Stochastic differential equations with variable structure driven by multiplicative Gaussian noise and sliding mode dynamic

Viorel Barbu[1], Stefano Bonaccorsi[2] and Luciano Tubaro[2]

[1]University Al. I. Cuza and Institute of Mathematics Octav Mayer, Iasi, Romania
[2]Department of Mathematics, University of Trento, Italy

July 6, 2016

Abstract

This work is concerned with existence of weak solutions to discontinuous stochastic differential equations driven by multiplicative Gaussian noise and sliding mode control dynamics generated by stochastic differential equations with variable structure, that is with jump nonlinearity. The treatment covers the finite dimensional stochastic systems and the stochastic diffusion equation with multiplicative noise.

1 Introduction

We consider here stochastic differential equations of the form

$$\begin{aligned} &dX + A\,X\,dt + f(X)\,dt = B(X)\,dW, \quad t \in (0,T) \\ &X(0) = x, \end{aligned} \tag{1.1}$$

in a real separable Hilbert space H, where $A\colon D(A) \subset H \to H$ is self-adjoint, positive definite such that $A^{-1+\delta}$ is of trace class for some $\delta \in (0,1)$, W is a cylindrical Wiener process of the form

$$W(t) = \sum_{j=1}^{\infty} \beta_j(t)\, e_j. \tag{1.2}$$

Here $\{e_j\}$ is an orthonormal basis in H, $Ae_j = \lambda_j e_j$ and $\{\beta_j\}_{j=1}^\infty$ is a mutually independent system of Brownian motions in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $(\mathcal{F}_t)_{t\geq 0}$. The operator $f\colon H \to H$ is Borel measurable and locally bounded while the operator $B\colon H \to \mathcal{L}^2(H)$ is Lipschitzian, where $\mathcal{L}^2(H)$ is the space of Hilbert-Schmidt operators on H.

It should be said that under these general conditions equation (1.1) is not well posed except the case of additive noise ($B(X) = I$) and f bounded, when (1.1) has a unique weak (martingale) solution, see [7]. Equation (1.1) has however a unique strong solution if f is Lipschitz or accretive and continuous, see [6], or more generally if f is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ with domain $D(f) = \mathbb{R}$, see [2].

Equations of the form (1.1) with discontinuous f describe systems with variable structure and, in particular, closed-loop control systems with "sliding" mode behaviour. Here we shall study from this perspective two special cases.

The first one is the finite dimensional system

$$\begin{aligned} &dX + f(X)\,dt = \sigma(X)\,dW \\ &X(0) = x \end{aligned} \tag{1.3}$$

where W is a n-dimensional Wiener process and $f \in L^\infty_{loc}(\mathbb{R}^n, \mathbb{R}^n)$, $\sigma \in \mathrm{Lip}(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^n))$. (In this case $H = \mathbb{R}^n$ and $B(X) \equiv \sigma(X)$.)

The second one is the stochastic partial differential equation

$$\begin{aligned} &dX - \Delta X\,dt + h(X)\,dt = b(X)\,dW, && \text{in } (0,T)\times\mathcal{O} \\ &X = 0, && \text{on } (0,T)\times\partial\mathcal{O} \\ &X(0,\xi) = x(\xi), && \xi \in \mathcal{O} \end{aligned} \tag{1.4}$$

in a bounded and open domain $\mathcal{O} \subset \mathbb{R}^d$, $d \geq 1$ with smooth boundary $\partial\mathcal{O}$, which is the special case of (1.1) in the space $H = L^2(\mathcal{O})$, where $A = -\Delta$, $D(A) = H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$, $f \equiv h$, W is a cylindrical Wiener process of the form (1.2) in $H = L^2(\mathcal{O})$ and the stochastic term $b(X)\,dW$ is a formal expression for $b(X)\,dW(t) = \sum_{j1}^\infty \mu_j\, b(X)\, e_j\, d\beta_j(t)$. In other words $B : H \to \mathcal{L}^2(H)$ is the realization of b in the space H, given by

$$B(x)y = b(x) \sum_{j=1}^\infty \mu_j \langle y, e_j\rangle_2\, e_j, \quad \forall y \in H = L^2(\mathcal{O})$$

and its Hilbert-Schmidt norm is

$$\|B(x)\|^2_{\mathcal{L}^2(H)} = \sum_{j=1}^\infty \mu_j^2\, |b(x)\, e_j|_2^2.$$

Like in deterministic case, in order to have existence for the solution of equation (1.4) one must extend it to a multivalued stochastic equation of the form

$$\begin{cases} dX - \Delta X\, dt + G(X)\, dt \ni b(X)\, dW, & \text{in } (0,T)\times\mathcal{O} \\ X = 0, & \text{on } (0,T)\times\partial\mathcal{O} \\ X(0,\xi) = x(\xi), & \xi\in\mathcal{O}, \end{cases} \tag{1.5}$$

here $G\colon L^2(\mathcal{O}) \to L^2(\mathcal{O})$ is the multivalued mapping defined by

$$G(X) = \{\eta \in L^2(\mathcal{O}); \eta(\xi) \in F(X(\xi)), \text{ for } \xi \text{ a.e. in } \mathcal{O}\}$$

where $F\colon \mathbb{R} \to 2^{\mathbb{R}}$ is the Filippov map associated with h, that is (see [8], [9]):

$$\begin{aligned} F(r) &= [m(h_r), M(h_r)], \qquad \forall r \in \mathbb{R} \\ m(h_r) &= \lim_{\delta\to 0} \operatorname*{ess\,inf}_{u\in[r-\delta,r+\delta]} h(u) \\ M(h_r) &= \lim_{\delta\to 0} \operatorname*{ess\,sup}_{u\in[r-\delta,r+\delta]} h(u). \end{aligned} \tag{1.6}$$

Roughly speaking, G is obtained from h by "filling" the jumps of h in discontinuity points. If $f \in L^\infty_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^n)$, where $n \geq 1$, as in the case of equation (1.3), the corresponding Filippov map $F\colon \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is defined as

$$F(r) = \bigcap_{\delta>0} \bigcap_{m(N)=0} \overline{\operatorname{conv} f(B_\delta(r) \setminus N)} \tag{1.7}$$

where m is the Lebesgue measure and $B_\delta(r)$ is the ball of centre r and radius δ. Of course $F(r_0) = f(r_0)$ in all continuity points r_0 of f. Then to get existence in (1.3) one should replace f by F given by (1.7). If f is monotone and measurable then F is maximal monotone in $\mathbb{R}^n \times \mathbb{R}^n$ and locally bounded in $\mathbb{R}^n$, see [1, Proposition 25], hence, as shown in [2, Theorem 2.2], equation (1.3) has a unique strong solution (see also [3]). In the general case we consider here, the best that we can however expect is only a martingale solution for (1.3) (see Theorem 2.1, in which in general we do not have the uniqueness of the solution).

Previously multivalued differential equations of this form with F of subgradient type were studied in context of existence theory for stochastic variational equations or in a more general setting, see [4, 5, 12] for a few recent works on this subject.

The main existence result for equation (1.3) is established in Section 2, where it's also given a "sliding mode" type result for this equation.

In Sections 3, 4 and 5 it is studied a similar problem for equation (1.5) and also for a stochastic parabolic system.

Systems of the form (1.1) with discontinuous nonlinear drift term f arises when one applies in system $dX + AX\,dt = B(X)\,dW$ a feedback controller $u = f(X)$ to force the trajectory of system to slide after some time on a given manifold Σ. In such a way the system is forced to move on a space of lower dimension than the original one and, as documented in literature, see [15], a major advantage of this approach is the robustness of sliding mode controller. In this paper we prove that there is a feedback controller for which the corresponding closed loop system has a martingale (weak) solution which moves on the manifold Σ after some time τ.

Notation We use the standard notation for the Sobolev spaces $H^k(\mathcal{O})$, $k = 1, 2$, $H_0^1(\mathcal{O})$ and the Lebesgue integrable function spaces on $\mathcal{O} \subset \mathbb{R}^n$. The norm of $H_0^1(\mathcal{O})$ is denoted by $\|\cdot\|_1$ and the norm of $L^p(\mathcal{O})$ by $|\cdot|_p$ $(1 \le p \le \infty)$. The scalar product of $L^2(\mathcal{O})$ and the duality pairing between $H_0^1(\mathcal{O})$ and the dual space $H^{-1}(\mathcal{O})$ is denoted by the same symbol $\langle \cdot, \cdot \rangle_2$. We denote by $C([0, T]; H)$ the space of all continuous H-valued functions on $[0, T]$ and we also refer to [6] for basic results pertaining stochastic processes with values in Hilbert spaces. Finally, we denote by $C_{\rm b}^k(\mathbb{R})$, $k = 0, 1$, the space of functions of class C^k on $\mathbb{R}$, with continuous and bounded derivatives up to order k. The norm in $\mathbb{R}$ or $\mathbb{R}^n$ is denoted by the same symbol $|\cdot|$, the difference being clear from the context.

2 Weak solution and "sliding" mode for equation (1.3)

We shall study here system (1.3) where W is a n-dimensional Wiener process defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ and $f \in L^\infty_{\rm loc}(\mathbb{R}^n; \mathbb{R}^n)$, $\sigma \in {\rm Lip}(\mathbb{R}^n; \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n))$.
We consider the Filippov map $F\colon \mathbb{R}^n \to \mathbb{R}^n$ associated with f which was introduced in (1.7).

Definition 2.1. *The system $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\ge 0}, W, X)$ is said to be a martingale solution to* (1.3) *if $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ is a filtered probability space on which it is defined an $(\mathcal{F}_t)_{t\ge 0}$-Wiener process W and X is an $(\mathcal{F}_t)_{t\ge 0}$-adapted, $\mathbb{R}^n$-valued, continuous process that satisfies $\mathbb{P}$-a.s. the equation*

$$X(t) + \int_0^t \eta(s)\,ds = x + \int_0^t \sigma(X(s))\,dW(s), \qquad \forall t \ge 0 \tag{2.1}$$

where $\eta = \eta(t)$ is an $\mathbb{R}^n$-valued $(\mathcal{F}_t)_{t\ge 0}$-adapted process such that

$$\eta \in L^2((0,T) \times \Omega), \quad \forall T > 0, \qquad \eta \in F(X), \quad \text{a.e. in } (0,\infty) \times \Omega.$$

This definition extends verbatim to infinite dimensional equation (1.1), see Definition 3.1 below. In literature such a solution is also called **weak solution**. A martingale solution which is $\bar{\mathcal{F}}_t^W$-adapted, where $\bar{\mathcal{F}}_t^W$ is the completed natural filtration of W is called strong solution, see [7].
We have

Theorem 2.1. *Assume that $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is measurable and that*

$$|f(r)| \le a_1\,|r| + a_2, \qquad \forall r \in \mathbb{R}^n \tag{2.2}$$

where $a_1,\ a_2 \ge 0$. Then for each $x \in \mathbb{R}^n$ there is at least one martingale solution $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\mathcal{F}_t)_{t\ge 0}, \tilde{\mathbb{P}}, \tilde{W}, \tilde{X})$ to (1.3) which satisfies the estimate

$$\tilde{\mathbb{E}} \sup_{t\in[0,T]} |\tilde{X}(t)|^2 \le C_T\,(|x|^2+1), \quad x \in R^n. \tag{2.3}$$

Proof. Consider the approximating equation

$$\begin{aligned} &dX_\varepsilon + f_\varepsilon(X_\varepsilon)\,dt = \sigma(X_\varepsilon)\,dW, \qquad t\in[0,T] \\ &X(0) = x \end{aligned} \tag{2.4}$$

where f_ε is a smooth approximation of f given by

$$f_\varepsilon(r) = \int_{\mathbb{R}^n} f(r-\varepsilon\,\theta)\,\rho(\theta)\,d\theta, \quad \forall \varepsilon>0, r\in\mathbb{R}^n. \tag{2.5}$$

Here $\rho \in C_0^\infty(\mathbb{R}^n)$ is any mollifier such that

$$\rho(r)\ge 0, \quad \rho(r)=\rho(-r), \quad \rho(r)=0 \text{ for } |r|\ge 1, \quad \int_{\mathbb{R}^n}\rho(r)\,dr = 1. \tag{2.6}$$

Let $X_\varepsilon \in L^2(\Omega; C([0,T];\mathbb{R}^n))$ be the strong solution to (2.4) (See Lemma 6.1). By (2.2) and Itô's formula it follows that

$$\begin{aligned} \tfrac12\,|X_\varepsilon(t)|^2 &\le \tfrac12\,|x|^2 + C\int_0^t (1+|X_\varepsilon(s)|^2)\,ds \\ &+ \tfrac12\int_0^t \mathrm{Tr}(\sigma(X_\varepsilon(s))\sigma^*(X_\varepsilon(s)))\,ds + \int_0^t X_\varepsilon(s)\cdot\sigma(X_\varepsilon(s))\,dW_s, \quad t\ge 0 \end{aligned}$$

and so by the Burkholder-Davis-Gundy theorem (see e.g. [6]) we have

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_\varepsilon(t)|^2\Big] \le C\,(1+|x|^2), \qquad \forall\varepsilon>0 \tag{2.7}$$

(Here and everywhere in the following we shall denote by C several positive constants independent of ε.)

We set $Y_\varepsilon = (X_\varepsilon, W)$ and we consider $\nu_\varepsilon = \mathcal{L}(Y_\varepsilon)$ (the law of Y_ε) that is $\nu_\varepsilon(\Gamma) = \mathbb{P}[Y_\varepsilon \in \Gamma]$ for each Borelian set $\Gamma \subset C([0,T];\mathbb{R}^n) \times C([0,T];\mathbb{R}^n)$. Let us show that $\{\nu_\varepsilon\}$ is tight in $(C([0,T];\mathbb{R}^n))^2 = C([0,T];\mathbb{R}^n) \times C([0,T];\mathbb{R}^n)$. This means that for each $\delta > 0$ there is a compact subset Γ of $(C([0,T];\mathbb{R}^n))^2$ such that $\nu_\varepsilon(\Gamma^c) \leq \delta$ for all $\varepsilon > 0$. We take for $r > 0$, $\gamma > 0$,

$$\Gamma = B_{r,\gamma} = \{y \in (C([0,T];\mathbb{R}^n))^2 : |y(t)| \leq r, \forall t \in [0,T], \\ |y(t) - y(s)| \leq \gamma \, |t-s|^{\frac{1}{2}}, \quad \forall t, s \in [0,T]\}$$

Clearly, by the Ascoli-Arzelà theorem, $B_{r,\gamma}$ is compact in $(C([0,T];\mathbb{R}^n))^2$. On the other hand, by (2.4) we have, via Itô's formula applied to the process $t \to |X_\varepsilon(t) - X_\varepsilon(s)|^2$,

$$\tfrac{1}{2}\,\mathbb{E}|X_\varepsilon(t) - X_\varepsilon(s)|^2 + \mathbb{E}\int_s^t f_\varepsilon(X_\varepsilon(\theta)) \cdot (X_\varepsilon(\theta) - X_\varepsilon(s))\,d\theta \\ \leq C\,\mathbb{E}\int_s^t |X_\varepsilon(\theta)|^2\,d\theta, \quad 0 \leq s \leq t \leq T.$$

Taking into account estimate (2.7), we obtain via Gronwall's lemma that

$$\mathbb{E}|X_\varepsilon(t) - X_\varepsilon(s)|^2 \leq C\int_s^t |X_\varepsilon(\theta)|^2\,d\theta \leq C\;|t-s|. \tag{2.8}$$

By estimates (2.7), (2.8) and by the well known inequality

$$\rho\,\mathbb{P}[|Y| \geq \rho] \leq \mathbb{E}|Y|, \quad \forall \rho > 0,$$

we see that there are γ, r independent of ε such that $\nu_\varepsilon(B^c_{r,\gamma}) \leq \delta$, as desired. Then by the Skorohod's theorem there exist a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and random variables $\tilde{X}$, $\tilde{X}_\varepsilon$, $\tilde{W}_\varepsilon$, $\tilde{W}$ such that $\mathcal{L}(\tilde{X}_\varepsilon, \tilde{W}_\varepsilon) = \mathcal{L}(X_\varepsilon, W_\varepsilon)$ and for $\tilde{\mathbb{P}}$-almost every $\omega \in \tilde{\Omega}$

$$\begin{array}{ll} \tilde{W}_\varepsilon \to \tilde{W},\ \tilde{X}_\varepsilon \to \tilde{X} & \text{in } C([0,T];\mathbb{R}^n) \\ \sigma(\tilde{X}_\varepsilon) \to \sigma(\tilde{X}) & \text{in } C([0,T];\mathbb{R}^n) \end{array} \tag{2.9}$$

as $\varepsilon \to 0$. We have also $\mathcal{L}(f_\varepsilon(\tilde{X}_\varepsilon)) = \mathcal{L}(f_\varepsilon(X_\varepsilon))$ and so by (2.2), (2.5) it follows that on a subsequence, denoted by $\{\varepsilon_n\}$,

$$f_{\varepsilon_n}(\tilde{X}_{\varepsilon_n}) \to \tilde{\eta} \qquad \text{weakly in } L^2((0,T) \times \tilde{\Omega}). \tag{2.10}$$

Let us show that

$$\tilde{\eta} \in F(\tilde{X}), \qquad dt \times d\tilde{\mathbb{P}}\text{-a.e. in } (0,T) \times \tilde{\Omega}. \tag{2.11}$$

We have by (2.5),

$$f_\varepsilon(\tilde{X}_\varepsilon(t,\omega)) = \int_{\mathbb{R}^n} f(\tilde{X}_\varepsilon(t,\omega) - \varepsilon\,\theta)\,\rho(\theta)\,d\theta \in \overline{\mathrm{conv} f(B_\varepsilon(\tilde{X}_\varepsilon(t,\omega)))},$$
$$dt \times d\tilde{\mathbb{P}}\text{-a.e. in } \in [0,T]\times\tilde{\Omega},$$

and this implies that

$$\Sigma(t,\omega) = \big\{w - \lim_{\varepsilon_n\to 0} f_{\varepsilon_n}(\tilde{X}_\varepsilon(t,\omega))\big\} \subset F(\tilde{X}(t,\omega)), \quad dt\times d\tilde{\mathbb{P}}\text{-a.e. in } [0,T]\times\tilde{\Omega}.$$

By (2.10) and by Mazur's theorem (see e.g., [17, pag. 120]) it follows that there is a convex combination of f_{ε_n}, that is

$$\varphi_n(t,\omega) = \sum_{i=1}^{k_n} \alpha_i^{(n)} f_{\varepsilon_i}(\tilde{X}_{\varepsilon_i}(t,\omega)),$$

$\sum_{i=1}^{k_n}\alpha_i^{(n)} = 1$, $0 \le \alpha_i^{(n)} \le 1$, which is strongly convergent in $L^2((0,T)\times\tilde{\Omega})$ to $\tilde{\eta}$ and so on a subsequence again denoted by $\{n\}$, we have

$$\lim_{n\to\infty} \varphi_n(t,\omega) = \tilde{\eta}(t,\omega), \qquad \text{a.e. } (t,\omega)\in(0,T)\times\tilde{\Omega}.$$

Since $\lim_{n\to\infty}\varphi_n(t,\omega) \in F(\tilde{X}(t,\omega))$ we obtain (2.11) as claimed.

If we define

$$\tilde{\mathcal{F}}_t^\varepsilon = \sigma\big(\tilde{X}_\varepsilon(s), \tilde{W}_\varepsilon(s); 0\le s\le t\big), \qquad t\ge 0,$$

$$\tilde{\mathcal{F}}_t = \sigma\big(\tilde{X}(s), \tilde{W}(s); 0\le s\le t\big), \qquad t\ge 0,$$

then it follows that $(\tilde{W}_\varepsilon, \tilde{\mathcal{F}}_t^\varepsilon)$ and $(\tilde{W}, \tilde{\mathcal{F}}_t)$ are Wiener processes and that $\mathbb{P}$-a.s.,

$$\tilde{X}_\varepsilon(t) + \int_0^t f_\varepsilon(\tilde{X}_\varepsilon(s))\,ds = x + \int_0^t \sigma(\tilde{X}_\varepsilon(s))\,d\tilde{W}_\varepsilon(s), \qquad \forall t\in[0,T].$$

Taking into account (2.10) and that $\mathbb{P}$-a.s (see Lemma 3.1 in [10])

$$\lim_{\varepsilon\to 0}\int_0^t \sigma(\tilde{X}_\varepsilon(s))\,d\tilde{W}_\varepsilon(s) = \int_0^t \sigma(\tilde{X}(s))\,d\tilde{W}(s), \qquad \forall t\in[0,T],$$

we obtain that $\tilde{\mathbb{P}}$-a.s.

$$\tilde{X}(t) + \int_0^t \tilde{\eta}(s)\,ds = x + \int_0^t \sigma(\tilde{X}(s))\,d\tilde{W}(s), \qquad \forall t\in[0,T].$$

This means that the system $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\geq 0}, \tilde{\mathbb{P}}, \tilde{W}(t), \tilde{X}(t))$ is a martingale solution to (1.3). The estimate (2.3) follows by (2.7) which in turn implies that

$$\tilde{\mathbb{E}} \sup_{t\in[0,T]} |\tilde{X}_\varepsilon(t)|^2 \leq C\,(1+|x|^2), \quad \forall \varepsilon > 0,\ t\in[0,T].$$

Such a process $\tilde{X}$ can be extended to all of $(0,\infty)$. □

Remark 2.1. If f and σ are in $L^\infty(\mathbb{R}^n)$, σ is Lipschitzian and $\sigma^*\sigma$ is uniformly elliptic, that is

$$\sum_{i,j=1}^{n} (\sigma^*\sigma)_{ij}(x)\,\xi_i\xi_j \geq \alpha \sum_{i=1}^{n} \xi_i^2, \qquad \forall \xi = (\xi_i) \in \mathbb{R}^n, \tag{2.12}$$

for some $\alpha > 0$, then, as shown by A. Yu. Veretennikov [16], equation (1.3) has a unique strong solution X; on these lines see also [10]. It should be said however that for the applications we have in mind and more precisely for existence of a sliding mode, the nondegeneracy condition (2.12) is too restrictive.

Remark 2.2. If besides (2.2) one assumes that f is monotone from $\mathbb{R}^n$ to itself, that is

$$(f(r) - f(\bar{r}))\cdot(r-\bar{r}) \geq 0, \qquad \forall\, r,\, \bar{r} \in \mathbb{R}^n,$$

then the corresponding Filippov mapping F is maximal monotone (see [1, pag. 46]) and by a standard argument it follows that equation (1.3) has a unique strong solution X obtained as

$$X = \lim_{\lambda\to 0} X_\lambda \qquad \text{in } L^2(\Omega; C([0,T];\mathbb{R}^n)),$$

where X_λ is the solution to approximating equation

$$\begin{aligned} &dX_\lambda + F_\lambda(X_\lambda)\,dt = \sigma(X_\lambda)\,dW, \quad t\in(0,T),\\ &X_\lambda(0) = x \end{aligned}$$

and $F_\lambda = \frac{1}{\lambda}(I - (I+\lambda F)^{-1})$ is the Yosida approximation of F.

A typical example of differential systems with variable structure of the form (1.3) is

$$f(r) = \begin{cases} f_1(r) & \text{if } g(r) \geq 0 \\ f_2(r) & \text{if } g(r) < 0 \end{cases} \qquad \forall r \in \mathbb{R}^n \tag{2.13}$$

where $g \in C^2(\mathbb{R}^n)$, $f_1, f_2 \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. We assume here that f_i, $i = 1, 2$ and $\sigma \in \mathrm{Lip}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))$ satisfy the following conditions

$$|f_i(r)| \le a_{1i}\,|r| + a_{2i}, \quad \forall r \in \mathbb{R}^n, \quad i = 1, 2 \tag{2.14}$$

$$\sup_{r \in \mathbb{R}^n} \{|\nabla g(r)| + |D^2 g(r)|\} < \infty \tag{2.15}$$

$$\nabla g(r) \cdot f_1(r) \ge \alpha \quad \text{in } \{r \in \mathbb{R}^n : g(r) > 0\} \tag{2.16}$$

$$\nabla g(r) \cdot f_2(r) \le -\alpha \quad \text{in } \{r \in \mathbb{R}^n : g(r) < 0\} \tag{2.17}$$

$$|\sigma^*(r)\sigma(r)| \left(|\nabla g(r)|^2 + |g(r)||D^2 g(r)|\right) \le C^*\,|g(r)|^2, \quad \forall r \in \mathbb{R}^n \tag{2.18}$$

where $\alpha > 0$ and $C^* > 0$. In particular (2.18) implies that $|\sigma^*\sigma||\nabla g(r)|^2 = 0$ on $\{r : g(r) = 0\}$, which in general does not imply $\sigma = 0$ on $[g = 0]$. We note also that by (2.15)–(2.17) it follows that $\Sigma = \{r \in \mathbb{R}^n : g(r) = 0\}$ is a $n - 1$ dimensional C^2-manifold. We have

Theorem 2.2. *Under assumptions* (2.14)–(2.18) *for each $x \in \mathbb{R}^n$ there is a martingale solution $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \ge 0}, \tilde{\mathbb{P}}, \tilde{W}, \tilde{X})$ to* (1.3) *with the following properties:*
(i) if $g(x) = 0$ then $\tilde{\mathbb{P}}$-a.s. $g(\tilde{X}(t)) = 0, \quad \forall t \ge 0$;
(ii) if $g(x) \ne 0$ and $\tau = \inf\{t > 0 : g(\tilde{X}(t)) = 0\}$ then

$$\tilde{\mathbb{P}}(\tau > t) \le \frac{\tilde{C}}{\alpha}\,(1 - e^{-\tilde{C}\,t})^{-1}\,|g(x)|, \qquad \forall t > 0, \tag{2.19}$$

where $\tilde{C} = C_1\,C^*$, C_1 a positive constant independent of g and σ. If $C^* = 0$ then

$$\tilde{\mathbb{P}}(\tau > t) \le (\alpha\,t)^{-1}|g(x)|, \qquad \forall t > 0.$$

Theorem 2.2 amounts to say that the switching manifold $\Sigma = \{x : g(x) = 0\}$ is invariant for stochastic system (1.3) with f given by (2.13) and that for $x \notin \Sigma$ the solution $\tilde{X}$ have reached the manifold Σ by time t with a probability greater or equal to $1 - (\alpha\,t)^{-1}\,|g(x)|$. In the classical automatic control terminology (see, e.g., [15]) this means that $g(x) = 0$ is a "sliding mode" equation for system (1.3) and Σ is a switching surface for this system. As a matter of fact this is typical "sliding" mode behaviour for the solution $X = \tilde{X}(t)$ and its dynamics has two phases: the first phase is on time interval $(0, \tau)$ until X reaches the surface Σ and the second one for $t \ge \tau$ in which $X(t)$ evolves on the sliding surface Σ. The reaching time $\tau = \tau(\omega)$ is a stopping time determined by (2.19).

Proof of Theorem 2.2. We note first that the function f can be written as

$$f(r) = f_1(r)H(g(r)) + f_2(r)H(-g(r)), \qquad \forall r \in \mathbb{R}^n,\ g(r) \ne 0$$

where H is the Heaviside function, while the corresponding Filippov multi-valued function F corresponding to f (see (1.7)) is just

$$F(y) = \begin{cases} f(y) & \text{for } g(y) \neq 0 \\ \overline{\text{conv}}\big[f_1(y), \lim\limits_{\substack{z\to y\\ g(z)<0}} f_2(z)\big] & \text{for } g(y) = 0 \end{cases} \tag{2.20}$$

Let $\tilde{X}$ be the martingale solution to (1.3) given by (2.9) where f is as in (2.13). In order to prove the theorem we need a few apriori estimates on the solution X_ε to (2.4) which will be obtained by applying Itô's formula to the function $\phi_\lambda(u) = \varphi_\lambda(g(u))$, where $\varphi_\lambda \in C^2(\mathbb{R})$ is defined as

$$\varphi_\lambda(r) = (r^2 + \lambda^2)^{\frac{1}{2}}, \qquad \lambda > 0, \quad \forall r \in \mathbb{R}. \tag{2.21}$$

Taking into account that, $\forall u, v \in \mathbb{R}^n$, one has

$$D\phi_\lambda(u) = \varphi'_\lambda(g(u))\nabla g(u) = (|g(u)|^2 + \lambda^2)^{-\frac{1}{2}} g(u)\nabla g(u),$$

$$\begin{aligned} &D^2\phi_\lambda(u)(v) = \varphi''_\lambda(g(u))(\nabla g(u)\cdot v)\nabla g(u) + \varphi'_\lambda(g(u))D^2 g(u)(v) \\ &= -(|g(u)^2 + \lambda^2)^{-\frac{3}{2}}|g(u)|^2(\nabla g(u)\cdot v)\nabla g(u) + \\ &\qquad (|g(u)|^2 + \lambda^2)^{-\frac{1}{2}}((\nabla g(u)\cdot v)\nabla g(u) + g(u)D^2 g(u)(v)) \end{aligned} \tag{2.22}$$

and therefore

$$\begin{aligned} &d\varphi_\lambda(g(X_\varepsilon(t))) + (|g(X_\varepsilon(t))|^2 + \lambda^2)^{-\frac{1}{2}} g(X_\varepsilon(t)) f_\varepsilon(X_\varepsilon(t)) \cdot \nabla g(X_\varepsilon(t))\, dt \\ &\quad = \tfrac{1}{2}\operatorname{Tr}[\sigma^*(X_\varepsilon(t))\sigma(X_\varepsilon(t))D^2\phi_\lambda(X_\varepsilon(t))]\, dt \\ &\qquad + (|g(X_\varepsilon(t))|^2 + \lambda^2)^{-\frac{1}{2}} g(X_\varepsilon(t))\sigma(X_\varepsilon(t))\, dW(t) \cdot \nabla g(X_\varepsilon(t)). \end{aligned}$$

We note that in virtue of (2.18), (2.22) we have

$$\begin{aligned} \operatorname{Tr}[\sigma^*(X_\varepsilon)\sigma(X_\varepsilon)D^2\phi_\lambda(X_\varepsilon)] &\le C\, |\sigma^*(X_\varepsilon)\sigma(X_\varepsilon)| \big(|\nabla g(X_\varepsilon)|^2 + \\ &\quad |g(X_\varepsilon)||D^2 g(X_\varepsilon)|\big)\, (|g(X_\varepsilon(t))|^2 + \lambda^2)^{-\frac{1}{2}} \le C\, C^*\, |g(X_\varepsilon)| \end{aligned}$$

Letting $\lambda \to 0$ we obtain that for $0 \le s \le t < \infty$

$$\begin{aligned} &|g(X_\varepsilon(t))| + \int_s^t f_\varepsilon(X_\varepsilon(\theta)) \cdot \nabla g(X_\varepsilon(\theta))\, \operatorname{sgn}(g(X_\varepsilon(\theta))\mathbb{1}_{[|g(X_\varepsilon(\theta))|>0]}\, d\theta \\ &\qquad \le |g(X_\varepsilon(s))| + \tfrac{1}{2}\, C\, C^* \int_s^t |g(X_\varepsilon(\theta))|\, d\theta \\ &\quad + \int_s^t \sigma(X_\varepsilon(\theta))dW(\theta) \cdot \nabla g(X_\varepsilon(\theta))\, \operatorname{sgn}(g(X_\varepsilon(\theta)))\mathbb{1}_{[|g(X_\varepsilon(\theta))|>0]}\, d\theta. \end{aligned} \tag{2.23}$$

By (2.5) and (2.13) we have

$$
\begin{aligned}
&f_\varepsilon(X_\varepsilon(t)) \cdot \nabla g(X_\varepsilon(t)) \,\mathrm{sgn}(g(X_\varepsilon(t))) \\
&\quad = \int_{[g(X_\varepsilon(t)-\varepsilon\theta)>0]} f_1(X_\varepsilon(t) - \varepsilon\theta) \cdot \nabla g(X_\varepsilon(t))\rho(\theta)\, d\theta \\
&\qquad + \int_{[g(X_\varepsilon(t)-\varepsilon\theta)<0]} f_2(X_\varepsilon(t) - \varepsilon\theta) \cdot \nabla g(X_\varepsilon(t))\rho(\theta)\, d\theta
\end{aligned}
$$

and so taking into account (2.14)–(2.16) we get

$$
f_\varepsilon(X_\varepsilon(t)) \cdot \nabla g(X_\varepsilon(t)) \,\mathrm{sgn}(g(X_\varepsilon(t))) \geq \alpha - \delta(\varepsilon)(1 + |X_\varepsilon(t)|), \qquad \forall t \geq 0
$$

where $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Taking into account (2.23) this yields

$$
\begin{aligned}
&|g(X_\varepsilon(t))| + \alpha \int_s^t \mathbb{1}_{[|g(X_\varepsilon(\theta))|>0]}(1 - \delta(\varepsilon)|X_\varepsilon(\theta)|)\, d\theta \\
&\qquad \leq |g(X_\varepsilon(s))| + \int_s^t \mathbb{1}_{[|g(X_\varepsilon(\theta))|>0]} \mathrm{sgn}(g(X_\varepsilon(\theta))) \nabla g(X_\varepsilon(\theta)) \cdot \sigma(X_\varepsilon(\theta))\, dW(\theta) \\
&\qquad\quad + C\,C^* \int_s^t |g(X_\varepsilon(\theta))|\, d\theta.
\end{aligned}
$$

The same inequality remains of course true for $(\tilde{X}_\varepsilon, \tilde{W}_\varepsilon)$ and so letting $\varepsilon \to 0$ we get ~~that~~ again the same inequality for $\tilde{X}$ given by (2.9). Taking into account (2.18) we get

$$
\begin{aligned}
&|g(\tilde{X}(t))| + \alpha \int_s^t \mathbb{1}_{[|g(\tilde{X}(\theta))|>0]}\, d\theta \leq |g(\tilde{X}(s))| + C\,C^* \int_s^t |g(\tilde{X}(\theta))|\, d\theta \\
&+ \int_s^t \mathbb{1}_{[|g(\tilde{X}(\theta))|>0]} \mathrm{sgn}(g(\tilde{X}(\theta))) \nabla g(\tilde{X}(\theta)) \cdot \sigma(\tilde{X}(\theta))\, d\tilde{W}(\theta), \\
&\qquad\qquad\qquad\qquad\qquad\qquad 0 \leq s \leq t < \infty, \ \tilde{\mathbb{P}}\text{-a.s.}
\end{aligned}
$$

and so (see Lemma 6.3)

$$
\begin{aligned}
&e^{-\tilde{C}t}|g(\tilde{X}(t))| + \alpha \int_s^t e^{-\tilde{C}\theta} \mathbb{1}_{[|g(\tilde{X}(\theta))|>0]}\, d\theta \leq e^{-\tilde{C}s}|g(\tilde{X}(s))| + \\
&\quad \int_s^t e^{-\tilde{C}\theta} \mathbb{1}_{[|g(\tilde{X}(\theta))|>0]} \mathrm{sgn}(g(\tilde{X}(\theta))) \nabla g(\tilde{X}(\theta)) \cdot \sigma(\tilde{X}(\theta))\, d\tilde{W}(\theta), \\
&\qquad\qquad\qquad\qquad\qquad\qquad 0 \leq s \leq t < \infty.
\end{aligned} \tag{2.24}
$$

In particular, it follows by (2.24), with $s = 0$ and taking the expectation, that if $g(x) = 0$ then $g(\tilde{X}(t)) = 0$ $\tilde{\mathbb{P}}$-a.s. for all $t \geq 0$. Moreover, by (2.24) it follows that $Z(t) = |g(\tilde{X}(t))|e^{-\tilde{C}\,t}$ is a nonnegative super-martingale and therefore for any couple of stopping times $\tau_1 < \tau_2$ we have $Z(\tau_1) \geq Z(\tau_2)$. This implies that if $\tau = \inf\{t > 0 : |Z(t)| = 0\}$ we have that $Z(t) = Z(\tau)$, $\tilde{\mathbb{P}}$-a.s. for $t > \tau$. On the other hand, by (2.24) it follows that

$$\tilde{\mathbb{E}}Z(t) + \alpha \int_0^t e^{-\tilde{C}\,s}\tilde{\mathbb{P}}(\tau > s)\,ds \leq |g(x)| + \tilde{C}\int_0^t \tilde{\mathbb{E}}Z(s)\,ds, \qquad \forall t \geq 0$$

and therefore

$$\tilde{\mathbb{P}}(\tau > t) \leq \frac{\tilde{C}}{\alpha}\,(1 - e^{-\tilde{C}\,t})^{-1}\,|g(x)|, \qquad \forall t > 0,$$

which is just (2.19). This shows that $\tilde{X}(t)$ reaches the manifold Σ in stopping time τ and remains there for $t > \tau$ with a probability $\tilde{\mathbb{P}}$ greater or equal $\frac{\tilde{C}}{\alpha}\,(1 - e^{-\tilde{C}\,t})^{-1}\,|g(x)|$. The proof is complete. □

Remark 2.3. If conditions (2.16), (2.17) are satisfied with $\alpha = 0$ in Theorem 2.2 then only the invariance part (i) follows. We note also that if $f_i,\ i = 1, 2$ are monotone then so is f and so, as noted earlier in Remark 2.2, equation (1.3) has a unique strong solution X for which the conclusions of Theorem 2.2 hold.

Remark 2.4. As follows by the proof, assumption (2.18) were imposed by the Itô formula and can be avoided if take the stochastic differential equation (1.3) in the Stratonovich sense, that is if one replaces $\sigma(X)\,dW$ by $\sigma(X)\circ dW$.

Theorem 2.2 can be used to design feedback controllers for stochastic differential systems with a sliding mode dynamics on a given surface $\Sigma = \{x : g(x) = 0\}$. Such an example is presented below.

Example 2.5. Consider the controlled stochastic second order system

$$\ddot{X} + a\,\dot{X} = \sigma_0(X, \dot{X})\dot{\beta} + u \qquad \text{in } (0, \infty). \tag{2.25}$$

We assume that $\sigma_0 \in \text{Lip}(\mathbb{R}^2)$.

Our aim is to find a feedback controller $u = -f_0(X, \dot{X})$ such that the corresponding closed loop system

$$\begin{aligned} &\ddot{X} + a\,\dot{X} + f_0(X, \dot{X}) = \sigma_0(X, \dot{X})\dot{\beta} \\ &X(0) = x_0, \quad \dot{X}(0) = x_1 \end{aligned} \tag{2.26}$$

has the sliding mode equation

$$aX + \dot{X} = 0. \tag{2.27}$$

Here β is a Brownian motion in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\dot{\beta}$ is the associated white noise.

We choose

$$f_0(r_1, r_2) = \alpha \operatorname{sgn}(a\, r_1 + r_2), \qquad \forall (r_1, r_2) \in \mathbb{R}^2 \tag{2.28}$$

where $\alpha > 0$ and rewrite equation (2.26) as

$$\begin{aligned} & dX_1 - X_2\, dt = 0 \\ & dX_2 + a\, X_2\, dt + \alpha \operatorname{sgn}(a\, X_1 + X_2)\, dt = \sigma_0(X_1, X_2)\, d\beta \end{aligned} \tag{2.29}$$

for $t \geq 0$, where as usually $\operatorname{sgn}(u) = \frac{u}{|u|}$ for $u \neq 0$.
Equation (2.29) is a "jump" system of the form (1.3) where

$$f(r_1, r_2) = \begin{pmatrix} -r_2 \\ a\, r_2 + \alpha \operatorname{sgn}(a\, r_1 + r_2) \end{pmatrix}, \qquad \forall (r_1, r_2) \in \mathbb{R}^2,$$

$$\sigma(r_1, r_2) = \begin{pmatrix} 0 \\ \sigma_0(r_1, r_2) \end{pmatrix}, \qquad \forall (r_1, r_2) \in \mathbb{R}^2,$$

and so f is of the form (2.13) where

$$f_1(r) = \begin{pmatrix} -r_2 \\ a\, r_2 + \alpha \end{pmatrix}, \quad f_2(r) = \begin{pmatrix} -r_2 \\ a\, r_2 - \alpha \end{pmatrix}, \qquad r = (r_1, r_2) \in \mathbb{R}^2$$

$$g(r) = a\, r_1 + r_2, \qquad r = (r_1, r_2).$$

We assume that

$$\sigma_0^2(r_1, r_2) \leq C\, (a\, r_1 + r_2)^2, \qquad \forall (r_1, r_2) \in \mathbb{R}^2.$$

It is easily seen that conditions (2.14)–(2.18) hold and so Theorem 2.2 is applicable to the present case. We get

Corollary 2.3. *The stochastic closed loop system* (2.29)*, equivalently* (2.26)*,* (2.28)*, has the "sliding mode"* (2.27)*. More precisely, for every* $(x_0, x_1) \in \mathbb{R}^2$ *there is a martingale solution* $(X_1(t), X_2(t))$ *which reaches the surface* $\Sigma = \{(x_1, x_2) : a\, x_1 + x_2 = 0\}$ *in time* t *with a probability* $\geq 1 - (\alpha\, t)^{-1} |a\, x_0 + x_1|$*, and remains* $\tilde{\mathbb{P}}$*-a.s. on this surface after that time.*

This describes a typical "sliding-mode" behaviour for solutions X to (2.26), namely

$$a\, X(t) + \dot{X}(t) = 0$$

on $(t_0, \infty) \times \Omega_0$ where $\tilde{\mathbb{P}}(\Omega_0) \geq 1 - (\alpha\, t_0)^{-1} |a\, x_0 + x_1|$. We refer to [11], [13], [14], for references and other significant results on "sliding-mode" behaviour of stochastic differential systems.

3 Existence of a weak solution to heat equation (1.4)

The following hypotheses will be assumed throughout in the sequel.

(i) $h \in L^\infty_{\rm loc}(\mathbb{R})$ and $|h(r)| \le a_1|r| + b_1$, $\forall r \in \mathbb{R}$

(ii) W is the cylindrical Wiener process (1.2) where $\{e_j\}_{j=1}^\infty$ is an orthonormal basis in $L^2(\mathcal{O})$ given by $-\Delta e_j = \lambda_j e_j$ in $\mathcal{O}$; $e_j = 0$ on $\partial\mathcal{O}$ and

$$\sum_{j=1}^\infty \mu_j^2\, |e_j|_\infty^2 < \infty \tag{3.1}$$

(iii) $b \in C^2(\mathbb{R}) \cap \mathrm{Lip}(\mathbb{R})$, $b(0) = 0$.

It follows by (3.1) that in this case the operator B defined in section 1 is Hilbert-Schmidt.

Definition 3.1. *Let $x \in L^2(\mathcal{O})$. We call weak (martingale) solution to (1.4) a tuple $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P}, W, X)$, where $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ is a filtered probability space where there are defined a $(\mathcal{F}_t)_{t\ge0}$-Wiener process W and a continuous $(\mathcal{F}_t)_{t\ge0}$-adapted $L^2(\mathcal{O})$-valued process $X = (X(t))_{t\ge0}$ such that, $\mathbb{P}$-a.s.,*

$$X(t) = e^{-tA}x + \int_0^t e^{-(t-s)A}\eta(s)\,ds + \int_0^t e^{-(t-s)A}\, b(X(s))\,dW(s), \tag{3.2}$$

where $\eta \in L^2((0,T) \times \mathcal{O} \times \Omega)$ is an $L^2(\mathcal{O})$-valued $(\mathcal{F}_t)_{t\ge0}$-adapted process such that

$$\eta \in G(X), \textit{a.e in } (0,T) \times \mathcal{O} \times \Omega, \quad \forall T > 0. \tag{3.3}$$

Here $A = -\Delta$ with $D(A) = H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$, e^{-At} is the C_0-semigroup on $L^2(\mathcal{O})$ generated by $-A$ and $G\colon \mathbb{R} \to 2^{\mathbb{R}}$ is the Filippov map (1.6).

The construction of a weak (martingale) solution. We consider the approximating equation

$$\begin{array}{ll} dX_\varepsilon - \Delta X_\varepsilon\,dt + h_\varepsilon(X_\varepsilon)\,dt = b(X_\varepsilon)\,dW, & \text{in } (0,T)\times\mathcal{O} \\ X_\varepsilon = 0, & \text{on } (0,T)\times\partial\mathcal{O} \\ X_\varepsilon(0,\xi) = x(\xi), & \xi \in \mathcal{O} \end{array} \tag{3.4}$$

where $\varepsilon > 0$ and, as in the finite dimensional case (see (2.5)),

$$h_\varepsilon(r) = \frac{1}{\varepsilon}\int_{-\infty}^\infty h(s)\,\rho(\tfrac{r-s}{\varepsilon})\,ds = \int_{-\infty}^\infty h(r - \varepsilon\,\theta)\,\rho(\theta)\,d\theta, \quad \forall r \in \mathbb{R}. \tag{3.5}$$

Here $\rho \in C_0^\infty(\mathbb{R})$ is such that

$$\rho(\theta) \geq 0 \quad \rho(\theta) = \rho(-\theta), \quad \rho(\theta) = 0 \text{ for } |\theta| \geq 1, \quad \int_{-\infty}^{\infty} \rho(\theta)\, d\theta = 1. \tag{3.6}$$

Clearly by (i) we have

$$h_\varepsilon \in C^1(\mathbb{R}), \qquad |h_\varepsilon(r)| \leq a_1\,|r| + b_1 + a_1\,\varepsilon, \quad \forall r \in \mathbb{R}, \varepsilon > 0. \tag{3.7}$$

By Lemma 6.2 equation (3.4) has a unique strong solution

$$X_\varepsilon \in L^2(\Omega; C([0,T]; L^2(\mathcal{O}))) \cap L^2(\Omega; L^2(0,T; H_0^1(\mathcal{O}))). \tag{3.8}$$

By Itô's formula we get $\mathbb{P}$-a.s.

$$\begin{aligned}
&\tfrac{1}{2}\,|X_\varepsilon(t)|_2^2 + \int_0^t \|X_\varepsilon(s)\|_1^2\, ds + \int_0^t \langle h_\varepsilon(X_\varepsilon(s)), X_\varepsilon(s)\rangle_2\, ds \\
&\quad = \tfrac{1}{2}\,|x|_2^2 + \tfrac{1}{2}\int_0^t \sum_{j=1}^{\infty} \mu_j^2 |b(X_\varepsilon(s))\, e_j|_2^2\, ds + \int_0^t \langle b(X_\varepsilon(s)) dW(s), X_\varepsilon(s)\rangle_2, \\
&\hspace{30em} \forall t \in [0,T],
\end{aligned}$$

and so by the Burkholder-Davis-Gundy formula we obtain by some calculation involving (i)–(iii)

$$\mathbb{E} \sup_{t\in[0,T]} |X_\varepsilon(t)|_2^2 + \mathbb{E}\int_0^t \|X_\varepsilon(s)\|_1^2\, ds \leq C\,(|x|_2^2 + 1), \quad \forall \varepsilon > 0, \tag{3.9}$$

where C is independent of ε. By (3.7) we also have

$$\mathbb{E} \sup_{t\in[0,T]} |h_\varepsilon(X_\varepsilon(t))|_2^2 \leq C\,(|x|_2^2 + 1).$$

Then on a subsequence, again denoted in the same way, we have for $\varepsilon \to 0$

$$\begin{aligned} X_\varepsilon \to X \quad &\text{weak-star in } L^\infty(0,T; L^2(\Omega; L^2(\mathcal{O})) \\ &\text{weakly in } L^2(0,T; L^2(\Omega; H_0^1(\mathcal{O})) \end{aligned} \tag{3.10}$$

$$h_\varepsilon(X_\varepsilon) \to \eta \quad \text{weakly in } L^2((0,T); L^2(\Omega; L^2(\mathcal{O}))) \tag{3.11}$$

$$b(X_\varepsilon) \to b^* \quad \text{weakly in } L^2(\Omega; L^2((0,T) \times \mathcal{O})) \tag{3.12}$$

and

$$\begin{cases} dX - \Delta X\, dt + \eta\, dt = b^*\, dW & \text{in } (0,T) \times \mathcal{O}, \\ X(0) = x & \text{in } \mathcal{O}, \\ X = 0 & \text{on } (0,T) \times \partial\mathcal{O}, \end{cases} \tag{3.13}$$

that is

$$X(t) - \int_0^t \Delta X(s)\,ds + \int_0^t \eta(s)\,ds = x + \int_0^t b^*(s)\,dW(s), \quad \forall t \in [0,T],\ \mathbb{P}\text{-a.s.} \tag{3.14}$$

where Δ is taken in sense of distributions on $\mathcal{O}$, hence by (3.10)

$$\Delta X \in L^2(\Omega; L^2(0,T; H^{-1}(\mathcal{O}))).$$

Since the weak convergences (3.10)-(3.12) are not sufficient to conclude that (3.3) holds, then proceeding as in the proof of Theorem 2.1 we shall replace $\{X_\varepsilon\}$ by a sequence $\{\tilde{X}_\varepsilon\}$ of processes defined in a probability space $\{\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \tilde{W}\}$ such that $\mathcal{L}(X_\varepsilon) = \mathcal{L}(\tilde{X}_\varepsilon)$ where $\mathcal{L}$ is the law of the process.

To this end, consider the sequence $\{\nu_\varepsilon\}_{\varepsilon \geq 0}$ of probability measures, $\nu_\varepsilon = \mathcal{L}(X_\varepsilon)$, that is $\nu_\varepsilon(B) = \mathbb{P}(X_\varepsilon \in B)$ for any Borelian set $B \subset C([0,T]; L^2(\mathcal{O}))$. We have

Lemma 3.1. *Let $x \in H_0^1(\mathcal{O})$. Then the sequence $\{\nu_\varepsilon\}_{\varepsilon>0}$ is tight in the space $C([0,T]; L^2(\mathcal{O}))$.*

Proof. This means that for each $\delta > 0$ there is a compact subset B of $C([0,T]; L^2(\mathcal{O}))$ such that $\nu_\varepsilon(B^c) \leq \delta$ for all $\varepsilon > 0$. We take for $r > 0$, $\gamma > 0$,

$$\begin{aligned} B = B_{r,\gamma} = \{ & y \in C([0,T]; L^2(\mathcal{O})) : |y(t)|_2 \leq r, \forall t \in [0,T], \\ & \|y\|_{L^\infty(0,T;H_0^1(\mathcal{O}))} \leq r, |y(t) - y(s)|_2 \leq \gamma\,|t-s|^{\frac{1}{2}}, \quad \forall t,s \in [0,T]\} \end{aligned}$$

On the other hand, by (3.4) we have via Itô's formula applied to the process $t \to |X_\varepsilon(t) - X_\varepsilon(s)|_2^2$

$$\begin{aligned} \tfrac{1}{2}\,\mathbb{E}|X_\varepsilon(t) - X_\varepsilon(s)|_2^2 + \mathbb{E}\int_s^t \langle \nabla X_\varepsilon(\theta), \nabla(X_\varepsilon(\theta) - X_\varepsilon(s))\rangle_2\,d\theta \\ + \mathbb{E}\int_s^t \langle h_\varepsilon(X_\varepsilon(\theta)), X_\varepsilon(\theta) - X_\varepsilon(s)\rangle_2\,d\theta \\ \leq C\,\mathbb{E}\int_s^t |X_\varepsilon(\theta)|_2^2\,d\theta \quad 0 \leq s \leq t \leq T. \end{aligned}$$

Taking into account estimates (3.7), (3.9) we obtain via Gronwall's lemma that

$$\mathbb{E}|X_\varepsilon(t) - X_\varepsilon(s)|_2^2 \leq C\int_s^t (|X_\varepsilon(\theta)|_2^2 + |\nabla X_\varepsilon(\theta)|_2^2)\,d\theta \leq C\,|t-s|. \tag{3.15}$$

Clearly, by Ascoli-Arzelà theorem, $B_{r,\gamma}$ is compact in $C([0,T];L^2(\mathcal{O}))$.
By estimate (3.9), taking into account that

$$\rho\,\mathbb{P}[|Y|\ge\rho]\le\mathbb{E}|Y|,\quad\forall\rho>0,$$

we infer that there are γ, r independent of ε such that $\nu_\varepsilon(B^c_{r,\gamma})\le\delta$, as desired. □

Then by the Skorohod theorem (see, e.g., Theorem 2.4 in [6]) there are a probability space $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ and stochastic processes $\tilde X$, $\{\tilde X_\varepsilon\}_{\varepsilon>0}$ on $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ such that the law $\mathcal{L}(\tilde X_\varepsilon)$ of $\tilde X_\varepsilon$ coincides with $\mathcal{L}(X_\varepsilon)$ and $\tilde{\mathbb{P}}$-a.s.

$$\tilde X_\varepsilon\to\tilde X\quad\text{in } C([0,T];L^2(\mathcal{O}))\tag{3.16}$$

as $\varepsilon\to0$. We have also $\mathcal{L}(X)=\mathcal{L}(\tilde X)$. Since $\mathcal{L}(h_\varepsilon(X_\varepsilon))=\mathcal{L}(h_\varepsilon(\tilde X_\varepsilon))$, $\mathcal{L}(\sigma(X_\varepsilon))=\mathcal{L}(\sigma(\tilde X_\varepsilon))$ by (3.16) and (3.5) we see that

$$\begin{aligned}&h_\varepsilon(\tilde X_\varepsilon)\to\tilde\eta,\\ &b(\tilde X_\varepsilon)\to b^*(\tilde X),\end{aligned}\qquad\text{a.e. in } (0,T)\times\mathcal{O}\times\tilde\Omega,\tag{3.17}$$

where $\mathcal{L}(\tilde\eta)=\mathcal{L}(\eta)$ and

$$\tilde\eta\in G(\tilde X),\qquad\text{a.e. in } (0,T)\times\mathcal{O}\times\tilde\Omega.\tag{3.18}$$

The latter follows as in the proof of Theorem 2.1 taking into account that in this case G is given by (1.5), but we omit the details.
We set

$$M_\varepsilon(t)=X_\varepsilon(t)-x-\int_0^t\Delta X_\varepsilon(s)\,ds+\int_0^t h_\varepsilon(X_\varepsilon(s))\,ds,\quad t\in[0,T]\tag{3.19}$$

and

$$\tilde M_\varepsilon(t)=\tilde X_\varepsilon(t)-x-\int_0^t\Delta\tilde X_\varepsilon(s)\,ds+\int_0^t h_\varepsilon(\tilde X_\varepsilon(s))\,ds,\quad t\in[0,T].\tag{3.20}$$

It turns out that $\tilde M_\varepsilon$ is a square integrable martingale on $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ with respect to the filtration $\mathcal{F}_t=\sigma\{\tilde X_s;s\le t\}$. Actually, since $\mathcal{L}(\tilde M_\varepsilon)=\mathcal{L}(M_\varepsilon)$ and M_ε is a square integrable martingale on $(\Omega,\mathcal{F},\mathbb{P})$ we have

$$\mathbb{E}\big[(\tilde X_\varepsilon(t)-\tilde X_\varepsilon(s)-\int_s^t\Delta\tilde X_\varepsilon(\theta)\,d\theta+\int_s^t h_\varepsilon(\tilde X_\varepsilon(\theta))\,d\theta\big]=0.\tag{3.21}$$

Passing to the limit in (3.20) and taking into account (3.16)–(3.18) one obtains that the process

$$\tilde M(t)=\tilde X(t)-x-\int_0^t\Delta\tilde X(s)\,ds+\int_0^t\tilde\eta(s)\,ds,\quad t\ge0,$$

where $\Delta\tilde{X} \in L^2(\Omega; L^2(0,T; H^{-1}(\mathcal{O})))$, is an $L^2(\mathcal{O})$-valued martingale with respect to filtration $\tilde{\mathcal{F}}_t = \sigma\{\tilde{X}(s), s \le t\}$, $t \in [0,T]$, with finite quadratic variation, see [6, pag. 234]. Then by the representation theorem 8.2 in [6] there is a larger probability space $(\tilde{\tilde{\Omega}}, \tilde{\tilde{\mathcal{F}}}, \tilde{\tilde{\mathbb{P}}})$, a filtration $\{\tilde{\tilde{\mathcal{F}}}_t\}_{t\ge 0}$ and an $L^2(\mathcal{O})$-cylindrical Wiener process $\tilde{\tilde{W}}(t)$ on it such that, $\tilde{\tilde{\mathbb{P}}}$-a.s.,

$$\tilde{M}(t) = \int_0^t b^*(\tilde{X}(s))\, d\tilde{\tilde{W}}(t), \qquad t \in [0,T].$$

This means that the system $(\tilde{\tilde{\Omega}}, \tilde{\tilde{\mathcal{F}}}, \{\tilde{\tilde{\mathcal{F}}}_t\}_{t\ge 0}, \tilde{\tilde{\mathbb{P}}}, \tilde{\tilde{W}}(t), \tilde{X}(t))$ is a martingale solution to (1.1). In the following we shall denote again this probability basis by $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\ge 0}, \tilde{\mathbb{P}}, \tilde{W}, \tilde{X})$. We have proved therefore

Theorem 3.2. *Under Hypotheses* (i), (ii)*, for each $x \in H_0^1(\mathcal{O})$, there is at least one martingale solution $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\ge 0}, \tilde{\mathbb{P}}, \tilde{W}, \tilde{X})$ to equation* (1.1) *and $\tilde{X}$ is given by* (3.16)*. Moreover, we have*

$$\tilde{X} \in L^2(\tilde{\Omega}; L^\infty(0,T; L^2(\mathcal{O}))) \cap L^2(\tilde{\Omega}; L^2(0,T; H_0^1(\mathcal{O}))). \tag{3.22}$$

We note that (3.22) follows by (3.9) and (3.16).

Remark 3.1. Under additional assumptions on b (for instance if it is independent of X) it turns out that the martingale solution $\tilde{X}$ is the unique strong solution, see [7]. Compare also Remark 2.1.

4 Sliding mode control of the stochastic heat equation

For parabolic stochastic equations of the form (1.1) a "sliding" mode dynamic arises for discontinuous ("jump") functions $h\colon \mathbb{R} \to \mathbb{R}$ of the form (2.13), that is

$$h(r) = \begin{cases} f_1(r) & \text{for } g(r) \ge 0 \\ f_2(r) & \text{for } g(r) < 0 \end{cases}, \qquad r \in \mathbb{R} \tag{4.1}$$

where g, f_1, f_2 are given continuous functions.

As in the previous finite dimensional case, the objective of the "sliding-mode" control is to design for the linear time invariant system

$$\begin{array}{ll} dX - \Delta X\, dt = du & \text{in } (0,T) \times \mathcal{O} \\ X = 0 & \text{on } (0,T) \times \partial\mathcal{O} \\ X(0) = x & \text{in } \mathcal{O} \end{array} \tag{4.2}$$

a stochastic feedback controller of the form

$$du = -h(X)\,dt + b(X)\,dW \tag{4.3}$$

such that the "sliding" motion occurs on the manifold $\Sigma = \{X : g(X) = 0\}$ which is also referred as "sliding" or "switching" manifold. Roughly speaking, this means that there is a trajectory of the closed loop system (4.2)-(4.3), which starts from initial state x, reaches the sliding manifold Σ at a certain random time t_0 and remains there for $t \geq t_0$. As a matter of fact, this last phase of the dynamics is called "sliding mode". The sliding mode is due to both f_1 and f_2: if the solution enters the region of influence of f_1, then this function pushes the solution to cross the manifold; then, as it enters the region of influence of f_2, the solution is pushed back towards the manifold again.

Of course in virtue of Theorem 3.2 a weak solution X to (4.2) in the sense of Definition 3.1 exists for the extended multivalued closed loop system

$$\begin{array}{ll} dX - \Delta X\,dt + F(X)\,dt = b(X)\,dW & \text{in } (0,T)\times\mathcal{O} \\ X = 0 & \text{on } (0,T)\times\partial\mathcal{O} \\ X(0) = x & \text{in } \mathcal{O}. \end{array} \tag{4.4}$$

Here $F\colon \mathbb{R} \to 2^{\mathbb{R}}$ is the multivalued Heaviside function (2.20) on $\mathbb{R}$. To begin with we shall prove first an invariance result for the manifold $\Sigma = \{X : g(X) = 0\}$.

Theorem 4.1. *Let $g \in C^2(\mathbb{R})$, f_1, f_2 be continuous functions which satisfy assumption* (i) *and let b satisfies* (iii). *Assume further that*

$$b^2(r)(g\,g'' + (g')^2)(r) \leq C^*\,g^2(r), \qquad \forall r \in \mathbb{R} \tag{4.5}$$

$$g(r)\,g''(r) + (g'(r))^2 \geq 0 \qquad \forall r \in \mathbb{R} \tag{4.6}$$

$$f_1(r)\,g'(r) \geq 0 \qquad \textit{for } g(r) > 0 \tag{4.7}$$

$$f_2(r)\,g'(r) \leq 0 \qquad \textit{for } g(r) < 0. \tag{4.8}$$

for some $C^* > 0$. Then, for all $x \in H^1_0(\mathcal{O})$ such that $g(x) = 0$ on $\mathcal{O}$, there is a martingale solution $(\tilde\Omega, \tilde{\mathcal{F}}, \{\{\tilde{\mathcal{F}}_t\}_{t\geq 0}, \tilde{\mathbb{P}}, \tilde W, \tilde X)$ to system (4.4) *such that*

$$g(\tilde X(t)) = 0, \qquad \forall t \in [0,T],\ \tilde{\mathbb{P}}\text{-}a.s.. \tag{4.9}$$

Proof. We start with the approximating equation (3.4). We apply the Itô formula to function $x \to g^2(x)$ and get

$$\begin{aligned}
&\int_{\mathcal{O}} g^2(X_\varepsilon(t,\xi))\,d\xi \\
&\quad + 2\int_0^t \int_{\mathcal{O}} (g\,g'' + (g')^2)(X_\varepsilon(s,\xi))|\nabla X_\varepsilon(s,\xi)|^2\,d\xi\,ds \\
&\quad + 2\int_0^t \int_{\mathcal{O}} h_\varepsilon(X_\varepsilon(s,\xi))g(X_\varepsilon(s,\xi))g'(X_\varepsilon(s,\xi))\,d\xi\,ds \\
&= \int_{\mathcal{O}} g^2(x)\,d\xi \\
&\quad + \sum_{j=1}^\infty \mu_j^2 \int_0^t \int_{\mathcal{O}} |b(X_\varepsilon(s,\xi))\,e_j|^2 |(g\,g'' + (g')^2)(X_\varepsilon(s,\xi))|\,d\xi\,ds \\
&\quad + \sum_{j=1}^\infty \mu_j \int_0^t \int_{\mathcal{O}} b(X_\varepsilon(s,\xi))g(X_\varepsilon(s,\xi))g'(X_\varepsilon(s,\xi))\,e_j\,d\xi\,d\beta_j(s)
\end{aligned}$$

Taking into account (3.5),(3.7), we obtain that

$$\begin{aligned}
&\int_{\mathcal{O}} h_\varepsilon(X_\varepsilon)g(X_\varepsilon)g'(X_\varepsilon)\,d\xi = \int \rho(\theta) \\
&\quad \Big(\int_{[g(X_\varepsilon-\varepsilon\theta)>0]} f_1(X_\varepsilon - \varepsilon\theta)(gg')(X_\varepsilon - \varepsilon\theta)\,d\xi + \\
&\quad \int_{[g(X_\varepsilon-\varepsilon\theta)<0]} f_2(X_\varepsilon - \varepsilon\theta)(gg')(X_\varepsilon - \varepsilon\theta)\,d\xi \Big)\,d\theta + \zeta_\varepsilon(t), \qquad \forall t \in [0,T],
\end{aligned}$$

where

$$\zeta_\varepsilon(t) \le \tilde{C}\,\varepsilon \int_{\mathcal{O}} (|X_\varepsilon(t,\xi)| + 1)\,d\xi$$

with $\tilde{C} = C^*\,C$, $C > 0$; thus, by (4.5)–(4.7), this yields

$$\begin{aligned}
\mathbb{E}\int_{\mathcal{O}} g^2(X_\varepsilon(t,\xi))\,d\xi &\le C\,\mathbb{E}\int_0^t \int_{\mathcal{O}} g^2(X_\varepsilon(t,\xi))\,d\xi\,ds \\
&\quad + \delta(\varepsilon)\,\mathbb{E}\int_0^t \int_{\mathcal{O}} (|X_\varepsilon(t,\xi)|^2 + 1)\,d\xi\,ds, \quad \forall t \in [0,T]
\end{aligned}$$

where $\lim_{\varepsilon\to 0}\delta(\varepsilon) = 0$ and the constant C is independent of ε.
This yields, via Gronwall's lemma,

$$\mathbb{E}\int_{\mathcal{O}} g^2(X_\varepsilon(t,\xi))\,d\xi \le \delta(\varepsilon)\ \exp(C\,t), \qquad \forall t \in [0,T]. \tag{4.10}$$

If $\tilde{X}_\varepsilon$ is defined as in the proof of Theorem 3.1, that is $\mathcal{L}(\tilde{X}_\varepsilon) = \mathcal{L}(X_\varepsilon)$ and (3.16) holds, we get by (4.10) that

$$\tilde{\mathbb{E}} \int_{\mathcal{O}} g^2(\tilde{X}_\varepsilon(t,\xi))\, d\xi \le \delta(\varepsilon) \exp(C\,t), \qquad \forall t \in [0,T],\ \forall \varepsilon > 0$$

where $\tilde{\mathbb{E}}$ is the expectation in probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Hence, letting ε tend to zero we get $g^2(\tilde{X}) = 0$, $dt \times d\xi \times \tilde{\mathbb{P}}$-a.e. in $(0,T) \times \mathcal{O} \times \tilde{\Omega}$ as claimed. □

Remark 4.1. In the particular case where the function

$$F(r) \equiv f_1(r) H(g(r)) + f_2(r) H(-g(r)), \quad r \in \mathbb{R}$$

is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ with $R(F) = D(F) = \mathbb{R}$, equation (4.4) has a unique strong solution X. This happens for instance if f_i, $i = 1, 2$, are monotonically nondecreasing continuous functions such that $f_1 \ge f_2$ on $\mathbb{R}$ and $g(x) = x$ (see [2]). Then the corresponding system (4.4)

$$\begin{aligned} &dX - \Delta X\, dt + (f_1(X) H(X) + f_2(X) H(-X))\, dt = \\ &\qquad\qquad\qquad\qquad b(X)\, dW, \quad \text{in } (0,T) \times \mathcal{O} \\ &X = 0, \qquad\qquad\qquad\qquad \text{on } (0,T) \times \partial\mathcal{O}, \end{aligned} \tag{4.11}$$

has the invariant manifold $X = 0$.

Under stronger assumptions on g and b it turns out that the closed loop system (4.2)–(4.3) (equivalently (4.11)) has a "sliding" mode dynamics with the switching manifold $\Sigma = \{X : g(X) = 0\}$. Namely, we assume that

Hypothesis 4.2. *f_i, $i = 1, 2$ satisfy assumption* (i) *of page* 14 *and $g \in C^2(\mathbb{R})$, $b \in C^2(\mathbb{R}) \cap Lip(\mathbb{R})$ are such that*

$$((g'(r))^2 + |g''(r) g(r)|)\, |b(r)|^2 \le C^* \, |g(r)|^2, \qquad \forall r \in \mathbb{R}, \tag{4.12}$$

$$f_1(r)\, g'(r) \ge \alpha \quad \textit{if } g(r) > 0;\ f_2(r)\, g'(r) \le -\alpha \quad \textit{if } g(r) < 0 \tag{4.13}$$

$$g', g'' \in L^\infty(\mathbb{R}), \quad g'' \operatorname{sgn} g \ge 0 \quad \textit{on } \mathbb{R}, \tag{4.14}$$

where $\alpha > 0$, $C^* \ge 0$.

We note that by Theorem 3.2, equation (4.11) has a martingale solution $\tilde{X}$ given by (3.16).

Theorem 4.3. *Under Hypothesis 4.2, for each $x \in H_0^1(\mathcal{O})$ there is a martingale solution $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \tilde{W}, \tilde{X})$ to* (4.2), (4.3) *such that for $\tau = \inf\{t : |g(\tilde{X}(t))| = 0\}$ we have*

$$\tilde{\mathbb{P}}(\tau > t) \leq \frac{\tilde{C}}{\alpha}(1 - e^{-\tilde{C}\,t})^{-1}\,|g(x)|_2^2. \tag{4.15}$$

where $\tilde{C} = C\,C^*$ and C is independent of g, σ. Moreover, if $g(x) = 0$ a.e. in $\mathcal{O}$, then there is a martingale solution $\tilde{X}$ such that $g(\tilde{X}(t)) = 0$ for all $t \geq 0$.

Proof. The proof is very similar to that of Theorem 2.2, so it will be sketched only. If X_ε is the solution to equation (3.4) and $\tilde{X}_\varepsilon$ such that $\mathcal{L}(X_\varepsilon) = \mathcal{L}(\tilde{X}_\varepsilon)$, we get via Itô's formula applied to function $x \to \psi(g(x))$ where

$$\psi_\lambda(u) = (|u|_2^2 + \lambda^2)^{\frac{1}{2}}, \qquad \forall u \in L^2(\mathcal{O})$$

and

$$D\psi_\lambda(u) = (|u|_2^2 + \lambda^2)^{-\frac{1}{2}}u, \qquad \forall u \in L^2(\mathcal{O})$$

$$D^2\psi_\lambda(u)(v) = (|u|_2^2 + \lambda^2)^{-\frac{1}{2}}v - (|u|_2^2 + \lambda^2)^{-\frac{3}{2}}u\langle u, v\rangle_2$$

We get

$$\begin{aligned}
&d\psi_\lambda(g(\tilde{X}_\varepsilon(t)))+\\
&(|g(\tilde{X}_\varepsilon(t))|_2^2 + \lambda^2)^{-\frac{1}{2}}\langle g(\tilde{X}_\varepsilon(t))g'(\tilde{X}_\varepsilon(t)), -\Delta\tilde{X}_\varepsilon(t) + h_\varepsilon(\tilde{X}_\varepsilon(t))\rangle_2\,dt =\\
&\qquad\langle b(\tilde{X}_\varepsilon(t))\,dW(t), g(\tilde{X}_\varepsilon(t))g'(\tilde{X}_\varepsilon(t))\rangle_2(|g(\tilde{X}_\varepsilon(t))|_2^2 + \lambda^2)^{-\frac{1}{2}}+\\
&\qquad\tfrac{1}{2}\sum_{j=1}^{\infty}\mu_j^2\int_{\mathcal{O}}|D^2\psi_\lambda(g(\tilde{X}_\varepsilon))e_j|^2\,d\xi\,dt
\end{aligned}$$

Taking into account that by (4.14)

$$-\langle g(\tilde{X}_\varepsilon(t))g'(\tilde{X}_\varepsilon(t)), \Delta\tilde{X}_\varepsilon(t)\rangle \geq \int_{\mathcal{O}} g(\tilde{X}_\varepsilon(t))g''(\tilde{X}_\varepsilon(t))|\nabla\tilde{X}_\varepsilon(t)|^2\,d\xi$$

and letting $\lambda \to 0$ we obtain that

$$\begin{aligned}
&d\,|g(\tilde{X}_\varepsilon(t))|_2 + \alpha\mathbb{1}_{[|g(\tilde{X}_\varepsilon(t))|>0]}dt \leq\\
&\qquad\langle b(\tilde{X}_\varepsilon(t))dW(t), g'(\tilde{X}_\varepsilon(t))\,\mathrm{sgn}(g(\tilde{X}_\varepsilon(t))\rangle_2\,dt+\\
&\tfrac{1}{2}\sum_{k=1}^{\infty}\mu_j^2\int_O|b(\tilde{X}_\varepsilon(t))\,e_j|^2((g'(\tilde{X}_\varepsilon(t)))^2 + |g''(\tilde{X}_\varepsilon(t))g(\tilde{X}_\varepsilon(t))|)\,d\xi\,dt
\end{aligned} \tag{4.16}$$

We have used here assumption (4.13) which, as we have seen in the proof of Theorem 2.2, implies that

$$f_\varepsilon(\tilde X_\varepsilon)\, g'(\tilde X_\varepsilon)\,\mathrm{sgn}\, g(\tilde X_\varepsilon) \ge \alpha - \delta(\varepsilon)(1+|\tilde X_\varepsilon|).$$

Now using (4.12) and letting $\varepsilon \to 0$ we get

$$|g(\tilde X(t))|_2 + \alpha \int_s^t \mathbb{1}_{[|g(\tilde X(\theta))|>0]} d\theta \le |g(\tilde X(s))|_2 + \tilde C \int_s^t |g(\tilde X(\theta))|_2\, d\theta + \int_s^t \langle b(\tilde X(\theta)) dW(\theta), g'(\tilde X(\theta))\,\mathrm{sgn}(g(\tilde X(\theta))\rangle_2 \qquad \text{for } 0 \le s \le t$$

for some constant $\tilde C > 0$.
This yields (see (2.24))

$$e^{-\tilde C t}|g(\tilde X(t))|_2 + \alpha \int_s^t e^{-\tilde C\theta} \mathbb{1}_{[|g(\tilde X(\theta))|>0]}\, d\theta \le e^{-\tilde C s}|g(\tilde X(s))|_2 + \int_s^t e^{-\tilde C\theta}\mathbb{1}_{[|g(\tilde X(\theta))|>0]}\, \langle g'(\tilde X(\theta))\mathrm{sgn}(g(\tilde X(\theta))), \sigma(\tilde X(\theta))\, dW(\theta)\rangle_2, \quad 0 \le s \le t < \infty.$$

Here $\tilde Z = |g(\tilde X(t))|_2\, e^{-\tilde C t}$ is a nonnegative supermartingale and so $\tilde Z(t) = \tilde Z(\tau)$ $\tilde{\mathbb{P}}$-a.s. for $t > \tau = \inf\{t > 0 : |\tilde Z(t)| = 0\}$. Taking expectation, we get

$$\tilde{\mathbb{E}}\tilde Z(t) + \alpha \int_0^t e^{-\tilde C s}\tilde{\mathbb{P}}[\tau > s]\, ds \le |g(x)|_2 + \tilde C \int_0^t \tilde{\mathbb{E}}\tilde Z(s)\, ds$$

which implies the desired estimate (4.15). □

Remark 4.2. By (4.14) we see that g is convex on $[g > 0]$, concave on $[g < 0]$ and so $[r : g(r) = 0] = [\alpha_1, \alpha_2]$ is a closed interval.
However, by (4.13) we see that $\alpha_1 = \alpha_2 = g^{-1}(0)$ and so the switching manifold $\Sigma = \{X : g(X) = 0\}$ reduces to the point $g^{-1}(0)$. Hence under the assumptions (4.12)–(4.14) the closed loop system (4.2)–(4.3) has for each $x \in H_0^1(\mathcal{O})$ a martingale solution which reaches the state $X = g^{-1}(0)$ in time t with probability estimated by (4.15) and remains there after that time.

5 "Sliding" mode control of a stochastic parabolic systems

Consider here the parabolic system

$$\begin{array}{ll} dX - \Delta X\,dt + f_1(X,Y)\,dt = b_1(X,Y)\,dW_1, & \text{in } (0,T)\times\mathcal{O}\cap\{g>0\} \\ dY - \Delta Y\,dt + f_2(X,Y)\,dt = b_2(X,Y)\,dW_2, & \text{in } (0,T)\times\mathcal{O}\cap\{g<0\} \\ X(0)=x(\xi), \quad Y(0)=y(\xi), & \xi\in\mathcal{O} \\ X=Y=0, & \text{on } (0,T)\times\partial\mathcal{O} \end{array} \tag{5.1}$$

where $f_i \in C(\mathbb{R}^2)$ satisfy assumption (i), $b_i \in C^2(\mathbb{R}) \cap \mathrm{Lip}(\mathbb{R})$, $i = 1,2$ and W_1, W_2 are Wiener processes of the form (1.2) in the space $H = L^2(\mathcal{O}) \times L^2(\mathcal{O})$. Let $g\colon \mathbb{R}^2 \to \mathbb{R}$, $g \in C^2(\mathbb{R}^2)$ be given.

System (5.1) is of the form (1.1) where $H = L^2(\mathcal{O}) \times L^2(\mathcal{O})$, $W = \binom{W_1}{W_2}$

$$A = \begin{pmatrix} -\Delta & 0 \\ 0 & -\Delta \end{pmatrix}, \quad D(A) = \big(H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})\big)^2$$

$$B = \begin{pmatrix} b_1(X,Y) & 0 \\ 0 & b_2(X,Y) \end{pmatrix}$$

$$f(X,Y) = \begin{cases} f_1(X,Y) & \text{if } g(X,Y) \ge 0, \\ f_2(X,Y) & \text{if } g(X,Y) < 0. \end{cases}$$

The corresponding Filippov map is

$$G(X,Y) = \begin{cases} f_1(X,Y) & \text{if } g(X,Y) \ge 0, \\ f_2(X,Y) & \text{if } g(X,Y) < 0, \\ [f_1(X,Y), f_2(X,Y)] & \text{if } g(X,Y) = 0. \end{cases}$$

Arguing as in the proof of Theorem 3.2 it follows that for each $(x,y) \in H_0^1(\mathcal{O}) \times H_0^1(\mathcal{O})$, system (5.1) has a martingale solution $(\tilde{X},\tilde{Y})$ obtained as limit of solutions $(\tilde{X}_\varepsilon,\tilde{Y}_\varepsilon)$ to corresponding approximating system

$$\begin{aligned} & d\begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} -\Delta & 0 \\ 0 & -\Delta \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} dt + G_\varepsilon\begin{pmatrix} X \\ Y \end{pmatrix} dt = \begin{pmatrix} b_1(X,Y)\,dW_1 \\ b_2(X,Y)\,dW_2 \end{pmatrix} \\ & \begin{pmatrix} X \\ Y \end{pmatrix}(0) = \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned} \tag{5.2}$$

where

$$G_\varepsilon(r_1,r_2) = \int_{\mathbb{R}^2} \rho(r-\varepsilon\theta)\, f(\theta)\, d\theta, \quad (r_1,r_2)\in\mathbb{R}^2, \ \ \varepsilon>0.$$

(Here ρ is a mollifier function in $\mathbb{R}^2$.)
Assume further that

$$\big(|\nabla g(r)|^2+|D^2 g(r)||g(r)|\big)\big(|b_1(r)|^2+|b_2(r)|^2\big)\le C^*\,|g(r)|^2,\quad \forall r\in\mathbb{R}^2 \tag{5.3}$$

$$f_1(r)\,g_{r_1}(r)\ge\alpha\quad\text{in } [r:g(r)>0] \tag{5.4}$$

$$f_2(r)\,g_{r_2}(r)\le-\alpha\quad\text{in } [r:g(r)<0] \tag{5.5}$$

where $\alpha>0$ and $(g_{r_1},g_{r_2})=\nabla g$,

$$g_{r_1r_1}\operatorname{sgn} g\ge 0,\quad (g^2_{r_1r_2}-g_{r_1r_1}g_{r_2r_2})\operatorname{sgn} g\le 0. \tag{5.6}$$

We have

Theorem 5.1. *Under assumptions* (5.3)–(5.6) *for each* $(x,y)\in H^1_0(\mathcal{O})\times H^1_0(\mathcal{O})$ *there is a martingale solution* $\{\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{P}},\tilde W,(\tilde X,\tilde Y)\}$ *to* (5.1) *such that if* τ *is the stopping time* $\tau=\inf\{t: g(\tilde X(t),\tilde Y(t))=0\}$ *then*

$$\tilde{\mathbb{P}}[\tau>t]\le\frac{\tilde C}{\alpha}(1-e^{-\tilde C t})^{-1}|g(x,y)|_{(L^2(\mathcal{O}))^2} \tag{5.7}$$

for some constant $\tilde C>0$.

The proof is exactly the same as that of Theorem 4.3 where the approximating equation (3.4) is replaced by (5.2). We note that in this case the corresponding inequality (4.16) is a consequence of hypothesis (5.6). The details are omitted.

A particular example is

$$g(r_1,r_2)=\alpha_1\,r_1+\alpha_2\,r_2,\qquad \forall r_1,r_2\in\mathbb{R}$$

which, for $f_i,\ b_i,\ i=1,2$ satisfying the condition

$$\begin{aligned}&\alpha_1\,f_1(r)\ge\alpha\qquad \text{in } \{\alpha_1\,r_1+\alpha_2\,r_2>0\}\\&\alpha_2\,f_2(r)\le-\alpha\quad\ \text{in } \{\alpha_1\,r_1+\alpha_2\,r_2<0\}\\&|b_1(r)|^2+|b_2(r)|^2\le C^*\,(\alpha_1\,r_1+\alpha_2\,r_2)^2\end{aligned}$$

where $\alpha>0$, imply that system (5.1) has a martingale solution $(\tilde X,\tilde Y)$ that reaches the linear manifold

$$\Sigma=\{\alpha_1\,\tilde X+\alpha_2\,\tilde Y=0\}$$

in a time t with probability $\tilde{\tilde{\mathbb{P}}}\ge 1-C\,t^{-1}|\alpha_1\,x+\alpha_2\,y|_{(L^2(\mathcal{O}))^2}$ and remains on this manifold afterwards.

As in Example 2.5 we may view the feedback controller $u=G(X,Y)$ as a sliding mode controller which forces the trajectory of system (5.1) to move on a lower order manifold $\{(X,Y):g(X,Y)=0\}$. Of course the above treatment is applicable to a n-dimensional stochastic system of the form (5.1).

6 Appendix

The results given below are without any doubt known in literature but we mention them for reader's convenience.

Lemma 6.1. *Let W be a n-dimensional Wiener process in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\sigma \in \mathrm{Lip}(\mathbb{R}^n; L(\mathbb{R}^n; \mathbb{R}^n))$ and $\ell\colon \mathbb{R}^n \to \mathbb{R}^n$ be a local Lipschitz function such that*

$$|\ell(r)| \le C\,(|r| + 1), \qquad \forall r \in \mathbb{R}^n. \tag{6.1}$$

Then for each $x \in \mathbb{R}^n$ the equation

$$\begin{aligned} &dX + \ell(X)\,dt = \sigma(X)\,dW, \qquad t \in (0,T)\\ &X(0) = x \end{aligned} \tag{6.2}$$

has a unique strong solution X.

Proof. We set

$$\ell_N(r) = \ell(r) \quad \text{for } |r| \le N, \quad \ell_N(r) = \ell(\tfrac{N\,r}{|r|}) \quad \text{for } |r| > N.$$

Since ℓ_N is Lipschitz, the equation

$$\begin{aligned} &dX_N + \ell_N(X_N)\,dt = \sigma(X_N)\,dW, \qquad t \in (0,T),\\ &X_N(0) = x \end{aligned} \tag{6.3}$$

has a unique strong solution X_N. By Itô's formula we have

$$\tfrac{1}{2} d|X_N(t)|^2 + \ell_N(X_n(t)) \cdot X_N(t)\,dt = \tfrac{1}{2}\mathrm{Tr}\,[\sigma^*\sigma]\,dt + \sigma(X_N)\,dW \cdot X_N$$

and taking into account (6.1) we get via Burkholder-Davis-Gundy theorem

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_N(t)|^2\Big] \le C(1 + |x|^2) \tag{6.4}$$

where C is independent of N. Consider the stopping time

$$\tau_j = \inf\{t \in (0,T) : |X_N(t)|^2 \ge j^2\}, \quad j \in \mathbb{N}. \tag{6.5}$$

We have

$$d(X_N - X_M) + (\ell_N(X_N) - \ell_M(X_M))\,dt = (\sigma(X_N) - \sigma(X_M))\,dW$$

and this yields

$$\frac{1}{2}|X_N(t)-X_M(t)|^2+ \int_0^t (\ell_N(X_N(s))-\ell_M(X_M(s)))\cdot(X_N(s)-X_M(s))\,ds = \frac{1}{2}\int_0^t \mathrm{Tr}[(\sigma(X_N(s))-\sigma(X_M(s)))(\sigma^*(X_N(s))-\sigma^*(X_M(s)))]\,ds + \int_0^t (\sigma(X_N(s))-\sigma(X_M(s)))\,dW(s)\cdot(X_N(s)-X_M(s)) \tag{6.6}$$

Taking into account that for $M,\, N \geq j$

$$\int_0^{t\wedge\tau_j} (\ell_N(X_N)-\ell_M(X_M))\cdot(X_N-X_M)\,ds \leq C\,j \int_0^{t\wedge\tau_j} |X_N(s)-X_M(s)|^2\,ds$$

we find by (6.6) via BDG inequality that

$$X_N(s)-X_M(s)=0, \qquad \forall s\in(0,\tau_j).$$

We set

$$X(t)=X_N(t), \quad \forall t\in(0,\tau_j),\ \forall N\geq j.$$

We have

$$X(t\wedge\tau_j)+\int_0^{t\wedge\tau_j}\ell(X(s))\,ds = x+\int_0^{t\wedge\tau_j}\sigma(X(s))\cdot dW(s)$$

and since by (6.4) and (6.5) $\lim_{j\to\infty}\tau_j = T$, $\mathbb{P}$-a.s. we infer that X is a solution to (6.2).
The uniqueness of solution X follows in a similar way taking into account that suach a solution satisfies, by Itô's formula and BDG theorem, the estimate (6.4) and that ℓ is locally Lipschitz. □

Lemma 6.2. *Let W be the Wiener process* (1.2), $\ell\colon \mathbb{R}\to\mathbb{R}$ *be locally Lipschitz,* $b\in \mathrm{Lip}(\mathbb{R})$ *and* (6.1) *holds. Then for each* $x\in L^2(\mathcal{O})$ *the equation*

$$\begin{array}{ll} dX-\Delta X\,dt+\ell(X)\,dt=b(X)\,dW, & \text{in } (0,T)\times\mathcal{O} \\ X(0)=x \quad \text{in } \mathcal{O} & \\ X=0 \quad \text{on } (0,T)\times\partial\mathcal{O} & \end{array} \tag{6.7}$$

has a unique strong solution

$$X\in L^2(\Omega;C([0,T];L^2(\mathcal{O})))\cap L^2(\Omega;L^2(0,T;H_0^1(\mathcal{O}))) \tag{6.8}$$

Proof. If ℓ is Lipschitz the result is well known (se e.g., [6]). Here we replace in (6.7) ℓ by ℓ_N, previously defined at pag. 25. Then the corresponding equation (6.7) has a unique strong solution

$$X_N \in L^2(\Omega; C([0,T]; L^2(\mathcal{O}))) \cap L^2(\Omega; L^2(0,T; H_0^1(\mathcal{O}))).$$

By Itô's formula combined with the BDG theorem we get

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_N(t)|_2^2\Big] + \mathbb{E}\int_0^t \|X_N(s)\|_1^2\, ds \le C\,(1+|x|_2^2)$$

where C is independent of N.

We define as in previous case the stopping time

$$\tau_j = \inf\{t\in[0,T] : |X_N(t)|_2^2 \ge j^2\}, \qquad \forall j\in\mathbb{N},$$

and obtain as above that for $M,\ N\ge j$

$$X_N(t) = X_M(t) \qquad \mathbb{P}\text{-a.s. } t\in(0,\tau_j).$$

Then $X\colon [0,T]\to L^2(\mathcal{O})$ defined by

$$X(t) = X_N(t) \qquad \forall t\in(0,\tau_j), \quad \forall N\ge j,$$

is a solution to (6.7). It follows also that (6.8) holds. □

Lemma 6.3. *Let $\varphi,\ \alpha : [0,T]\to\mathbb{R}$, $\psi : [0,T]\to\mathbb{R}^n$ be $(\mathcal{F}_t)_{t\ge0}$-adapted processes with continuous paths such that*

$$\varphi(t) + \int_s^t \alpha(\theta)\,d\theta \le A + C\int_s^t \varphi(\theta)\,d\theta + \int_s^t \psi(\theta)\cdot dW_\theta,$$

where A is a $\mathcal{F}_s$-random variable. Then

$$\varphi(t) + \int_s^t e^{C(t-\theta)}\alpha(\theta)\,d\theta \le e^{C(t-s)}A + \int_s^t e^{C(t-\theta)}\psi(\theta)\cdot dW_\theta$$

Proof. Define

$$\gamma(t) = -\int_s^t \alpha(\theta)\,d\theta + A + C\int_s^t \varphi(\theta)\,d\theta.$$

Then we have

$$\varphi(t) \le \gamma(t) + \int_s^t \psi(\theta)\cdot dW_\theta$$

$$\gamma'(t) = -\alpha(t) + C\varphi(t), \qquad \gamma(s) = A$$

$$\gamma'(t) \leq -\alpha(t) + C\,\gamma(t) + C\int_s^t \psi(\theta)\cdot dW_\theta$$

$$\left(e^{-Ct}\gamma(t)\right)' \leq -e^{-Ct}\alpha(t) + Ce^{-Ct}\int_s^t \psi(\theta)\cdot dW_\theta$$

$$e^{-Ct}\gamma(t) - e^{-C\,s}A \leq -\int_s^t e^{-Cu}\alpha(u)\,du + \int_s^t Ce^{-Cu}\,du\int_s^u \psi(\theta)\cdot dW_\theta$$

Now, by integration by parts, we get

$$\int_s^t Ce^{-Cu}\,du\int_s^u \psi(\theta)\cdot dW_\theta = -e^{-Ct}\int_s^t \psi(\theta)\cdot dW_\theta + \int_s^t e^{-C\theta}\psi(\theta)\cdot dW_\theta$$

Hence

$$e^{-Ct}\gamma(t) + \int_s^t e^{-Cu}\alpha(u)\,du \leq e^{-C\,s}A + \int_s^t e^{-C\theta}\psi(\theta)\cdot dW_\theta - e^{-Ct}\int_s^t \psi(\theta)\cdot dW_\theta$$

which implies the desired inequality. □

Acknowledgement This work was done while V. Barbu was visiting the Mathematics Department of University of Trento and was partially supported by a grant of CNCS UEFCSCDI (Romania) Project PN-II-ID-PCE-2012-4-0456.

References

[1] V. Barbu, Nonlinear Differential Equations of Monotone Type in Banach Spaces, Springer (2010)

[2] V. Barbu, A variational approach to stochastic nonlinear parabolic problems, J. Math Anal. Appl. 384, 2-15 (2011)

[3] V. Barbu, Optimal control approach to nonlinear diffusion equations driven by Wiener noise, J. Optim. Theory Appl. 153, 1-26 (2012)

[4] R. Buckdahn, L. Maticiuc, E. Pardoux, A. Rascanu, Stochastic variational inequalities in non convex domains, Journal of Differential Equations vol. 259, no. 12, 7332-7374 (2015)

[5] G. Da Prato, H. Frankowska, A stochastic Filippov theorem, Stochastic Anal. Appl. 12, 409-426 (1994)

[6] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions (second edition), Cambridge (2014)

[7] G. Da Prato, F. Flandoli, E. Priola, M. Röckner, Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift, Annals of Probability 41 (5) 3306–3344 (2013)

[8] A. F. Filippov, Differential Equations with Discontinuous Right-hand Sides, Mat. Sbornik (N.S.) 51 (93), 99-128 (1960), Amer. Math. Soc. Transl. 42, 199-231 (1964)

[9] A. F. Filippov, Differential Equations with Discontinuous Righthand Sides, Kluwer Academic Publishers, Norwel, (1988)

[10] I. Gyöngy, N. Krylov, Existence of strong solutions for Itô's stochastic equations via approximations, Probab. Theory Related Fields, 105 n. 2, 143-158 (1996)

[11] K-C. Hsu, Sliding mode controllers for uncertain systems with input nonlinearities, Journal of Guidance, Control, and Dynamics, July, Vol. 21, No. 4 : pp. 666-669 (doi: 10.2514/3.22105)

[12] A. Rascanu, E. Rotenstein, A non-convex setup for multivalued differential equations driven by oblique subgradients, Nonlinear Anal. 111, 82-104 (2014)

[13] P. Shi, Y. Xia, G. P. Liu, D. Rees, On designing of sliding-mode control for stochastic jump systems, IEEE Transactions on Automatic Control, vol. 51, no. 1, pp. 97-103, 2006

[14] D.J.W. Simpson, R. Kuske, Stochastically perturbed sliding motion in piecewise-smooth systems, Discrete Contin. Dyn. Syst. Ser. B, 19 (9), 2889-2913 (2014)

[15] V. Utkin, A control engineer's guide to sliding mode control, IEEE Transactions on Control Systems Technology, vol. 47, n. 3, 328-342 (1999)

[16] A. Yu. Veretennikov, On strong solution and explicit formulas for solutions of stochastic integral equations, Math. USSR Sbornik, 39, 387-403 (1981)

[17] K. Yosida, Functional Analysis, Springer-Verlag, Berlin-Heidelberg-New York (1980)