Stochastic Modelling and Computational Sciences
Vol. 3 No. 2 (December, 2023)

PROBABILISTIC INTERPRETATIONS OF NONLINEAR SECOND ORDER PARABOLIC EQUATIONS AND SYSTEMS AND NUMERICAL ALGORITHMS BASED ON THEM

YANA BELOPOLSKAYA

Abstract. We construct probabilistic representations of classical and /or viscosity solutions of the Cauchy problem for quasilinear and fully nonlinear parabolic equations and systems. In addition we develop numerical schemes to construct approximate solutions of the considered problems based on their probabilistic representations.

1. Introduction

In this paper we discuss a number of probabilistic interpretations for systems of nonlinear second order parabolic equations. Namely, we construct probabilistic representations of classical and viscosity solutions of the backward Cauchy problem for nonlinear parabolic equations and systems and develop numerical algorithms based on them.

The existence of connections between solutions of the Cauchy problem for linear parabolic equations and stochastic processes was first revealed by A.N. Kolmogorov [1]. The correspondent PDEs associated with Markov stochastic processes are called now the forward and backward Kolmogorov equations.

Connections between solutions of the Cauchy problem for semilinear parabolic equations and solutions of stochastic differential equations (SDEs) were revealed in pioneer papers by H. McKean [2] and M. Freidlin [3],[4]. Actually, H.McKean constructed a Markov process associated with the forward nonlinear Kolmogorov equation while M.Freidlin constructed a Markov process associated with the backward nonlinear Kolmogorov equation.

Let us give a bit more precise description of these results.

Consider the backward Cauchy problem

$$u_s + \frac{1}{2} TrA(s,x,u)\nabla^2 uA^*(s,x,u) + \langle a(s,x,u), \nabla u\rangle + c(s,x,u)u = 0, \qquad (1.1)$$

$$u(T,x) = h(x), \quad x \in R^d, s \in [0,T].$$

Date: Date of Submission September 01, 2023; Date of Acceptance September 25, 2023; Communicated by Igor V. Pavlov.

2000 *Mathematics Subject Classification.* Primary 60H30, 60H35; Secondary 35K55.

Key words and phrases. Nonlinear parabolic equations and systems, the Cauchy problem, probabilistic representation, stochastic differential equations, numerical solutions.

* This research is supported by the RSF Grant 22-21-00016.

Here and below we use notations $\langle a, b\rangle = \sum_{j=1}^d a_j b_j$ for an inner product of $a, b \in R^d$, ∇u for a gradient of u and $TrA\nabla^2 uA^* = \sum_{i,j,k=1}^d A_{ik}\frac{\partial^2 u}{\partial x_i \partial x_j}A_{kj}$.

Denote by $(\Omega, \mathcal{F}, P)$ a probability space and let $w(t) \in R^d$ be the standard Wiener process defined on it. A Markov process associated with the backward Cauchy problem (1.1) was constructed by Freidlin [3] as a solution of a stochastic system

$$d\xi(t) = a(t, \xi(t), u(t, \xi(t)))dt + A(t, \xi(t), u(t, \xi(t)))dw(t), \quad \xi(s) = x, \tag{1.2}$$

$$u(s, x) = E\left[exp\left\{\int_s^T c(\tau, \xi_{s,x}(\tau), u(\tau, \xi_{s,x}(\tau)))d\tau\right\} h(\xi_{s,x}(T))\right]. \tag{1.3}$$

To be more precise it was proved that if coefficients a, A, c and the Cauchy data $h(x)$ in (1.1) are smooth enough and bounded then there exists a unique solution of the system (1.2), (1.3) and the function $u(s, x)$ given by (1.3) (provided it is twice differentiable) is a unique classical solution of the problem (1.1).

A Markov process associated with the forward Cauchy problem

$$\frac{\partial \mu}{\partial t} + \sum_{i=1}^d \frac{\partial(a_i[y, \mu]\mu)}{\partial y_i} = \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2(M_{ik}[y, \mu]\mu)}{\partial y_i \partial y_j}, \tag{1.4}$$

$$\mu(0, dy) = \mu_0(dy),$$

was constructed in [2] as a solution of a stochastic equation

$$d\xi(\theta) = a[\xi(\theta), \mu(\theta)]d\theta + A[\xi(\theta), \mu(\theta)]dw(\theta), \; \xi(s) = \xi_0 \tag{1.5}$$

where

$$\mu(\theta, dy) = P\{\xi(\theta) \in dy\}, \quad \mu(0, dy) = \mu_0(dy). \tag{1.6}$$

Here $M_{ij}[y, \mu] = \sum_{k=1}^d A_{ik}[y, \mu]A_{kj}[y, \mu]$, $A[y, \mu] = \int_{R^d} A(y-x)\mu(dx)$ and $\xi_0 \in R^d$ is a random variable independent of $w(t)$ such that $P\{\xi_0 \in dy\} = \mu_0(dy)$. It was shown that under some conditions there exists a unique solution of the system (1.5), (1.6) and besides the measure $\mu(t, dy)$ given by (1.6) is a weak solution to (1.4). If there exists a density $\mu(t, y)$ of the measure $\mu(t, dy)$ with respect to the Lebesgue measure then $\mu(t, y)$ also solves (1.4) in the weak sense with initial data $\mu(0, y) = \mu_0(y)$, where $\mu_0(dy) = \mu_0(y)dy$.

It should be mentioned that Freidlin in [3] also started with the forward Cauchy problem

$$v_s = \frac{1}{2}TrA(x, v)\nabla^2 vA^*(x, v) + \langle a(x, v), \nabla v\rangle + c(x, v)v, \quad v(0, x) = u_0(x), \tag{1.7}$$

but immediately reduced (1.7) to the backward Cauchy problem

$$u_s + \frac{1}{2}TrA(x, u)\nabla^2 uA^*(x, u) + \langle a(x, u), \nabla u\rangle + c(x, u)u = 0, \quad u(T, x) = u_0(x), \tag{1.8}$$

with respect to a function $u(T - s, x) = v(s, x)$ and then proceeded as it was mentioned above.

Thus, in the framework of a stochastic interpretation one can consider the equation (1.1) as a backward Kolmogorov equation for the Markov process $\xi(t)$

satisfying (1.2), (1.3) while the equation (1.4) can be considered as the forward Kolmogorov equation for the Markov process $\xi(t)$ satisfying (1.5), (1.6).

We say that a stochastic model associated with the Cauchy problem for a nonlinear parabolic equation or system is constructed if we have an independent description of a Markov process associated with this Cauchy problem.

As a rule the required Markov process is constructed as a solution of a certain stochastic differential equation with coefficients depending on the unknown solution of the original Cauchy problem. To obtain a closed independent of the PDE description of the required process one needs to derive an additional relation that serves as a probabilistic representation of a solution to the Cauchy problem.

If we consider a linear parabolic equation of the form (1.1) with coefficients $a(x,u) = a(x), A(x,u) = A(x)$ independent of u then (1.1) comes to be a classical backward Kolmogorov equation and the corresponding forward Kolmogorov equation has the form

$$\frac{\partial \mu}{\partial s} + \langle \nabla, [a(y)\mu] \rangle = \frac{1}{2} Tr \nabla^2 [M(y)\mu], \quad \mu(0, dy) = \mu_0(dy). \tag{1.9}$$

One can easily notice that linear versions of (1.1) and (1.9) are dual with respect to the duality

$$\langle\langle f, \mu \rangle\rangle = \int_{R^d} f(y)\mu(dy)$$

between the space of bounded Borel measures $\mu(dy)$ and bounded Borel functions $f(y)$. Obviously, (1.4) coincides with (1.9) provided coefficients in (1.4) do not depend on μ. Unfortunately this duality is ruined in a nonlinear case but it still helps to understand the nature of a solution to (1.4) and to find the generator of the required Markov process especially if one deals with systems with cross-diffusion [5].

We say that an equation

$$u_s + \frac{1}{2} Tr A^u(s,x) \nabla^2 u [A^u]^*(s,x) + \langle a^u(s,x), \nabla u \rangle + c^u(s,x)u = 0, \tag{1.10}$$

is semilinear if its coefficients a^u, A^u, c^u have the form $A^u(s,x) = A(s,x,u(s,x))$, quasilinear if $A^u(s,x) = A(s,x,u(s,x),\nabla u(s,x))$ and fully nonlinear if $A^u(s,x) = A(s,x,u(s,x),\nabla u(s,x),\nabla^2 u(s,x))$.

Probabilistic interpretation of semilinear parabolic equation of the form (1.1) was extended to the case of quasilinear and fully nonlinear equations by Yu.Dalecky and Ya. Beloposkaya [6], [7]. To extend the approach to deal with quasilinear or even with fully nonlinear PDEs within the framework of this approach one needs to consider (1.1) as the first equation in a system of PDEs called a differential prolongation of the original system [8].

An alternative probabilistic approach to backward quasilinear parabolic equations was suggested by Pardoux and Peng [9], [10]. This approach was extended by a number of authors (see [11], [12] and references there). Note that the BSDE approach allows to construct a so called viscosity solution [13] of the Cauchy problem for a quasilinear PDE in the case when a classical solution does not exists. Later this approach was extended to deal with fully nonlinear parabolic equations and weakly coupled systems of nonlinear equations [14]– [16]. Note that a combination of the two approaches allows as well to define and study viscosity solutions of strongly coupled systems of nonlinear parabolic equations [17].

It should be mentioned that among a number of results the probabilistic point of view allows to reveal some new features of systems of nonlinear parabolic equations. First, we show that the probabilistic representation allows to understand that some classes of the systems can be considered as scalar equations with respect to functions with a changed phase space [18],[19]. On the other hand one may see that there are exist classes of nonlinear equations and systems such that a solution $v(t, x)$ of the forward Cauchy problem for them can be easily reduced to the solution $u(T-t, x) = v(t, x)$ of the backward one and vice versa. Actually, this transformation is rather formal and should be used carefully since it may lead to losing important properties of the required solutions. Nevertheless we may apply it to develop effective numerical algorithms to obtain an approximate solution of the problem under consideration. At the same time there are some classes of systems which do not admit such reduction at all. As a rule these systems admit an interpretation as systems of nonlinear forward Kolmogorov equations (see [20] – [22]) and the McKean's approach can be extended to them. In general there is a great number of papers devoted to the extension of the McKean's approach (see recent book [23] and references there). Due to volume limitations we do not discuss here this very interesting topic.

Our special attention will be paid instead to discussion of numerical schemes based on the probabilistic representations of solutions to the backward Cauchy problem. We discuss numerical schemes based both on probabilistic representations of the backward Cauchy problem solutions based on forward SDEs and their multiplicative functionals and those based on BSDEs. Note that the numerical schemes of the type discussed in the article recently appears to be rather hot topic since they pretend to give a possibility to overcome the curse of dimensionality in constructing numerical solutions of parabolic equations combining probabilistic representations and the possibilities of neural network theory [24] –[26].

Finally, it should be mentioned that some probabilistic models associated with nonlinear parabolic equations and systems can be considered as underlying microscopical models describing the "physical nature" of the phenomena under study while others can serve only as convenient artificial tools to construct efficient numerical schemes. We put the words physical nature in quotation marks since it may concern not only physical but also chemical or biological problems as well as problems of financial mathematics.

The main aim of this article is to present principal ideas which allow to construct stochastic models for nonlinear parabolic equations and systems and develop effective numerical algorithms to construct approximations of both classical and viscosity solutions of the Cauchy problem. Because of volume restrictions in some cases we give only a sketch of a proof mentioning papers where the full proof can be found.

The remainder of this article is structured as follows:

In section 2 we construct stochastic processes associated with the Cauchy problem for a semilinear parabolic equation and a system of such equations.

In section 3 we extend this approach to quasilinear and fully nonlinear parabolic equations and systems. Note that the models considered in sections 2 and 3 allow to construct a classical solution of the original Cauchy problem.

In section 4 we discuss an alternative approach to quasilinear and fully nonlinear parabolic equations and systems based on the BSDE theory. This approach allows to construct viscosity solutions both to quasilinear and fully nonlinear scalar parabolic equations. Moreover a combination of approaches of sections 2 and 4 allows to define and construct a viscosity solution of a system of nonlinear parabolic equations treating it as a scalar equation with a changed phase space.

In section 5 we consider numerical methods to solve the Cauchy problem for nonlinear parabolic equations and systems based on the probabilistic representations of their solutions. We start with probabilistic representations described in sections 2 and 3 consider probabilistic representations of nonlinear PDE solutions in terms of suitable diffusion processes and their multiplicative operator functionals and use them to derive numerical algorithms. Next we deal with numerical algorithms based on the results of section 4. Namely, we consider an FBSDE associated with a fully nonlinear parabolic equation and reduce it to a certain optimal control problem. To solve numerically this problem we apply the neural network theory and derive the required numerical algorithms. As a result we obtain a numerical solution of the original PDE problem.

2. Probabilistic model of the Cauchy problem for semilinear backward parabolic equations and systems

Let us start with an exposition of a probabilistic model for the Cauchy problem for a semilinear parabolic equation of the form

$$u_s + \frac{1}{2}TrA^u(s,x)\nabla^2 u[A^u]^*(s,x) + \langle a^u(s,x), \nabla u\rangle = 0, \quad u(T,x) = u_0(x). \quad (2.1)$$

We fix a probability space $(\Omega, \mathcal{F}, P)$ and denote by $w(t) \in R^d$ the Wiener process and by $\mathcal{F}_t \subset \mathcal{F}$ a flow of σ-subalgebras generated by $w(t)$.

To construct a probabilistic model of (2.1) we consider a stochastic equation of the form

$$d\xi(t) = a(t,\xi(t),u(t,\xi(t)))dt + A(t,\xi(t),u(t,\xi(t)))dw(t), \quad \xi(s) = x, \quad (2.2)$$

and notice that (2.2) includes two unknown objects, namely, the process $\xi(t) \in R^d$ and a function $u(s,x) \in R$. To obtain a closed system we add to (2.2) a relation

$$u(s,x) = E[h(\xi_{s,x}(T))] = E[h(\xi(T))|\xi(s) = x] \quad (2.3)$$

and study (2.2), (2.3) as a closed system with respect to $\xi(t)$ and $u(s,x)$.

Introduce some necessary notations.

Let $C^k(R^d)$ be the Banach space of k-differentiable bounded functions defined on R^d with the norm $\|u\|_\infty = \sup_{x\in R^d}|u(x)|$ and $\mathcal{L}(R^d)$ be a set of bounded Lipschitz continuous functions with the norm $\|u\|_{\mathcal{L}} = \sup_{x\in R^d}|u(x)|$.

We say that condition **C 2.1** holds if there exist positive constants C, L, L_0, K_0, K_0^1 such that the functions $a(t,x,u) \in R^d, A(t,x,u) \in R^d \otimes R^d, x \in R^d, u \in R, t \in [0,T]$ satisfy estimates

$$\|a(t,x,u)\|^2+\|A(t,x,u)\|^2\le C[1+\|x\|^2+K_u\|u\|^2],$$
$$\|a(t,x,u_1)-a(t,y,u_2)\|^2+\|A(t,x,u_1)-A(t,y,u_2)\|^2\le L\|x-y\|^2+C_u\|u_1-u_2\|^2,$$
$$\|u_0(x)\|^2_\infty\le K_0,\quad \|u_0(x)-u_0(y)\|^2\le L_0\|x-y\|^2,\quad \|\nabla u_0(x)\|^2_\infty\le K_0^1,$$

where $C_u, K_u>0$.

Let $v(s,x)$ satisfies inequalities

$$\|v(s)\|^2_\infty=K_v(s)<\infty,\ |v(s,x)-v(s,y)|^2\le L_v(s)\|x-y\|^2$$

and $L_v(s)<\infty$ for $s\in[0,T]$. Consider an SDE

$$d\xi(t)=a(t,\xi(t),v(t,\xi(t)))dt+A(t,\xi(t),v(t,\xi(t)))dw(t),\quad \xi(s)=x,\ s\le t. \tag{2.4}$$

Applying the Ito formula and standard estimates we may prove the following assertion.

Lemma 2.1. *Let* **C 2.1** *hold. Then there exists a unique solution $\xi(t)=\xi_{s,x,v}(t)$ of* (2.4) *satisfying estimates*

$$E\|\xi(t)\|^2\le\left[\|x\|^2+C(T-s)+C\int_s^t K_v(\tau)d\tau\right]e^{C(T-s)},$$
$$E\|\xi_{s,x,v}(t)-\xi_{s,y,v}(t)\|^2\le\|x-y\|^2L_0e^{\int_s^t[L+C_vL_v(\tau)]d\tau}, \tag{2.5}$$
$$E\|\xi_{s,x,v}(t)-\xi_{s,x,v_1}(t)\|^2\le C_v\int_s^t\|v(\tau)-v_1(\tau)\|^2_{\mathcal{L}}d\tau e^{\int_s^t[L+C_vL_v(\tau)]d\tau}. \tag{2.6}$$

In addition the function $u(s,x)=E[u_0(\xi_{s,x}(T))]$ satisfies the estimates

$$\|u(s)\|^2_\infty\le K_0$$

and

$$|u(s,x)-u(s,y)|^2\le L_0\|x-y\|^2exp\left[\int_s^T L[1+KL_v(\tau)]d\tau\right], \tag{2.7}$$

where $K=C_vL^{-1}$.

Lemma 2.2. *Let* **C 2.1** *hold. Then there exists an interval $[T_1,T]$ and bounded functions $\alpha(s)$, $\beta(s)$, $s\in[T_1,T]$ such that the function $u(s,x)=E[u_0(\xi_{s,x,v}(T))]$ satisfies estimates*

$$\|u(s)\|^2_\infty\le\alpha(s),\quad |u(s,x)-u(s,y)|^2\le\beta(s)\|x-y\|^2, \tag{2.8}$$

if $\|v(s)\|^2_\infty\le\alpha(s)$ and $|v(s,x)-v(s,y)|\le\beta(s)\|x-y\|$.

Proof. Under **C 2.1** one may choose $\alpha(s)=K_0$. To derive the second estimate in (2.8) we note that an estimate

$$L_u(s)\le L_0e^{\int_s^T L[1+KL_v(\tau)]d\tau} \tag{2.9}$$

can be deduced from (2.5).

Choosing β as a solution of the equation

$$\beta(s)=L_0exp[\int_s^T L[1+K\beta(\tau)]d\tau] \tag{2.10}$$

we note that β satisfies as well the Cauchy problem

$$\frac{d\beta(s)}{ds} = -L[1 + K\beta(s)]\beta(s), \quad \beta(T) = L_0.$$

A solution to this Cauchy problem is unique and has the form

$$\beta(s) = \frac{L_0 e^{L(T-s)}}{1 + KL_0[1 - e^{L(T-s)}]}. \tag{2.11}$$

Hence for $s \in [T_1, T]$, with $\Delta_1 = |T - T_1|$ such that

$$\Delta_1 < \frac{1}{L} ln\,[1 + \frac{1}{KL_0}], \tag{2.12}$$

the function $\beta(s)$ given by (2.11) is a required function. □

To construct a solution of the system (2.2), (2.3) we consider processes $\xi^k(t)$ and functions $u^k(s,x)$ defined by

$$d\xi^k(\tau) = a_{f^k}(\tau, \xi^k(\tau))d\tau + A_{f^k}(\tau, \xi^k(\tau))dw(\tau), \quad \xi^k(s) = x, \tag{2.13}$$

$$u^0(s,x) = u_0(x), \quad u^{k+1}(s,x) = E[u_0(\xi^k_{s,x}(T))]. \tag{2.14}$$

Theorem 2.3. *Let* **C 2.1** *hold. Then the family $u^k(s,x)$ defined by* (2.14) *uniformly in x converges to a limit function $u(s,x)$ for any $s \in [T_1, T]$, with $\Delta_1 = |T - T_1|$ satisfying* (2.12)*. In addition the family of processes $\xi^k(t)$ defined by* (2.13) *converges in mean square to a limit process $\xi(t)$.*

Proof. From lemma 2.1 we can deduce that $\Phi(s,x,u) = E[u_0(\xi_{s,x}(t))] = u(s,x)$ defines a contraction map in $\mathcal{L}$. Denote by

$$\kappa^k(s,x) = |u^{k+1}(s,x) - u^k(s,x)|^2$$

and by

$$\zeta^k(s) = sup_x \kappa^k(s,x).$$

The estimates of lemma 2.2 yield

$$\kappa^k(s,x) \le L_{u_0} \int_s^t \|u^k(\tau) - u^{k-1}(\tau)\|^2_\infty d\tau e^{L_f(t-s)}$$

and hence the estimate

$$\zeta^k(s) \le \delta^k \int_s^t \cdots \int_s^{t_2} \|u^1(\tau_1) - u^0\|^2_\infty d\tau_1 \dots d\tau_k$$

holds for $\delta = K_0 L_{u_0} exp[L_f(T - s)]$.

As far as functions u^k are uniformly bounded and

$$\|u^1(s) - u^0\|^2_{\mathcal{L}} \le const < \infty,$$

we get an estimate

$$\|u^k(s) - u^{k-1}(s)\|^2_{\mathcal{L}} \le \frac{N^k}{k!} const,$$

where $N = \delta(T - s)$. It addition the limit function $u(s,x)$ is Lipschitz continuous in x since for any $s \in [T_1, T]$ we deduce the estimate

$$|u^k(s,x) - u^k(s,y)|^2 \le \beta(s)\|x - y\|^2,$$

where $\beta(s)$ given by (2.11) is bounded on $[T_1, T]$ and this estimate is uniform in k. □

Now we can state the following assertion.

Theorem 2.4. *Let* **C 2.1** *hold and u_0 is bounded and Lipschitz continuous. Then there exists an interval $[T_1, T]$ such that for any $s \in [T_1, T]$ there exists a unique solution $(\xi_{s,x}(t), u(s,x))$ of the system (2.2), (2.3). The function $u(s,x) \in R$ is bounded and Lipschitz continuous while $\xi(t) \in R^d$ is an $\mathcal{F}_t$-measurable Markov process such that $E\|\xi_{s,x}(t)\|^2 < \infty$ for $T_1 < s < t \leq T$.*

Proof. We deduce from theorem 2.3 that there exists a bounded Lipschitz continuous in $x \in R^d$ limit function $u(s,x)$. Then by lemma 2.1 we can prove that for each $s \in [T_1, T]$ there exists a solution of the system (2.2), (2.3) To prove uniqueness of the solution we assume on the contrary that there exists two solutions $u_1(s,x)$ and $u_2(s,x)$, satisfying (2.2), (2.3) such that $u_1(0,x) = u_2(0,x) = u_0(x)$. By lemma 2.1 we know that

$$\|u_1(s) - u_2(s)\|^2_{\mathcal{L}} \leq \int_s^T \|u_1(\tau) - u_2(\tau)\|^2_{\mathcal{L}} d\tau,$$

and deduce applying the Gronwall lemma that $\|u_1(s) - u_2(s)\|^2_{\mathcal{L}} = 0$. Uniqueness of a solution to (2.2) under **C 2.1** results from the classical results of the SDE theory as well as the Markov property of $\xi_{s,x}(t)$. Lipschitz continuity of $u(s,x)$ results from estimates in the proof of theorem 2.4. □

Remark 2.5. The family $u^k(s,x)$ converges to $u(s,x)$ uniformly in $s \in t_1, T$ and $x \in R^d$ since there exists a positive constant M such that

$$\sup_{s\in[T_1,T]} \|u_1(s) - u_2(s)\|^2_{\mathcal{L}} \leq M \int_{T_1}^T \|u_1(\tau) - u_2(\tau)\|^2_{\mathcal{L}} d\tau.$$

Then by the Gronwall lemma $\sup_{s\in[T_1,T]} \|u_1(s) - u_2(s)\|^2_{\mathcal{L}} = 0$.

To expose the link between the solution of the system (2.2), (2.3) and the Cauchy problem (2.1) let us assume that there exists a unique classical solution $u(s) \in C^2(R^d)$ of (2.1). Then by the integral version of the Ito formula we have that the process $u(t, \xi(t))$ satisfies the equation

$$u(T, \xi(T)) = u(s,x) + \int_s^T [u_t(t, \xi(t)) + \mathcal{L}^u u(t, \xi(t))]dt + \tag{2.15}$$

$$+ \int_s^T A^u(t, \xi(t)) \nabla u(t, \xi(t)) dw(t),$$

where $\mathcal{L}^v u(t,x) = \frac{1}{2} Tr A^v(t,x) \nabla^2 u [A^v]^*(t,x) + \langle a^v(t,x), \nabla u(t,x) \rangle$. Evaluating expectation of both parts of (2.15) and keeping in mind that $u(s,x)$ solves (2.1) we obtain

$$u(s,x) = E[u_0(\xi_{s,x}(T))].$$

To prove the inverse assertion we have to prove that under some suitable assumptions the functions $u(s,x)$ satisfying (2.3) is twice differentiable. We prove the corresponding results as a consequence of the next section assertions.

To extend this approach to systems of parabolic equations we need some additional functional spaces.

Given $\phi \in C_b(R^d; R^{d_1})$ we denote by Θ the set of functions $\Phi(z) = \langle h, \phi(x)\rangle$ defined on $Z = R^d \times R^{d_1}$, $z = (x, h) \in Z$ with the norm

$$\|\Phi\|_\Theta = \sup_{\|h\|=1} \sup_{x\in R^d} |\langle h, \phi(x)\rangle|.$$

Let $\Theta_1 = C_b(R^d; R^{d_1})$ with the norm $\|\phi\|_{\Theta_1} = \sup_{x\in R^d}\|\phi(x)\|$ and $\mathcal{L}$ denote its subset consisting of Lipschitz continuous functions. One can easily check that

$$\|\Phi\|_\Theta = \|\phi\|_{\Theta_1}.$$

Let $c^u(s,x) \in R^{d_1}, C^u(s,x)y \in L(R^{d_1})$, $x, y \in R^d$, $u \in R^{d_1}$, $a^u(s,x), A^u(s,x)$ be as above and we denote by $[C^u(s,x)(h,y)]_m = \sum_{i=1}^d \sum_{l=1}^{d_1} C^i_{ml}(s,x,u(s,x))h_l y_i$.

To simplify formulas below we use Einstein convention about summing up over the repeating indices if the contrary is not mentioned.

Consider the Cauchy problem for a system of semilinear parabolic equations

$$\frac{\partial u_m}{\partial s} + \langle a(s,x,u), \nabla\rangle u_m + \frac{1}{2} Tr A(s,x,u)\nabla^2 u_m A^*(s,x,u)+$$
$$+B^i_{ml}(s,x,u)\nabla_i u_l + c_{ml}(s,x,u)u_l = 0, \quad u_m(T,x) = u_{0m}(x), \tag{2.16}$$
$$m, l = 1, 2, \ldots, d_1, \quad i = 1, \ldots, d.$$

To construct a stochastic model associated with (2.16) consider a system of SDEs

$$d\xi = a^u(\tau, \xi(\tau))d\tau + A^u(\tau, \xi(\tau))dw(\tau), \quad \xi(s) = x, \tag{2.17}$$

$$d\eta(t) = c^u(\tau, \xi(\tau))\eta(\tau)d\tau + C^u(\tau, \xi(\tau))(\eta(\tau), dw(\tau)), \quad \eta(s) = h, \tag{2.18}$$

and a closing equation

$$\langle h, u(s,x)\rangle = E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle]. \tag{2.19}$$

As above we use notations

$$E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle] = E[\langle \eta(T), u_0(\xi(T))\rangle | \xi(s) = x, \eta(s) = h].$$

and besides $[C^u(h,w)]_m = \sum_{q=1}^{d_1} \sum_{i=1}^d C^i_{mq} h_q w_i$.

Remark 2.6. Note that coefficients $c(s,x,u), C(s,x,u)$ in (2.18) are dual to coefficients $c^*(s,x,u)$ and $C^*(s,x,u)$ in (2.16) where $C^*(s,x,u)A(s,x,u) = B(s,x,u)$ that is

$$\sum_{m,l=1}^{d_1} \left[\sum_{m=1}^{d_1} c_{lm} h_m\right] u_l = \sum_{m=1}^{d_1} h_m \left[\sum_{l=1}^{d_1} c_{ml} u_l\right],$$

$$\sum_{m,l=1}^{d_1} \left[\sum_{i=1}^{d} B^i_{ml} \nabla_i u_l\right] h_m = \sum_{m=1}^{d_1} h_m \left[\sum_{l=1}^{d_1}\sum_{k=1}^{d} C^k_{ml} \nabla_i u_l A^{ki}\right].$$

To construct a solution to (2.17) –(2.19) we consider a system

$$u^0(s,x) = u_0(x), \quad \xi^0(s) = x, \tag{2.20}$$

$$d\xi^k(\tau) = a^{u^k}(\tau, \xi^k(\tau))d\tau + A^{u^k}(\tau, \xi^k(\tau))dw(\tau), \quad \xi^k(s) = x, \tag{2.21}$$

$$d\eta^k(t) = c^{u^k}(\tau, \xi^k(\tau))\eta^k(\tau)d\tau + C^{u^k}(\tau, \xi^k(\tau))(\eta^k(\tau), dw(\tau)), \quad \eta^k(s) = h, \tag{2.22}$$

$$\langle h, u^{k+1}(s,x)\rangle = E[\langle \eta^k_{s,h}(T), u_0(\xi^k_{s,x}(T))\rangle] \tag{2.23}$$

and prove that the family $u^k(s,x)$ defined by (2.23) is uniformly bounded and equicontinuous in the norm of Θ_1, or equivalently that the family $\Phi^k(s,z) = \langle h, u^k(s,x)\rangle$ is uniformly bounded and equicontinuous in the norm of Θ.

We say that **C 2.3** holds if **C 2.1** is valid for $u \in R^{d_1}$ and there exist constants $L, C_1 > 0$ and C_0 such that

$$\langle c(s,x,u)h, h\rangle + \|C(s,x,u)h\|^2 \le [C_0 + C_1\|u\|^2]\|h\|^2,$$

$$\|[c(s,x,u) - c(s,y,u)]h\|^2 + \|[C(s,x,u) - C(s,y,u)]h\|^2 \le$$
$$\le [L\|x-y\|^2 + M_u\|u-u_1\|^2]\|h\|^2,$$

where $M_u = M(u,u_1)$ is a positive constant depending on u and u_1.

We say that **C 2.4** holds if **C 2.2** holds and **C 2.3** is valid both for coefficients c^v, C^v and their derivatives up to the order 2.

It can be verified that conditions **C 2.3**, **C 2.4** ensure that conditions **C 2.1** and **C 2.2** hold for coefficients q and Q in (2.20) and hence we can apply the above results to prove the following assertion.

Let us describe more details of this approach.

Let $\phi \in C_b(R^d, R^{d_1})$ and $\Phi(z) = \langle h, \phi(s,x)\rangle$. Recall that $\|\Phi\|_\Theta = \|\phi\|_{\Theta_1}$.

As above we have to prove that the family of successive approximations $u^k(s,x)$ given by (2.23) converges in Θ_1 to a limit function $u(s,x)$. To this end given positive functions $\gamma_v(s)$ and $L_v(s)$ and a function $v(s,x) \in R^{d_1}$ such that

$$sup_{x\in R^d}\|v(s,x)\| = K_v(s) \le \gamma_v(s), \|v(s,x) - v(s,y)\|^2 \le L_v(s)\|x-y\|^2, \quad (2.24)$$

we consider stochastic processes

$$d\xi(\tau) = a^v(\tau,\xi(\tau))d\tau + A^v(\tau,\xi(\tau))dw(\tau), \quad \xi(s) = x, \quad (2.25)$$

$$d\eta(\tau) = c^v(\tau,\xi(\tau))\eta(\tau)d\tau + C^v(\tau,\xi(\tau))(\eta(\tau), dw(\tau)), \quad \eta(s) = h. \quad (2.26)$$

Lemma 2.7. *Let* **C 2.2** *hold,* $u_0 \in \Theta_1$ *and* $v(s) \in \Theta_1$ *satisfies* (2.24) *and processes* $\xi(t), \eta(t)$ *are governed by* (2.25), (2.26)*. Then there exists an interval* $\Delta_2 = [T_2, T]$ *such that a function* $g(s,x)$ *defined by*

$$\langle h, g(s,x)\rangle = E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle], \quad (2.27)$$

belongs to Θ_1 *for all* $s \in \Delta_2$.

Proof. Our aim is to verify that there exists an interval $\Delta_2 = [T_2, T]$, with $T_2 \le T$ such that $K_g(s) \le \gamma(s)$ for all $s \in \Delta_2$.

We deduce from estimates in **C 2.3** and standard stochastic integral estimates that

$$\|g(s)\|_{\Theta_1} \le K_{u_0} exp\left[\int_s^T [2C_0 + 3C_1K_v(\tau)]d\tau\right],$$

where $K_{u_0} = sup_x\|u_0(x)\|$. By arguments similar to those used in the proof of theorem 2.4 we verify that the function $\gamma(s)$, defined by

$$\gamma(\tau) = \frac{2C_0K_{u_0}e^{2C_0(T-\tau)}}{2C_0 + 3C_1K_{u_0} - 3C_1K_{u_0}e^{2C_0(T-\tau)}}, \quad (2.28)$$

has the required properties and if $K_v(\tau) \le \gamma(\tau)$, then $K_g(\tau) \le \gamma(\tau)$ for $\tau \in [T_2, T]$.

It follows from (2.28) that $\gamma(\tau)$ is a bounded function for all $\tau \in [0,T]$ if $2C_0+3C_1K_{u_0} < 0$. Otherwise $\gamma(\tau)$ is bounded over the set $\Delta_2 = [T_2, T]$ such that

$$|T_2 - T| < \frac{1}{2C_0} ln \left[1 + \frac{2C_0}{3C_1K_{u_0}}\right]. \tag{2.29}$$

□

We may apply the above results to successive approximations

$$f_k(s,x) = \nabla u_k(s,x), k = 1,\dots,N, \quad \text{of} \quad \nabla u(s,x),$$

where $u(s,x)$ is given by (2.14). This will allow to prove that $u_k(s,x)$ are uniformly in k equicontinuous in x for each $s \in [T_2, T]$.

Denote by $z = (x,y) \in R^d \times R^d$, $g = (v, \nabla v)$,

$$\tilde{a}^g(s,z) = (a^v(s,x), \nabla a^v(s,x)y), \quad \tilde{A}^g(s,z) = (A^v(s,x), \nabla A^v(s,x)y),$$

$$\tilde{c}^g(s,z)h = (c^v(x)h, \nabla c^v(s,x)(h,y)), \quad \tilde{C}^g(s,z)h = (C^v(s,x)h, \nabla C^v(s,x)(h,y))$$

and let $f(s,x) = \nabla v(s,x)$.

Next we consider the processes $\alpha(t) = \nabla \xi_{s,x}(t), \beta(t) = \nabla \eta(t)$ satisfying SDEs

$$d\alpha(\tau) = \nabla a^g(\tau, \xi(\tau))\alpha(\tau)d\tau + \nabla A^g(\tau,\xi(\tau))(\alpha(\tau), dw(\tau)), \quad \alpha(s) = I, \tag{2.30}$$

$$\begin{gathered} d\beta(\tau) = \nabla c^g(\tau,\xi(\tau))(\alpha(\tau),\eta(\tau))d\tau + \\ +c^g(\tau,\xi(\tau))(\alpha(\tau)\beta(\tau))d\tau + \nabla C^g(\tau,\xi(\tau))(\alpha(\tau),\eta(\tau))dw(\tau)+ \\ +C^g(\tau,\xi(\tau))(\beta(\tau),dw(\tau)), \quad \eta(s) = 0, \end{gathered} \tag{2.31}$$

where processes $\xi(t)$ and $\eta(t)$ satisfy (2.25) and (2.26) respectively. Here I is the unity matrix and we use notations of the type

$$\nabla a^g(s,x) = \nabla_x a(s,x,v(s,x)) + \nabla_v a(s,x,v(s,x))\nabla_x v(s,x).$$

Next we have to verify that the function $\nabla u(s,x)$ such that

$$\begin{gathered} \langle h, \nabla u(s,x)y\rangle = E[\langle \beta_{s,h}(T)y, u_0(\xi_{s,x}(T))\rangle + \\ +\langle \eta_{s,h}(T), \nabla u_0(\xi_{s,x}(T))\alpha(T)y\rangle], \end{gathered} \tag{2.32}$$

is bounded provided that function $\nabla v(s,x)$ is bounded.

As it is easy to see the system (2.17) – (2.19), (2.30) – (2.32) has a structure which is similar to the structure of the system (2.17) – (2.19) itself. This allows to apply the above speculations to this new more bulky system.

Coming back to to the Cauchy problem of the form (2.16) with respect to the function $q = (u, \nabla u)$, we assume that there are exists constants $C_0^1, C_1^1, K_{u_0}^1$ such that for each $h \in R_1^d$ with a finite norm coefficients $\tilde{a}^q, \tilde{c}^q$ and $\tilde{A}^q, \tilde{C}^q$ satisfy estimates **C 2.3**. Then we deduce that there exists an interval $\Delta_3 = [T_3, T]$, and a function $\gamma_2(s)$, bounded on this interval such that the inequality $\|\nabla_y v(s)\|_{\Theta_1} \le \gamma_2(s)\|y\|$ yields an estimate $\|\nabla_y u(s)\|_{\Theta_1} \le \gamma_2(s)\|y\|$.

To be more precise by arguments similar to those used in the proof of lemma 2.2 we can prove that a function $\gamma_2(s)$ of the form

$$\gamma_2(s) = \frac{2C_0^1K_{u_0}^1 e^{2C_0^1(T-s)}}{2C_0^1 + 3C_1^1K_{u_0}^1 - 3C_1^1K_{u_0}^1 e^{2C_0^1(T-s)}}$$

has the following property. If $\|\nabla_y v(t,x)\| \le \gamma_2(s)\|y\|$ over an interval $\Delta_4 = [\max(T_2, T_3), T]$ such that

$$|\Delta_4| < \frac{1}{2C_0^1} \ln\left[1 + \frac{2C_0^1}{3C_1^1 K_0^1}\right], \tag{2.33}$$

then $\|\nabla_y u(t,x)\| \le \gamma_2(s)\|y\|$ on this interval.

Recall that to prove that $u_k(s,x)$ are uniformly in k Lipschitz continuous in x it is enough to verify that they have bounded derivatives $\nabla u_k(s,x)$ uniformly in k. In order to prove that the function $f(s,x) = \nabla u(s,x)$ is bounded we can consider the stochastic system (2.17) – (2.19), (2.30) – (2.32) and repeat the above considerations.

Finally, having the above apriori estimates we can prove the following assertion.

Theorem 2.8. *Assume that* **C 3** *holds. Then there exists an interval* Δ_4 *satisfying* (2.33) *such that for* $s \in \Delta_4$ *there exists a unique solution of the system* (2.17) – (2.19)*,* (2.30) – (2.32)*.*

We may apply the above considerations to the second order differential prolongation of the system (2.17)–(2.19) which is obtained by adding to (2.17)– (2.19) the system (2.30)–(2.32) which governs the processes $\alpha(t) = \nabla\xi(t)$, $\beta(t) = \nabla\eta(t)$, as well as SDEs for $\alpha_1(t) = \nabla^2\xi(t), \beta_1(t) = \nabla^2\eta(t)$ having a form of nonuniform linear SDEs and a relation for $\nabla^2 u(s,x)$

$$\langle h, \nabla^2 u(s,x)(y,y_1)\rangle = \nabla_{y_1}\langle h, \nabla u(s,x)y\rangle = \nabla_{y_1} E[\langle \beta_{s,h}(T)y, u_0(\xi_{s,x}(T))\rangle +$$

$$+\langle \eta_{s,h}(T), \nabla u_0(\xi_{s,x}(T))\alpha(T)y\rangle],$$

where $\nabla_y u = \langle y, \nabla u\rangle$.

Theorem 2.9. *Let* **C 2.4** *hold. Then there exists an interval* $[T_4, T]$*, with the length satisfying (2.33) and for all* $s \in [T_4, T]$ *there exists a unique solution of* $(2.17) - (2.19), (2.30) - (2.32)$*. If coefficients of the system* (2.16) *and the initial function* u_0 *are 2 times continuously differentiable in* x *then the function* $u(s,x)$ *is also twice continuously differentiable possibly over a smaller interval* $[T_5, T] \subset [T_4, T]$*.*

Proof. Due to the above results it remains only to prove the last assertion of the theorem. Let us differentiate the system (2.17)-(2.19). As we have seen above as a result we obtain a more complicated system for the processes $\xi(t), \eta(t), \zeta(t), \kappa(t))$, where $\zeta(t) = \nabla\xi_{s,x}(t)$, $\kappa(t) = \nabla\eta_{s,x}(t)$ and functions $u(s,x)$, $\nabla u(s,x)$, though a structure of this new system is similar to the structure of the original system. Thus under the condition **C 4** we can verify that conditions of theorem 2.4 hold and its conclusion can be applied to the system under consideration. If we prove that there exists a solution to this new system we may deduce that the solution $u(s,x)$ of (2.17)-(2.19) has a bounded gradient. Repeating this procedure once more we can prove that the function $u(s,x)$ is twice differentiable under suitable conditions on coefficients and initial data of (2.16). □

To show links between (2.16) and (2.17)–(2.19) assume that $u(s,x)$ is a unique classical solution of (2.16) and apply the Ito formula to the function $\Phi(s,x,h) =$

$\langle h, u(s,x)\rangle$ and the two component process $\zeta(t) = (\xi(t), \eta(t))$ satisfying (2.17)–(2.19). As a result we can verify that $v(s,x) = E\langle \eta_u(T), u_0(\xi_u(T))\rangle$ satisfies (2.16). Finally we can show that $u(s,x) \equiv v(s,x)$. To prove the inverse assertion we have to verify that the function $u(s,x)$ defined by (2.19) have two bounded derivatives in x and then apply the above considerations.

Theorem 2.10. *Assume that* **C 2.4** *holds. Then there exists an interval* $[T_5, T]$ *such that for* $s \in [T_5, T]$ *there exists a unique solution* $\xi_{s,x}(t), \eta_{s,h}(t), u(s,x)$ *of the system* (2.17) $-$ (2.19). *The function* $u(s,x)$ *is twice differentiable and hence it defines a unique classical solution to* (2.16)

One can see more detailed proof of the above statements in [6], [19].

Remark 2.11. The probabilistic representation of a solution to (2.16) prompts that one can reduce this problem to an equivalent Cauchy problem for a scalar parabolic equation

$$\Phi_s + \frac{1}{2} Q^u \nabla_z^2 \Phi [Q^u]^* + \langle q^u, \nabla_z \Phi\rangle = 0, \quad \Phi(T,z) = \langle h, u_0(x)\rangle, \tag{2.34}$$

with respect to a scalar function $\Phi(s,z), s \in [0,T], z = (x,h) \in R^d \times R^{d_1}$. Here

$$q^u(s,z) = \begin{pmatrix} a(s,x,u) & 0 \\ 0 & c(s,x,u)h \end{pmatrix}, \quad Q^u(s,z) = \begin{pmatrix} A(s,x,u) & 0 \\ 0 & C(s,x,u)h \end{pmatrix}. \tag{2.35}$$

Note that one can rewrite the stochastic system (2.17) $-$ (2.19) in the form

$$d\gamma(t) = q^u(t,\gamma(t))dt + Q^u(t,\gamma(t))\, dW(t), \quad \gamma(s) = \gamma = (x,h), \tag{2.36}$$

$$\Phi(s,z) = E[\Phi_0(\gamma_{s,z}(T))], \tag{2.37}$$

where $\gamma(t) = (\xi(t), \eta(t))^*$, $W(t) = (w(t), w(t))^*$. Besides the function $u(s,x)$ defined by (2.19) can be presented in the form

$$\langle h, u(s,x) = E\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle = \langle h, E\,[\Gamma^*(s,T) u_0(\xi(T))]\rangle. \tag{2.38}$$

where $\Gamma(T,s)h = \eta_{s,h}(T)$.

Remark 2.12. Let us briefly mention one more class of systems of nonlinear parabolic equations called system with switching regimes or hybrid systems for which a probabilistic approach shows that the original system of parabolic equations might be considered as a scalar parabolic equations with a different phase space. In linear case systems of this type were studied by many authors (see [18] and references there). The results mentioned here were obtained in [19].

Let M be an integer and $V = \{1, \ldots, M\}$ be a fixed discrete set. Along with the Wiener process $w(t) \in R^d$ we will need a Markov chain $\gamma(t) \in V$ defined on the same probability space $(\Omega, \mathcal{F}, P)$. Let

$$[Q^u v]_m = \sum_{l=1}^M q_{ml}(x,u) v_l$$

and

$$\mathcal{L}_m^u v^m = \langle a^m(x,u), \nabla v^m\rangle + \frac{1}{2} Tr A^m(x,u) \nabla^2 v^m [A^m]^*(x,u).$$

Consider the Cauchy problem

$$\frac{\partial u_m}{\partial s} + \mathcal{L}^u_m u^m + [Q^u u]_m = 0, \quad u^m(s,x) = u^m_0(x), m = 1, \ldots, M, \tag{2.39}$$

and note that in this case we do not assume summation in m.

Let the matrix $Q(x,u) = (q^u_{lm}(x))$ possess the following properties:

1) $q^u_{lm}(x) = q_{lm}(x,u) \in R$ are bounded in x polylinear in u for all $l, m \in V$ and $x \in R^d, u \in R^M$;

2) $q_{lm}(x,u) \geq 0$ for all $x \in R^d, u \in R^M$ and $l \neq m$;

3) $q_{mm}(x,u) = -\sum_{l \neq m} q_{ml}(x,u)$ for all $x \in R^d, u \in R^M, m \in V$.

To construct a stochastic model of a solution to (2.39) we consider an SDE

$$d\xi(t) = a^u(\xi(t), \gamma(t))dt + A^u(\xi(t), \gamma(t))dw(t), \quad \xi(s) = x, \gamma(s) = m, \tag{2.40}$$

where $a : R^d \times V \times R^M \to R^d, A : R^d \times V \times R^M \to R^d \otimes R^d$ and

$$P(\gamma(t+\Delta t) = l | \gamma(t) = j, \xi(\theta), \gamma(\theta), \theta \leq t\} = q^u_{jl}(\xi(t))\Delta t + o(\Delta t), \quad \text{if} \quad l \neq j. \tag{2.41}$$

To get a closing relation we think of u as a function defined on a set $[0,T] \times R^d \times V$, that is $u_m(s,x) = u(s,x,m)$ and set

$$u(s,x,m) = E[u_0(\xi(T), \gamma(T)) | \xi(s) = x, \gamma(s) = m]. \tag{2.42}$$

Note that the evolution of the discrete component $\gamma(t)$ can be represented via a stochastic integral withe respect to a Poisson random measure [27]. In order to construct this representation for $x \in R^d$ and $i, j \in V$ with $i \neq j$ we define a set $\Delta_{ij}(x)$ of the consecutive (with respect to the lexicographic ordering on $V \times V$) closed from the left and open from the right intervals of the real line having the length $q_{ij}(x)$. Let $g : R^d \times V \times R \to R$ has the form

$$g(x,i,y) = \sum_{j=1}^{M} (j-i) I_{y \in \Delta_{ij}(x)}. \tag{2.43}$$

Then having the partition $\{\Delta_{ij}(x), i, j \in V\}$ we obtain $g(x,u,y) = j - i$ if $y \in \Delta_{ij}(x)$, otherwise $g(x,i,y) = 0$ and (2.43) is equivalent to

$$d\gamma(t) = \int_R g(\xi(t), \gamma(t-), y) p(dt, dy),$$

where $p(dt, dy)$ is the Poisson measure with intensity $dt \times dy$ which is independent of the Wiener process $w(t)$.

Connections between stochastic system of the form (2.40) – (2.42) and a classical solution of the Cauchy problem (2.39) were investigated in [19].

3. Systems of quasilinear and fully nonlinear PDEs

Consider a system of quasilinear parabolic equations

$$\frac{\partial u_m}{\partial s} + \frac{1}{2} F^{ij}(x,u) \nabla^2_{x_i x_j} u_m + B^i_{ml}(x,u,\nabla u) \nabla_{x_i} u_l + c_{ml}(x,u,\nabla u) u_l = 0,$$

$$u_m(T,x) = u_{0m}(x), \quad m = 1, 2, \ldots, d_1, \quad i, j = 1, \ldots, d, \tag{3.1}$$

where $F(x,u) = A(x,u)A^*(x,u)$.

To extend the above approach to (3.1) we include it to a larger system adding to (3.1) the system of parabolic equations for functions $v_m^j(t,x) = \nabla_j u_m(t,x)$ and denoting by $g_n(t,x) = u_n(t,x)$ for $n = 1,\dots,d_1$ and $g_n(t,x) = \nabla_{x_j} g_m(t,x)$ for $j = 1,\dots,d$, $m = 1,\dots,d_1$ and $n = d_1+1,\dots,d_1 \times d$. One can find at least at the formal level that $v_m^j(t,x)$ satisfies the Cauchy problem

$$\begin{gathered}\frac{\partial v_m^j}{\partial s} + \frac{1}{2}F^{ik}\nabla^2_{x_i x_k} v_m^j + B^i_{ml}\nabla_i v_l^j + c_{ml}v_l^j + \frac{1}{2}[\nabla_{x_j}F^{ik} + \nabla_{u_q}F^{ik}v_q^j]\nabla_i v_m^k \\ +[\nabla_{x_j}B^i_{ml} + \nabla_{u_q}B^i_{ml}v_q^j + \nabla_{v_q^k}B^i_{ml}\nabla_{x_j}v_q^k]v_l^i + [\nabla_{x_j}c_{ml} + \nabla_{u_q}c_{ml}v_q^j + \\ +\nabla_{v_q^k}c_{ml}\nabla_{x_j}v_q^k]u_l = 0, \quad u_m(T,x) = u_{0m}(x), \quad m,l = 1,\dots,d_1, i,j,k = 1,\dots,d.\end{gathered} \tag{3.2}$$

Finally we can write the system (3.1), (3.2) as a system with respect to components $g_n = u_m, n = 1,\dots d_1, g_n = \nabla_i u_l, n = d_1+1,\dots,d\times d_1$, of a vector valued function $g(t,x)$,

$$\frac{\partial g_n}{\partial s} + G^i_{np}(x,g)\nabla_i g_p + g_{nl}(x,g)g_l + \frac{1}{2}F^{ik}(x,u)\nabla^2_{x_i x_k} g_n = 0, \tag{3.3}$$

$$g_n(T,x) = u_{0m}(x), \quad n = 1,2,\dots,d_1,$$

$$g_n(T,x) = \nabla_i u_{0m}, i = 1,\dots,d,\ n = d_1+1,\dots,d_1 + d\times d_1.$$

Here

$$\begin{gathered}G^i_{np}\nabla_i g_p = B^i_{np}\nabla_{x_i} g_p^j, \text{ if } \quad n = 1,\dots,d_1, \\ G^i_{np}\nabla_i g_p = B^i_{np}\nabla_{x_i} g_p + \frac{1}{2}[[\nabla_{x_j}F^{ik} + \nabla_{u_q}F^{ik}v_q^j]\nabla_{x_i}v_m^k \\ g_{np}g_p = c_{np}g_p, n,p = 1,\dots d_1, \\ g_{np}g_p = [\nabla_{x_j}B^i_{ml} + \nabla_{u_q}B^i_{ml}v_q^j + \nabla_{v_q^k}B^i_{ml}\nabla_{x_j}v_q^k]v_l^i + \\ +[\nabla_{x_j}c_{ml} + \nabla_{u_q}c_{ml}v_q^j + \nabla_{v_q^k}c_{ml}\nabla_{x_j}v_q^k]u_l.\end{gathered}$$

Analyzing (3.1) and (3.3) it easy to see that they make a system with a structure similar to the structure of the system (3.1) itself. Hence we can apply to it the considerations of the previous section. The correspondent stochastic system will include an SDE for a basic process $\xi(t)$ of the form

$$d\xi = A(\xi(\tau), u(\tau,\xi(\tau)))dw(\tau), \quad \xi(s) = x, \tag{3.4}$$

and the following SDEs

$$d\eta(t) = c(\xi(\tau), u(\tau,\xi(\tau)))\eta(\tau)d\tau + C_u(\xi(\tau), u(\tau,\xi(\tau)))(\eta(\tau), dw(\tau)), \quad \eta(s) = h, \tag{3.5}$$

$$d\alpha(t) = [\nabla A(\xi(\tau), u(\tau,\xi(\tau))) + \nabla_{u_q}A(\xi(\tau), u(\tau,\xi(\tau)))\nabla u_q]\alpha(\tau)dw(\tau), \alpha(s) = I \tag{3.6}$$

$$\begin{gathered}d\beta(t) = c(\xi(\tau), u(\tau,\xi(\tau)))\beta(\tau)d\tau + \lambda(\xi(\tau), u(\tau,\xi(\tau)))(\alpha(\tau),\eta(\tau))d\tau + \\ +C(\xi(\tau), u(\tau,\xi(\tau)))\beta(\tau)dw(\tau) + \Lambda(\xi(\tau), u(\tau,\xi(\tau))(\alpha(\tau),\eta(\tau)))dw(\tau), \quad \beta(s) = 0.\end{gathered} \tag{3.7}$$

Here

$$\lambda^j_{qm}(x,u) = \nabla_{x_j}c_{qm}(x,u) + \nabla_{u_l}c_{qm}v_l^j,$$

$$\Lambda^{ij}_{qm} = \nabla_{x_i}C^j_{qm} + \nabla_{u_l}C^j_{qm}v_l^i, \quad i,j = 1,\dots,d, \quad q,m = 1,\dots,d_1.$$

Besides to obtain a closed system we need two more equations

$$\langle h, u(s,x)\rangle = E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle], \tag{3.8}$$

$$\langle h, \nabla u(s,x)\rangle = E\left[\langle \beta(T), u_0(\xi(T))\rangle + \langle \eta(T), \nabla u_0(\xi(T))\alpha(T)\rangle\right]. \tag{3.9}$$

Next we have to state conditions on the coefficients of (3.1) to ensure that coefficients itself and their derivatives satisfy conditions **C 2** and apply the results of the previous section.

Theorem 3.1. *Assume that coefficients A, c, C in (3.1) satisfy* **C 2.4**. *Then there exists and interval $[T_5, T], 0 \le T_5 < T$ such that for all $s, t \in [T_5, T]$ there exists a unique stochastic processes $\xi(t), \eta(t), \alpha(t), \beta(t)$ and functions $u(s,x), \nabla u(s,x)$ which satisfy the system* (3.4) – (3.9).

In a similar way we can treat the Cauchy problem for a fully nonlinear parabolic equation of the form

$$u_t + \Phi(x, u, \nabla u, \nabla^2 u) = 0, \quad u(T,x) = u_0(x). \tag{3.10}$$

To explain our approach we consider a bit more simple case, namely the system of the form

$$u^m_t + \Phi(u, \nabla^2 u^m) = 0, \quad u^m(T,x) = u_{0m}(x), m = 1, \dots, d_1. \tag{3.11}$$

As above we consider a differential prolongation of the system (3.10) which allows to include this equation into a system of semilinear parabolic equations. Denote by $v^0 = u \in R^{d_1}, v^1 = \nabla u \in R^{d_1} \otimes R^d$, $v^2 = \nabla^2 u \in R^{d_1} \otimes R^d \otimes R^d, v^3 = \nabla^3 u \in R^{d_1} \otimes R^d \otimes R^d \otimes R^d$ and derive equations to govern $v^{mq}, m = 1, \dots, d_1, q = 0, \dots, 3$. Let $\nabla_{v^{m2}}\Phi(v^0, v^2) = \frac{1}{2}A(v^0, v^{m2})A^*(v^0, v^{m2})$ and

$$TrA(v^0, v^{m2})\nabla^2 v^m A^*(v^0, v^{m2}) = A_{ik}(v)\nabla^2_{x_i x_j} v A_{kj}(v).$$

By formal differentiation with respect to spatial argument $x \in R^d$ we derive the following system

$$v^{m0}_t + Tr\nabla_{v^{m2}}\Phi(v^0, v^{m2})\nabla^2 v^{m0} + \Phi(v^0, v^{m2}) - Tr\nabla_{v^{m2}}\Phi(v^0, v^{m2})v^{m2} = 0, \tag{3.12}$$
$$v^0(T,x) = u_0(x).$$
$$v^{m1}_t + \nabla_{v^{q0}}\Phi(v^0, v^{m2})v^{q1} + Tr\nabla_{v^{m2}}\Phi(v^0, v^{m2})\nabla^2 v^{m1} = 0, \tag{3.13}$$
$$v^{m1}(T,x) = \nabla u^m_0(x).$$

$$\begin{aligned} v^{m2}_t &+ [\nabla_{v^{q0}}\Phi(v^0, v^2)\nabla v^{q0} + \nabla_{v^2}\Phi(v^0, v^2)\nabla v^{q2}]v^{q1} + \nabla_{v^{q0}}\Phi(v^0, v^{m2})v^{q2} \\ &+ Tr\nabla_{v^{m2}}\Phi(v^0, v^{m2})\nabla^2 v^{m2} + [\nabla_{v^{q0}}\Phi(v^0, v^{m2})\nabla v^{q0} \\ &+ \nabla_{v^{m2}}\Phi(v^0, v^{m2})\nabla v^{q2}]\nabla v^{m2} = 0, \quad v^{m2}(T,x) = \nabla^2 u^m_0(x). \end{aligned} \tag{3.14}$$

Note that in (3.12)–(3.14) summation over m is not supposed.

Unfortunately, the last term in (3.13) is still nonlinear in ∇v and hence we need an equation for v^{m3}. We leave it to a reader to verify that as a result we obtain a system of semilinear parabolic equations which has a form

$$v^l_t + \frac{1}{2}TrA(v)\nabla^2 v^l A^*(v) + C^{sl}_i(v)A_{ik}(v)\nabla_k v^{sl} + c^{ls}(v)v^{sl} = 0, \tag{3.15}$$
$$v^l(T) = v^l_0(x), \quad l = 1, \dots, M = d_1[1 + d + d^2 + d^3].$$

Thus, if $\nabla_{u^{m2}}\Phi(u, u^{m2}) > 0$ and coefficients $A(v), c(v)$ and $C(v)$ in (3.15) satisfy condition **C.2.4** then we can apply the results of the previous section to the system (3.15).

4. FBSDE and quasilinear and fully nonlinear parabolic equations and systems

In this section we consider an alternative probabilistic approach to quasilinear and fully nonlinear parabolic equations and systems. In addition we show the way to give a probabilistic interpretation of a viscosity solution of the Cauchy problem for systems of nonlinear parabolic equations.

Consider a diagonal system of PDEs of the form

$$u^m_s + \frac{1}{2}TrA(x,u)\nabla^2 u^m A^*(x) + \langle a^u(x), \nabla u^m\rangle + f^m(x,u,\nabla u) = 0, \tag{4.1}$$

$$u^m(T,x) = u_{0,m}(x) \in R,\ m = 1, \dots d_1.$$

Here $A : R^d \times R^{d_1} \to R^d \otimes R^d, a : R^d \times R^{d_1} \to R^d, f : R^d \times R^{d_1} \times R^d \otimes R^{d_1} \to R^{d_1}$. Let there exists a classical solution $u(s,x) \in R^{d_1}$ of (4.1) and a stochastic process $\xi(t)$ satisfies an SDE

$$d\xi(t) = a(\xi(t), u(t,\xi(t)))dt + A(\xi(t), u(t,\xi(t)))dw(t), \quad \xi(s) = 0. \tag{4.2}$$

Keeping in mind (4.1) and applying Ito's formula we derive an expression for a stochastic differential $dy(t)$ of the stochastic process $y(t) = u(t,\xi(t))$

$$dy(t) = -g(\xi(t), y(t), z(t))dt + \langle z(t), dw(t)\rangle, \quad y(T) = u_0(\xi(T)) = \zeta \in R^{d_1}, \tag{4.3}$$

where $z(t) = A^*(\xi(t)u(t,\xi(t)))\nabla u(t,\xi(t))$ and $f(x,u,\nabla u) = g(x,u,[A^u]^*\nabla u)$. The couple (4.2), (4.3) is called a forward-backward stochastic equation (FBSDE).

Denote by $\mathcal{F}_t$ a flow of sigma-subalgebras of $\mathcal{F}$, generated by the Wiener process $w(t)$. To explain specific features of an equation of the type (4.3) note that even in the case when $\xi(t)$ is a known $\mathcal{F}_t$ adapted stochastic process an equation of the form (4.3) includes two unknown processes $y(t)$ and $z(t)$ and thus its solution is defined as a couple of $\mathcal{F}_t$-adapted stochastic process such that $y(T) = \zeta$. Thus, one needs an additional relation to construct a solution to such an equation and the Ito martingale representation theorem can be used to obtain the required relation.

Namely, given $\mathcal{F}_T$-measurable random variable $\zeta \in R^d \otimes R^{d_1}$ under suitable conditions on g one may consider a square integrable martingale

$$\chi = E[\zeta + \int_0^T g(\xi(\tau), y(\tau), z(\tau))d\tau|\mathcal{F}_t].$$

By the Ito theorem χ^m could be presented in the form

$$\chi = E[\chi] + \int_0^T z(\tau)dw(\tau), \tag{4.4}$$

where $z(t) \in R^d \otimes R^{d_1}$ is $\mathcal{F}_t$-measurable uniquely defined process such that

$$E\left[\int_0^T \|z(t)\|^2 dt\right] \leq \infty$$

in a matrix norm. Finally, one can verify that a couple $(y(t), z(t))$ satisfies (4.3) if

$$y(t) = E\left[\zeta + \int_t^T g(\xi(\tau), y(\tau), z(\tau))d\tau|\mathcal{F}_t\right] \tag{4.5}$$

and $z(t)$ is defined by (4.4).

To be more precise we need some spaces of stochastic processes. Denote by:

$L^2_T(R^{d_1})$ be the space of $\mathcal{F}_T$-measurable random variables $\chi \in R^{d_1}$ such that $E\|\xi\|^2 < \infty$;

$\mathcal{S}^2(R^{d_1})$ the set of continuous stochastic processes $y(t) \in R^{d_1}$ such that

$$E[\sup_{0\le t\le T} \|y(t)\|^2] < \infty;$$

$\mathcal{H}^2(M)$ the set of $R^d \otimes R^{d_1} = M$ valued processes $z(t)$ such that for $h \in R^d$ with $\|h\| < \infty$,

$$E\left[\int_0^T \|z(t)h\|^2 dt\right] \le \infty.$$

We say that processes $\xi(t) \in R^d, y(t) \in R^{d_1}, z(t) \in M$ solve the FBSDE (4.2), (4.3) if they are $\mathcal{F}_t$ measurable, $\xi(t) \in \mathcal{S}^2(R^d)$, $y(t) \in \mathcal{S}^2(R^{d_1})$, $z(t) \in \mathcal{H}^2$ and for given $\mathcal{F}_T$ - measurable random variable $\zeta \in R^{d_1}$ such that $E[\|\zeta\|^2] < \infty$

$$y(t) = \zeta + \int_t^T g(\xi(\tau), y(\tau), z(\tau))d\tau - \int_t^T z(\tau)dw(\tau) \tag{4.6}$$

holds with probability 1.

We generalize this approach to apply it to a nondiagonal system of nonlinear parabolic equations

$$u^m_s + \frac{1}{2} TrA(x,u,\nabla u)\nabla^2 u^m A^*(x,u,\nabla u) + \langle a(x,u,\nabla u), \nabla u^m\rangle + \tag{4.7}$$
$$+B^{ml}_i(x,u,\nabla u)\nabla_{x_i} u_l + c^{ml}(x,u,\nabla u)u_l + g^m(x,u,\nabla u) = 0,$$
$$u^m(T,x) = u_{0,m}(x) \in R,\ m = 1,\dots d_1.$$

First we extend to the system (4.7) the relations obtained in the Remark (2.12). Namely, we rewrite (4.7) as a scalar equation

$$\frac{\partial \Phi}{\partial s} + \frac{1}{2} TrQ^*(x,h)\nabla^2\Phi Q(x,h) + \langle q(x,h), \nabla\Phi\rangle + G(s,h,x,\Phi, Q^*\nabla\Phi) = 0, \tag{4.8}$$
$$\Phi(T,x) = \Phi_0(x,h) = \langle h, u_0(x)\rangle.$$

with respect to a scalar function $\Phi(s,x,h) = \langle h, u(s,x)\rangle$.

Here

$$TrQ^*\nabla^2\Phi(s,x,h)Q = A^*_{ki}\frac{\partial^2\Phi_l(s,x,h)}{\partial x_i\partial x_j}A_{jk} + 2C^{lm}_k\frac{\partial^2\Phi(s,x,h)}{\partial x_j\partial h_m}A_{jk} +$$
$$+C^{lm}_k h_m\frac{\partial^2\Phi(s,x,h)}{\partial h_l\partial h_p}C^{mp}_k = A^*_{ki}\frac{\partial^2 u_l(s,x)}{\partial x_i\partial x_j}A_{jk}h_l + 2C^{qm}_k\frac{\partial u_q(s,x)}{\partial x_j}A_{jk}h_m,$$

since, due to linearity of $\Phi(s,x,h)$ in h, we have $\frac{\partial^2\Phi(s,x,h)}{\partial h_q\partial h_p} \equiv 0$. In addition

$$\langle q, \nabla\Phi(s,x,h)\rangle = a_j\frac{\partial\Phi(s,x,h)}{\partial x_j} + c^{lm}h_m\frac{\partial\Phi(s,x,h)}{\partial h_l} = a_j\frac{\partial u_l}{\partial x_j}h_l + c^{lm}h_m u_l,$$
$$G(s,x,h) = h_l g_l(s,x,u,A^*\nabla u).$$

Next we consider a stochastic process $\zeta(t) = (\xi(t), \eta(t))$ given by

$$d\xi = a^u(\tau,\xi(\tau))d\tau + A^u(\tau,\xi(\tau))dw(\tau), \quad \xi(s) = x, \tag{4.9}$$

$$d\eta(t) = c^u(\tau,\xi(\tau))\eta(\tau)d\tau + C^u(\tau,\xi(\tau))(\eta(\tau), dw(\tau)), \quad \eta(s) = h, \tag{4.10}$$

as a solution of a stochastic equation

$$d\zeta(\tau) = q(\zeta(\tau), \Phi(t, \zeta(\tau)))d\tau + Q(\zeta(\tau), \Phi(t, \zeta(\tau)))dw(\tau), \quad \zeta(s) = (x, h). \tag{4.11}$$

In addition we notice that a solution of (4.10) (provided it exists) gives rise to a multiplicative operator functional $\Gamma(t, s, \xi(\cdot)) \equiv \Gamma(t, s)$ of the process $\xi(t)$ satisfying (4.9), that is $\eta(t) = \Gamma(t, s)h$ and $\Gamma(t, s)h = \Gamma(t, \theta)\Gamma(\theta, s)h$ a.s. for $0 \le s \le \theta \le t \le T$. Hence to derive an FBSDE associated with (4.8) we can proceed as follows.

Assume that there exists a classical solution to the Cauchy problem (4.7) or what is equivalent suppose that there exists a classical solution to (4.8) and compute a stochastic differential of a scalar stochastic process $Y(t) = \langle \eta(t), u(t, \xi(t))\rangle$. We can verify that

$$dY(t) = \langle d\eta(t), u(t, \xi(t))\rangle + \langle \eta(t), du(t, \xi(t))\rangle + \langle d\eta(t), du(t, \xi(t))\rangle.$$

Taking into account (4.9), (4.10) and applying the Ito formula we derive

$$dY(t) = -F(\xi(t), Y(t), Z(t))dt + \langle Z(t), dW(t)\rangle, \quad Y(T) = \zeta = \langle h, u_0(\xi(T))\rangle, \tag{4.12}$$

where $W(t) = (w(t), w(t))^*$, $\Gamma(t)h \equiv \Gamma(t, s)h = \eta_{s,h}(t)$ and

$$\begin{aligned}\langle Z(t), dW(t)\rangle &= \langle C(\Gamma(t)h, dw(t)), u(t, \xi(t))\rangle + \langle \Gamma(t)h, \nabla u(t, \xi(t))Adw\rangle = \\ &= \langle h, \Gamma^*(t)[C^*u(t, \xi(t)) + A^*\nabla u(t, \xi(t))]dw(t)\rangle.\end{aligned}$$

Keeping in mind that

$$Y(t) = \langle \eta(t), u(t, \xi(t))\rangle = \langle h, \Gamma^*(t)u(t, \xi(t))\rangle$$

it is easy to deduce from (4.9) a BSDE

$$dy(t) = -f(\xi(t), y(t), z(t))dt + z(t)dw(t), \quad y(T) = u_0(\xi(T)), \tag{4.13}$$

where

$$\begin{aligned} &f(\xi(t), y(t), z(t)) = \\ &= \Gamma^*(t)g\left(\xi(t), u(t, \xi(t)), C^*(t, \xi(t))u(t, \xi(t)) + A^*(t, \xi(t))\nabla u(t, \xi(t))\right) = \\ &= \Gamma^*(t)g\big(\xi(t), [\Gamma^*]^{-1}(t)y(t), C^*(\xi(t))[\Gamma^*]^{-1}(t)y(t) + A^*(\xi(t))[\Gamma^*]^{-1}(t)z(t)\big)\end{aligned} \tag{4.14}$$

and from $\langle h, z(t)dw(t)\rangle = \langle Z(t), dW(t)\rangle$. and

$$Z(t) = ([\Gamma^*]^{-1}(t)C^*(t, \xi(t))u(t, \xi(t)), [\Gamma^*]^{-1}(t)A^*(\xi(t))\nabla u(t, \xi(t)))^*,$$

we deduce

$$z(t) = [\Gamma^*]^{-1}(t)[C^*(t, \xi(t))u(t, \xi(t)) + A^*(t, \xi(t))u(t, \xi(t))\nabla u] \in \mathcal{H}^2(M).$$

When the solution $y(t)$ is a scalar process and a comparison theorem holds one can prove that the function $u(s, x)$ defined by $y(s) = u(s, x)$ is a viscosity solution of the Cauchy problem for a corresponding quasilinear parabolic equation [14].

The above interpretation of the system (4.7) as a scalar equation (4.8) prompts that one can call $y(s) = u(s, x)$ a viscosity solution of (4.7) if $Y(s) = \langle h, u(s, x)\rangle$ is a viscosity solution of (4.8). The proof of existence and uniqueness of a solution to the FBSDE (4.10)–(4.12). one can find in [19].

Let us discuss a probabilistic approach to interpretation of the Cauchy problem for a fully nonlinear parabolic equation based on the theory of FBSDE.

We consider the Cauchy problem

$$u_t + \tilde{\Phi}(x, u, \nabla u, \nabla^2 u) = 0, \quad u(0, x) = u_0(x) \in R,\ x \in R^d,\ t \in [0, T]. \tag{4.15}$$

We assume that the function $\tilde{\Phi}(x, y, z, r)$, $x \in R^d, y \in R, p \in R^d, r \in R^d \otimes R^d$ has a positive definite derivative with respect to $r \in R^d \otimes R^d$ that is $M_{ij} = \nabla_{r_ij}\tilde{\Phi}$ possesses the property $x^* M x \geq 0$ for all $x \in R^d \setminus \{0\}$. Let $\nabla_r \tilde{\Phi}(x, y, v, r) = \frac{1}{2}\tilde{A}(x, y, p, r)\tilde{A}^*(x, y, p, r)$.

As in the previous section along with (4.15) we consider an equation with respect to the function $v = \nabla u \in R^d$

$$v_t + \nabla\tilde{\Phi}(x, u, v, \nabla v) + \nabla_u \tilde{\Phi}(x, u, v, \nabla v)v + \nabla_v \tilde{\Phi}(x, u, v, \nabla v)\nabla v + \tag{4.16}$$

$$+\nabla_r \tilde{\Phi}(x, u, v, \nabla v)\nabla^2 v = 0, \quad v(0, x) = \nabla u_0(x) \in R^d,\ x \in R^d,\ t \in [0, T].$$

and rewrite (4.15) in the form

$$u_t + \frac{1}{2} Tr \tilde{A}(x, u, v, \nabla v)\nabla^2 u \tilde{A}^*(x, u, v, \nabla v) + \tilde{\Phi}(x, u, v, \nabla v) - \tag{4.17}$$

$$-\nabla_r \tilde{\Phi}(x, u, v, \nabla v)\nabla v = 0, \quad u(0, x) = u_0(x).$$

It is easy to see that the system (4.16), (4.17) is a system of quasilinear parabolic equations.

Set $A(x, y^1, y^2, z^2) = \tilde{A}(x, y, p, [\tilde{A}^*]^{-1} z^2)$, where $y^1 = y, y^2 = p, z^2 = \tilde{A}^* r$, and consider an SDE

$$d\xi(t) = A(\xi(t), y^1(t), y^2(t), z^2(t))dw(t), \quad \xi(s) = \xi_0. \tag{4.18}$$

Assume that $u(t, x)$ is a C^3 smooth function and (u, v) is a solution of (4.15) (4.16). Given a process $\xi(t)$ satisfying (4.18) we consider processes

$$y^1(t) = u(t, \xi(t)), y^2(t) = v(t, \xi(t))$$

and

$$z^1(t) = A^*(\xi(t), y^1(t), y^2(t), \nabla^2 u(t, \xi(t)))\nabla u(t, \xi(t)),$$
$$z^2(t) = A^*(\xi(t), y^1(t), y^2(t), \nabla^2 u(t, \xi(t))).$$

By the Ito formula we derive

$$dy^1(t) = [u_t(t, \xi(t)) + \mathcal{A}^u u(t, \xi(t))]dt + \langle z^1(t), dw(t)\rangle,$$
$$dy^2(t) = [v_t(t, \xi(t)) + \mathcal{A}^u v(t, \xi(t))]dt + z^2(t)dw(t),$$

where $\mathcal{A}^g u(t, x) = \frac{1}{2} Tr A(x, g, \nabla g, \nabla^2 g)\nabla^2 u(t, x) A^*(x, g, \nabla g, \nabla^2 g)$. Setting

$$f^1(x, y^1, y^2, z^2) = \tilde{\Phi}(x, y^1, y^2, [\tilde{A}^*]^{-1} z^2) - \nabla_r \tilde{\Phi}(x, y^1, y^2, [\tilde{A}^*]^{-1} z^2) y^2,$$
$$f^2(x, y^1, y^2, z^2) = \nabla\tilde{\Phi}(x, y^1, y^2, [\tilde{A}^*]^{-1} z^2) + \nabla_u \tilde{\Phi}(x, y^1, y^2, [\tilde{A}^*]^{-1} z^2) v +$$
$$+\nabla_v \tilde{\Phi}(x, y^1, y^2, [\tilde{A}^*]^{-1} z^2)[\tilde{A}^*]^{-1} z^2$$

we obtain a system of BSDEs

$$dy^1(t) = -f^1(\xi(t), y^1(t), y^2(t), z^2(t))dt + z^1 dw(t), \quad y^1(T) = u_0(\xi(T)) = \zeta^1, \tag{4.19}$$

$$dy^2(t) = -f^2(\xi(t), y^1(t), y^2(t), z^2(t))dt + z^2 dw(t), \quad y^2(T) = v_0(\xi(T)) = \zeta^2. \tag{4.20}$$

To introduce a notion of a solution to the FBSDE (4.18)– (4.20) let

$$\mathcal{B}^3 = S^2_T(R^d) \times S^2_T(R \times R^d) \times H^2_T(R^d \times R^d \otimes R^d).$$

We say that a triple $(\xi(t), y(t), z(t)) \in \mathcal{B}^3$ is a solution of the FBSDE (4.18) – (4.20) if the processes $\xi(t), y(t), z(t)$ are $\mathcal{F}_t$- measurable and with probability 1 we have

$$\xi(t) = x + \int_s^t A(\xi(\tau), y^1(\tau), y^2(\tau), z^2(\tau))dw(\tau), \tag{4.21}$$

$$y(t) = \zeta + \int_t^T f(\xi(\tau), y^1(\tau), y^2(\tau), z^2(\tau))d\tau - \int_t^T z(\tau)dW(\tau). \tag{4.22}$$

Here

$$y = (y^1, y^2) \in R \times R^d, f = (f^1, f^2) \in R \times R^d, z = (z^1, z^2) \in R^d \times (R^d \otimes R^d),$$
$$W(t) = (w(t), w(t)), \in R^d \times R^d, \zeta = (\zeta^1, \zeta^2) \in R \times R^d,$$
$$\zeta^1 = h(\xi(T)), \zeta^2 = \nabla h(\xi(T)).$$

Setting

$$\kappa = (x, y, z), \quad \alpha(x, y, z) = A(x, y^1, y^2, z^2), \quad \beta(x, y, z) = f(x, y^1, y^2, z^2)$$

we rewrite (4.21), (4.22) as a system

$$\xi(t) = x + \int_s^t \alpha(\xi(\tau), y(\tau), z(\tau))dw(\tau), \tag{4.23}$$

$$y(t) = \zeta + \int_t^T \beta(\xi(\tau), y(\tau), z(\tau))d\tau - \int_t^T z(\tau)dW(\tau) \tag{4.24}$$

and note that this system is not closed.

To make the system closed we have to apply the Ito martingale representation theorem which states the following.

Theorem 4.1. *Let $\chi \in R \times R^d$ be an $\mathcal{F}_T$- local square integrable martingale then it admits a representation of the form*

$$\chi = E\chi + \int_0^T z(\tau)dw(\tau)$$

and $z(\cdot) \in M = H_T^2(R^d \times R^d \otimes R^d)$ is unique.

Below we state conditions to ensure that we can apply the Ito theorem to the local martingale

$$\chi = E[\zeta + \int_0^T f(\xi(\tau), y^1(\tau), y^2(\tau), z^2(\tau))d\tau|\mathcal{F}_t],$$

where $\zeta \in L_T^2(R \times R^d)$ and to prove existence and uniqueness of a solution to the system (4.23), (4.24).

Let $\kappa = (x, y, z)^*$, $G \in (R \times R^d) \otimes R^d, \hat{\alpha}(\kappa) = (-G^*\beta, G\alpha)^*(\kappa)$, where $G\alpha = (G\alpha_1, \dots, G\alpha_d)$.

We say that condition **C 4** holds if

1) $\hat{\alpha}(\kappa)$ is C^1- smooth and sublinear in κ; $h(x), \nabla h(x)$ are bounded and Lipschitz continuous.

2) there exist positive constants μ_1, μ_2, μ_3 such that given a full rank matrix G

$$\langle \hat{\alpha}(\kappa) - \hat{\alpha}(\kappa_1), \kappa - \kappa_1 \rangle \le -\mu_1 \|G\bar{x}\|^2 - \mu_2(\|G^*\bar{y}\|^2 + \|G^*\bar{z}\|^2) \quad \forall \kappa, \kappa_1 \in \mathcal{B}_3,$$

$$\langle u_0(x) - u_0(x_1), G(x - x_1)\rangle \geq \mu_3 \|G\bar{x}\|^2$$

for any $\kappa = (x, y, z)$, $u_1 = (x_1, y_1, z_1), \bar{x} = x - x_1, \bar{y} = y - y_1, \bar{z} = z - z_1$, and $\mu_1 + \mu_2 > 0$ and $\mu_2 + \mu_3 > 0$.

The following assertion is a consequence of the results in [28].

Theorem 4.2. *Assume that* **C 4** *holds. Then there exists a unique solution* $(\xi(t), y(t), z(t)) \in \mathcal{B}^3$ *of the FBSDE* (4.23), (4.24).

We consider a system of successive approximations (the Picard iteration) to a solution $\xi(t), y(t) = (y^1(t), y^2(t)), z(t) = (z^1(t), z^2(t))$ of (4.23), (4.24).

$$\xi_0(t) = x, \quad y_0(t) = (h(x), \nabla h(x)) = g(x), \quad z_0(t) = 0,$$

$$\xi_1(t) = x + \int_s^t \alpha(x, h(x), 0) dw(\tau), \tag{4.25}$$

$$y_1(t) = g(\xi_1(T)) + \int_t^T \beta(\xi_1(\tau), y_1(\tau), z_1(\tau)) d\tau - \int_t^T z_1(\tau) dW(\tau), \tag{4.26}$$

$$\dots$$

$$\xi_{n+1}(t) = x + \int_s^t \alpha(\xi_n(\tau), y_n(\tau), z_n(\tau)) dw(\tau), \tag{4.27}$$

$$y_{n+1}(t) = h(\xi_n(T)) + \int_t^T \beta(\xi_n(\tau), y_n(\tau), z_n(\tau)) d\tau - \int_t^T z_{n+1}(\tau) dW(\tau). \tag{4.28}$$

One can prove that the solution of the decoupled FBSDE (4.27), (4.28) converges to a solution of (4.23), (4.24) as $n \to \infty$. In other words setting

$$\Delta\xi_n(t) = \|\xi_{n+1}(t) - \xi_n(t)\|^2, \quad \Delta y_n(t) = \|y_{n+1}(t) - y_n(t)\|^2,$$
$$\Delta z_n(t) = \|z_{n+1}(t) - z_n(t)\|^2,$$

one can verify that

$$\lim_{n\to\infty} E\left[\sup_{s\leq t\leq T} [\Delta\xi_n(t) + \Delta y_n(t)] + \int_s^T \Delta z_n(t) dt\right] = 0.$$

We refer to [19], [29] for the proof of this theorem.

Note that one can reduce an FBSDE to the following optimal control problem [24], [34], [37] :

to find $y(s) = y_0$ and $z(\tau), \tau \in [s, T]$ such that

$$\inf_{y_0, z(t), s\leq t\leq T} E[\|h(\xi^{y_0, z(\cdot)}(T)) - y^{y_0, z(\cdot)}(T)\|^2] = 0,$$

where

$$\xi^{y_0,z}(t) = x + \int_s^t \alpha(\xi^{y_0,z}(\tau), y^{y_0,z}(\tau), z(\tau)) dw(\tau),$$

$$y^{y_0,z}(s) = y_0 + \int_s^T \beta(\xi^{y_0,z}(\tau), y^{y_0,z}(\tau), z(\tau)) d\tau - \int_s^T z(\tau) dW(\tau(s),$$

where $y_0 \in R^{d+1}$ is $\mathcal{F}_0$-measurable and $z(t) \in R^d \times R^d \otimes R^d$ is an $\mathcal{F}_t$ adapted square integrable process.

To verify this we check that if $y(s) = y_0 \in R^{d+1}$ and $\{z(t)\}_{s\le t\le T} \in H^2_T(H)$, where $H = R^d \times R^d \otimes R^d$ satisfy

$$\xi(t) = x + \int_s^t \alpha(\xi(\tau), y(\tau), z(\tau))dw(\tau), \tag{4.29}$$

$$y(s) = g(\xi(T)) + \int_s^T \beta(\xi(\tau), y(\tau), z(\tau))d\tau - \int_s^T z(\tau)dW(\tau) \tag{4.30}$$

then obviously

$$E\left[\|g(\xi(T)) - y(T)\|^2\right] = 0.$$

On the other hand if we choose $y_0 = y(s)$ and $z(t)$ as control parameters of the optimal control problem

$$\inf_{\tilde{y}_0, \{\tilde{z}(t)\}_{s\le t\le T}} E\left[\|g(\tilde{\xi}(T)) - \tilde{y}(T)\|^2\right], \tag{4.31}$$

where

$$\tilde{\xi}(t) = x + \int_s^t \alpha(\tilde{\xi}(\tau), \tilde{y}(\tau), \tilde{z}(\tau))dw(\tau), \tag{4.32}$$

$$\tilde{y}(t) = g(\tilde{\xi}(T)) + \int_t^T \beta(\tilde{\xi}(\tau), \tilde{y}(\tau), \tilde{z}(\tau))d\tau - \int_t^T \tilde{z}(\tau)dW(\tau) \tag{4.33}$$

then we deduce that

$$\inf_{\tilde{y}_0, \{\tilde{z}(t)\}_{s\le t\le T}} E\left[\|g(\tilde{\xi}(T)) - \tilde{y}(T)\|^2\right] = 0.$$

As a result we conclude that one may reduce the system (4.23), (4.24) to variational problem (4.31)– (4.33).

5. Numerical algorithms for solution of nonlinear parabolic equations and systems

In this section we describe numerical algorithms to construct approximate solution of the Cauchy problem for nonlinear parabolic equations and systems based on probabilistic representations of these solutions.

5.1. Numerical SDE schemes for solution of semilinear equations and systems. We start with probabilistic representations of a classical solution to the Cauchy problem

$$\begin{gathered}\frac{\partial u_m}{\partial s} + \langle a(s,x,u), \nabla\rangle u_m + B^i_{ml}(s,x,u)\nabla_i u_l + c_{ml}(s,x,u)u_l + \\ + \frac{1}{2}TrA(s,x,u)\nabla^2 u_m A^*(s,x,u) = 0, \quad u_m(T,x) = u_{0m}(x), \\ m = 1,2,\ldots,d_1,\end{gathered} \tag{5.1}$$

constructed in section 2 and then derive an algorithm based on the representation of section 3.

Consider a probabilistic counterpart of the Cauchy problem (5.1)

$$d\xi = a^u(\tau,\xi(\tau))d\tau + A^u(\tau,\xi(\tau))dw(\tau), \quad \xi(s) = x \in R^d, \tag{5.2}$$

$$d\eta(t) = c^u(\tau,\xi(\tau))\eta(\tau)d\tau + C^u(\tau,\xi(\tau))(\eta(\tau), dw(\tau)), \quad \eta(s) = h \in R^{d_1}, \tag{5.3}$$

$$\langle h, u(s,x)\rangle = E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle]. \tag{5.4}$$

Our aim is to construct a numerical solution of (5.2)–(5.4) .

Let $s = t_0 < t_1 \cdots < t_N = T$, with $t_k = kh, \Delta t = \frac{T-s}{N}$ be a given partition, $\Delta t = t_{k+1} - t_k$ and $\Delta_k w = w(t_{k+1}) - w(t_k)$ for $k = 0, \ldots, N-1$.

To construct a numerical scheme based on the stochastic system (5.2)–(5.4) we present the relation (5.4) in the form

$$\langle h, u(s,x)\rangle = E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle] = \langle h, E[S^*(s,T)u_0(\xi_{s,x}(T))\rangle],$$

where $S(T,s)h = \eta_{s,h}(T)$. From the Markov property of the stochastic process $\xi_{s,x}(t)$ and uniqueness of the solution $\eta_{s,h}(t)$ to (5.3) we deduce that

$$\langle h, E[S^*(s,T)u_0(\xi_{s,x}(T))]\rangle = \langle h, U(s,T)u_0(x)\rangle, \tag{5.5}$$

where $U(s,T)$ is an evolution family acting in the space $C_b(R^d; R^{d_1})$ of continuous bounded functions valued in R^{d_1}. Due to the evolution property of $U(s,T)$ we have an equality $U(s,T) = \prod_{k=0}^{N-1} U(t_k, t_{k+1})$.

To construct an approximation of $U(s,T)$ we apply the Euler scheme to approximate the processes $\xi_{s,x}(t)$ and $\eta_{s,h}(t)$.

$$\xi_{t_k,x}(t_{k+1}) \sim \bar{\xi}_{t_k,x}(t_{k+1}) = x + a(x, u(t_k,x))\Delta t + A(x, u(t_k,x))\Delta_k w, \tag{5.6}$$

$$\eta_{t_k,h}(t_{k+1}) \sim \bar{\eta}_{t_k,h}(t_{k+1}) = h + c(x, u(t_k,x))h\Delta t + C(x, u(t_k,x))(h, \Delta_k w), \tag{5.7}$$

and define the function $\bar{u}(t,x)$ by a relation

$$\langle h, \bar{u}(t,x)\rangle = \langle h, \bar{U}(t_k, t_{k+1})\bar{u}(t_{k+1},x)\rangle = E[\langle \bar{\eta}_{t_k,h}(t_{k+1}), \bar{u}(t_{k+1}, \bar{\xi}_{t_k,x}(t_{k+1}))\rangle]. \tag{5.8}$$

In addition we approximate $\Delta_k w$ by $\epsilon\sqrt{\Delta t}$, where ϵ is the Bernoulli random variable valued in $\{-1, 1\}$ with a distribution $P\{\epsilon = -1\} = P\{\epsilon = 1\} = \frac{1}{2}$ and denote by $\hat{U}(s,t)$ the map defined by (5.4) with $\hat{\xi}_{t_k,x}(t_{k+1})$ and $\hat{\eta}_{t_k,h}(t_{k+1})$ instead of $\bar{\xi}_{t_k,x}(t_{k+1})$ and $\bar{\eta}_{t_k,h}(t_{k+1})$, where $\hat{\xi}_{t_k,x}(t_{k+1})$ and $\hat{\eta}_{t_k,h}(t_{k+1})$ are defined by

$$\hat{\xi}_{t_k,x}(t_{k+1}) = x + a(x, \hat{u}(t_k,x))\Delta t + A(x, \hat{u}(t_k,x))\epsilon\sqrt{\Delta t}, \tag{5.9}$$

$$\hat{\eta}_{t_k,h}(t_{k+1}) = h + c(x, \hat{u}(t_k,x))h\Delta t + C(x, \hat{u}(t_k,x))(h, \epsilon\sqrt{\Delta t}), \tag{5.10}$$

$$\langle h, \hat{u}(t_k,x)\rangle = E\langle \hat{\eta}_{t_k,h}(t_{k+1}), \hat{u}(t_{k+1}, \hat{\xi}_{t_k,x}(t_{k+1}))\rangle. \tag{5.11}$$

From the Markov property of $\xi_{s,x}(t)$ and properties of a solution to the linear SDE (5.3) we can deduce the following assertion.

Lemma 5.1. *Assume that coefficients in* (5.2), (5.3) *and the function $u_0(x)$ satisfy* **C 2.4**. *Then the map $U(s,T)$ defined by*

$$\begin{aligned}\langle h, U(s,T)u_0(x)\rangle &= E[\langle \eta_{s,h}(T), u_0(\xi_{s,x}(T))\rangle] = \\ &= \langle h, E[S^*(s,T)u_0(\xi_{s,x}(T))]\rangle\end{aligned} \tag{5.12}$$

is an evolution family.

Note that unlike $U(s,t)$ both $\bar{U}(s,t)$ and $\hat{U}(s,t)$ cease to be evolution families that is $\bar{U}(s,t) \neq \bar{U}(s,\theta)\bar{U}(\theta,t)$ for $s < \theta < t$.

Given $\bar{\eta}(t_{k+1}) = \bar{S}(t_k, t_{k+1})h, \hat{\eta}(t_{k+1}) = \hat{S}(t_k, t_{k+1})h$ we have $\langle \bar{S}h, v\rangle = \langle h, \bar{S}^* v\rangle$ and

$$\begin{aligned}\langle h, [\bar{U}(t_k, t_{k+1})\bar{v}](t_k,x)\rangle &= E\langle \bar{\eta}(t_{k+1}), \bar{v}(t_{k+1}, \bar{\xi}_{t_k,x}(t_{k+1}))\rangle = \\ &= \langle h, E[\bar{S}^*(t_k, t_{k+1})\bar{v}(t_{k+1}\bar{\xi}_{t_k,x}(t_{k+1}))\rangle\end{aligned} \tag{5.13}$$

where $\bar{v}(t_k, x) = \bar{U}(t_k, t_{k+1})\bar{v}(t_{k+1}, x)$ and similar relations for $\hat{U}(t_k, t_{k+1}), \hat{v}(t_k, x)$.

Let

$$\hat{U}_n(s, T) = \prod_{k=0}^{n-1} \hat{U}(t_k, t_{k+1}). \tag{5.14}$$

Keeping in mind (5.9) – (5.11) we obtain

$$[\hat{U}(t_k, t_{k+1})v]^m(t_k, x) = \sum_{l=1}^{d_1} \hat{U}(t_k, t_{k+1})^{ml} v^l(t_{k+1}, x) =$$

$$\begin{aligned}
&= \frac{1}{2}\sum_{l=1}^{d_1}[\delta^{ml} + c^{ml}(x, v(t_{k+1}, x))\Delta t + C^{ml}(x, v(t_{k+1}, x))\sqrt{\Delta t}]\times \\
&\quad \times v^l(t_{k+1}, x + a(x, \bar{v}(t_{k+1}, x)))\Delta t + A(x, v(t_{k+1}, x))\sqrt{\Delta t} + \\
&+ \frac{1}{2}\sum_{l=1}^{d_1}[[\delta^{ml} + c^{ml}(x, \bar{u}(t_{k+1}, x))\Delta t - C^{ml}(x, v(t_{k+1}, x))\sqrt{\Delta t}]\times \\
&\quad \times v^l(t_{k+1}, x + a(x, v(t_{k+1}, x))\Delta t - A(x, v(t_{k+1}, x))\sqrt{\Delta t})],
\end{aligned} \tag{5.15}$$

where $\delta^{lm} = 1$, if $l = m$ and $\delta^{lm} = 0$, if $l \neq m$.

Finally, we apply $\hat{U}_n(s, T) = \prod_{k=0}^{n-1} \hat{U}(t_k, t_{k+1})$ to approximate $U_n(s, T)u_0 = \prod_{k=0}^{n-1} U(t_k, t_{k+1})$ in spite of the lack of evolution property of the map $\hat{U}(s, T)$. Based on Marsden results [33] we prove that $lim_{n\to\infty}\hat{U}_n(s, T)u_0 = U(s, T)u_0$.

Theorem 5.2. *Let $V_1(s, T)$ and $V_2(s, T)$ be bounded maps acting in $C([0, T] \times R^d)$ such that*

$$\|V_i(s, t)u\|_\infty \le K_i\|u\|_\infty, ||V_i(s, t)u - V_i(s, t)v\|_\infty \le e^{\gamma_i}t\|u - v\|_\infty,\ i = 1, 2.$$

Let $V_i(s, T)$ are differentiable in s and

$$\|V_1(t_k, t_{k+1})u - V_2(t_k, t_{k+1})u\|_\infty \le \gamma(\Delta)^{1+\alpha}\|v\|_\infty$$

for some $\alpha > 0$. Assume that there exists a limit

$$\lim_{n\to\infty} \prod_{k=0}^{n-1} V_1(t_k, t_{k+1})v = V_1(s, T)v,$$

then there exists a limit

$$\lim_{n\to\infty} \prod_{k=0}^{n-1} V_2(t_k, t_{k+1})v = V_2(s, T)v.$$

We apply the results of this assertion to the above maps $U(s, T)$ given by (5.12) and $\hat{U}(s, T)$ given by (5.15) .

Lemma 5.3. *Let the processes $\xi(t)$, $\eta(t)$ and the function $u(s, x)$ be defined by (5.7)– (5.9). Then maps*

$$U(s, T)u_0(x) = \mathbf{E}_{s,x}S^*(s, T)u_0(\xi(T))],$$

$$\bar{U}(t_k, t_{k+1})v(x) = \mathbf{E}_{t_k,x}[\bar{S}^*(t_k, t_{k+1})v(\bar{\xi}(t_{k+1}))],$$

$$\hat{U}(t_k, t_{k+1})v(x) = \mathbf{E}_{t_k,x}[\hat{S}^*(t_k, t_{k+1})v(\hat{\xi}(t_{k+1}))]$$

satisfy conditions of the previous assertion and thus there exists limits

$$\lim_{n\to\infty}\prod_{k=0}^{n-1}\bar U(t_k,t_{k+1})u_0 = U(s,T)u_0, \quad \lim_{n\to\infty}\prod_{k=0}^{n-1}\hat U(t_k,t_{k+1})u_0 = U(s,T)u_0.$$

Proof. Let $V_1(t_k,t_{k+1}) = U(t_k,t_{k+1})$, $V_2(t_k,t_{k+1}) = \bar U(t_k,t_{k+1})$, $V_3(t_k,t_{k+1}) = \hat U(t_k,t_{k+1})$. From the results of section 2 we deduce

$$\|V_q(t_k,t_{k+1})u(t_{k+1})\|_\infty \le \|u(t_{k+1})\|_\infty exp\{K\Delta\} + \beta\Delta, \quad q = 1,2,3$$

and

$$\|V_1(t_k,t_{k+1})v(t_{k+1}) - V_2(t_k,t_{k+1})v(t_{k+1})\| \le C(\Delta t)^{1+\alpha}\|v(t_{k+1})\|_\infty,$$

$$\|V_2(t_k,t_{k+1})v(t_{k+1}) - V_3(t_k,t_{k+1})v(t_{k+1})\| \le C(\Delta t)^{1+\alpha}\|v(t_{k+1})\|_\infty,$$

where γ, C, α are positive constants. It results from theorem 5.1 that

$$U(s,T)u_0(x) = \lim_{n\to\infty}\prod_{k=0}^{n-1}\bar U(t_k,t_{k+1})u_0(x).$$

By the triangle inequality we have

$$\sup_x\|[\bar U(s,T) - \hat U(s,T)]u_0(x)\| \le \sup_x\|[U(s,T) - \bar U(s,T)]u_0(x)\| +$$

$$+\sup_x\|[\bar U(s,T) - \hat U(s,T)]u_0(x)\|,$$

and in addition we deduce from proposition 5.2 that

$$\lim_{n\to\infty}\prod_{k=0}^{n-1}\bar U(t_k,t_{k+1})u_0(x) = \lim_{n\to\infty}\prod_{k=0}^{n-1}\hat U(t_k,t_{k+1})u_0(x).$$

This yields

$$U(s,T)u_0(x) = \lim_{n\to\infty}\prod_{k=0}^{n-1}\hat U(t_k,t_{k+1})u_0(x).$$

□

Remark 5.4. Under **C 2.4** the assertions of lemma 5.3 are valid as well in the case when we consider the Caauchy problem for a nonuniform PDE of the form

$$\frac{\partial u_m}{\partial s} + \langle a(s,x,u),\nabla\rangle u_m + B^i_{ml}(s,x,u)\nabla_i u_l + c_{ml}(s,x,u)u_l +$$

$$+\frac{1}{2}TrA(s,x,u)\nabla^2 u_m A^*(s,x,u) = g_m(x,u), \quad u_m(T,x) = u_{0m}(x),$$

where $g(x,u)$ is a smooth bounded function. In this case we have

$$U(s,T)u_0(x) = E[S^*(s,T)u_0(\xi_{s,x}(T))] + E\left[\int_s^T S^*(s,\theta)g(\xi_{s,x}(\theta),u(\theta,\xi_{s,x}(\theta)))d\theta\right]$$

and

$$\bar U(t_k,t_{k+1})u(t_{k+1},x) = E\left[\bar S^*(t_k,t_{k+1})u(t_{k+1},\bar\xi_{t_k,x}(t_{k+1}))\right] +$$

$$+E\left[\int_{t_k}^{t_{k+1}}\bar S^*(t_k,\theta)g(\bar\xi_{t_k,x}(\theta),\bar u(\theta,\bar\xi_{t_k,x}(\theta)))d\theta\right],$$

$$\hat{U}(t_k, t_{k+1})\hat{u}(t_{k+1}, x) = E\left[\hat{S}^*(t_k, t_{k+1})\hat{u}(t_{k+1}, \hat{\xi}_{t_k,x}(t_{k+1}))\right] +$$
$$+E\left[\int_{t_k}^{t_{k+1}} \hat{S}^*(t_k, \theta) g(\hat{\xi}_{t_k,x}(\theta), \hat{u}(\theta, \hat{\xi}_{t_k,x}(\theta)))d\theta\right]$$

As a result we get

$$\hat{u}(t_k, x) = \frac{1}{2}[(I + c(x, \hat{u}(t_{k+1}, x))h + C(x,\ \hat{u}(t_{k+1}, x))h^{\frac{1}{2}})\times \tag{5.16}$$
$$\times\hat{u}(t_{k+1}, x + a(x, \hat{u}(t_{k+1}, x))h + A(x, \hat{u}(t_{k+1}, x))h^{\frac{1}{2}})+$$
$$+\frac{1}{2}[(I + c(x, \hat{u}(t_{k+1}, x))h - C(x, \hat{u}(t_{k+1}, x))h^{\frac{1}{2}})\times$$
$$\times\hat{u}(t_{k+1}, x + a(x, \hat{u}(t_{k+1}, x))h - A(x, \hat{u}(t_{k+1}, x))h^{\frac{1}{2}})+$$
$$+g(x, \hat{u}(t_{k+1}, x))h.$$

As a final step to obtain a numerical scheme in the case $d = 1$ we discretize the space variable x, compute $u(t_k, x_j)$ and use linear interpolation to define $u(t_k, x)$ for $x_j < x < x_{j+1}$ (see details in [30] – [32]).

To apply the above approach to a quasilinear or a fully nonlinear parabolic equation we include it to a system of semilinear parabolic equations as it was done in section 3 and then proceed as above.

5.2. Numerical FBSDE schemes for solution of the Cauchy problem for fully nonlinear equations and systems. Here we apply the results of section 4 to construct a numerical solution of a nonlinear parabolic equation based on the so called deep BSDE theory developed in recently in a number of papers [24]–[26], [34] – [37]. The deep BSDE theory combines probabilistic representations of a solution to the Cauchy problem for a nonlinear parabolic equation with the neural network theory. In section 4 we have shown that the fully nonlinear parabolic equation of the form (4.15) can be reduced to the system of quasilinear parabolic equations (4.16), (4.17). Next it was shown that to solve the resulting quasilinear system one can consider the correspondent FBSDE of the form (4.18) – (4.20). Slightly changing our construction we consider the Cauchy problem

$$u_t + \Phi(x, u, \nabla u, \nabla^2 u) = 0, \quad u(0, x) = u_0(x) \in R,\ x \in R^d,\ t \in [0, T]. \tag{5.17}$$

and reduce it to a suitable FBSDE

$$d\xi(t) = A(\xi(t), y^1(t), y^2(t), z^2(t))dw(t), \quad \xi(s) = x, \tag{5.18}$$

$$dy^j(t) = -f^j(\xi(t), y^1(t), y^2(t), z^2(t))dt+ \tag{5.19}$$
$$+z^j(t)dw(t), \quad y^j(T) = g^j(\xi(T)),$$

where coefficients A and $f = (f^1, f^2)$ coincide with coefficients in (4.19), (4.20). FBSDE (5.18), (5.19) is associated with (5.17) provided

$$y^1(t) = u(t, \xi(t)), y^2(t) = v(t, \xi(t)) = \nabla u(t, \xi(t)), \quad z^1(t) = \nabla u(t, \xi(t)),$$
$$z^2(t) = \nabla v(t, \xi(t)), \quad g^1(x) = h(x), g^2(x) = \nabla h(x).$$

Next we discuss numerical schemes which allow to obtain an approximate solution of the FBSDE (5.18), (5.19) based on the deep FBSDE theory.

To construct numerically a viscosity solution u of (5.17) presented in the form $u(s,x) = y^1(s)$, where $y^1(t) = u(t,\xi(t))$, $y^2(t) = \nabla u(t,\xi(t))$, and $\xi(t), y^j(t), j = 1,2$, satisfy (5.18), (5.19).

The main idea here can be stated as follows. As we have seen in section 4 one can reduce solution of an FBSDE to solution of the following control problem :

to find

$$\inf_{y_0, z(\cdot)} E\left[\|g(\xi^{y_0,z(\cdot)}(T)) - y^{y_0,z(\cdot)}(T)\|^2\right] \tag{5.20}$$

such that

$$\xi^{y_0,z(\cdot)}(t) = x + \int_s^t A(\xi^{y_0,z(\cdot)}(\tau), y^{y_0,z(\cdot)}(\tau), z(\tau))dw(\tau), \tag{5.21}$$

$$y^{y_0,z(\cdot)}(t) = y_0 - \int_s^t f(\xi^{y_0,z(\cdot)}(\tau), y^{y_0,z(\cdot)}(\tau), z^{y_0,z(\cdot)}(\tau))d\tau + \int_s^t z(\tau)dW(\tau), \tag{5.22}$$

where $y_0 = y(s)$- $\mathcal{F}_0$ adapted random variable valued in $R \times R^d$ and $z(t)$ is $\mathcal{F}_t$-adapted matrix-valued square integrable stochastic process. The couple $(y_0, z(\cdot))$ is a control variable of the considered control problem.

Within this framework

$$\inf_{y_0, z(\cdot)} E\left[\|g(\xi^{y_0,z(\cdot)}(T)) - y^{y_0,z(\cdot)}(T)\|^2\right] = 0$$

and infimum is achieved when $\xi^{y_0,z(\cdot)}(t), y^{y_0,z(\cdot)}(t), z(t))$ satisfy (5.21), (5.22).

To solve the control problem of the form (5.20)-(5.22) it comes to be very effective to apply the neural network technique. We discuss it while considering a more advanced control problem below.

Let us consider one more scheme suggested in recent paper [37].

Changing the loss function we choose the whole process $y(\cdot)$ as a control together with $z(\cdot)$.

Then the control problem has the form :

to find

$$\inf_{u(\cdot), Z(\cdot)} E\left[\|g(\xi^{u,Z}(T)) - y^{u,Z}(T)\|^2 + \int_s^T \|y^{u,Z}(t) - u(t)\|^2 dt\right], \tag{5.23}$$

where

$$\xi^{u,Z}(t) = x + \int_s^t A(\xi^{u,Z}(\tau), y^{u,Z}(\tau), Z(\tau))dw(\tau), \tag{5.24}$$

$$y^{u,Z}(t) = y_0 + \int_s^t f(\xi^{u,Z}(\tau), y^{u,Z}(\tau), z^{y_0,z(\cdot)}(\tau))d\tau - \int_s^t Z(\tau)dW(\tau) \tag{5.25}$$

and solution

$$\inf_{u(\cdot), Z(\cdot)} E\left[\|g(\xi^{u,Z}(T)) - y^{u,Z}(T)\|^2 + \int_s^T \|y^{u,Z}(t) - u(t)\|^2 dt\right] = 0,$$

is achieved when $\xi^{u,Z}(t), y^{u,Z}(t), Z(t))$ satisfy (5.24), (5.25).

To solve this optimal problem effectively one can apply the neural network technique. Let us recall some notion and results from the neural network theory.

For an integer $n \in N$ consider a partition $s = t_0 < t_1 < \cdots < t_N = T$ of the interval $[s,T]$ and define neural networks $\mathcal{S}_k^{j,\beta}(\cdot) j = 1,2$

$$\mathcal{S}_k^{1,\beta} = \Gamma_{d,d}^{\beta,[(2K+k)d+1](d+1)} \circ \psi_d \circ \Gamma_{d,d}^{\beta,[(K+k)d+1](d+1)} \circ \psi_d \circ \Gamma_{d,d}^{\beta,(kd+1)(d+1)}, \quad (5.26)$$

and

$$\mathcal{S}_k^{2,\beta} = \Gamma_{d^2,d}^{\beta,[(5Kd+kd^2+1)](d+1)} \circ \psi_d \circ \Gamma_{d,d}^{\beta,[(4K+k)d+1](d+1)} \circ \psi_d \circ \Gamma_{d,d}^{\beta,[(3K+k)d+1](d+1)}. \quad (5.27)$$

to approximate functions $\nabla u(t_{k+1},\cdot), \nabla v(t_{k+1},\cdot)$. Here each network has one input layer with dimension d, two hidden layers with dimensions d and one output layer where dimensions of output layers are d and d^2 respectively.

For an activation function we choose

$$\psi_d(x) = (\max(x_1,0),\ldots,\max(x_d,0)),\ x \in R^d,$$

and affine transformations $\Gamma_{q,l}^{\beta,\alpha} : R^l \to R^q$ in (5.19), (5.20) are chosen to have a form

$$\Gamma_{q,l}^{\beta,\alpha}(x) = \begin{pmatrix} \beta_{\alpha+1} & \beta_{\alpha+2} & \cdots & \beta_{\alpha+l} \\ \beta_{\alpha+l+1} & \beta_{\alpha+l+2} & \cdots & \beta_{\alpha+2l} \\ \beta_{\alpha+2l+1} & \beta_{\alpha+2l+2} & \cdots & \beta_{\alpha+3l} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{\alpha+(q-1)l+1} & \beta_{\alpha+(q-1)l+2} & \cdots & \beta_{\alpha+ql} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_l \end{pmatrix} + \begin{pmatrix} \beta_{\alpha+ql+1} \\ \beta_{\alpha+ql+m} \\ \beta_{\alpha+ql+3} \\ \vdots \\ \beta_{\alpha+ql+q} \end{pmatrix}, \quad (5.28)$$

where $\alpha = d+1$.

Note that the condition $\rho \geq (5Kd + Nd^2 + 1)(d+1)$ ensures that $\Gamma_{q,l}^{\beta,\alpha} : R^l \to R^q$ acts correctly. Finally, we arrive to an optimization problem which could be approximately solved by applying stochastic descent gradient (SGD) method.

To construct the required solution for (5.23) – (5.25) we apply the Euler-Maruyama scheme to discretize (5.23), (5.24) setting $\bar\xi(t_0) = x, \bar y(t_0) = y_0$,

$$\bar\xi(t_{k+1}) = \bar\xi(t_k) + A(\bar\xi(t_k), \bar u(t_k), \bar v(t_k), \bar z^2(t_k))\Delta_k w, \quad (5.29)$$

$$\bar y^j(t_{k+1}) = \bar y^j(t_k) - f^j(\bar\xi(t_k), \bar y^1(t_k), \bar y^2(t_k), \bar z^2(t_k))\Delta t + \quad (5.30)$$

$$+\bar z^j(t_k)\Delta_k w, \quad j = 1,2.$$

Here we wrote the integral form of the backward SDE in the form similar to the integral form of a forward SDE.

We present $\bar\xi(t_0) = \xi_0, \bar y(t_0) = (u_0, v_0)$,

$$\bar u(t_k) = \phi_1(\bar\xi(t_k), \bar u(t_k); \beta_k^1), \quad \bar v(t_k) = \phi_2(\bar\xi(t_k), \bar u(t_k), \bar v(t_k); \beta_k^2),$$

$$\bar z^1(t_k) = \phi_3(\bar\xi(t_k), \bar u(t_k); \beta_k^3), \quad \bar z^2(t_k) = \phi_4(\bar\xi(t_k), \bar u(t_k), \bar v(t_k); \beta_k^4).$$

Note that solving the last equations we can obtain

$$\bar u(t_k) = \tilde\phi_1(\bar\xi(t_k; \beta_k^1), \quad \bar v(t_k) = \tilde\phi_2(\bar\xi(t_k; \beta_k^2),$$

$$\bar z^1(t_k) = \tilde\phi_3(\bar\xi(t_k; \beta_k^3), \quad \bar z^2(t_k) = \tilde\phi_4(\bar\xi(t_k; \beta_k^4).$$

Hence, we have to construct four neural networks at the sane time. All parameters of these networks are represented as β. The loss function has the form

$$\mathcal{L}(\beta) = \inf_{\beta} E\left[\|g(\bar{\xi}^{\beta}(T)) - \bar{y}^{\beta}(T)\|^2 + \sum_{k=0}^{N-1} \|\bar{y}^{\beta}(t_k) - \bar{u}(t_k)\|^2 \right]$$

where $\beta = (y_0, z(\cdot))$.

Similar to [37] we may write the correspondent algorithm as follows.

Choosing as an input the Wiener process increments $\Delta_k w$, initial parameters β^0 and learning rate ϱ we have to obtain as an output the couple $(\bar{\xi}^q(T), \bar{y}^q(t_k)), k = 1\dots, N$. On each interval $[t_k, t_{k+1}]$ we solve the optimization problem applying SGD method (stochastic decent gradient) with q iterations, $q = 1, 2, \dots$.

1. For $q = 1, \dots,$ set $L = 0, \bar{\xi}^q(t_0) = x$, $\bar{y}^{j,q}(t_0) = \phi^j(x; \beta_0^{q-1})$;
2. For $k = 0, \dots, N-1$ set

$$\bar{u}^q(t_k) = \phi^1(\bar{\xi}^q(t_k), \beta_k^{1,q-1}), \quad \bar{v}^q(t_k) = \phi^2(\bar{\xi}^q(t_k), \beta_k^{2,q-1}),$$
$$\bar{z}^{1,q}(t_k) = \phi^3(\bar{\xi}^q(t_k), \beta_k^{3,q-1}),\ \bar{z}^{2,q}(t_k) = \phi^4(\bar{\xi}^q(t_k), \beta_k^{4,q-1}).$$

3. By Euler -Maruyama schemes (5.22), (5.23) we calculate $\xi^q(t_{k+1}), y^q(t_{k+1})$ and $z^q(t_{k+1})$ on each time interval $[t_k, t_{k+1}]$

$$\bar{\xi}^q(t_{k+1}) = \bar{\xi}^q(t_k) + A(\bar{\xi}^q(t_k), \bar{u}^q(t_k), \bar{v}^q(t_k), \bar{z}^{2q}(t_k))\Delta_k w,$$
$$y^{j,q}(t_{k+1}) = y^{j,q}(t_k) + f^j(\bar{\xi}^q(t_k), \bar{y}^{1,q}(t_k), \bar{y}^{2,q}(t_k), z^{2,q}(t_k))\Delta t -$$
$$- z^{j,q}(t_k) A(\bar{\xi}^q(t_k), \bar{y}^{1,q}(t_k), \bar{y}^{2,q}(t_k), \bar{z}^{2q}(t_k))\Delta_k w, \quad j = 1, 2,$$

4.
$$(\beta^{q+1}, \bar{y}_0^{q+1}) = (\beta^q, \bar{y}_0^q) - \varrho\nabla \frac{1}{M}\sum_{m=1}^{M} \|\bar{y}^q(T) - g(\bar{\xi}^q(T))\|^2.$$

References

1. Kolmogorov A. N. : Uber die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math.Ann.* **104** (1931) 415–458.
2. McKean H. :A class of Markov processes associated with nonlinear parabolic equations, *Proc. Nat. Acad. Sci. USA* **59** 6, (1966) 1907–1911.
3. Freidlin M.: Quasilinear parabolic equations and measures in function space, *Functional Analysis and Its Applications* **1** (1967) 234–240.
4. Freidlin M.: *Functional integration and partial differential equations*, Annals of Mathematics Studies, vol. 109, Princeton University Press, Princeton, 1985.
5. Belopolskaya Ya.: Probabilistic interpretations of quasilinear parabolic systems, in: *AMS Contemporary Mathematics CONM* **734** (2019) 39–56.
6. Belopolskaya Ya.,Dalecky Yu.: Investigation of the Cauchy problem for systems of quasilinear equations via Markov processes, *Izv. VUZ Matematika.* N 12, (1978) 6–17.
7. Belopolskaya Ya., Dalecky Yu. : *Stochastic Equations and Differential Geometry*, Kluwer, Dordrecht, 1990.
8. Pommaret J.F.: *Systems of partial differential equations and Lie pseudogroups*, Gordon and Breach Sci. Publ. N.Y., 1978.
9. Pardoux E., Peng S.: Adapted solution of a backward stochastic differential equation, *Systems and Control Letters*, **14** 1, (1990) 55–61.
10. Pardoux E., Peng S.: Backward Stochastic Differential equations and Quasilinear Parabolic Partial Differential Equations, *Lecture Notes in CIS* **176** Springer (1992) 200–217.

11. Pardoux E., Rascanu A.: *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, Springer 2014.
12. Carmona R.: *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications* SIAM 2016.
13. Crandall M., Ishii H., Lions P-L.: User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (N.S.)* **27** (1992) 1– 67.
14. Pardoux E.: Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, in: *Stochastic Analysis and Relates Topics: The Geilo Workshop* Birkhäser (1996) 79–127.
15. Cheridito P., Soner M., Touzi N., Victoir N.: Second- order backward stochastic differential equations and fully nonlinear parabolic PDEs, *Communications on Pure and Applied Mathematics.* **60** 7 (2007) 1081–1110.
16. Belopolskaya Ya, Woyczinski W.: SDEs, FBSDEs and fully nonlinear parabolic systems, *Rendiconti del Seminario Matematico Torino* **71** 2 (2013) 209 – 219.
17. Belopolskaya Ya.: Probabilistic counterparts of nonlinear parabolic PDE systems *Modern Stochastics and Applications - Springer Optimization and Its Applications* **90** (2014) 71–94.
18. Yin G., Zhu C.: *Hybrid switching diffusion*, Springer 2010.
19. Belopolskaya Ya.: Probabilistic Models of the Conservation and Balance Laws in Switching Regimes, *J. Math. Sci.*, **229** 6 (2018) 601–625.
20. Belopolskaya Ya.: Probabilistic Interpretation of the Cauchy Problem Solution for Systems of Nonlinear Parabolic Equations, *Lobachevskii J. Math,* **41** 4 (2020) 597–612.
21. Belopolskaya Ya.: Stochastic interpretation of quasilinear parabolic systems with cross diffusion, *Theory of Probability and its Applications* **61** 2 (2017), 208–234.
22. Belopolskaya Ya.: Probabilistic interpretations of quasilinear parabolic systems. Differential Equations, Mathematical Physics, and Applications*AMS Contemporary Mathematics, CONM* **734** (2019) 39–56.
23. Carmona R, Delarue F. : *Probabilistic Theory of Mean Field Games with Applications I, II.* Springer 2018.
24. E W., Han J., Jentzen A.: Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, **5** 4 (2017) 349–380.
25. Hutzenthaler M, Jentzen A., Kruse T.: Overcoming the curse of dimensionality in the numerical approximation of parabolic partial differential equations with gradient-dependent nonlinearities, *Foundations of Computational Mathematics* **22** (2022) 905–966.
26. Raissi M.: Forward-backward stochastic neural networks: deep learning of high- dimensional partial differential equations, *arXiv preprint, arXiv:1804.07010*, Apr 2018.
27. Skorokhod A.V.: *Studies in the Theory of Random Processes,* Addison Wesley 1965.
28. Peng S., Wu Zh. : Fully coupled forward-backward stochastic differential equations and applications to optimal control. *Siam J. Contr. Opt.,* **37** 3. (1999) 825–843.
29. Abraham R., Riviere O.: Forward-backward stochastic differential equations and PDE with gradient dependent second order coefficients, *ESAIM: Probability and Statistics***10** (2006) 184–205.
30. Milstein G., Tretyakov M. : *Stochastic Numerics for Mathematical Physics*, Springer, Berlin, 2004.
31. Belopolskaya Ya., Nemchenko E.: Probabilistic Representations and Numerical Algorithms for Classical and Viscosity Solutions of the Cauchy Problem for Quasilinear Parabolic Systems, *J. Math. Sci.* **225** 5 (2017) 733–750.
32. Belopolskaya Ya., E. Nemchenko E.: Probabilistic Models and Numerical Algorithms for Quasilinear Parabolic Equations and Systems, *Comm. Stoch. Anal.* **14**1-2 (2020) 1–12.
33. Marsden J.: On product formulas fo nonlinear semigroups. *J.Funct. Anal.* **13** (1973) 51 – 72.
34. Han J., Long J., Convergence of the deep BSDE method for coupled FBSDEs. *arXiv:1811.01165* (2018).

35. Raissi M., Perdikaris P. Karniadakis G.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comp. Physics* **378** 1 (2019) 686–707.
36. Zhang W., Cai W.: FBSDE based neural network algorithms for high-dimensional quasilinear parabolic PDEs, *Journal of Computational Physics*, **470** C (2022) 111557.
37. Ju S.,Peng S., Peng Y., Zhang X.: Three Algorithms for Solving High-Dimensional Fully Coupled FBSDEs Through Deep Learning *IEEE Intelligent Systems.* **35** (2020) 71–84.
38. Baydin A. G., Pearlmutter B. A., Radul A. A., Siskind J. M.,: Automatic differentiation in machine learning: a survey, *arXiv preprint arXiv:1502.05767* (2015).

Yana Belopolskaya: Sirius University of Science and Technology, Sochi, 354340, PDMI RAS, St. Petersburg, 191023, Russia

Email address: yana.belopolskaya@gmail.com