ROUGH ASYMPTOTICS OF FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS*

Jin Ma

Department of Mathematics
Purdue University
West Lafayette, IN 47907, USA
majin@math.purdue.edu

Tim Zajic

Department of Mathematics
University of Minnesota
Minneapolis, MN 55455, USA
zajic@math.umn.edu

Abstract: This note is concerned with the asymptotic behavior of adapted solutions to forward-backward SDEs when the forward diffusion contains small noises. We prove the sample path Large Deviation Principle (LDP) for the adapted solutions to the FBSDEs under appropriate conditions stated in terms of a certain type of convergence of the solutions to the associated quasilinear PDEs. As an application we apply the LDP results to a problem of rare event simulation. We provide a necessary condition for the existence of an asymptotically efficient estimator amongst the class of importance sampling estimators.

1 INTRODUCTION

Let $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{t \geq 0})$ be a complete filtered probability space on which is defined a Brownian motion W such that $\{\mathcal{F}_t\}$ is the augmented natural filtration generated by W . We consider the following *forward-backward stochastic differ-*

*The first author is supported in part by the Office of Naval Research grant #N00014-96-1-0262; and the second author is supported in part by NSF Grant DMS-9508709.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35359-3_40](https://doi.org/10.1007/978-0-387-35359-3_40)

S. Chen et al. (eds.), *Control of Distributed Parameter and Stochastic Systems*
© IFIP International Federation for Information Processing 1999

ential equation (FBSDE) on a given time duration $[0, T]$:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \sqrt{\varepsilon} \int_0^t \sigma(s, X_s, Y_s) dW_s, \\ Y_t = g(X_T) + \int_t^T \hat{b}(s, X_s, Y_s, Z_s) ds + \int_t^T Z_s dW_s, \end{cases} \quad t \in [0, T], \quad (1.1)$$

where $0 < \varepsilon \ll 1$.

We refer the readers to El Karoui-Quenez-Peng (1997) [7] for an overview of backward SDEs, and to Ma-Protter-Yong (1994) [13], Hu-Peng (1995) [12], Duffie-Ma-Yong (1994) [6], and Cvitanic-Ma (1996) [2] for the theory and applications involving FBSDEs. Throughout this note we denote the *adapted solution* of (1.1) by $\Theta^\varepsilon \triangleq (X^\varepsilon, Y^\varepsilon, Z^\varepsilon)$, and we assume that X^ε , Y^ε , and Z^ε are all one-dimensional. Also, we shall make use of the following assumptions:

(H1) The functions σ , b , $\hat{b}$ and g are continuous in (t, x, y, z) , and continuously differentiable in (x, y, z) with uniformly bounded first order partial derivatives;

(H2) The function g belongs boundedly to $C^{2+\alpha}(\mathbf{R})$ for some $\alpha \in (0, 1)$;

(H3) There exists positive numbers $\mu > 0$, and $K > 0$ such that

$$\begin{cases} \mu \leq \sigma(t, x, y) \leq K, & \forall(t, x, y); \\ |b(t, x, 0, 0)| + |\hat{b}(t, x, 0, z)| \leq K, & \forall(t, x, z). \end{cases}$$

The main purpose of this note is to study the asymptotic behavior of $\{\Theta^\varepsilon\}$ as $\varepsilon \rightarrow 0$. In particular we would like to show that the large deviations principle (LDP) holds for the distribution of $\{\Theta^\varepsilon\}$, as $\varepsilon \rightarrow 0$. It turns out that the following type of convergence of a family of functions will be essential in our future discussion.

Definition 1.1 (HYPOTHESIS A) Let $\varphi^\varepsilon \in C^{1,2}([0, T] \times \mathbf{R})$, $\varepsilon > 0$, be a family of functions, and $\varphi^0 \in C([0, T] \times \mathbf{R})$. We say that the family $\{\varphi^\varepsilon\}_{\varepsilon > 0}$ satisfies “Hypothesis A” with limiting function φ^0 , if the following conditions hold:

(1) The family $\{\varphi^\varepsilon\}$ converges uniformly on compacts to φ^0 , as $\varepsilon \rightarrow 0$; and, for each $M > 0$, there exists $K^M > 0$ such that

$$|\varphi^0(t, x) - \varphi^0(t, y)| \leq K^M |x - y|, \quad \forall(t, x) \in [0, T] \times [-M, M].$$

(2) $\lim_{\varepsilon \rightarrow 0} |\sqrt{\varepsilon} \varphi_x^\varepsilon| = 0$, uniformly on compacts.

Combining the LDP results with the simulation of the adapted solution of FBSDEs (cf. Douglas-Ma-Protter [5]), we then discuss the simulation of rare events for the paths of FBSDEs. Rare event simulation has been extensively studied in queuing and reliability models (cf. Heidelberg [11]). In the setting of diffusion processes, it can be used to assess the probability of exercising a deep out of the money option when the underlying stock possesses small volatility (see Fournié-Lasry-Touzi [9]). We consider *importance sampling* estimators and provide necessary conditions for the existence of an unbiased estimator that is *asymptotically efficient* (or *asymptotically optimal*) in the sense of Sadowsky-Bucklew [16].

2 LARGE DEVIATION FOR FBSDES

Let us recall some definitions concerning large deviations (cf. e.g. [4]). Let $\mathcal{X}$ be a topological space equipped with σ -field $\mathcal{B}$, and $I : \mathcal{X} \mapsto [0, \infty]$ be a lower semicontinuous function such that the level sets $\Psi_\alpha^I = \{x : I(x) \leq \alpha\}$ are compact for all $\alpha < \infty$. We say that a family of $\mathcal{X}$ -valued random variables satisfies the LDP with *good rate function* I , if their distributions on $(\mathcal{X}, \mathcal{B})$, denoted by $\{\mu_\varepsilon\}_{\varepsilon>0}$, satisfy

$$-\inf_{\gamma \in \Gamma^\circ} I(\gamma) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leq -\inf_{\gamma \in \bar{\Gamma}} I(\gamma),$$

for all $\Gamma \in \mathcal{B}$, where Γ° and $\bar{\Gamma}$ denote the interior and closure of Γ , respectively. In this note we consider the case when $\mathcal{X} = C^3([0, T]; \mathbf{R})$ with sup-norm; denote

$$H_1(0, T) \triangleq \left\{ g_t = \int_0^t f_s ds, \ t \in [0, T] : \int_0^T |f_s|^2 ds < \infty \right\},$$

and endow H_1 with the norm $\|g\|_{H_1} \triangleq \sqrt{\int_0^T |f_s|^2 ds}$. Also, we denote, for functions $b, \sigma : [0, T] \times \mathbf{R} \rightarrow \mathbf{R}$ that satisfy (H1) and (H3),

$$F_{b, \sigma}(f) \triangleq \int_0^t \{[\dot{f}_s - b(s, f_s)]\sigma^{-1}(s, f_s)\} ds, \quad \forall f \in H_1. \quad (2.1)$$

Following the *Four Step Scheme*, initiated by Ma-Protter-Yong in [13], we know that under the assumptions (H1)-(H3) (1.1) can be solved “explicitly” as follows:

$$\begin{cases} X_t^\varepsilon = x + \int_0^t \tilde{b}^\varepsilon(s, X_s^\varepsilon) ds + \int_0^t \tilde{\sigma}^\varepsilon(s, X_s^\varepsilon) dW_s, \\ Y_t^\varepsilon = \theta^\varepsilon(t, X_t^\varepsilon), \quad Z_t^\varepsilon = -\sqrt{\varepsilon} \tilde{\sigma}^\varepsilon(t, X_t^\varepsilon) \theta_x^\varepsilon(t, X_t^\varepsilon), \end{cases} \quad (2.2)$$

where $\theta^\varepsilon \in C^{1,2}([0, T] \times \mathbf{R})$ is the classical solution to the following quasi-linear parabolic PDE:

$$\begin{cases} 0 = \theta_t^\varepsilon + \frac{1}{2} \varepsilon \tilde{\sigma}^\varepsilon(t, x)^2 \theta_{xx}^\varepsilon + \tilde{b}^\varepsilon(t, x) \theta_x^\varepsilon + \hat{b}(t, x, \theta^\varepsilon, -\sqrt{\varepsilon} \sigma(t, x, \theta^\varepsilon) \theta_x^\varepsilon), \\ \theta^\varepsilon(T, x) = g(x), \end{cases} \quad (2.3)$$

and $\tilde{b}^\varepsilon(t, x) \triangleq b(t, x, \theta^\varepsilon(t, x), -\sqrt{\varepsilon} \tilde{\sigma}^\varepsilon(t, x) \theta_x^\varepsilon(t, x))$, $\tilde{\sigma}^\varepsilon(t, x) \triangleq \sigma(t, x, \theta^\varepsilon(t, x))$.

Our principle result is the following

Theorem 2.1 *Suppose that the solutions of (2.3), $\{\theta^\varepsilon\}_{\varepsilon>0}$, satisfy Hypothesis A with limiting function θ^0 . Then the solution family $\{(X^\varepsilon, Y^\varepsilon, Z^\varepsilon)\}_{\varepsilon>0}$ satisfies the LDP in $C^3([0, T]; \mathbf{R})$ with good rate function given by*

$$I_x(f, g, h) = \inf \left\{ \frac{1}{2} \int_0^T |\dot{F}_{\tilde{b}, \tilde{\sigma}}(f)_s|^2 ds \mid f \in H_1^x; \ g = \theta^0(\cdot, f), \ h = 0 \right\}$$

where

$$\tilde{b}(t, x) = b(t, x, \theta^0(t, x), 0), \quad \tilde{\sigma}(t, x) = \sigma(t, x, \theta^0(t, x)); \quad (2.4)$$

and θ^0 is the (unique) viscosity solution to the first order PDE

$$\begin{cases} \theta_t^0 + b(t, x, \theta^0, 0)\theta_x^0 + \hat{b}(t, x, \theta^0, 0) = 0 \\ \theta^0(T, x) = g(x). \end{cases} \quad (2.5)$$

3 VERIFICATION OF HYPOTHESIS A

1. Small Duration Case.

We first show that if the duration T and the coefficients b , $\hat{b}$, σ , and g satisfy a certain “compatibility condition”, then the family $\{\theta^\varepsilon\}$ will satisfy Hypothesis A. Such a condition contains in particular the case when the duration T is small.

To begin with, note that using a maximum-principle type of argument one can show that there exists a constant $M > 0$, independent of ε , such that

$$\sup_{\varepsilon > 0} \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} |\theta^\varepsilon(t, x)| \leq M. \quad (3.1)$$

Next, with $\varepsilon > 0$, consider the FBSDE on $0 \leq t \leq s \leq T$:

$$\begin{cases} X_s = x + \int_t^s b(r, X_r, Y_r, Z_r) dr + \sqrt{\varepsilon} \int_t^s \sigma(r, X_r, Y_r) dW_r, \\ Y_s = g(X_T) + \int_s^T \hat{b}(r, X_r, Y_r, Z_r) dr + \int_s^T Z_r dW_r, \quad s \in [t, T]. \end{cases} \quad (3.2)$$

Denote the adapted solution of (3.2) on $[t, T]$ by $(X^{t, x, \varepsilon}, Y^{t, x, \varepsilon}, Z^{t, x, \varepsilon})$. Since the coefficients are deterministic, it follows that $\theta^\varepsilon(t, x) \triangleq Y_t^{t, x, \varepsilon}$ is also deterministic, and is in fact the classical solution to the quasilinear PDE (2.3), thanks to the Four Step Scheme. The following result is more or less well understood in the theory of backward SDEs, and is a modification of those of Pardoux-Tang [15] and Cvitanic-Ma [3] in the FBSDE case.

Lemma 3.1 *Assume (H1)–(H3). Let $T > 0$ be given. Suppose that there exist positive constants C_1, C_2, C_3, C_4 and $\lambda > 0$, such that $0 < 1 - KC_4 < 1$, and the following “compatibility condition” holds:*

$$1 - \mu(T)K^2 > 0, \quad \lambda_1 - \mu(T)KC_3 > 0, \quad (3.3)$$

where

$$\begin{cases} \lambda_1 = \lambda - 2K - K(C_1^{-1} + C_2^{-1}) - \varepsilon^2 K^2; \\ \lambda_2 = -\lambda - 2K - KC_3^{-1} - KC_4^{-1}, \\ \mu(T) = (KC_1 + \varepsilon K^2)B(\lambda_2, T) + \frac{1}{1 - KC_4}A(\lambda_2, T)KC_2; \\ A(\lambda, t) = e^{-(\lambda \wedge 0)t}; \\ B(\lambda, t) = \frac{1 - e^{-\lambda t}}{\lambda}. \end{cases} \quad (3.4)$$

Then there exists a constant $C > 0$, depending only on K and T , such that for all $(t_1, x_1), (t_2, x_2) \in [0, T] \times \mathbb{R}$, and $\varepsilon \in (0, 1)$, it holds that

$$|\theta^\varepsilon(t_1, x_1) - \theta^\varepsilon(t_2, x_2)|^2 \leq C\{|x_1 - x_2|^2 + (1 + |x_1|^2 + |x_2|^2)|t_1 - t_2|\}. \quad (3.5)$$

Remark. It is easy to check that (3.3) is solvable at $T = 0$. A “continuous dependence” argument then shows that there exists a constant $T_0 > 0$, independent of ε , such that (3.3) holds for all $0 < T < T_0$. Consequently (3.5) holds when $T > 0$ is small. ■

Lemma 3.1, together with (3.1), essentially says that $\{\theta^\varepsilon\}$ is uniformly bounded and (locally) equicontinuous. Thus in view of the Arzelà-Ascoli theorem we derive

Theorem 3.2 *Assume (H1)–(H3). Suppose that $T > 0$ is such that the compatibility conditions (3.3) and (3.4) are solvable. Then the family of solutions $\{\theta^\varepsilon\}_{0 < \varepsilon \leq 1}$ of the PDE (2.3) satisfies Hypothesis A with limiting function θ^0 , where θ^0 is a viscosity solution to the first order PDE (2.5). Moreover, the solution is unique among those that are uniform Lipschitz in the x variable. In particular, there exists a constant $T_0 > 0$, such that for all $0 < T < T_0$, the above statement holds true.*

2. Large Duration Case.

This case is more complicated, since the PDE theory tells us that it is quite remote to establish Hypothesis A for the solutions to nonlinear PDEs in general. Our purpose here is to find reasonable sufficient conditions.

To establish Hypothesis A-(1) we follow the so-called Barles-Perthame procedure (cf. Barles-Perthame [1], or Fleming-Soner [8]). Namely, we define two functions

$$\theta^*(t, x) = \limsup_{\substack{(\varepsilon, y) \rightarrow (t, x) \\ \varepsilon \rightarrow 0}} \theta^\varepsilon(t, x) \quad \text{and} \quad \theta_*(t, x) = \liminf_{\substack{(\varepsilon, y) \rightarrow (t, x) \\ \varepsilon \rightarrow 0}} \theta^\varepsilon(t, x).$$

Clearly, both θ^* and θ_* are well defined thanks to (3.1); $\theta_* \leq \theta^*$; and θ_* (resp. θ^*) is upper (resp. lower) semicontinuous. Next, we use standard arguments to show that, under (H1)–(H3), θ^* is a viscosity subsolution and θ_* is a viscosity supersolution to the first order PDE (2.5). A “comparison principle” is then established, showing $\theta^* \leq \theta_*$, and consequently that $\theta^* = \theta_*$ ($\triangleq \theta^0$) is continuous and the convergence must be uniform on compact sets. Thus Hypotheses A-(1) will follow, provided we can show that θ^0 is uniform Lipschitz in x .

Continuing, using gradient estimates for nonlinear PDEs, together with (H1)–(H3), as well as an extra condition that $b(t, x, y, 0) = 0$, we can show that there exists a constant $M > 0$, independent of ε , such that

$$\sup_{(t, x)} |\theta_x^\varepsilon(t, x)| \leq M(1 + \frac{1}{\varepsilon^r}), \quad \forall \varepsilon \in (0, 1). \quad (3.6)$$

which clearly leads to Hypothesis A-(2). Therefore we have the following

Theorem 3.3 *Assume (H1)–(H3) and that $b(t, x, y, 0) = 0$, $\forall (t, x, y) \in [0, T] \times \mathbb{R}^2$. Then the solution sequence $\{\theta^\varepsilon\}$ of (2.3) satisfies Hypothesis A with limiting function θ^0 which is the unique viscosity solution to (2.5).*

We remark here that the assumption $b(t, x, y, 0) = 0$ in Theorem 3.3 is almost necessary for (3.6) to hold, which is obviously stronger than Hypothesis A-(2). However, in the case when Hypothesis A-(2) can be verified independently, we can replace the assumption $b(t, x, y, 0) = 0$ by the requirement that $b(t, x, y, z)$ be independent of y , a slightly different situation.

4 RARE EVENT SIMULATION

In this section we turn to the problem of simulating the probability of certain events $\{\Theta^\varepsilon \in A\} \in \mathcal{F}$, where A is some Borel set of $C^3([0, T]; \mathbb{R})$, such that $P\{\Theta^\varepsilon \in A\} \rightarrow 0$, as $\varepsilon \rightarrow 0$ (i.e., A is a “rare event”). Noting that (2.2) shows that Y^ε and Z^ε are functions of X^ε , we shall restrict our attention to estimating $\gamma_A(\varepsilon) \triangleq P\{X^\varepsilon \in A\}$, where A is now some Borel set of $C([0, T]; \mathbb{R})$.

Consider the standard Monte Carlo estimator of $\gamma_A(\varepsilon)$: let $X^\varepsilon(1), \dots, X^\varepsilon(N)$ be N independent copies of sample paths X^ε , which can be simulated using the method of [5], and define $\hat{\gamma}_N^\varepsilon := \frac{1}{N} \sum_{n=1}^N 1_A(X^\varepsilon(n))$. Clearly, $\hat{\gamma}_N^\varepsilon$ is an unbiased estimator of $\gamma_A(\varepsilon)$ such that $\hat{\gamma}_N^\varepsilon \rightarrow \gamma_A(\varepsilon)$ almost surely as $N \rightarrow \infty$, thanks to the Strong Law of Large Numbers. A drawback of this estimator, however, is that as the event gets rarer (i.e., $\gamma_A(\varepsilon) \ll 1$) the *relative error* of the estimator (the quotient of the standard deviation of the estimator $\hat{\gamma}_N^\varepsilon$ and $\gamma_A(\varepsilon)$ itself), given by $\mathcal{E}_r(\hat{\gamma}_N^\varepsilon) \triangleq \sqrt{1 - \gamma_A(\varepsilon)} / \sqrt{N \gamma_A(\varepsilon)}$, behaves like $1/\sqrt{N \gamma_A(\varepsilon)}$, which makes the formation of confidence intervals overly time consuming. In an attempt to overcome this difficulty we turn to the so-called *importance sampling estimators*, which we now describe.

For any $h \in H_1$ and $\varepsilon > 0$, define a process

$$M_t^{h, \varepsilon} = \exp \left\{ \frac{1}{\sqrt{\varepsilon}} \int_0^t \dot{h}_s dW_s + \frac{1}{2\varepsilon} \int_0^t |\dot{h}_s|^2 ds \right\}, \quad t \in [0, T].$$

It can be easily checked that $(M^{h, \varepsilon})^{-1}$ is a martingale. By the Girsanov theorem, under a new probability measure $P^{h, \varepsilon}$ defined by $dP^{h, \varepsilon}/dP = (M_T^{h, \varepsilon})^{-1}$, $W_t^{h, \varepsilon} = W_t + \frac{1}{\sqrt{\varepsilon}} h_t$ will be a Brownian motion, and the pair $(X^\varepsilon, M^{h, \varepsilon})$ will satisfy the SDE:

$$\begin{cases} X_t^\varepsilon = x + \int_0^t [\tilde{b}^\varepsilon(s, X_s^\varepsilon) - \tilde{\sigma}^\varepsilon(s, X_s^\varepsilon) \dot{h}_s] ds + \int_0^t \sqrt{\varepsilon} \tilde{\sigma}^\varepsilon(s, X_s^\varepsilon) dW_s^{h, \varepsilon} \\ M_t^{h, \varepsilon} = 1 + \frac{1}{\sqrt{\varepsilon}} \int_0^t \dot{h}_s M_s^{h, \varepsilon} dW_s^{h, \varepsilon}. \end{cases} \quad (4.1)$$

Now if we simulate N independent copies of the sample paths of the solution to (4.1), $\{(\tilde{X}^{h, \varepsilon}(n), \tilde{M}^{h, \varepsilon}(n))\}_{n=1}^N$, on any probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, with $W^{h, \varepsilon}$

in (4.1) being replaced by any other $\tilde{P}$ -Brownian motion(!), and define

$$\tilde{\gamma}_N^{h,\varepsilon} \triangleq \frac{1}{N} \sum_{n=1}^N 1_A(\tilde{X}^{h,\varepsilon}(n)) \tilde{M}^{h,\varepsilon}(n),$$

then by the uniqueness in law of the SDE, it is seen that $\tilde{\gamma}_N^{h,\varepsilon}$ is again an unbiased estimator of $\gamma_A(\varepsilon)$ such that $\tilde{\gamma}_N^{h,\varepsilon} \rightarrow \gamma_A(\varepsilon)$, $\tilde{P}$ -almost surely, as $N \rightarrow \infty$. However, the relative error of the estimator $\tilde{\gamma}_N^{h,\varepsilon}$ now becomes

$$\mathcal{E}_r(\tilde{\gamma}_N^{h,\varepsilon}) = \frac{\sqrt{\tilde{E}\{\tilde{\gamma}_N^{h,\varepsilon} - E(\tilde{\gamma}_N^{h,\varepsilon})\}^2}}{E\{\tilde{\gamma}_N^{h,\varepsilon}\}} = \frac{\sqrt{\eta^{h,\varepsilon} - \gamma_A^2(\varepsilon)}}{\sqrt{N}\gamma_A(\varepsilon)}.$$

where $\eta^{h,\varepsilon} \triangleq E^{h,\varepsilon}\{1_A(X^\varepsilon)M_T^{h,\varepsilon}\}^2$. Note that Jensen's inequality tells us that $\eta^{h,\varepsilon} \geq \gamma_A^2(\varepsilon)$. Thus if the LDP holds for X^ε with good rate function I_x , and if A is an I -continuity set (i.e., $\inf_{f \in A^\circ} I_x(f) = \inf_{f \in A} I_x(f) \triangleq I_A$) such that $0 < I_A < \infty$, then we must have

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log \eta^{h,\varepsilon} \geq 2 \lim_{\varepsilon \rightarrow 0} \varepsilon \log \gamma_A(\varepsilon) = -2I_A.$$

If, in particular, the equality $\lim_{\varepsilon \rightarrow 0} \varepsilon \log \eta^{h,\varepsilon} = -2I_A$ holds, then we say that the importance sampling estimator $\tilde{\gamma}_N^{h,\varepsilon}$ is *asymptotically efficient* or *asymptotically optimal* (cf. Sadowsky-Bucklew [16]). One of the main features of an estimator being asymptotically efficient is that the *critical sample size* $N_\alpha(\varepsilon) \triangleq \inf\{N \in \mathbb{N} : \mathcal{E}_r(\tilde{\gamma}_N^{h,\varepsilon}) \leq \alpha\}$ will grow at most sub-exponentially as $\varepsilon \rightarrow 0$, an extremely desirable property in rare event simulation. In other words we have

Theorem 4.1 *An importance sampling estimator $\tilde{\gamma}_N^{h,\varepsilon}$ is asymptotically optimal if and only if $\lim_{\varepsilon \rightarrow 0} N_\alpha(\varepsilon)e^{-\frac{\varepsilon}{\kappa}} \leq 0$ for any $\kappa > 0$.*

The remaining question is then how to choose $h \in H_1$ so that the corresponding importance sampling estimator $\tilde{\gamma}_N^{h,\varepsilon}$ is asymptotically efficient. To this end let us recall the mapping $F_{\tilde{b},\tilde{\sigma}}$ defined by (2.1), and $\tilde{b}$ and $\tilde{\sigma}$ defined by (2.4). Define, for $f \in H_1^x$, $I_x(f) = \frac{1}{2} \int_0^T |\dot{F}_{\tilde{b},\tilde{\sigma}}(f)_s|^2 ds$. Let us now simply denote $F \triangleq F_{\tilde{b},\tilde{\sigma}}$. The following definition of F -dominating point is motivated by the notion of “dominating point” from large deviation theory (see, e.g., Ney [14] and Sadowsky-Bucklew[16]), adapted to our setting. For $A \in \mathcal{B}(C[0,T])$, we say that $f^* \in \bar{A} \cap H_1^x$ is an F -dominating point of A if

$$\int_0^T \dot{F}(f^*)_s (\dot{F}(f^*)_s - \dot{F}(g)_s) ds \leq 0, \quad \forall g \in \bar{A} \cap H_1^x.$$

We then have

Theorem 4.2 *Suppose that A is an I -continuity set such that $0 < I_A < \infty$. Suppose also that there exists an F -dominating point, f^* , of A . Then the importance sampling estimator $\tilde{\gamma}_N^{h^*,\varepsilon}$, where $h^* \triangleq F(f^*)$, is asymptotically efficient.*

References

- [1] Barles, G. and Perthame, B., Exit time problems in optimal control and vanishing viscosity solutions of Hamilton-Jacobi equations, *SIAM J. Control Optim.*, Vol. 26, 1988, pp. 1133-1148.
- [2] Cvitanović, J. and Ma, J., Hedging options for a large investor and forward-backward SDE's, *Ann. Appl. Prob.*, Vol. 6, No. 2, 1996, pp. 370-398.
- [3] Cvitanović, J. and Ma, J., Reflected forward-backward SDEs and obstacle problems with boundary conditions, Preprint, 1997.
- [4] Dembo, A. and Zeitouni, O., *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, 1993.
- [5] Douglas, J., Ma, J. and Protter, P., Numerical methods for forward-backward stochastic differential equations, *Ann. Appl. Prob.*, Vol. 6, No. 3, 1996, pp. 940-968.
- [6] Duffie, D., Ma, J. and Yong, J., Black's console rate conjecture, *Ann. Appl. Prob.*, Vol. 5, No. 2, pp. 356-382, 1995.
- [7] El Karoui, N, S. Peng, and M.C. Quenez, Backward stochastic differential equation in finance, *Mathematical Finance*, Vol. 7, No. 1, 1997, pp. 1-71.
- [8] Fleming, W. and Soner, M., *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York, 1993.
- [9] Fournié, E., Lasry, J.M. and Touzi, N., Monte carlo methods for stochastic volatility models, *Numerical Methods in Finance*, pp. 146-164, Cambridge Univ. Press, Cambridge, 1997.
- [10] Fournié, E., Lasry, J.M. and Lions, P.L., Some nonlinear methods for studying far-from-the-money contingent claims. *Numerical methods in finance*, pp. 115-145, Cambridge Univ. Press, Cambridge, 1997.
- [11] Heidelberger, P., Fast simulation of rare events in queueing and reliability models, *ACM Trans. on Mod. and Comp. Simulation*, Vol. 5, No. 1, 1995, pp. 43-85.
- [12] Hu, Y. and Peng, S., Solution of forward-backward stochastic differential equations, *Prob. Th. Rel. Fields*, Vol. 103, 1995, pp. 273-283.
- [13] Ma, J., Protter, P. and Yong, J., Solving forward-backward stochastic differential equations explicitly - a four step scheme, *Prob. Th. Rel. Fields.*, Vol. 98, 1994, pp. 339-359.
- [14] Ney, P., Dominating points and the asymptotics of large deviations for random walk on $\mathbb{R}^d$, *Ann. Proba.*, Vol. 11, No. 1, 1983, pp. 158-167.
- [15] Pardoux, E. and Tang, S., Forward-backward SDEs and quasilinear parabolic PDEs, Preprint, 1997.
- [16] Sadowsky, J.S. and Bucklew, J.A., On large deviations theory and asymptotically efficient Monte Carlo estimation, *IEEE Trans. Inf. Theor.*, Vol. 36, 1990, pp. 579-588.