - X_t^{\varepsilon_2}\|_{H^{-1}}^2 \leq & 2E\|x_0^1 - x_0^2\|_{H^{-1}}^2 \\
& + C(\varepsilon_1+\varepsilon_2)(\mathbb{E}\|x_0^1\|_2^2 + \mathbb{E}\|x_0^2\|_2^2 + 1),
\end{aligned}
\tag{2.7}
$$

for $K > 0$ large enough.

Let now $X^{\varepsilon_1}, X^{\varepsilon_2}$ be strong solutions to (2.5) with the same initial condition $x_0 \in L^2(\Omega; L^2)$. Then (2.7) implies

$$
\mathbb{E}\sup_{t\in[0,T]} e^{-Kt}\|X_t^{\varepsilon_1} - X_t^{\varepsilon_2}\|_{H^{-1}}^2 \leq C(\varepsilon_1+\varepsilon_2)(\mathbb{E}\|x_0\|_2^2 + 1)
$$

and thus

$$
\mathbb{E}\sup_{t\in[0,T]} \|X_t^{\varepsilon} - X_t\|_{H^{-1}}^2 \to 0 \quad \text{for } \varepsilon \to 0
$$

for some $\mathcal{F}_t$-adapted process $X \in L^2(\Omega; C([0,T]; H^{-1}))$ with $X_0 = x_0$. For $x_0^1, x_0^2 \in L^2(\Omega; L^2)$ taking the limit $\varepsilon \to 0$ in (2.7) yields

$$
\mathbb{E}\sup_{t\in[0,T]} e^{-Kt}\|X_t^1 - X_t^2\|_{H^{-1}}^2 \leq 2E\|x_0^1 - x_0^2\|_{H^{-1}}^2. \tag{2.8}
$$

Let now $X^{\varepsilon,n}$ be the unique strong solution (cf. Lemma B.1) to

$$
\begin{aligned}
dX_t^{\varepsilon,n} &= \varepsilon\Delta X_t^{\varepsilon,n}dt + \Delta\phi^{\varepsilon}(X_t^{\varepsilon,n})dt + B(t, X_t^{\varepsilon,n})dW_t, \\
X_0^{\varepsilon,n} &= x_0^n,
\end{aligned}
\tag{2.9}
$$

for some sequence $x_0^n \to x_0$ in $L^2(\Omega; H^{-1})$ with $x_0^n \in L^2(\Omega; L^2)$. Using (2.7) and (2.8) we obtain the existence of a sequence of $\mathcal{F}_t$-adapted processes $X^n \in L^2(\Omega; C([0,T]; H^{-1}))$ with $X_0^n = x_0^n$ and an $\mathcal{F}_t$-adapted process $X \in L^2(\Omega; C([0,T]; H^{-1}))$ with $X_0 = x_0$ such that

$$
\mathbb{E}\sup_{t\in[0,T]} \|X_t^{\varepsilon,n} - X_t^n\|_{H^{-1}}^2 \to 0 \quad \text{for } \varepsilon \to 0
$$

and

$$
\mathbb{E}\sup_{t\in[0,T]} \|X_t^n - X_t\|_{H^{-1}}^2 \to 0 \quad \text{for } n \to \infty.
$$

Let now G, Z be as in Definition 2.1. Itô's formula implies

$$
\begin{aligned}
\mathbb{E}\|X_t^{\varepsilon,n} - Z_t\|_{H^{-1}}^2 = & \mathbb{E}\|x_0^n - Z_0\|_{H^{-1}}^2 \\
& + 2\mathbb{E}\int_0^t (\varepsilon\Delta X_t^{\varepsilon,n} + \Delta\phi^{\varepsilon}(X_t^{\varepsilon,n}) - G_r, X_r^{\varepsilon,n} - Z_r)_{H^{-1}} dr \\
& + \mathbb{E}\int_0^t \|B(r, X_r^{\varepsilon,n}) - B(r, Z_r)\|_{L_2}^2 dr.
\end{aligned}
$$

For $v \in H^{-1}$ we set

$$\varphi^\varepsilon(v) = \begin{cases} \int_{\mathcal{O}} \psi^\varepsilon(v)dx & , \ v \in L^2 \\ +\infty & , \ \text{otherwise.} \end{cases}$$

Using convexity of ψ^ε we have

$$(\Delta\phi^\varepsilon(X_r^{\varepsilon,n}), X_r^{\varepsilon,n} - Z_r)_{H^{-1}} + \varphi^\varepsilon(X_r^{\varepsilon,n}) \le \varphi^\varepsilon(Z_r)$$

$dt \otimes \mathbb{P}$-a.e.. Moreover,

$$\begin{aligned}(\varepsilon\Delta X_r^{\varepsilon,n}, X_r^{\varepsilon,n} - Z_r)_{H^{-1}} &\le \varepsilon\|\Delta X_r^{\varepsilon,n}\|_{H^{-1}}\|X_r^{\varepsilon,n} - Z_r\|_{H^{-1}} \\ &\le \varepsilon^{\frac{4}{3}}\|\Delta X_r^{\varepsilon,n}\|_{H^{-1}}^2 + \varepsilon^{\frac{2}{3}}\|X_r^{\varepsilon,n} - Z_r\|_{H^{-1}}^2\end{aligned}$$

$dt \otimes \mathbb{P}$-a.e.. Hence,

$$\begin{aligned}&\mathbb{E}\|X_t^{\varepsilon,n} - Z_t\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi^\varepsilon(X_r^{\varepsilon,n})dr \\ \le&\mathbb{E}\|x_0^n - Z_0\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi^\varepsilon(Z_r)dr \\ &- 2\mathbb{E}\int_0^t (G_r, X_r^{\varepsilon,n} - Z_r)_{H^{-1}}dr + C\mathbb{E}\int_0^t \|X_r^{\varepsilon,n} - Z_r\|_{H^{-1}}^2 dr \\ &+ 2\mathbb{E}\int_0^t \varepsilon^{\frac{4}{3}}\|\Delta X_r^{\varepsilon,n}\|_{H^{-1}}^2 + \varepsilon^{\frac{2}{3}}\|X_r^{\varepsilon,n} - Z_r\|_{H^{-1}}^2 dr.\end{aligned} \tag{2.10}$$

Due to the definition of φ^ε and (C.4) we have

$$\begin{aligned}|\varphi^\varepsilon(v) - \varphi(v)| &\le C\varepsilon(1 + \varphi(v)) \\ &\le C\varepsilon(1 + \|v\|_2^2),\end{aligned} \tag{2.11}$$

for all $v \in L^2$. Hence,

$$\mathbb{E}\int_0^t \varphi^\varepsilon(X_r^{\varepsilon,n})dr \ge \mathbb{E}\int_0^t \varphi(X_r^{\varepsilon,n})dr - C\varepsilon\mathbb{E}\int_0^t (1 + \|X_r^{\varepsilon,n}\|_2^2)dr$$

and thus

$$\liminf_{\varepsilon\to 0} \mathbb{E}\int_0^t \varphi^\varepsilon(X_r^{\varepsilon,n})dr \ge \mathbb{E}\int_0^t \varphi(X_r^n)dr.$$

Using $\varphi^\varepsilon \le \varphi$, due to (C.3), and (2.6) we may thus let $\varepsilon \to 0$ and then $n \to \infty$ in (2.10) to obtain

$$\begin{aligned}&\mathbb{E}\|X_t - Z_t\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi(X_r)dr \\ \le&\mathbb{E}\|x_0 - Z_0\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi(Z_r)dr \\ &- 2\mathbb{E}\int_0^t (G_r, X_r - Z_r)_{H^{-1}}dr + C\mathbb{E}\int_0^t \|X_r - Z_r\|_{H^{-1}}^2 dr,\end{aligned}$$

where $\varphi(X) \in L^1([0,T] \times \Omega)$ follows from lower-semicontinuity of φ on H^{-1}.

Step 2: Uniqueness

Let X be an SVI solution to (2.1) and let $Y^{\varepsilon,n}$ be the (strong) solution to (2.9) with initial condition $y_0^n \in L^2(\Omega; L^2)$ satisfying $y_0^n \to y_0$ in $L^2(\Omega; H^{-1})$. Then (2.4) with $Z = Y^{\varepsilon,n}$ and $G = \varepsilon\Delta Y^{\varepsilon,n} + \Delta\phi^\varepsilon(Y^{\varepsilon,n})$ yields

$$
\begin{aligned}
&\mathbb{E}\|X_t - Y_t^{\varepsilon,n}\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi(X_r)dr \\
&\le \mathbb{E}\|x_0 - y_0^n\|_{H^{-1}}^2 + 2\mathbb{E}\int_0^t \varphi(Y_r^{\varepsilon,n})dr \\
&\quad - 2\mathbb{E}\int_0^t (\varepsilon\Delta Y_r^{\varepsilon,n} + \Delta\phi^\varepsilon(Y_r^{\varepsilon,n}), X_r - Y_r^{\varepsilon,n})_{H^{-1}}dr \\
&\quad + C\mathbb{E}\int_0^t \|X_r - Y_r^{\varepsilon,n}\|_{H^{-1}}^2 dr.
\end{aligned}
\tag{2.12}
$$

For $x \in L^2$ we have

$$-(\Delta\phi^\varepsilon(Y^{\varepsilon,n}), x - Y^{\varepsilon,n})_{H^{-1}} + \varphi^\varepsilon(Y^{\varepsilon,n}) \le \varphi^\varepsilon(x) \quad dt \otimes d\mathbb{P} - \text{a.e.}.$$

Due to (2.11) we obtain

$$-(\Delta\phi^\varepsilon(Y^{\varepsilon,n}), x - Y^{\varepsilon,n})_{H^{-1}} + \varphi(Y^{\varepsilon,n}) \le \varphi(x) + C\varepsilon(1 + \varphi(Y^{\varepsilon,n})) \quad dt \otimes d\mathbb{P} - \text{a.e.}.$$

Since $\varphi(X) \in L^1([0,T] \times \Omega)$ and since φ is the lower-semicontinuous hull of $\varphi_{|L^2}$ on H^{-1}, for a.e. $(t,\omega) \in [0,T] \times \Omega$ we can find a sequence $x^m \in L^2$ such that $x^m \to X_t(\omega)$ in H^{-1} and $\varphi(x^m) \to \varphi(X_t(\omega))$. Hence,

$$-(\Delta\phi^\varepsilon(Y^{\varepsilon,n}), X - Y^{\varepsilon,n})_{H^{-1}} + \varphi(Y^{\varepsilon,n}) \le \varphi(X) + C\varepsilon(1 + \varphi(Y^{\varepsilon,n})) \quad dt \otimes d\mathbb{P} - \text{a.e.}$$

and (2.12) implies

$$
\begin{aligned}
\mathbb{E}\|X_t - Y_t^{\varepsilon,n}\|_{H^{-1}}^2 &\le \mathbb{E}\|x_0 - y_0^n\|_{H^{-1}}^2 \\
&\quad + 2\mathbb{E}\int_0^t \varepsilon^{\frac{4}{3}}\|\Delta Y_r^{\varepsilon,n}\|_{H^{-1}}^2 + \varepsilon^{\frac{2}{3}}\|X_r - Y_r^{\varepsilon,n}\|_{H^{-1}}^2 dr \\
&\quad + C\mathbb{E}\int_0^t \|X_r - Y_r^{\varepsilon,n}\|_{H^{-1}}^2 dr + C\varepsilon\mathbb{E}\int_0^t (1 + \varphi(Y_r^{\varepsilon,n}))dr.
\end{aligned}
$$

Taking $\varepsilon \to 0$ then $n \to \infty$ yields

$$\mathbb{E}\|X_t - Y_t\|_{H^{-1}}^2 \le \mathbb{E}\|x_0 - y_0\|_{H^{-1}}^2 + C\mathbb{E}\int_0^t \|X_r - Y_r\|_{H^{-1}}^2 dr,$$

which by Gronwall's inequality concludes the proof. □

3. Regularity and Strong solutions

We consider SPDE of the form

$$
\begin{aligned}
dX_t &\in \Delta(|X_t|^{m-1}X_t)dt + \sum_{k=1}^\infty g^k X_t d\beta_t^k, \\
X_0 &= x_0
\end{aligned}
\tag{3.1}
$$

with $m \in [0,1]$ and zero Dirichlet boundary conditions on a smooth, bounded domain $\mathcal{O} \subseteq \mathbb{R}^d$, in arbitrary dimension $d \in \mathbb{N}$. Here, β^k are independent Brownian motions on a normal, filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ and

(B) $g^k \in C^1(\bar{\mathcal{O}})$ with

$$\sum_{k=1}^\infty \|g^k\|_{C^1(\bar{\mathcal{O}})}^2 < \infty.$$

For $v \in H^{-1}$ we set

$$B(v)(h) = \sum_{k=1}^\infty g^k v(e_k, h)_{H^{-1}},$$

where $e_k \in H^{-1}$ is an orthonormal basis of H^{-1}. Then $B : H^{-1} \to L_2(H^{-1}, H^{-1})$ is Lipschitz continuous, i.e.

$$\begin{aligned}\|B(u) - B(v)\|^2_{L_2(H^{-1},H^{-1})} &= \sum_{k=1}^{\infty} \|g^k(u-v)\|^2_{H^{-1}} \\ &\le \sum_{k=1}^{\infty} \|g^k\|^2_{C^1(\bar{\mathcal{O}})} \|u-v\|^2_{H^{-1}}, \quad \text{for all } u, v \in H^{-1}\end{aligned}$$

and

$$\begin{aligned}\|B(v)\|^2_{L_2(H^{-1},L^2)} &= \sum_{k=1}^{\infty} \|g^k v\|^2_{L^2} \\ &\le \sum_{k=1}^{\infty} \|g^k\|^2_{C^0(\bar{\mathcal{O}})} \|v\|^2_{L^2}, \quad \text{for all } v \in L^2.\end{aligned}$$

We define ψ, ϕ, φ and $\mathcal{M}, L^{m+1} \cap H^{-1}, \mathcal{M} \cap H^{-1}$ as in Section 2 and rewrite (3.1) in the *relaxed* form

$$\begin{aligned}dX_t &\in -\partial\varphi(X_t)dt + B(X_t)dW_t, \\ X_0 &= x_0.\end{aligned} \tag{3.2}$$

(Generalized) strong solutions to (3.1) are then defined as in Definition A.1 (with $H = H^{-1}$).

For initial conditions $x_0 \in L^2(\Omega; H^{-1})$ satisfying $\mathbb{E}\varphi(x_0) < \infty$ we prove the existence of strong solutions to (3.1). Moreover, we will prove regularizing properties with respect to the initial condition due to the subgradient structure of the drift. This allows to characterize solutions for initial conditions $x_0 \in L^2(\Omega; H^{-1})$ as generalized strong solutions.

Theorem 3.1. *Let $x_0 \in L^2(\Omega; H^{-1})$.*

i. There is a unique generalized strong solution (X, η) to (3.1) and X satisfies

$$\mathbb{E}t\varphi(X_t) + \mathbb{E}\int_0^t r\|\eta_r\|^2_{H^{-1}} dr \le C\left(\mathbb{E}\|x_0\|^2_{H^{-1}} + 1\right).$$

ii. If $\mathbb{E}\varphi(x_0) < \infty$, then there is a unique strong solution (X, η) to (3.1) satisfying

$$\mathbb{E}\varphi(X_t) + \mathbb{E}\int_0^t \|\eta_r\|^2_{H^{-1}} dr \le \mathbb{E}\varphi(x_0) + C,$$

The (generalized) strong solution (X, η) coincides with the limit solution constructed in [11].

The proof of Theorem 3.1 proceeds in several steps. In particular, the singularity of ψ causes the need for a non-singular regularization of ψ. This approximation has to be carefully chosen in order to obtain uniform bounds. Passing to the limit will in turn rely on Mosco-convergence of the regularized potentials. We will first consider the case of non-singular potentials, keeping careful track of the arising constants in Section 3.0.1, the limit will then be taken in Section 3.0.2.

3.0.1. *Non-singular potential ψ.* In this section we restrict to the approximating case of a smooth, non-singular nonlinearity ψ. We assume that $\psi \in C^3(\mathbb{R}; \mathbb{R}_+)$ is convex with Lipschitz continuous derivatives $\phi = \dot{\psi}$, $\dot{\phi}$ satisfying $\psi(0) = \phi(0) = 0$ and

$$\begin{aligned}\phi(r)r &\ge c_\psi\psi(r) - C_\psi \\ \dot{\phi}(r)r^2 &\le C_\psi\psi(r), \quad \forall r \in \mathbb{R},\end{aligned} \tag{3.3}$$

for some constants $C_\psi, c_\psi > 0$.

We consider the following non-degenerate, non-singular approximation of (3.1) (cf. Appendix B):

$$
\begin{aligned}
dX_t^\varepsilon &= \varepsilon \Delta X_t^\varepsilon dt + \Delta\phi(X_t^\varepsilon)dt + B(X_t^\varepsilon)dW_t, \\
X_0^\varepsilon &= x_0.
\end{aligned} \tag{3.4}
$$

For $v \in H^{-1}$ we define

$$
\varphi^\varepsilon(v) := \begin{cases} \frac{\varepsilon}{2}\int_{\mathcal{O}} |v|^2 dx + \int_{\mathcal{O}} \psi(v)dx, & \text{for } v \in L^2 \\ +\infty, & \text{otherwise.} \end{cases}
$$

Note that $\varphi^\varepsilon \in C^1(L^2)$ with Lipschitz continuous derivative given by

$$
D\varphi^\varepsilon(v)(h) = \varepsilon \int_{\mathcal{O}} vh dx + \int_{\mathcal{O}} \phi(v)h dx.
$$

To check the claimed continuity we note that

$$
\begin{aligned}
D\varphi^\varepsilon(v)(h) - D\varphi^\varepsilon(w)(h) &= \varepsilon \int_{\mathcal{O}} (v-w)h dx + \int_{\mathcal{O}} (\phi(v) - \phi(w))h dx \\
&\le \varepsilon \|v-w\|_2 \|h\|_2 + \|\phi\|_{Lip} \|v-w\|_2 \|h\|_2 \\
&\lesssim (\varepsilon + 1)\|v-w\|_2 \|h\|_2.
\end{aligned}
$$

Moreover, for $k \in \mathbb{N}$ large enough, we have $\varphi^\varepsilon \in C^2(H_0^1 \cap H^{2k+1})$ with Lipschitz continuous second derivative given by

$$
D^2\varphi^\varepsilon(v)(g,h) = \varepsilon \int_{\mathcal{O}} hg dx + \int_{\mathcal{O}} h\dot{\phi}(v)g dx.
$$

Indeed:

$$
\begin{aligned}
D^2\varphi^\varepsilon(v)(g,h) - D^2\varphi^\varepsilon(w)(g,h) &= \int_{\mathcal{O}} h\left(\dot{\phi}(v) - \dot{\phi}(w)\right) g dx \\
&\le \|h\|_3 \|\dot{\phi}(v) - \dot{\phi}(w)\|_3 \|g\|_3 \\
&\le \|\dot{\phi}\|_{Lip} \|h\|_{H^{2k+1}} \|v-w\|_{H^{2k+1}} \|g\|_{H^{2k+1}},
\end{aligned}
$$

where we used the Sobolev embedding $H^{2k+1} \hookrightarrow L^3$ for $k \in \mathbb{N}$ large enough. Moreover, φ^ε is a convex, lower-semicontinuous function on H^{-1} with subgradient given by

$$
A^\varepsilon(v) := -\partial\varphi^\varepsilon(v) = \varepsilon\Delta v + \Delta\phi(v) \in H^{-1}, \quad \text{for } v \in H_0^1.
$$

Hence, we may write (3.4) as

$$
dX_t^\varepsilon = -\partial\varphi^\varepsilon(X_t^\varepsilon)dt + B(X_t^\varepsilon)dW_t.
$$

(Generalized) strong solutions to (3.4) are then defined according to Definition A.1. By [13], for each $x_0 \in L^2(\Omega; H^{-1})$ there is a unique variational solution X^ε to (3.4) with respect to the Gelfand triple

$$
L^2 \hookrightarrow H^{-1} \hookrightarrow (L^2)^*.
$$

Due to Lemma B.1, X^ε is a strong solution if $x_0 \in L^2(\Omega; L^2)$.

Lemma 3.2. *Let $x_0 \in L^2(\Omega; L^2)$. For each $\varepsilon > 0$ we have $\varphi^\varepsilon(X^\varepsilon) \in L^1([0,T] \times \Omega)$ with*

$$
\mathbb{E}\int_0^T \varphi^\varepsilon(X_r^\varepsilon)dr \le C(\mathbb{E}\|x_0\|_{H^{-1}}^2 + 1),
$$

for some constant C independent of $\varepsilon > 0$ and depending on ψ via the constants c_ψ, C_ψ only.

Proof. Note that, using (3.3)

$$\begin{aligned}(A^\varepsilon(v),v)_{H^{-1}} &= -\int_{\mathcal{O}} \left(\varepsilon|v|^2+\phi(v)v\right)dx\\ &\le -\int_{\mathcal{O}} \left(\varepsilon|v|^2+c\psi(v)-C\right)dx\\ &\le -c\varphi^\varepsilon(v)+C,\end{aligned}$$

for all $v\in H_0^1$. By Itô's formula we have

$$\begin{aligned}&\mathbb{E}e^{-Kt}\|X_t^\varepsilon\|_{H^{-1}}^2\\ =&\mathbb{E}\|x_0\|_H^2+2\mathbb{E}\int_0^t e^{-Kr}(A^\varepsilon(X_r^\varepsilon),X_r^\varepsilon)_{H^{-1}}+e^{-Kr}\|B(X_r^\varepsilon)\|_{L_2(U,H^{-1})}^2dr\\ &-K\int_0^t e^{-Kr}\|X_r^\varepsilon\|_{H^{-1}}^2dr\\ \le&\mathbb{E}\|x_0\|_{H^{-1}}^2-2\mathbb{E}\int_0^t ce^{-Kr}\varphi^\varepsilon(X_r^\varepsilon)+Ce^{-Kr}\|X_r^\varepsilon\|_{H^{-1}}^2dr\\ &-K\int_0^t e^{-Kr}\|X_r^\varepsilon\|_{H^{-1}}^2dr+C.\end{aligned}$$

Choosing K large enough yields the claim. □

Based on the strong solution property of X^ε we derive the key estimate in the following

Lemma 3.3. *Let $x_0\in L^2(\Omega;L^2)$. For all $\varepsilon>0$ we have*

$$\mathbb{E}t\varphi^\varepsilon(X_t^\varepsilon)+\mathbb{E}\int_0^t r\|\varepsilon X_r^\varepsilon+\phi(X_r^\varepsilon)\|_{H_0^1}^2dr\le C\left(\mathbb{E}\|x_0\|_{H^{-1}}^2+1\right).$$

and

$$\mathbb{E}\varphi^\varepsilon(X_t^\varepsilon)+\mathbb{E}\int_0^t \|\varepsilon X_r^\varepsilon+\phi(X_r^\varepsilon)\|_{H_0^1}^2dr\le C\mathbb{E}\varphi^\varepsilon(x_0), \tag{3.5}$$

for some constant C independent of $\varepsilon>0$ and depending on ψ via the constants c_ψ, C_ψ only.

Proof. Let J^λ be the resolvent of $-\Delta$ on H^{-1}. We let $G^\lambda=G^{\lambda,k}:=J^\lambda\circ\cdots\circ J^\lambda$ the k-th iteration of J^λ. Then $G^\lambda: H^{-1}\to H_0^1\cap H^{2k+1}$ is linear and continuous. Moreover,

$$\begin{aligned}\|G^\lambda v\|_2&\le\|v\|_2,\quad\forall v\in L^2\\ \|G^\lambda v\|_{H_0^1}&\le\|v\|_{H_0^1},\quad\forall v\in H_0^1\end{aligned}$$

and

$$\begin{aligned}\|G^\lambda v-v\|_2&\le\|J^\lambda G^{\lambda,k-1}v-J^\lambda v\|_2+\|J^\lambda v-v\|_2\\ &\le\|G^{\lambda,k-1}v-v\|_2+\|J^\lambda v-v\|_2.\end{aligned}$$

Iterating this yields

$$\|G^\lambda v-v\|_2\le k\|J^\lambda v-v\|_2\to 0\quad\text{for }\lambda\to 0$$

for all $v\in L^2$. Analogously,

$$\|G^\lambda v-v\|_{H_0^1}\to 0\quad\text{for }\lambda\to 0$$

for all $v\in H_0^1$.

We define $\varphi^{\varepsilon,\lambda} := \varphi^\varepsilon \circ G^\lambda$. Since $G^\lambda : H^{-1} \to H_0^1 \cap H^{2k+1}$ is a linear, continuous operator we have $\varphi^{\varepsilon,\lambda} \in C^2(H^{-1})$ with Lipschitz continuous derivatives given by

$$D\varphi^{\varepsilon,\lambda}(v)(h) = -\left(\varepsilon\Delta G^\lambda v + \Delta\phi(G^\lambda v), G^\lambda h\right)_{H^{-1}},$$
$$D^2\varphi^{\varepsilon,\lambda}(v)(g,h) = -\varepsilon\int_{\mathcal{O}}(G^\lambda h)(G^\lambda g)dx + \int_{\mathcal{O}}(G^\lambda h)\dot{\phi}(G^\lambda v)(G^\lambda g)dx.$$

By Lemma B.1 there is a unique strong solution X^ε to (3.4) with $X^\varepsilon \in L^2(\Omega; L^\infty([0,T]; L^2)) \cap L^2([0,T]\times\Omega; H_0^1)$. We apply Itô's formula to $t\varphi^{\varepsilon,\lambda}(X_t^\varepsilon)$ to get:

$$\begin{aligned}
&\mathbb{E}t\varphi^{\varepsilon,\lambda}(X_t^\varepsilon)\\
&= -\mathbb{E}\int_0^t r(\varepsilon\Delta G^\lambda X_r^\varepsilon + \Delta\phi(G^\lambda X_r^\varepsilon), G^\lambda(\varepsilon\Delta X_r^\varepsilon + \Delta\phi(X_r^\varepsilon)))_{H^{-1}}dr\\
&+ \frac{\varepsilon}{2}\sum_{k=1}^\infty \mathbb{E}\int_0^t r\int_{\mathcal{O}}|G^\lambda(g^k X_r^\varepsilon)|^2 dxdr\\
&+ \frac{1}{2}\sum_{k=1}^\infty \mathbb{E}\int_0^t r\int_{\mathcal{O}} G^\lambda(g^k X_r^\varepsilon)\dot{\phi}(G^\lambda X_r^\varepsilon)G^\lambda(g^k X_r^\varepsilon)dxdr\\
&+ \mathbb{E}\int_0^t \varphi^{\varepsilon,\lambda}(X_r^\varepsilon)dr.
\end{aligned}$$

We first note that

$$\begin{aligned}
\int_{\mathcal{O}}|G^\lambda(g^k X_r^\varepsilon)|^2 dx &\le \int_{\mathcal{O}}|g^k X_r^\varepsilon|^2 dx\\
&\le \|g^k\|_{C^0}^2\|X_r^\varepsilon\|_2^2.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\int_{\mathcal{O}} G^\lambda(g^k X_r^\varepsilon)\dot{\phi}(G^\lambda X_r^\varepsilon)G^\lambda(g^k X_r^\varepsilon)dx &= \int_{\mathcal{O}}(G^\lambda(g^k X_r^\varepsilon))^2\dot{\phi}(G^\lambda X_r^\varepsilon)dx\\
&\le \|\dot{\phi}\|_\infty\|g^k\|_{C^0}^2\|X_r^\varepsilon\|_2^2.
\end{aligned}$$

We note $G^\lambda v \to v$ in H_0^1 for $v \in H_0^1$. Since $\dot{\phi}$ is Lipschitz we have $\dot{\phi}(G^\lambda X^\varepsilon) \to \dot{\phi}(X^\varepsilon)$ in $L^2([0,T]\times\Omega; H_0^1)$. Moreover, $G^\lambda g^k X_r^\varepsilon \to g^k X_r^\varepsilon$ in $L^2([0,T]\times\Omega; H_0^1)$. Using (3.3) this implies

$$\begin{aligned}
&\lim_\lambda \mathbb{E}\int_0^t r\int_{\mathcal{O}} G^\lambda(g^k X_r^\varepsilon)\dot{\phi}(G^\lambda X_r^\varepsilon)G^\lambda(g^k X_r^\varepsilon)dxdr\\
&= \mathbb{E}\int_0^t r\int_{\mathcal{O}}(g^k X_r^\varepsilon)^2\dot{\phi}(X_r^\varepsilon)dxdr\\
&\le \|g^k\|_{C^0}^2 C_\psi \mathbb{E}\int_0^t r\int_{\mathcal{O}}\psi(X_r^\varepsilon)dxdr\\
&= \|g^k\|_{C^0}^2 C_\psi \mathbb{E}\int_0^t r\varphi(X_r^\varepsilon)dr.
\end{aligned}$$

Hence, by dominated convergence we obtain

$$\begin{aligned}
&\lim_{\lambda\to 0}\frac{\varepsilon}{2}\sum_{k=1}^\infty \mathbb{E}\int_0^t r\int_{\mathcal{O}}|G^\lambda(g^k X_r^\varepsilon)|^2 dxdr\\
&+ \lim_{\lambda\to 0}\frac{1}{2}\sum_{k=1}^\infty \mathbb{E}\int_0^t r\int_{\mathcal{O}} G^\lambda(g^k X_r^\varepsilon)\dot{\phi}(G^\lambda(X_r^\varepsilon))G^\lambda(g^k X_r^\varepsilon)dxdr\\
&\le C\left(1 + \mathbb{E}\int_0^t r\varphi^\varepsilon(X_r^\varepsilon)dr\right),
\end{aligned}$$

for some constant C independent of $\varepsilon > 0$ and depending on ψ via the constants c_ψ, C_ψ only. We note that

$$\begin{aligned}&(\varepsilon\Delta G^\lambda X_r^\varepsilon + \Delta\phi(G^\lambda X_r^\varepsilon), G^\lambda(\varepsilon\Delta X_r^\varepsilon + \Delta\phi(X_r^\varepsilon)))_{H^{-1}}\\ &= (\varepsilon G^\lambda X_r^\varepsilon + \phi(G^\lambda X_r^\varepsilon), G^\lambda(\varepsilon X_r^\varepsilon + \phi(X_r^\varepsilon)))_{H_0^1}.\end{aligned}$$

Since $X^\varepsilon \in L^2([0,T]\times\Omega; H_0^1)$ we have (using dominated convergence)

$$G^\lambda X^\varepsilon \to X^\varepsilon \text{ in } L^2([0,T]\times\Omega; H_0^1) \text{ for } \lambda \to 0.$$

Since $\phi, \dot\phi$ are Lipschitz continuous this implies $\phi(G^\lambda X^\varepsilon) \to \phi(X^\varepsilon)$ for $\lambda \to 0$ in $L^2([0,T]\times\Omega; H_0^1)$. Hence, we obtain

$$\begin{aligned}&\lim_{\lambda\to 0} -\mathbb{E}\int_0^t r(\varepsilon G^\lambda X_r^\varepsilon + \phi(G^\lambda X_r^\varepsilon), G^\lambda(\varepsilon X_r^\varepsilon + \phi(X_r^\varepsilon)))_{H_0^1}dr\\ &= -\mathbb{E}\int_0^t r\|\varepsilon X_r^\varepsilon + \phi(X_r^\varepsilon)\|_{H_0^1}^2 dr.\end{aligned}$$

Since φ^ε is continuous on L^2,

$$|\varphi^{\varepsilon,\lambda}(v)| \le C(1 + \|v\|_2^2)$$

and $G^\lambda X^\varepsilon \to X^\varepsilon$ in $L^2([0,T]\times\Omega; H_0^1)$, $G^\lambda X_t^\varepsilon \to X_t^\varepsilon$ in $L^2(\Omega; L^2)$ for all $t \in [0,T]$, by dominated convergence we get

$$\lim_{\lambda\to 0}\mathbb{E}t\varphi^{\varepsilon,\lambda}(X_t^\varepsilon) = \mathbb{E}t\varphi^\varepsilon(X_t^\varepsilon)$$

and

$$\lim_{\lambda\to 0}\mathbb{E}\int_0^t \varphi^{\varepsilon,\lambda}(X_r^\varepsilon)dr = \mathbb{E}\int_0^t \varphi^\varepsilon(X_r^\varepsilon)dr.$$

Putting these estimates together yields

$$\begin{aligned}\mathbb{E}t\varphi^\varepsilon(X_t^\varepsilon) \le &-\mathbb{E}\int_0^t r\|\varepsilon X_r^\varepsilon + \phi(X_r^\varepsilon)\|_{H_0^1}^2 dr + C\left(1 + \mathbb{E}\int_0^t r\varphi^\varepsilon(X_r^\varepsilon)dr\right)\\ &+ \mathbb{E}\int_0^t \varphi^\varepsilon(X_r^\varepsilon)dr.\end{aligned}$$

By Lemma 3.2 we conclude

$$\mathbb{E}t\varphi^\varepsilon(X_t^\varepsilon) + \mathbb{E}\int_0^t r\|\varepsilon X_r^\varepsilon + \phi(X_r^\varepsilon)\|_{H_0^1}^2 dr \le C\left(\mathbb{E}\|x_0\|_{H^{-1}}^2 + 1\right),$$

for some constant C independent of $\varepsilon > 0$ and depending on ψ via the constants c_ψ, C_ψ only.

To prove (3.5) we proceed as above but applying Itô's formula for $\varphi^{\varepsilon,\lambda}(X_t^\varepsilon)$ instead of $t\varphi^{\varepsilon,\lambda}(X_t^\varepsilon)$. □

3.0.2. *Proof of Theorem 3.1.*

Proof of Theorem 3.1: The proof proceeds via a three-step approximation. First, we approximate the singular potential $\psi(r) = \frac{1}{m+1}|r|^{m+1}$ by the smooth functions

$$\psi^\delta(r) = \frac{1}{m+1}\left(\left(r^2+\delta\right)^{\frac{m+1}{2}} - \delta^{\frac{m+1}{2}}\right).$$

Note

$$\begin{aligned}\phi^\delta(r) &:= (\psi^\delta)'(r) = \left(r^2+\delta\right)^{\frac{m-1}{2}} r\\ \dot\phi^\delta(r) &:= \left(r^2+\delta\right)^{\frac{m-1}{2}} + (m-1)\left(r^2+\delta\right)^{\frac{m-3}{2}} r^2\\ &= \left(r^2+\delta\right)^{\frac{m-3}{2}}(\delta + mr^2).\end{aligned}$$

Thus, for $\delta \in [0,1]$

$$\begin{aligned}\phi^\delta(r)r = \left(r^2+\delta\right)^{\frac{m-1}{2}} r^2 &= \left(r^2+\delta\right)^{\frac{m+1}{2}} - \left(r^2+\delta\right)^{\frac{m-1}{2}}\delta \\ &\geq (m+1)\psi^\delta(r) - 1\end{aligned}$$

and

$$\begin{aligned}\dot\phi^\delta(r)r^2 &= \left(r^2+\delta\right)^{\frac{m-3}{2}}(\delta+mr^2)r^2 \\ &\leq (m+1)\left(r^2+\delta\right)^{\frac{m+1}{2}} \\ &= (m+1)^2\psi^\delta(r).\end{aligned}$$

That is, (3.3) is satisfied with constants c_ψ, C_ψ independent of $\delta > 0$. Moreover, we observe

$$|\psi^\delta(r)-\psi(r)| = \frac{2}{m+1}\delta^{\frac{m+1}{2}}. \tag{3.6}$$

Next we consider vanishing viscosity

$$\begin{aligned}dX_t^{\varepsilon,\delta} &= \varepsilon\Delta X_t^{\varepsilon,\delta}dt + \Delta\phi^\delta(X_t^{\varepsilon,\delta})dt + B(X_t^{\varepsilon,\delta})dW_t, \\ X_0^{\varepsilon,\delta} &= x_0.\end{aligned} \tag{3.7}$$

In a third approximating step we consider smooth approximations of the initial condition, i.e. we first assume $x_0 \in L^2(\Omega; L^2)$.

Step 1: $\delta \to 0$

Let $\varepsilon > 0$, $x_0 \in L^2(\Omega; L^2)$ and for $v \in H^{-1}$ set

$$\varphi^{\varepsilon,\delta}(v) := \begin{cases} \frac{\varepsilon}{2}\int_{\mathcal{O}}|v|^2dx + \int_{\mathcal{O}}\psi^\delta(v)dx, & \text{for } v \in L^2 \\ +\infty, & \text{otherwise.}\end{cases}$$

Since

$$\partial\varphi^{\varepsilon,\delta}(v) = \begin{cases} -\varepsilon\Delta v - \Delta\phi^\delta(v), & \text{for } v \in H_0^1 \\ \emptyset, & \text{otherwise.}\end{cases}$$

we may rewrite (3.7) as

$$dX_t^{\varepsilon,\delta} = -\partial\varphi^{\varepsilon,\delta}(X_t^{\varepsilon,\delta})dt + B(X_t^{\varepsilon,\delta})dW_t$$

and define strong solutions to (3.7) as in Definition A.1. By Lemma B.1 there is a (unique) strong solution $X^{\varepsilon,\delta}$ to (3.7). From Lemma B.1 and Lemma 3.3 we have

$$\mathbb{E}\sup_{t\in[0,T]}\|X_t^{\varepsilon,\delta}\|_2^2 + 2\varepsilon\mathbb{E}\int_0^T\|X_r^{\varepsilon,\delta}\|_{H_0^1}^2dr \leq C(\mathbb{E}\|x_0\|_2^2+1) \tag{3.8}$$

$$\mathbb{E}t\varphi^{\varepsilon,\delta}(X_t^{\varepsilon,\delta}) + \mathbb{E}\int_0^t r\|\varepsilon X_r^{\varepsilon,\delta}+\phi^\delta(X_r^{\varepsilon,\delta})\|_{H_0^1}^2dr \leq C\left(\mathbb{E}\|x_0\|_{H^{-1}}^2+1\right)$$

and

$$\begin{aligned}\mathbb{E}\varphi^{\varepsilon,\delta}(X_t^{\varepsilon,\delta}) + \mathbb{E}\int_0^t\|\varepsilon X_r^{\varepsilon,\delta}+\phi^\delta(X_r^{\varepsilon,\delta})\|_{H_0^1}^2dr &\leq C\mathbb{E}\varphi^{\varepsilon,\delta}(x_0) \\ &\leq C(\mathbb{E}\|x_0\|_2^2+1),\end{aligned} \tag{3.9}$$

with a constant $C > 0$ independent of $\varepsilon, \delta > 0$. Hence, we may extract a subsequence $\delta_n \to 0$ such that

$$\begin{aligned}X^{\varepsilon,\delta_n} &\rightharpoonup \bar X^\varepsilon, \quad \text{in } L^2([0,T]\times\Omega; H_0^1) \\ X^{\varepsilon,\delta_n} &\rightharpoonup^* \bar X^\varepsilon, \quad \text{in } L^2(\Omega; L^\infty([0,T]; L^2)) \\ \Delta\phi^{\delta_n}(X^{\varepsilon,\delta_n}) &\rightharpoonup \eta^\varepsilon, \quad \text{in } L^2([0,T]\times\Omega; H^{-1}).\end{aligned}$$

Since $B : H^{-1} \to L_2(H^{-1}; H^{-1})$ is continuous, linear we have $B(X^{\varepsilon,\delta_n}) \rightharpoonup B(\bar{X}^\varepsilon)$ in $L^2([0,T] \times \Omega; L_2(H^{-1}; H^{-1}))$ and thus $dt \otimes d\mathbb{P}$-almost everywhere

$$\bar{X}_t^\varepsilon = x_0 + \int_0^t (\varepsilon \Delta \bar{X}_r^\varepsilon + \eta_r^\varepsilon) dr + \int_0^t B(\bar{X}_r^\varepsilon) dW_r.$$

Hence, defining

$$X_t^\varepsilon := x_0 + \int_0^t (\varepsilon \Delta \bar{X}_r^\varepsilon + \eta_r^\varepsilon) dr + \int_0^t B(\bar{X}_r^\varepsilon) dW_r$$

we have $X^\varepsilon = \bar{X}^\varepsilon$, $dt \otimes d\mathbb{P}$-almost everywhere and due to [13, Theorem 4.2.5] we have $X^\varepsilon \in L^2(\Omega; C([0,T]; H^{-1}))$. We aim to prove that X^ε is a strong solution to

$$dX_t^\varepsilon \in -\partial \varphi^\varepsilon(X_t^\varepsilon) dt + B(X_t^\varepsilon) dW_t, \tag{3.10}$$

where (for $v \in H^{-1}$)

$$\varphi^\varepsilon(v) := \begin{cases} \frac{\varepsilon}{2} \int_{\mathcal{O}} |v|^2 dx + \int_{\mathcal{O}} \psi(v) dx, & \text{for } v \in L^2 \\ +\infty, & \text{otherwise.} \end{cases}$$

Note that

$$\partial \varphi^\varepsilon(v) = \begin{cases} \{\varepsilon \Delta v + \Delta \zeta |\ \zeta \in H_0^1 \text{ with } \zeta \in \phi(v) \text{ a.e.}\}, & \text{for } v \in H_0^1 \\ \emptyset, & \text{otherwise.} \end{cases} \tag{3.11}$$

It thus remains to identify $\varepsilon \Delta X^\varepsilon + \eta^\varepsilon \in -\partial \varphi^\varepsilon(X^\varepsilon)$, $dt \otimes d\mathbb{P}$-almost everywhere. Itô's formula yields

$$\begin{aligned} \mathbb{E} e^{-Ct} \|X_t^\varepsilon\|_{H^{-1}}^2 =& \mathbb{E}\|x_0\|_{H^{-1}}^2 + \mathbb{E} \int_0^t e^{-Cr} (\varepsilon \Delta X_r^\varepsilon + \eta_r^\varepsilon, X_r^\varepsilon)_{H^{-1}} dr \\ &+ \mathbb{E} \int_0^t e^{-Cr} \|B(X_r^\varepsilon)\|_{L_2}^2 dr - C\mathbb{E} \int_0^t e^{-Cr} \|X_r^\varepsilon\|_{H^{-1}}^2 dr \quad \forall t \in [0,T]. \end{aligned} \tag{3.12}$$

Using Itô's formula for $X^{\varepsilon,\delta}$ yields (for $C > 0$ large enough)

$$\begin{aligned} &\mathbb{E} e^{-Ct} \|X_t^{\varepsilon,\delta_n}\|_{H^{-1}}^2 \\ =& \mathbb{E}\|x_0\|_{H^{-1}}^2 + \mathbb{E} \int_0^t e^{-Cr} (\varepsilon \Delta X_r^{\varepsilon,\delta_n} + \Delta \phi^{\delta_n}(X_r^{\varepsilon,\delta_n}), X_r^{\varepsilon,\delta_n})_{H^{-1}} dr \\ &+ \mathbb{E} \int_0^t e^{-Cr} \|B(X_r^{\varepsilon,\delta_n}) - B(X_r^\varepsilon)\|_{L_2}^2 dr - \mathbb{E} \int_0^t e^{-Cr} \|B(X_r^\varepsilon)\|_{L_2}^2 dr \\ &+ 2\mathbb{E} \int_0^t e^{-Cr} (B(X_r^{\varepsilon,\delta_n}), B(X_r^\varepsilon))_{L_2} dr - C\mathbb{E} \int_0^t e^{-Cr} \|X_r^{\varepsilon,\delta_n} - X_r^\varepsilon\|_{H^{-1}}^2 dr \\ &+ C\mathbb{E} \int_0^t e^{-Cr} \|X_r^\varepsilon\|_{H^{-1}}^2 dr - 2C\mathbb{E} \int_0^t e^{-Cr} (X_r^{\varepsilon,\delta_n}, X_r^\varepsilon)_{H^{-1}} dr \\ \leq& \mathbb{E}\|x_0\|_{H^{-1}}^2 + \mathbb{E} \int_0^t e^{-Cr} (\varepsilon \Delta X_r^{\varepsilon,\delta_n} + \Delta \phi^{\delta_n}(X_r^{\varepsilon,\delta_n}), X_r^{\varepsilon,\delta_n})_{H^{-1}} dr \\ &- \mathbb{E} \int_0^t e^{-Cr} \|B(X_r^\varepsilon)\|_{L_2}^2 dr + 2\mathbb{E} \int_0^t e^{-Cr} (B(X_r^{\varepsilon,\delta_n}), B(X_r^\varepsilon))_{L_2} dr \\ &+ C\mathbb{E} \int_0^t e^{-Cr} \|X_r^\varepsilon\|_{H^{-1}}^2 dr - 2C\mathbb{E} \int_0^t e^{-Cr} (X_r^{\varepsilon,\delta_n}, X_r^\varepsilon)_{H^{-1}} dr. \end{aligned}$$

Taking $\liminf_{\delta_n\to 0}$ we obtain (first in distributional sense in t then a.e. by the Lebesgue Theorem)

$$\begin{aligned}
&\mathbb{E}e^{-Ct}\|X_t^\varepsilon\|_{H^{-1}}^2\\
\le&\liminf_{\delta_n\to 0}\mathbb{E}e^{-Ct}\|X_t^{\varepsilon,\delta_n}\|_{H^{-1}}^2\\
\le&\mathbb{E}\|x_0\|_{H^{-1}}^2+\liminf_{\delta_n\to 0}\mathbb{E}\int_0^t e^{-Cr}(\varepsilon\Delta X_r^{\varepsilon,\delta_n}+\Delta\phi^{\delta_n}(X_r^{\varepsilon,\delta_n}),X_r^{\varepsilon,\delta_n})_{H^{-1}}dr\\
&-\mathbb{E}\int_0^t e^{-Cr}\|B(X_r^\varepsilon)\|_{L_2}^2dr+2\mathbb{E}\int_0^t e^{-Cr}(B(X_r^\varepsilon),B(X_r^\varepsilon))_{L_2}dr\\
&+C\mathbb{E}\int_0^t e^{-Cr}\|X_r^\varepsilon\|_{H^{-1}}^2dr-2C\mathbb{E}\int_0^t e^{-Cr}(X_r^\varepsilon,X_r^\varepsilon)_{H^{-1}}dr\\
\le&\mathbb{E}\|x_0\|_{H^{-1}}^2+\liminf_{\delta\to 0}\mathbb{E}\int_0^t e^{-Cr}(\varepsilon\Delta X_r^{\varepsilon,\delta_n}+\Delta\phi^{\delta_n}(X_r^{\varepsilon,\delta_n}),X_r^{\varepsilon,\delta_n})_{H^{-1}}dr\\
&+\mathbb{E}\int_0^t e^{-Cr}\|B(X_r^\varepsilon)\|_{L_2}^2dr-C\mathbb{E}\int_0^t e^{-Cr}\|X_r^\varepsilon\|_{H^{-1}}^2dr\quad\text{a.e. } t\in[0,T].
\end{aligned}$$

Subtracting (3.12) we obtain

$$\begin{aligned}
&\mathbb{E}\int_0^t e^{-Cr}(\varepsilon\Delta X_r^\varepsilon+\eta_r^\varepsilon,X_r^\varepsilon)_{H^{-1}}dr\\
&\le\liminf_{\delta_n\to 0}\mathbb{E}\int_0^t e^{-Cr}(\varepsilon\Delta X_r^{\varepsilon,\delta_n}+\Delta\phi^{\delta_n}(X_r^{\varepsilon,\delta_n}),X_r^{\varepsilon,\delta_n})_{H^{-1}}dr.
\end{aligned}\tag{3.13}$$

We now consider the convex, lower-semicontinuous functionals $\bar\varphi^\varepsilon,\bar\varphi^{\varepsilon,\delta}:L^2([0,T]\times\Omega;H^{-1})\to\bar{\mathbb{R}}$ defined by

$$\begin{aligned}
\bar\varphi^\varepsilon(v):=&\begin{cases}\mathbb{E}\int_0^T e^{-Cr}\int_{\mathcal{O}}(\frac{\varepsilon}{2}v^2+\psi(v))dxdr & \text{, if } v\in L^2([0,T]\times\Omega;L^2)\\ +\infty & \text{, otherwise}\end{cases}\\
=&\mathbb{E}\int_0^T e^{-Cr}\varphi^\varepsilon(v_r)dr.
\end{aligned}$$

and $\bar\varphi^{\varepsilon,\delta}$ being defined analogously, where we endow $L^2([0,T]\times\Omega;H^{-1})$ with the equivalent norm

$$\|v\|_{L^2([0,T]\times\Omega;H^{-1})}^2:=\mathbb{E}\int_0^T e^{-Cr}\|v_r\|_{H^{-1}}^2dr.$$

Due to the characterization of subgradients of integral functionals proved in [15, Theorem 21] we have

$$\partial\bar\varphi^\varepsilon(v)=\{\eta\in L^2([0,T]\times\Omega;H^{-1})|\ \eta\in\partial\varphi^\varepsilon(v),\ dt\otimes d\mathbb{P}\text{-a.e.}\}\tag{3.14}$$

and

$$\partial\bar\varphi^{\varepsilon,\delta}(v)=\{-\varepsilon\Delta v-\Delta\phi^\delta(v)\}\quad\text{for } v\in L^2([0,T]\times\Omega;H_0^1).\tag{3.15}$$

Since $\bar\varphi^{\varepsilon,\delta}\to\bar\varphi^\varepsilon$ uniformly for $\delta\to 0$ (cf. (3.6)), we also have $\bar\varphi^{\varepsilon,\delta}\to\bar\varphi^\varepsilon$ in Mosco sense. Due to (3.15) we have

$$(-\varepsilon\Delta X^{\varepsilon,\delta_n}-\Delta\phi^{\delta_n}(X^{\varepsilon,\delta_n}),Y-X^{\varepsilon,\delta_n})_{L^2([0,T]\times\Omega;H^{-1})}+\bar\varphi^{\varepsilon,\delta_n}(X^{\varepsilon,\delta_n})\le\bar\varphi^{\varepsilon,\delta_n}(Y),$$

for all $Y\in L^2([0,T]\times\Omega;H^{-1})$. Using (3.13) and Mosco convergence of $\bar\varphi^{\varepsilon,\delta}$ to $\bar\varphi^\varepsilon$ we may take the $\liminf_{n\to\infty}$ to get

$$(-\varepsilon\Delta X^\varepsilon-\eta^\varepsilon,Y-X^\varepsilon)_{L^2([0,T]\times\Omega;H^{-1})}+\bar\varphi^\varepsilon(X^\varepsilon)\le\bar\varphi^\varepsilon(Y).$$

Hence, $\varepsilon\Delta X^\varepsilon + \eta^\varepsilon \in -\partial\bar{\varphi}^\varepsilon(X^\varepsilon)$ and we conclude $\varepsilon\Delta X^\varepsilon + \eta^\varepsilon \in -\partial\varphi^\varepsilon(X^\varepsilon)\ dt \otimes d\mathbb{P}$-almost everywhere due to (3.14). Then, (3.11) yields

$$\eta^\varepsilon = \Delta\zeta^\varepsilon \tag{3.16}$$

with $\zeta^\varepsilon \in H_0^1$ and $\zeta^\varepsilon \in \phi(X^\varepsilon)$ a.e.. In conclusion, X^ε is a strong solution to (3.10). Passing to the limit in (3.8), (3.9) yields

$$\mathbb{E}\sup_{t\in[0,T]}\|X_t^\varepsilon\|_2^2 + 2\varepsilon\mathbb{E}\int_0^T \|X_r^\varepsilon\|_{H_0^1}^2 dr \le C(\mathbb{E}\|x_0\|_2^2 + 1) \tag{3.17}$$

$$\mathbb{E}t\varphi^\varepsilon(X_t^\varepsilon) + \mathbb{E}\int_0^t r\|\varepsilon\Delta X^\varepsilon + \eta_r^\varepsilon\|_{H^{-1}}^2 dr \le C\left(\mathbb{E}\|x_0\|_{H^{-1}}^2 + 1\right)$$

and

$$\begin{aligned}\mathbb{E}\varphi^\varepsilon(X_t^\varepsilon) + \mathbb{E}\int_0^t \|\varepsilon\Delta X^\varepsilon + \eta_r^\varepsilon\|_{H^{-1}}^2 dr &\le \mathbb{E}\varphi^\varepsilon(x_0)\\ &\le C(\mathbb{E}\|x_0\|_2^2 + 1).\end{aligned} \tag{3.18}$$

Step 2: $\varepsilon \to 0$

For $\varepsilon_1, \varepsilon_2 > 0$ let $(X^{\varepsilon_1}, \eta^{\varepsilon_1})$, $(X^{\varepsilon_2}, \eta^{\varepsilon_2})$ be two strong solutions to (3.10) with initial conditions $x_0^1, x_0^2 \in L^2(\Omega; L^2)$ respectively. Itô's formula implies

$$\begin{aligned}&e^{-Kt}\|X_t^{\varepsilon_1} - X_t^{\varepsilon_2}\|_{H^{-1}}^2\\ &= \|x_0^1 - x_0^2\|_{H^{-1}}^2 + \int_0^t 2e^{-Kr}(\varepsilon_1\Delta X^{\varepsilon_1} + \eta^{\varepsilon_1} - (\varepsilon_2\Delta X^{\varepsilon_2} + \eta^{\varepsilon_2}), X_r^{\varepsilon_1} - X_r^{\varepsilon_2})_{H^{-1}} dr\\ &+ \int_0^t e^{-Kr}(X_r^{\varepsilon_1} - X_r^{\varepsilon_2}, B(X_r^{\varepsilon_1}) - B(X_r^{\varepsilon_2}))_{H^{-1}} dW_r\\ &+ \int_0^t e^{-Kr}\|B(X_r^{\varepsilon_1}) - B(X_r^{\varepsilon_2})\|_{L_2}^2 dr - K\mathbb{E}\int_0^t e^{-Kr}\|X_r^{\varepsilon_1} - X_r^{\varepsilon_2}\|_{H^{-1}}^2 dr.\end{aligned}$$

Due to (3.16) we have

$$(\eta^{\varepsilon_1} - \eta^{\varepsilon_2}, X^{\varepsilon_1} - X^{\varepsilon_2})_{H^{-1}} \le 0$$

and we note that

$$(\varepsilon_1\Delta X_r^{\varepsilon_1} - \varepsilon_2\Delta X_r^{\varepsilon_2}, X_r^{\varepsilon_1} - X_r^{\varepsilon_2})_{H^{-1}} \le 2(\varepsilon_1 + \varepsilon_2)(\|X_r^{\varepsilon_1}\|_2^2 + \|X_r^{\varepsilon_2}\|_2^2).$$

Hence, the Burkholder-Davis-Gundy inequality and Lemma B.1 imply

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}\|X_t^{\varepsilon_1} - X_t^{\varepsilon_2}\|_{H^{-1}}^2 \le & C\mathbb{E}\|x_0^1 - x_0^2\|_{H^{-1}}^2\\ &+ C(\varepsilon_1 + \varepsilon_2)\left(\mathbb{E}\|x_0^1\|_2^2 + \mathbb{E}\|x_0^2\|_2^2 + 1\right).\end{aligned} \tag{3.19}$$

Let now $x_0 \in L^2(\Omega; L^2)$ and for each $\varepsilon > 0$ let $(X^\varepsilon, \eta^\varepsilon)$ be a solution to (3.10) with initial condition x_0. Due to (3.19) there is an $\mathcal{F}_t$-adapted process $X \in L^2([0,T] \times \Omega; H^{-1})$ with $X_0 = x_0$ such that

$$X^\varepsilon \to X \quad \text{in } L^2(\Omega; C([0,T]; H^{-1})) \quad \text{for } \varepsilon \to 0.$$

Using (3.18) we can extract a weakly convergent subsequence

$$\varepsilon_n\Delta X^{\varepsilon_n} + \eta^{\varepsilon_n} \rightharpoonup \eta, \quad \text{in } L^2([0,T] \times \Omega; H^{-1}).$$

By step one we have $\varepsilon_n\Delta X^{\varepsilon_n} + \eta^{\varepsilon_n} \in -\partial\varphi^{\varepsilon_n}(X^{\varepsilon_n})$ a.e., hence

$$(\varepsilon_n\Delta X^{\varepsilon_n} + \eta^{\varepsilon_n}, X^{\varepsilon_n} - Y)_{L^2([0,T]\times\Omega;H^{-1})} + \bar{\varphi}^{\varepsilon_n}(X^{\varepsilon_n}) \le \bar{\varphi}^{\varepsilon_n}(Y),$$

for all $Y \in L^2([0,T] \times \Omega; H^{-1})$. For $v \in L^2([0,T] \times \Omega; H^{-1})$ we define

$$\bar{\varphi}(v) = \mathbb{E}\int_0^T \varphi(v_r) dr.$$

Again, due to [15, Theorem 21] we have

$$\partial\bar{\varphi}(v) = \{\eta \in L^2([0,T]\times\Omega; H^{-1}) |\ \eta \in \partial\varphi(v),\ dt\otimes d\mathbb{P}\text{-a.e.}\}. \tag{3.20}$$

For $v \in L^2$ we observe that $\varphi^\varepsilon(v) = \frac{\varepsilon}{2}\|v\|_2^2 + \varphi(v)$ and thus

$$\begin{aligned}(\varepsilon_n\Delta X^{\varepsilon_n} + \eta^{\varepsilon_n}, X^{\varepsilon_n} - Y)_{L^2([0,T]\times\Omega;H^{-1})} + \bar{\varphi}(X^{\varepsilon_n})\\ \leq \bar{\varphi}(Y) + C\varepsilon_n\|Y\|^2_{L^2([0,T]\times\Omega;L^2)},\end{aligned} \tag{3.21}$$

for all $Y \in L^2([0,T]\times\Omega; L^2)$.

Let $J^\lambda = (1-\lambda\Delta)^{-1}$ be the resolvent of $-\Delta$ on H^{-1}. Then

$$\|J^\lambda v\|_{H^{-1}} \leq \|v\|_{H^{-1}}$$

and $J^\lambda v \to v$ in H^{-1} for $\lambda \to 0$. Moreover,

$$\|J^\lambda v\|_{L^2} \leq \frac{C}{\lambda}\|v\|_{H^{-1}}$$

for all $v \in H^{-1}$. For $v \in L^{m+1}\cap H^{-1}$ we have

$$\varphi(J^\lambda v) = \|J^\lambda v\|_{L^{m+1}} \leq \|v\|_{L^{m+1}} = \varphi(v).$$

For the case $m = 0$, in addition: Let $v = v_\mu \in \mathcal{M}\cap H^{-1}$. Since φ is the lower-semicontinuous hull of φ restricted to $L^1\cap H^{-1}$ (cf. Appendix D), there is a sequence $v^n \in L^1\cap H^{-1}$ such that $v^n \to v$ in H^{-1} and $\varphi(v^n)\to\varphi(v)$. By lower-semicontinuity of φ we conclude

$$\begin{aligned}\varphi(J^\lambda v) &\leq \lim_{n\to\infty}\varphi(J^\lambda v^n)\\ &\leq \lim_{n\to\infty}\varphi(v^n)\\ &= \varphi(v), \quad \forall v\in\mathcal{M}\cap H^{-1}.\end{aligned} \tag{3.22}$$

Given $Y \in L^2([0,T]\times\Omega; H^{-1})$ we set $Y^\varepsilon := J^{\varepsilon^{\frac{1}{4}}}Y \in L^2([0,T]\times\Omega; L^2)$. By dominated convergence $Y^\varepsilon \to Y$ in $L^2([0,T]\times\Omega;H^{-1})$. From (3.21) we obtain

$$\begin{aligned}&(\varepsilon_n\Delta X^{\varepsilon_n} + \eta^{\varepsilon_n}, X^{\varepsilon_n} - Y^{\varepsilon_n})_{L^2([0,T]\times\Omega;H^{-1})} + \bar{\varphi}(X^{\varepsilon_n})\\ &\leq \bar{\varphi}(Y^{\varepsilon_n}) + C\varepsilon_n\|Y^{\varepsilon_n}\|^2_{L^2([0,T]\times\Omega;L^2)}\\ &\leq \bar{\varphi}(Y) + C\sqrt{\varepsilon_n}\|Y\|^2_{L^2([0,T]\times\Omega;H^{-1})}.\end{aligned}$$

Taking $n\to\infty$ and using lower semicontinuity of $\bar{\varphi}$ we arrive at

$$(\eta, X - Y)_{L^2([0,T]\times\Omega;H^{-1})} + \bar{\varphi}(X) \leq \bar{\varphi}(Y)$$

for all $Y \in L^2([0,T]\times\Omega;H^{-1})$ and thus $\eta\in\partial\bar{\varphi}(X)$, which implies $\eta\in\partial\varphi(X)$ a.e. by (3.20). In conclusion, X is a strong solution to

$$dX_t \in -\partial\varphi(X_t)dt + B(X_t)dW_t. \tag{3.23}$$

Taking the limit in (3.17), (3.18) yields

$$\mathbb{E}\sup_{t\in[0,T]}\|X_t\|_2^2 \leq C(\mathbb{E}\|x_0\|_2^2 + 1) \tag{3.24}$$

$$\mathbb{E}t\varphi(X_t) + \mathbb{E}\int_0^t r\|\eta_r\|^2_{H^{-1}}dr \leq C\left(\mathbb{E}\|x_0\|^2_{H^{-1}} + 1\right)$$

and

$$\begin{aligned}\mathbb{E}\varphi(X_t) + \mathbb{E}\int_0^t \|\eta_r\|^2_{H^{-1}}dr &\leq C\mathbb{E}\varphi(x_0)\\ &\leq C(\mathbb{E}\|x_0\|_2^2 + 1).\end{aligned} \tag{3.25}$$

Moreover, from (3.19) we obtain

$$\mathbb{E}\sup_{t\in[0,T]}\|X_t^1-X_t^2\|_{H^{-1}}^2 \le C\mathbb{E}\|x_0^1-x_0^2\|_{H^{-1}}^2, \tag{3.26}$$

where X^1, X^2 are the corresponding limits for the initial conditions x_0^1, x_0^2 respectively.

Step 3: Proof of (i)

Suppose $x_0 \in L^2(\Omega; H^{-1})$ satisfying $\mathbb{E}\varphi(x_0) < \infty$. We consider the case $m = 0$, the case $m > 0$ can be treated analogously. Let $J^\lambda = (1-\lambda\Delta)^{-1}$ be the resolvent of $-\Delta$ on H^{-1}, set $x_0^n = J^{\frac{1}{n}}x_0$ and let (X^n, η^n) be the corresponding strong solution to (3.23) constructed in step two. By dominated convergence we have

$$x_0^n \to x_0, \quad \text{in } L^2(\Omega; H^{-1}).$$

Due to (3.25), (3.26) and (3.22) we may extract (weakly) convergent subsequences

$$\begin{aligned} X^n &\to X, \quad \text{in } L^2(\Omega; C([0,T]; H^{-1})) \\ \eta^n &\rightharpoonup \eta, \quad \text{in } L^2([0,T]\times\Omega; H^{-1}). \end{aligned}$$

Since $\eta^n \in -\partial\bar{\varphi}(X^n)$, strong-weak closedness of the subgradient $\partial\bar{\varphi}$ implies $\eta \in -\partial\bar{\varphi}(X)$ and thus $\eta \in -\partial\varphi(X)$ a.e.. It then easily follows that (X,η) is a strong solution to (3.1).

Step 3: Proof of (ii)

Let $x_0 \in L^2(\Omega; H^{-1})$ and $x_0^n \in L^2(\Omega; L^2)$ with $x_0^n \to x$ in $L^2(\Omega; H^{-1})$, $\mathbb{E}\|x_0^n\|_{H^{-1}}^2 \le \mathbb{E}\|x_0\|_{H^{-1}}^2$ and let (X^n, η^n) be the corresponding strong solutions to (3.23) constructed in step two. By (3.26) we have

$$\mathbb{E}\sup_{t\in[0,T]}\|X^n - X^m\|_{H^{-1}}^2 \le C\mathbb{E}\|x_n - x_m\|_{H^{-1}}^2.$$

Hence, $X^n \to X$ in $L^2(\Omega; C([0,T]; H^{-1}))$. Moreover,

$$\mathbb{E}t\varphi(X_t^n) + \mathbb{E}\int_0^t r\|\eta_r^n\|_{H^{-1}}^2 dr \le C\left(\mathbb{E}\|x_0\|_H^2 + 1\right).$$

Hence, there is a map η with $\eta \in L^2([\tau,T]\times\Omega; H_0^1)$ such that

$$\eta^n \rightharpoonup \eta, \quad \text{in } L^2([\tau,T]\times\Omega; H^{-1}),$$

for all $\tau > 0$. By strong-weak closedness of $\partial\bar{\varphi}$ we have $\eta \in -\partial\bar{\varphi}(X)$ and thus $\eta \in -\partial\varphi(X)$ a.e.. Hence, X is a generalized strong solution satisfying

$$\mathbb{E}t\varphi(X_t) + \mathbb{E}\int_0^t r\|\eta\|_{H^{-1}}^2 dr \le C\left(\mathbb{E}\|x_0\|_{H^{-1}}^2 + 1\right).$$

□

Appendix A. (Generalized) strong solutions to gradient type SPDE

Let $\varphi : H \to \mathbb{R}$ be a proper, lower-semicontinuous, convex function on a separable real Hilbert space H. We consider SPDE of the type

$$\begin{aligned} dX_t &\in -\partial\varphi(X_t)dt + B(t,X_t)dW_t, \\ X_0 &= x_0, \end{aligned} \tag{A.1}$$

where W is a cylindrical Wiener process in a separable Hilbert space U defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t\ge0}$ and $B : [0,T]\times H\times\Omega \to L_2(U,H)$ is Lipschitz continuous, i.e.

$$\|B(t,v) - B(t,w)\|_{L_2(U,H)}^2 \le C\|v-w\|_H^2 \quad \forall v,w \in H$$

and all $(t,\omega) \in [0,T]\times\Omega$. Furthermore, we assume that

$$\|B(\cdot,0)\|_{L_2(U,H)} \in L^2([0,T]\times\Omega).$$

We then define

Definition A.1. Let $x_0 \in L^2(\Omega; H)$. An H-continuous, $\mathcal{F}_t$-adapted process $X \in L^2(\Omega; C([0,T]; H))$ for which there exists a selection $\eta \in -\partial\varphi(X)$, $dt \otimes d\mathbb{P}$-a.e. is said to be a

i. strong solution to (A.1) if
$$\eta \in L^2([0,T] \times \Omega; H)$$
and $\mathbb{P}$-a.s.
$$X_t = x_0 + \int_0^t \eta_r dr + \int_0^t B(r, X_r) dW_r, \quad \forall t \in [0,T].$$

ii. generalized strong solution to (A.1) if
$$\eta \in L^2([\tau,T] \times \Omega; H), \quad \forall \tau > 0$$
and $\mathbb{P}$-a.s.
$$X_t = X_\tau + \int_\tau^t \eta_r dr + \int_\tau^t B(r, X_r) dW_r, \quad \forall t \in [\tau,T],$$
for all $\tau > 0$.

Appendix B. Non-degenerate, non-singular stochastic fast diffusion equations

In this section we consider non-degenerate, non-singular approximations to (1.1), that is

$$\begin{aligned} dX_t &= \varepsilon \Delta X_t dt + \Delta\phi(X_t) dt + B(t, X_t) dW_t, \\ X_0 &= x_0, \end{aligned} \tag{B.1}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous, monotone function satisfying $\phi(0) = 0$. We further assume that W is a cylindrical Wiener process in a separable Hilbert space U defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t\geq 0}$ and $B : [0,T] \times H^{-1} \times \Omega \to L_2(U, H^{-1})$ is Lipschitz continuous, i.e.
$$\|B(t,v) - B(t,w)\|^2_{L_2(U,H^{-1})} \le C\|v - w\|^2_{H^{-1}} \quad \forall v, w \in H^{-1},$$
for some constant $C > 0$ and all $(t, \omega) \in [0,T] \times \Omega$. We assume
$$\|B(t,v)\|^2_{L_2(U,L^2)} \le C(1 + \|v\|_2^2), \quad \forall v \in L^2, \tag{B.2}$$
and all $(t, \omega) \in [0,T] \times \Omega$. By [13] there is a unique variational solution X to (B.1) with respect to the Gelfand triple
$$L^2 \hookrightarrow H^{-1} \hookrightarrow (L^2)^*.$$
Under an additional regularity assumption on the diffusion coefficients B we prove that in fact, these solutions are strong solutions in H^{-1}.

Lemma B.1. *Let $x_0 \in L^2(\Omega; L^2)$. Then*
$$\mathbb{E} \sup_{t\in[0,T]} \|X_t\|_2^2 + \varepsilon \mathbb{E} \int_0^T \|X_r\|^2_{H_0^1} dr \le C(\mathbb{E}\|x_0\|_2^2 + 1),$$
with a constant $C > 0$ independent of ε and ϕ.

Proof. In the following we let $(e_i)_{i=1}^\infty$ be an orthonormal basis of eigenvectors of $-\Delta$ in H^{-1}. We further let $P^n : H^{-1} \to \operatorname{span}\{e_1, \dots, e_n\}$ be the orthogonal projection onto the span of the first n eigenvectors. We recall that the unique variational solution X to (B.1) is constructed in [13] as a (weak) limit in $L^2([0,T]\times\Omega;L^2)$ of the solutions to the following Galerkin approximation

$$\begin{aligned} dX^n_t &= \varepsilon P^n \Delta X^n_t dt + P^n\Delta\phi(X^n_t)dt + P^n B(t,X^n_t)dW^n_t,\\ X^n_0 &= P^n x_0. \end{aligned}$$

We observe that

$$\begin{aligned} \|X^n_t\|_2^2 &= \|P^n x_0\|_2^2 + 2\int_0^t (X^n_r, \varepsilon P^n\Delta X^n_r + P^n\Delta\phi(X^n_r))_2 dr\\ &\quad + 2\int_0^t (X^n_r, P^n B(r,X^n_r)dW^n_r)_2 dr + \int_0^t \|P^n B(r,X^n_r)\|^2_{L_2(U,L^2)}dr\\ &= \|P^n x_0\|_2^2 - 2\varepsilon\int_0^t \|X^n_r\|^2_{H^1_0}dr + 2\int_0^t {}_{H^1_0}\langle X^n_r, \Delta\phi(X^n_r)\rangle_{H^{-1}}dr\\ &\quad + 2\int_0^t (X^n_r, P^n B(r,X^n_r)dW^n_r)_2 dr + \int_0^t \|P^n B(r,X^n_r)\|^2_{L_2(U,L^2)}dr. \end{aligned}$$

Since ϕ can be approximated by a sequence of increasing, Lipschitz functions in $C^1(\mathbb{R})$ we observe that

$${}_{H^1_0}\langle X^n_r, \Delta\phi(X^n_r)\rangle_{H^{-1}} = -(\nabla X^n_r, \nabla\phi(X^n_r))_2 \le 0.$$

Using this, the Burkholder-Davis-Gundy inequality and (B.2) we obtain

$$\frac{1}{2}\mathbb{E}\sup_{t\in[0,T]} e^{-Ct}\|X^n_t\|_2^2 \le \mathbb{E}\|x_0\|_2^2 - 2\varepsilon\mathbb{E}\int_0^T e^{-Cr}\|X^n_r\|^2_{H^1_0}dr + C.$$

Passing to the (weak) limit yields the result. $\square$

Appendix C. Moreau-Yosida approximation of singular powers

In this section we collect and prove some facts about the Moreau-Yosida approximation of certain monomials (cf. e.g. [1, Section 2.2] for backgound on the Moreau-Yosida approximation).

For $m\in[0,1]$ let $\psi(r) := \frac{1}{m+1}|r|^{m+1}$, $r\in\mathbb{R}$ and let $\psi^\varepsilon:\mathbb{R}\to\mathbb{R}$ be the Moreau-Yosida approximation of ψ, i.e.

$$\psi^\varepsilon(r) := \inf_{s\in\mathbb{R}}\left\{\frac{|r-s|^2}{2\varepsilon} + \psi(s)\right\}.$$

Then

$$\partial\psi^\varepsilon = \phi^\varepsilon(r) := \frac{1}{\varepsilon}(r - J^\varepsilon r) \in \phi(J^\varepsilon r) \quad \forall r\in\mathbb{R},$$

is the Yosida approximation of $\phi = \partial\psi$, where $J^\varepsilon = (I+\varepsilon\phi)^{-1}$ is the resolvent of ϕ with I denoting the identity map on $\mathbb{R}$. We note that

$$|\phi^\varepsilon(r)| \le |\phi(r)| := \inf\{|\eta| : \eta\in\phi(r)\} \quad \forall r\in\mathbb{R}. \tag{C.1}$$

Moreover,

$$\begin{aligned} \psi^\varepsilon(r) &= \frac{1}{2\varepsilon}|r - J^\varepsilon r|^2 + \psi(J^\varepsilon r)\\ &= \frac{\varepsilon}{2}|\phi^\varepsilon(r)|^2 + \psi(J^\varepsilon r) \quad \forall r\in\mathbb{R} \end{aligned} \tag{C.2}$$

and thus

$$\psi(J^\varepsilon r) \le \psi^\varepsilon(r) \le \psi(r) \quad \forall r\in\mathbb{R}. \tag{C.3}$$

By the subgradient inequality we have

$$\eta(J^\varepsilon r - r) + \psi(r) \le \psi(J^\varepsilon r)$$

for all $\eta \in \phi(r)$. Hence, using the definition of ϕ^ε

$$\begin{aligned}\psi(r)-\psi(J^\varepsilon r) &\le -\eta(J^\varepsilon r - r)\\ &\le |\eta|\varepsilon|\phi^\varepsilon(r)|\end{aligned}$$

for every $\eta \in \phi(r)$. Hence, using (C.1) and (C.3) we obtain

$$\begin{aligned}|\psi(r)-\psi^\varepsilon(r)| &\le \varepsilon|\phi(r)|^2\\ &\le C\varepsilon(1+\psi(r)) \quad \forall r\in\mathbb{R}.\end{aligned} \tag{C.4}$$

We note that for all $a, b \in \mathbb{R}$

$$\begin{aligned}(\phi^{\varepsilon_1}(a)-\phi^{\varepsilon_2}(b))\cdot(a-b) =&(\phi^{\varepsilon_1}(a)-\phi^{\varepsilon_2}(b))\cdot(J^{\varepsilon_1}a-J^{\varepsilon_2}b)\\ &+(\phi^{\varepsilon_1}(a)-\phi^{\varepsilon_2}(b))\cdot(a-J^{\varepsilon_1}a-(b-J^{\varepsilon_2}b))\\ \ge&(\phi^{\varepsilon_1}(a)-\phi^{\varepsilon_2}(b))\cdot(\varepsilon_1\phi^{\varepsilon_1}(a)-\varepsilon_2\phi^{\varepsilon_2}(b))\\ \ge&-\frac{1}{2}(\varepsilon_1+\varepsilon_2)\left(|\phi^{\varepsilon_1}(a)|^2+|\phi^{\varepsilon_2}(b)|^2\right).\end{aligned}$$

Since

$$|\phi^{\varepsilon_1}(a)|^2 \le |\phi(a)|^2 \le C(1+|a|^2)$$

we conclude

$$(\phi^{\varepsilon_1}(a)-\phi^{\varepsilon_2}(b))\cdot(a-b) \ge -C(\varepsilon_1+\varepsilon_2)(1+|a|^2+|b|^2). \tag{C.5}$$

Appendix D. Relaxation of L^{m+1} norms on H^{-1}

For $m \ge 0$ we define

$$L^{m+1}\cap H^{-1} := \left\{ v\in L^{m+1} | \int vhdx \le C\|h\|_{H_0^1},\ \forall h\in C_c^1(\mathcal{O}) \text{ for some } C\ge 0\right\}.$$

By continuity, every $v \in L^{m+1}\cap H^{-1}$ the map $h \mapsto \int vhdx$ can be uniquely extended from C_c^1 to a bounded linear functional on H_0^1. Hence, $L^{m+1}\cap H^{-1} \subseteq H^{-1}$. We set

$$\varphi(v) := \begin{cases} \frac{1}{m+1}\|v\|_{m+1}^{m+1} &, \quad v\in L^{m+1}\cap H^{-1}\\ +\infty &, \quad H^{-1}\setminus(L^{m+1}\cap H^{-1}),\end{cases}$$

Lemma D.1. *Assume $m > 0$. Then φ is lower-semicontinuous on H^{-1}.*

Proof. Let $v^n \in L^{m+1}\cap H^{-1}$ with $\|v^n\|_{m+1}^{m+1} \le C$ and $v^n \to v$ in H^{-1}. Then $v^n \rightharpoonup v$ in L^{m+1} for some subsequence again denoted by v^n. Now $\|\cdot\|_{m+1}^{m+1}$ is weakly lower-semicontinuous on L^{m+1} and thus

$$\|v\|_{m+1}^{m+1} \le C.$$

□

Due to the lack of reflexivity of L^1 this argument fails in the case $m = 0$. We next provide a characterization of the corresponding lower-semicontinuous hull of $\varphi_0(\cdot) := \|\cdot\|_1$ on H^{-1}.

Let $\mathcal{M} = \mathcal{M}(\mathcal{O})$ be the space of signed Borel measures with finite total variation on $\mathcal{O} \subseteq \mathbb{R}^d$ and

$$\mathcal{M}\cap H^{-1} := \left\{\mu\in\mathcal{M} | \int_{\mathcal{O}} h(x)d\mu(x) \le C\|h\|_{H_0^1},\ \forall h\in C_c^1(\mathcal{O}) \text{ for some } C\ge 0\right\}.$$

By continuity, for every $v \in \mathcal{M}\cap H^{-1}$ the map $h \mapsto \int hdv$ can be extended from C_c^1 to a (uniquely determined) bounded linear functional on H_0^1. The resulting map $\iota : \mathcal{M}\cap H^{-1} \to H^{-1}$ thus is injective. Hence, $\mathcal{M}\cap H^{-1} \subseteq H^{-1}$ and in the following we identify $\mathcal{M}\cap H^{-1}$ with its embedding into H^{-1} , except for the proof of Lemma

D.2 below, where for $v \in \mathcal{M} \cap H^{-1}$ with $v = \iota(\mu)$ we write $v_\mu := \iota(\mu)$. We extend $\mu \in \mathcal{M}(\mathcal{O})$ to all of $\mathbb{R}^d$ by zero.

Lemma D.2. *Let $m = 0$ and φ be the lower-semicontinuous hull of $\varphi_0(\cdot) := \|\cdot\|_1$ on H^{-1}. Then*

$$\varphi(v) := \begin{cases} \mathrm{TV}(\mu) & , \quad v = v_\mu \in \mathcal{M} \cap H^{-1} \\ +\infty & , \quad \textit{otherwise.} \end{cases}$$

Proof. We first prove that φ is weakly (hence strongly) lower-semicontinuous on H^{-1}. We recall that $\mathcal{M}$ is the dual $C_0(\mathcal{O})^*$ of $C_0(\mathcal{O})$, i.e. the set of all continuous functions vanishing onf the boundary $\partial\mathcal{O}$, equipped with the sup-norm. Let $v_{\mu^n} \in \mathcal{M} \cap H^{-1}$ with $\varphi(v_{\mu^n}) \le C$ and $v_{\mu^n} \rightharpoonup v$ in H^{-1}. Since $C_0(\mathcal{O})$ is separable and μ^n has uniformly bounded total variation we have

$$\mu^n \rightharpoonup \tilde{\mu}$$

weakly* in $\mathcal{M}$ for some subsequence of μ^n and some $\tilde{\mu} \in \mathcal{M}$. Taking the limit $n \to \infty$ in

$$v_{\mu^n}(h) = \int_{\mathcal{O}} h(x) d\mu^n(x), \quad h \in C_c^1$$

and using $v_{\mu^n} \rightharpoonup v$ in H^{-1} yields

$$v(h) = \int_{\mathcal{O}} h(x) d\tilde{\mu}(x) \quad \forall h \in C_c^1$$

and thus $v = v_{\tilde{\mu}} \in \mathcal{M} \cap H^{-1}$. Since the total variation norm is lower semicontinuous with respect to weak convergence we obtain

$$\varphi(v) = \mathrm{TV}(\tilde{\mu}) \le \liminf_{n\to\infty} \mathrm{TV}(\mu^n) = \liminf_{n\to\infty} \varphi(v_{\mu^n}) \le C,$$

which proves φ to be weakly lower-semicontinuous on H^{-1}.

It remains to prove that φ is the lower-semicontinuous hull on H^{-1} of φ restricted to $L^1 \cap H^{-1}$. Let $v \in H^{-1}$ and let $v^\varepsilon := J^\varepsilon v \in L^2$, where $J^\varepsilon := (I - \varepsilon\Delta)^{-1}$ is the resolvent of $-\Delta$ in H^{-1}. Note that $v^\varepsilon \to v$ in H^{-1}, where we identify $v^\varepsilon \in L^2$ with its embedding into H^{-1} via

$$v^\varepsilon(h) = \int_{\mathcal{O}} v^\varepsilon h dx, \quad h \in H_0^1.$$

For $v \in L^2 \subseteq H^{-1}$ we observe

$$\begin{aligned} (J^\varepsilon v)(h) &= \int_{\mathcal{O}} (J^\varepsilon v) h dx \\ &= \int_{\mathcal{O}} v J^\varepsilon h dx \\ &= v(J^\varepsilon h) \quad \forall h \in H_0^1. \end{aligned}$$

By the density of L^2 in H^{-1} this yields

$$J^\varepsilon v = v \circ J^\varepsilon \quad \forall v \in H^{-1}.$$

Moreover, we recall $J^\varepsilon : C_0(\mathcal{O}) \to C_0(\mathcal{O})$, $\|J^\varepsilon h\|_{C_0} \le \|h\|_{C_0}$ for all $h \in C_0$ and for $h \in C_c^\infty(\mathcal{O})$ we have

$$J^\varepsilon h \to h \quad \text{in } C_0. \tag{D.1}$$

Let $\mu \in \mathcal{M} = (C_0(\mathcal{O}))^*$. We define $\mu^\varepsilon(h) = (J^\varepsilon\mu)(h) := \mu(J^\varepsilon h)$ for $h \in C_0$ and identify μ^ε with its representation in $\mathcal{M}$. Hence,

$$\int_{\mathcal{O}} h d\mu^\varepsilon = \int_{\mathcal{O}} J^\varepsilon h d\mu$$

for all $h \in C_0$ and thus

$$\begin{aligned}TV(\mu^\varepsilon) &= \sup_{\|h\|_{C_0 \le 1}} \mu^\varepsilon(h) \\ &\le TV(\mu) \sup_{\|h\|_{C_0 \le 1}} \|J^\varepsilon h\|_{C_0} \\ &\le TV(\mu).\end{aligned}$$

Consequently, there is a subsequence $\mu^{\varepsilon_n} \rightharpoonup \tilde{\mu}$ weakly* in $\mathcal{M}$ and due to (D.1) this means $\mu^{\varepsilon_n} \rightharpoonup \mu$ weakly* in $\mathcal{M}$ and

$$TV(\mu) \le \liminf_{n\to\infty} TV(\mu^{\varepsilon_n}) \le TV(\mu).$$

A standard contradiction argument then yields

$$\lim_{\varepsilon\to 0} TV(\mu^\varepsilon) = TV(\mu). \tag{D.2}$$

Let now $v_\mu \in \mathcal{M} \cap H^{-1}$. Then

$$\begin{aligned}(J^\varepsilon v_\mu)(h) &= v_\mu(J^\varepsilon h) \\ &= \int_{\mathcal{O}} J^\varepsilon h d\mu \\ &= \int_{\mathcal{O}} h d\mu^\varepsilon,\end{aligned}$$

for all $h \in C_c^1$ and thus $J^\varepsilon v_\mu \in \mathcal{M} \cap H^{-1}$ with $J^\varepsilon v_\mu = v_{J^\varepsilon \mu}$. Due to (D.2) this implies

$$\begin{aligned}\varphi(J^\varepsilon v_\mu) &= \varphi(v_{J^\varepsilon \mu}) \\ &= TV(J^\varepsilon \mu) \\ &\to TV(\mu) \\ &= \varphi(v_\mu).\end{aligned}$$

Since $J^\varepsilon v_\mu \in L^2 \subseteq L^1 \cap H^{-1}$ and $J^\varepsilon v_\mu \to v_\mu$ in H^{-1}, φ is the lower-semicontinuous hull on H^{-1} of φ restricted to $L^1 \cap H^{-1}$. $\square$

Lemma D.3. *Let $m \in [0,1]$ and $v \in L^{m+1} \cap H^{-1}$. Then*

$$\partial\varphi(v) \supseteq \{-\Delta w : \ w \in H_0^1, \ w \in \phi(v) \ a.e.\}.$$

Proof. **Case $m > 0$:** Let J_ε, $\varepsilon > 0$ be as in the proof of Lemma D.2. Assume that $w = \phi(v) = |v|^{m-1}v \in H_0^1$. Then $w \in L^{\frac{m+1}{m}}$ and we have to show that

$$\varphi(v) \le (-\Delta w, v - y) + \varphi(y) \quad \forall y \in L^{m+1} \cap H^{-1}.$$

As in the proof of Lemma D.2 we have for all $y \in L^{m+1} \cap H^{-1}$

$$\begin{aligned}&\varphi(v) - (-\Delta w, v - y)_{H^{-1}} \\ &= \varphi(v) - \lim_{\varepsilon\to 0}(-\Delta w, J_\varepsilon(v - y))_{H^{-1}} \\ &= \varphi(v) - \lim_{\varepsilon\to 0}(w, J_\varepsilon(v - y))_2 \\ &= \varphi(v) - \int_{\mathcal{O}} w(v - y) dx,\end{aligned}$$

where we used that $J_\varepsilon v \to v$ in L^p as $\varepsilon \to 0$ for every $v \in L^p$ and all $p \in [1, \infty)$. The last expression equals

$$\begin{aligned}
&\frac{1}{m+1}\int_{\mathcal{O}} |v|^{m+1}dx - \int_{\mathcal{O}} w(v-y)dx \\
&\leq \frac{1}{m+1}\int_{\mathcal{O}} |v|^{m+1}dx - \int_{\mathcal{O}} |v|^{m+1}dx + \int_{\mathcal{O}} |v|^{m-1}vydx \\
&\leq \varphi(y),
\end{aligned}$$

where we used Hölder's and Young's inequality in the last step. This finishes the proof.

Case $m = 0$: Assume $w \in \phi(v)$ a.e. with $w \in H_0^1$. Arguing as in the case $m > 0$ and using $|w| \leq 1$ a.e. we have

$$\begin{aligned}
\varphi(v) - (-\Delta w, v - y)_{H^{-1}} &= \varphi(v) - \int_{\mathcal{O}} w(v-y)dx \\
&\leq \int_{\mathcal{O}} |y|dx,
\end{aligned}$$

which finishes the proof. □

References

[1] Viorel Barbu. *Nonlinear differential equations of monotone types in Banach spaces*. Springer Monographs in Mathematics. Springer, New York, 2010.

[2] Viorel Barbu, Giuseppe Da Prato, and Michael Röckner. Finite time extinction for solutions to fast diffusion stochastic porous media equations. *C. R. Math. Acad. Sci. Paris*, 347(1-2):81–84, 2009.

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

[4] Viorel Barbu, Giuseppe Da Prato, and Michael Röckner. Stochastic porous media equations and self-organized criticality. *Comm. Math. Phys.*, 285(3):901–923, 2009.

[5] Viorel Barbu, Giuseppe Da Prato, and Michael Röckner. Finite time extinction of solutions to fast diffusion equations driven by linear multiplicative noise. *J. Math. Anal. Appl.*, 389(1):147–164, 2012.

[6] Viorel Barbu and Michael Röckner. Stochastic porous media equations and self-organized criticality: convergence to the critical state in all dimensions. *Comm. Math. Phys.*, 311(2):539–555, 2012.

[7] Viorel Barbu and Michael Röckner. Stochastic variational inequalities and applications to the total variation flow perturbed by linear multiplicative noise. *Arch. Ration. Mech. Anal.*, pages 1–38, 2013.

[8] Benjamin Gess. Strong solutions for stochastic partial differential equations of gradient type. *J. Funct. Anal.*, 263(8):2355–2383, 2012.

[9] Benjamin Gess. Finite time extinction for stochastic sign fast diffusion and self-organized criticality. *to appear in Comm. Math. Phys.*, 2014.

[10] Benjamin Gess and Michael Röckner. Stochastic variational inequalities and regularity for degenerate stochastic partial differential equations. *arXiv:1405.5866*, pages 1–26, 2014.

[11] Benjamin Gess and Jonas M. Tölle. Multi-valued, singular stochastic evolution inclusions. *J. Math. Pures Appl. (9)*, 101(6):789–827, 2014.

[12] Wei Liu. Invariance of subspaces under the solution flow of SPDE. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 13(1):87–98, 2010.

[13] Claudia Prévôt and Michael Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[14] Jiagang Ren, Michael Röckner, and Feng-Yu Wang. Stochastic generalized porous media and fast diffusion equations. *J. Differential Equations*, 238(1):118–152, 2007.

[15] R. Tyrrell Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974. Lectures given at the Johns Hopkins University, Baltimore, Md., June, 1973, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 16.

[16] Michael Röckner. A Dirichlet problem for distributions and specifications for random fields. *Mem. Amer. Math. Soc.*, 54(324):v+76, 1985.

[17] Michael Röckner and Feng-Yu Wang. General extinction results for stochastic partial differential equations and applications. *J. Lond. Math. Soc. (2)*, 87(2):545–560, 2013.

[18] Juan L. Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, Oxford, 2007. Mathematical theory.

[19] Juan Luis Vázquez. *Smoothing and decay estimates for nonlinear diffusion equations*, volume 33 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2006. Equations of porous medium type.