Interpolation for partly hidden diffusion processes*

Changsun Choi, Dougu Nam**

Division of Applied Mathematics, KAIST, Daejeon 305-701, Republic of Korea

Abstract

Let X_t be *n*-dimensional diffusion process and S(t) be a smooth set-valued function. Suppose X_t is invisible when $X_t \in S(t)$, but we can see the process exactly otherwise. Let $X_{t_0} \in S(t_0)$ and we observe the process from the beginning till the signal reappears out of the obstacle after t_0 . With this information, we evaluate the estimators for the functionals of X_t on a time interval containing t_0 where the signal is hidden. We solve related 3 PDE's in general cases. We give a generalized last exit decomposition for n-dimensional Brownian motion to evaluate its estimators. An alternative Monte Carlo method is also proposed for Brownian motion. We illustrate several examples and compare the solutions between those by the closed form result, finite difference method, and Monte Carlo simulations.

Key words: interpolation, diffusion process, excursions, backward boundary value problem, last exit decomposition, Monte Carlo method

1 Introduction

Usually, the interpolation problem for Markov process indicates that of evaluating the probability distribution of the process between the discretely observed times. One of the origin of the problem is Schrödinger's *Gedankenexperiment* where we are interested in the probability distribution of a diffusive particle in a given force field when the initial and final time density functions are given. We can obtain the desired density function at each time in a given time interval by solving the forward and backward Kolmogorov equations. The background and methods for the problem are well reviewed in [17]. The problem also attracts interests in signal processing. It is useful in restoration of lost

¹ *This work is supported by BK21 project.

 $^{^2}$ **corresponding author, Email - address: dunam@amath.kaist.ac.kr

signal samples [8]. In this paper, we consider a special type of interpolation problem for which similar applications are expected.

Let X_t be the signal process represented by some system of stochastic differential equations. Suppose X_t is invisible when $X_t \in S(t)$ for some set-valued function S(t), $0 \le t \le T < \infty$ and we can observe the process exactly otherwise i.e. the observation process is

$$Y_t = X_t I_{X_t \notin S(t)}, \ 0 \le t \le T.$$

In [10], 1-dimensional signal process and a fixed obstacle $S(t) = (0, \infty)$ are considered and the optimal estimator $E[f(X_1)I_{X_1>0}|Y_{[0,1]}]$ was derived for bounded Borel function f. The key identity is

$$E[f(X_1)I_{X_1>0}|Y_{[0,1]}] = E[f(X_1)I_{X_1>0}|\tau]$$
(1)

where $1 - \tau$ is the lastly observed time before 1 before the signal enters into the obstacle. This is not a stopping time but if we consider the reverse time process χ_t of X_t , τ is a stopping time and we can prove (1) by applying the strong Markov property to χ_t . This simplified estimator is again evaluated by the formula

$$E[f(X_1)I_{X_1>0}|\tau] = \int_0^\infty f(x)q(x|\tau)dx$$
 (2)

where $q(x|\tau) = \frac{p_1(x)u_x(\tau)}{\int_0^\infty p_1(z)u_z(\tau)dz}$, $p_1(x)$ is the density function of X_1 and $u_x(t)$ is the first hitting time density of χ_t starting at x to 0 at t.

In [11], this problem was considered in a more generalized setting, where X_t is n-dimensional diffusion process and the obstacle S(t) is a time-varying setvalued function. For this setting, we should consider also the first hitting state χ_{τ} as well as time τ and the corresponding identity becomes

$$E[f(X_1)I_{X_1 \in S(1)}|Y_{[0,1]}] = E[f(X_1)I_{X_1 \in S(1)}|\tau, \chi_{\tau}]$$
(3)

and q(x|t) is replaced by

$$q(x|t,y) = \frac{p_1(x)u_x(t,y)}{\int_0^\infty p_1(z)u_z(t,y)dz}.$$
(4)

where $u_x(t, y)$ is the first hitting density on the surface of the images of S(t). This problem can be thought as a kind of filtering problem (not a prediction), because we observe the hidden signal up to time 1(to be estimated) and the hidden process after time $1 - \tau$ still gives the information that the signal is in the obstacle. In this paper, we consider related extended problems. We continue observing the hidden signal process till it reappears out of the obstacle. With this additional information we can estimate the hidden state with more accuracy. An interesting fact is that even if the signal does not reappear in a finite fixed time T, this information also improves the estimator. We also evaluate extrapolator and backward filter. The interpolator developed in Section 3.4 in particular, can be applied to reconstructing the lost parts of random signals in a probabilistically optimal way.

In section 3, we prove the corresponding Bernstein-type [17] identity similar to (1) and (3) and derive the optimal estimator. Due to this identity, the problem reduces to that of computing the density functions of general excursions (or meander) in the solution of stochastic differential equations on given excursion intervals. It is well known that the transition density function for stochastic differential equation is evaluated through a parabolic PDE. For the interpolator, we should solve a Kolmogorov equation and two backward boundary value problems. We explain a numerical algorithm by finite difference method for general 1-dimensional diffusion processes. When the signal is a simple Brownian motion process, we can use the so called last exit decomposition principle and we can derive the estimators more conveniently, which will be confirmed by a numerical computation example. Alternatively, for Brownian motion, we can use Monte Carlo method for some boundaries by numerical simulation of various conditional Brownian motions. We give several examples and compare the results by each method.

2 Reverse Time Diffusion Process

Consider the following system of stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \quad X_0 = 0, \quad 0 \le t < T$$
(5)

where $a: [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $b: [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are measurable functions and W_t is *m*-dimensional Brownian motion. According to [10] and [11], to obtain the desired estimators, we need the reverse time diffusion process χ_t of X_t on [0,T). In [3], the conditions for the existence and uniqueness of the reverse time diffusion process are stated and they are abbreviated to

Assumption 1 the functions $a_i(t, x)$ and $b_{ij}(t, x)$, $1 \le i \le n$, $1 \le j \le m$ are bounded and uniformly Lipschitz continuous in (t, x) on $[0, T] \times \mathbb{R}^n$;

Assumption 2 the functions $(b b^{\top})_{ij}$ are uniformly Hölder continuous in x;

Assumption 3 there exists c > 0 such that

$$\sum_{i,j=1}^{n} (b \, b^{\top})_{ij} \, y_i y_j \ge c |y|^2, \quad for \ all \quad y \in \mathbb{R}^n, \quad (t,x) \in [0,T] \times \mathbb{R}^n;$$

Assumption 4 the functions a_x, b_x and b_{xx} satisfy Assumption 1 and 2.

We say that X_t in (5) is time reversible iff the coefficients satisfy the above 4 conditions.

Theorem 1 [3] Assume Assumption 1 - Assumption 4 hold. Let p(t, x) be the solution of the Kolmogorov equation for X_t . Then X_t can be realized as

$$X_t = X_T + \int_T^t \bar{a}(s, X_s) d(-s) + \int_T^t b(s, X_s) d\bar{W}_s, \quad 0 \le t < T$$

where W_s is a Brownian motion independent of X_T , and

$$\bar{a}(s, X_s) = -a(s, X_s) + \frac{\sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\sum_{k=1}^m b_{ik}(s, x) b_{\cdot k}(s, x) p(s, x) \right) \Big|_{x=X_s}}{p(s, X_s)}.$$
 (6)

3 **Derivation of The Optimal Estimators**

3.1Notations & Definitions

We will use the following notations throughout this paper.

- $P_{cc}(\mathbb{R}^n)$: the space of closed and connected subsets in \mathbb{R}^n .
- $S: [0,T] \to P_{cc}(\mathbb{R}^n)$: set-valued function. $\mathbb{S}_{(a,b)} := \bigcup_{a < t < b} S(t)$, $\partial \mathbb{S}_{(a,b)} := \bigcup_{a < t < b} \partial S(t)$, $\partial \mathbb{S} := \partial \mathbb{S}_{(0,T)}$.

We say S is smooth iff for each $x \in \partial \mathbb{S}$, there exists a unique tangent plane at x. We only consider smooth set-valued functions.

- X_t : the time reversible signal process satisfying (5).

- $Y_t := X_t I_{X_t \notin S(t)}, t \in [0, T], \quad Y_{[a,b]} := \{Y_t\}_{t \in [a,b]}.$ $X_t^{t_0} := X_{t+t_0}, \quad \chi_t^{t_0} := X_{t_0-t} I_{t \in [0,T]}, \text{ the reverse time process.}$ $\tilde{X}_t^{t_0} := X_t^{t_0} I_{X_t^{t_0} \notin S(t_0+t)}, \quad \tilde{\chi}_t^{t_0} := X_t^{t_0} I_{X_t^{t_0} \notin S(t_0-t)}.$

Note that $X_t^{t_0}$ and $\chi_t^{t_0}$ have the same initial distribution of $\xi \stackrel{d}{\sim} X_{t_0}$.

Let $\mathcal{F}_t = \sigma\{W_t : 0 \le t < T - t_0\}, \ \mathcal{G}_t = \sigma\{\bar{W}_t : 0 \le t < t_0\}, \ \text{and} \ \xi \ \text{be}$ all independent and we consider the filtrations $\sigma(\xi, \mathcal{F}_t), \ 0 \leq t < T - t_0$ and $\sigma(\xi, \mathcal{G}_t), \ 0 \leq t < t_0$. Let $X_{t_0} \in S(t_0)$ and we define two random times

• $\tau_{t_0} := t_0 - \sup\{s | s < t_0, X_s \notin S(s)\} \lor 0,$

• $\sigma_{t_0} := \inf\{s | s > t_0, X_s \notin S(s)\} \land T - t_0.$

We omit the subscript t_0 for convenience. τ and σ can be rephrased w.r.t. $X_t^{t_0}$ and $\chi_t^{t_0}$ as follows:

$$\tau = \tau(\xi)$$
: the first hitting time of $\chi_t^{t_0}$ to $S(t_0 - t)$ when $\xi \in S(t_0)$,
 $\sigma = \sigma(\xi)$: the first hitting time of $X_t^{t_0}$ to $S(t_0 + t)$ when $\xi \in S(t_0)$.

Let X_t^x and χ_t^x be the processes $X_t^{t_0}$ and $\chi_t^{t_0}$ which start at x and P_{X^x} and P_{χ^x} be probability measures induced by them, then with P_{ξ} induced by ξ , these three probability measures are independent. We denote E_{χ^0} , E_{X^0} and E_{ξ} for the corresponding integrations.

We define the first hitting density of $X_t^{t_0}$ on $\partial \mathbb{S}_{(t_0,T)}$ as follows: We first define two measures on $\partial \mathbb{S}_{(t_0,T)}$. Let \tilde{B} be open set in $\partial \mathbb{S}_{(t_0,T)}$ and $x \in S(t_0)$.

$$\nu_x(\tilde{B}) := P_{X^x}((\sigma(x), X^{t_0}_{\sigma(x)}) \in \tilde{B}),$$

$$\mu_n(\tilde{B}) : \text{n-dimensional Hausdorff measure on } \partial \mathbb{S}_{(t_0,T)}.$$

Since $\nu_x \ll \mu_n$, there exists the unique density function $v_x(t, y) := \partial \nu_x / \partial \mu_n(t, y)$, $(t, y) \in \partial \mathbb{S}_{(t_0,T)}$ s.t.

$$\nu_x(\tilde{B}) := \int_{\tilde{B}} v_x(t,y) d\mu_n(t,y).$$

We can similarly define $u_x(s,z) := \partial \lambda_x / \partial \mu_n(s,z), (s,z) \in \partial \mathbb{S}_{(0,t_0)}$ where $\lambda_x(\tilde{A}) := P\chi^x((\tau(x), \chi_{\tau(x)}) \in \tilde{A})$ for open set $\tilde{A} \subset \partial \mathbb{S}_{(0,t_0)}$.

3.2 Extrapolation

Suppose $X_{t_0} \in S(t_0)$ and we have observed Y_t for $0 \le t \le t_0$ and we want to estimate a future state $X_t^{t_0}$, $0 \le t \le T - t_0$. From the lastly observed point before t_0 , we obtain the conditional density $v(x) = q(x|\tau, \chi_{\tau}^{t_0})$ for the present estimator as in (4) and we use this density as the initial condition for the Kolmogorov equation;

$$\frac{\partial p^{t_0}(t,x)}{\partial t} = -\sum_{i=1}^n \frac{\partial (p^{t_0}(t,x)a(t+t_0,x))}{\partial x} + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 (b^\top (t+t_0,x)p^{t_0}(t,x)b(t+t_0,x))}{\partial x_i \partial x_j}$$
(7)
$$p^{t_0}(0,x) = v(x).$$

The solution $p^{t_0}(t, x)$ is the conditional density for the extrapolator.

3.3 Backward filtering

Suppose $X_{t_0} \in S(t_0)$ and we observe the hidden signal from the present time t_0 up to a future time and we want to estimate X_{t_0} .

Applying the strong Markov property, we can get a similar identity as (3)

$$E[f(X_{t_0})I_{X_{t_0}\in S(t_0)}|Y_{[t_0,T]}] = E[f(X_{t_0})I_{X_{t_0}\in S(t_0)}|\sigma, X_{\sigma}^{t_0}],$$
(8)

We can think of two cases. One is the case when the hidden signal reappears before T and the other is when the signal does not appear up to T.

3.3.1 Case 1: $t_0 < t_0 + \sigma \leq T$

Consider the joint density function

$$v(x; s, z) = \frac{P\left[\xi \in dx; \, (\sigma(\xi), X^{t_0}_{\sigma(\xi)}) \in d(s, z)\right]}{dx \, d(s, z)}, \quad t_0 < s \le T.$$

By conditioning on ξ , we can decompose this density such as

$$v(x;s,z) = p_{t_0}(x)v_x(s,z),$$

where $p_{t_0}(x)$ is the density function of X_{t_0} . Hence, by the identity (8), we obtain the formula

$$E[f(X_{t_0})I_{X_{t_0}\in S(t_0)}|Y_{[t_0,T]}] = \int_{S_{t_0}} f(x)v(x|\sigma, X_{\sigma}^{t_0})dx$$
(9)

where $v(x|t,y) = \frac{p_{t_0}(x)v_x(s,z)}{\int_{S_{t_0}} p_{t_0}(r)v_r(s,z)dr}$. We can get $p_{t_0}(x) = p(t_0,x)$ by solving the Kolmogorov equation

$$\frac{\partial p(t,x)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial (p(t,x)a(t,x))}{\partial x} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 (b^T(t,x)p(t,x)b(t,x))}{\partial x_i \partial x_j}, \quad (10)$$
$$p(0,x) = \delta(x),$$

where $\delta(x)$ is Dirac-delta function. $v_x(s, z)$ is obtained by the following procedure. We need to solve the following backward boundary value problem.

Lemma 2 We let $C_0^{\infty}(\partial \mathbb{S})$ be the set of infinitely differentiable function ψ 's such that $\operatorname{Supp} \psi \subset \subset \partial \mathbb{S}$. Consider the following partial differential equation

$$\frac{\partial v}{\partial t} + \frac{1}{2} \mathbf{Tr}[b^{\top}(t,x)v_{xx}b(t,x)] + v_x^{\top}a(t,x) = 0, \quad (t,x) \in \mathbb{S}_{(t_0,T)}$$
(11)

with the boundary conditions

$$v(t,x) = \psi(t,x), \quad for \quad (t,x) \in \partial \mathbb{S}_{(t_0,T)},$$

$$v(T,x) = 0, \quad for \quad x \in S(t),$$
(12)

where Tr is the trace of $m \times m$ matrix and $\psi \in C_0^{\infty}(\partial \mathbb{S}_{(t_0,T)})$. Then there exists a unique solution v(t,x) in $C^{1,2}([t_0,T) \times S(t))$ and L such that

$$v(t,x) = \int_{\partial \mathbb{S}_{(t_0,T)}} L(t,x;r,y)\psi(r,y) \, d\mu_n(r,y)$$

holds.

Proof. See [12] and references therein.

By applying Ito's formula, Lemma 2, and the standard localization technique for local martingale, we obtain the following lemma.

Lemma 3 Under the same assumptions in Lemma 2, we get

$$E[\psi(\sigma(x), X^{t_0}_{\sigma(x)})] = v(0, x)$$

In particular, $(\sigma(x), X^{t_0}_{\sigma(x)})$ has the joint density function $L(0, x; s, z) (= v_x(s, z))$.

Proof. See Lemma 3.2 [11].

Hence, we have

$$E[\psi(\sigma(x), X^{t_0}_{\sigma(x)})] = \int_{\partial \mathbb{S}_{(t_0,T)}} v_x(r, y)\psi(r, y) \, d\mu_n(r, y)$$

and by taking a sequence $\psi_n^{s,z} \in C_0^{\infty}(\partial \mathbb{S}_{(t_0,T)})$ converging to the Dirac- δ function at (s, z) on $\partial \mathbb{S}_{(t_0,T)}$, we obtain the joint density $v_x(s, z)$ such as,

$$\lim_{n \to \infty} E[\psi_n^{s,z}(\sigma(x), X_{\sigma(x)}^{t_0})] = v_x(s, z).$$

Let $v_n^{s,z}$ be the corresponding solution of (11) with the boundary conditions (12) for $\psi = \psi_n^{s,z}$, then by Lemma 3 we have

$$v_x(s,z) = \lim_{n \to \infty} v_n^{s,z}(0,x).$$

3.3.2 Case 2 : $T < t_0 + \sigma$

Just $v_x(s, z)$ in (9) is replaced by $P(t_0 + \sigma(x) > T)$.

3.4 Interpolation

Suppose we have observed Y_t from the beginning to the final time T and we want to estimate X_{t_0} , $0 < t_0 < T$, when $X_{t_0} \in S(t_0)$.

The following lemma asserts that we can apply the strong Markov property to two strong Markov processes simultaneously even if they start at same ξ .

Lemma 4 Let f and g be bounded Borel functions, then

$$E[f(\chi_{\tau+t}^{t_0}) \cdot g(X_{\sigma+s}^{t_0}) | \mathcal{F}_{\tau(\xi)}, \xi, \mathcal{G}_{\sigma(\xi)}] = E[f(\chi_{\tau+t}^{t_0}) | \mathcal{F}_{\tau(\xi)}, \xi] \cdot E[g(X_{\sigma+s}^{t_0}) | \xi, \mathcal{G}_{\sigma(\xi)}].$$

Proof. Let $F = F(\xi)$, H, and $G = G(\xi)$ be $\mathcal{F}_{\tau(\xi)}$, ξ , and $\mathcal{G}_{\sigma(\xi)}$ measurable sets respectively, then conditioning on ξ , we have

$$\begin{split} E\Big[E[f(\chi_{\tau+t}^{t_0})|\mathcal{F}_{\tau(\xi)},\xi] \cdot E[g(X_{\sigma+s}^{t_0})|\xi,\mathcal{G}_{\sigma(\xi)}] \cdot I_{FHG}\Big] \\ &= E_{\xi}\Big[I_H \cdot E\Big[E[f(\xi+\chi_{\tau+t}^0) \cdot I_F|\mathcal{F}_{\tau(\xi)},\xi] \cdot E[g(\xi+X_{\sigma+s}^0) \cdot I_G|\xi,\mathcal{G}_{\sigma(\xi)}] \,\Big|\,\xi\Big]\,\Big] \\ &= E_{\xi}\Big[I_H \cdot E_{\chi^0}\Big[f(\xi+\chi_{\tau+t}^0) \cdot I_F\Big] \cdot E_{X^0}\Big[g(\xi+X_{\sigma+s}^0) \cdot I_G\Big]\Big] \\ &= E_{\xi}\Big[I_H \cdot E_{\chi^0 \times X^0}\Big[f(\xi+\chi_{\tau+t}^0) \cdot g(\xi+X_{\sigma+s}^0) \cdot I_{FG}\Big]\Big] \\ &= E\Big[f(\chi_{\tau+t}^{t_0}) \cdot g(X_{\sigma+s}^{t_0}) \cdot I_{FHG}\Big]. \end{split}$$

Theorem 5 For bounded Borel function f, we have

$$E[f(X_{t_0})I_{X_{t_0}\in S(t_0)}|Y_{[0,T]}] = E[f(X_{t_0})I_{X_{t_0}\in S(t_0)}|\tau, \chi_{\tau}^{t_0}, \sigma, X_{\sigma}^{t_0}].$$
(13)

Remark 6 This identity means that if we observe the hidden process up to time T, it is enough for us to know the lastly observed time and position before t_0 and the first reappearing time and position after t_0 to estimate the hidden state X_{t_0} optimally.

Proof. Let $A_i, B_i, i = 1, 2, \dots, n$, and \tilde{A}, \tilde{B} be Borel sets in \mathbb{R}^n and $\partial \mathbb{S}$ respectively. We let $A := \prod_{i=1}^n A_i$ and $B := \prod_{i=1}^n B_i$ and denote

$$\{\tilde{\chi}_{\tau+\mathbf{t}} \in A\} := \{\tilde{\chi}_{\tau+t_1} \in A_1, \cdots, \tilde{\chi}_{\tau+t_n} \in A_n\},\\ \{\tilde{X}_{\sigma+\mathbf{s}} \in B\} := \{\tilde{X}_{\sigma+s_1} \in B_1, \cdots, \tilde{X}_{\sigma+s_n} \in B_n\}.$$

It is enough to prove that both sides of the equation (13) have the same expected values when they are multiplied by $I_{\{(\tau,\chi_{\tau}^{t_0})\in\tilde{A}, \ \tilde{\chi}_{\tau+t}\in A\}} \cdot I_{\{(\sigma,X_{\sigma}^{t_0})\in\tilde{B}, \ \tilde{X}_{\sigma+s}\in B\}}$.

By using Lemma 4 and the strong Markov property, we have

$$E\Big[f(X_{t_0})I_{X_{t_0}\in S(t_0)}\cdot I_{\{(\tau,\chi_{\tau}^{t_0})\in\tilde{A},\ \tilde{\chi}_{\tau+\mathbf{t}}\in A\}}I_{\{(\sigma,X_{\sigma}^{t_0})\in\tilde{B},\ \tilde{X}_{\sigma+\mathbf{s}}\in B\}}\Big]$$

= $E\Big[f(\xi)I_{\xi\in S(t_0)}I_{\{(\tau,\chi_{\tau}^{t_0})\in\tilde{A}\}}I_{\{(\sigma,X_{\sigma}^{t_0})\in\tilde{B}\}}E\Big[I_{\{\tilde{\chi}_{\tau+\mathbf{t}}\in A\}}I_{\{\tilde{X}_{\sigma+\mathbf{s}}\in B\}}\Big|\mathcal{F}_{\tau(\xi)},\xi,\mathcal{G}_{\sigma(\xi)}\Big]\Big]$
= $E\Big[f(\xi)I_{\xi\in S(t_0)}I_{\{(\tau,\chi_{\tau}^{t_0})\in\tilde{A}\}}I_{\{(\sigma,X_{\sigma}^{t_0})\in\tilde{B}\}}P^{\chi_{\tau}^{t_0}}(\tilde{\chi}_{\mathbf{t}}\in A)\cdot P^{X_{\sigma}^{t_0}}(\tilde{X}_{\mathbf{s}}\in B)\Big].$

On the other hand, similarly we have

$$\begin{split} E\Big[E\Big[f(X_{t_{0}})I_{X_{t_{0}}\in S(t_{0})}\Big|\tau,\chi_{\tau}^{t_{0}},\sigma,X_{\sigma}^{t_{0}}\Big]\cdot I_{\{(\tau,\chi_{\tau}^{t_{0}})\in\tilde{A},\ \tilde{\chi}_{\tau+\mathbf{t}}\in A\}}I_{\{(\sigma,X_{\sigma}^{t_{0}})\in\tilde{B},\ \tilde{X}_{\sigma+\mathbf{s}}\in B\}}\Big]\\ &= E\Big[E\Big[f(\xi)I_{\xi\in S(t_{0})}I_{\{(\tau,\chi_{\tau}^{t_{0}})\in\tilde{A}\}}I_{\{(\sigma,X_{\sigma}^{t_{0}})\in\tilde{B}\}}\Big|\tau,\chi_{\tau}^{t_{0}},\sigma,X_{\sigma}^{t_{0}}\Big]\\ &\quad \cdot E\Big[I_{\{\tilde{\chi}_{\tau+\mathbf{t}}\in A\}}I_{\{\tilde{X}_{\sigma+\mathbf{s}}\in B\}}\Big|\mathcal{F}_{\tau(\xi)},\xi,\mathcal{G}_{\sigma(\xi)}\Big]\Big]\\ &= E\Big[E\Big[f(\xi)I_{\xi\in S(t_{0})}I_{\{(\tau,\chi_{\tau}^{t_{0}})\in\tilde{A}\}}I_{\{(\sigma,X_{\sigma}^{t_{0}})\in\tilde{B}\}}\Big|\tau,\chi_{\tau}^{t_{0}},\sigma,X_{\sigma}^{t_{0}}\Big]\\ &\quad \cdot P^{\chi_{\tau}^{t_{0}}}(\tilde{\chi}_{\mathbf{t}}\in A)\cdot P^{X_{\sigma}^{t_{0}}}(\tilde{X}_{\mathbf{s}}\in B)\Big]\\ &= E\Big[f(\xi)I_{\xi\in S(t_{0})}I_{\{(\tau,\chi_{\tau}^{t_{0}})\in\tilde{A}\}}I_{\{(\sigma,X_{\sigma}^{t_{0}})\in\tilde{B}\}}P^{\chi_{\tau}^{t_{0}}}(\tilde{\chi}_{\mathbf{t}}\in A)\cdot P^{X_{\sigma}^{t_{0}}}(\tilde{X}_{\mathbf{s}}\in B)\Big].\end{split}$$

This completes the proof.

Let q(x; t, y; s, z) be the joint density function such that

$$q(x;t,y;s,z) = \frac{P\Big[\xi \in dx; \, (\tau(\xi), \chi^{t_0}_{\tau(\xi)}) \in d(t,y); \, (\sigma(\xi), X^{t_0}_{\sigma(\xi)}) \in d(s,z)\Big]}{dx \, d(t,y) \, d(s,z)},$$
(14)

By conditioning on ξ , we can decompose the joint density such as

$$q(x; t, y; s, z) = p_{t_0}(x)u_x(t, y)v_x(s, z)$$

where $p_{t_0}(x)$ is the density function of X_{t_0} , $u_x(t, y)$ and $v_x(s, z)$ are the first hitting densities defined in section 3.1. The density function $p_{t_0}(x) = p(t_0, x)$ is obtained by solving the Kolmogorov's equation (10), $v_x(s, z)$ is given in section 3.3 and $u_x(t, y)$ is obtained by solving the following backward boundary value problem for the reverse time process χ^{t_0}

$$\frac{\partial u}{\partial t} + \frac{1}{2} \mathbf{Tr}[d^{\mathsf{T}}(t,x)u_{xx}d(t,x)] + u_x^{\mathsf{T}}c(t,x) = 0, \quad (t_0 - t,x) \in \mathbb{S}_{(0,t_0)}$$

with the boundary conditions

$$u(t,x) = \varphi(t,x), \quad \text{for} \quad (t,x) \in \partial \mathbb{S}_{(0,t_0)},$$
$$u(t_0,x) = 0, \quad \text{for} \quad x \in S(t),$$

where $c(t, x) := \bar{a}(t_0 - t, x), d(t, x) := b(t_0 - t, x)$, and $\varphi \in C_0^{\infty}(\partial \mathbb{S})$ is close to a Dirac-delta function on $\partial \mathbb{S}_{(0,t_0)}$ - see [11].

Hence, by the identity (13), we obtain the formula for the interpolator such as

Theorem 7

$$E[f(X_{t_0})I_{X_{t_0}\in S(t_0)}|Y_{[0,T]}] = \int_{S_{t_0}} f(x)q(x|t,y;s,z)dx$$

where $q(x|t, y; s, z) = \frac{p_{t_0}(x)u_x(t, y)v_x(s, z)}{\int_{S_{t_0}} p_{t_0}(r)u_r(t, y)v_r(s, z)dr}$.

Example 1 (Brownian excursion and meander)

Let W_t be standard 1-dimensional Brownian motion starting at 0 and hidden by $S(t) = (0, \infty), t > 0$. Suppose $W_{t_0} \in S(t_0)$ and $\tau = t, \sigma = s$, then the joint density in (14) is

$$\frac{P(X_{t_0} \in dx, \tau(X_{t_0}) \in dt, \sigma(X_{t_0}) \in ds)}{dx dt ds} = p_{t_0}(x)u_x(t, 0)v_x(s, 0),$$
(15)

where

$$u_x(t,0) = \frac{x}{\sqrt{2\pi(t_0 - t)t^3}} e^{-(t_0 - t)x^2/2t}$$
 and $v_x(s,0) = \frac{x}{\sqrt{2\pi s^3}} e^{-x^2/2s}$.

Then, the conditional density for W_{t_0} is evaluated as

$$P[W_{t_0} \in dx | \tau = t, \sigma = s] / dx = \frac{p_{t_0}(x)u_x(t, 0)v_x(s, 0)}{\int_{S_{t_0}} p_{t_0}(z)u_z(t, 0)v_z(s, 0)dz}$$
$$= \sqrt{\frac{2}{\pi}} \sqrt{\frac{(s+t)^3}{s^3 t^3}} x^2 \exp\left\{-\frac{x^2(s+t)}{2st}\right\}.$$

We can derive the same result using Corollary of Theorem 6 in [4].

When $T < t_0 + \sigma$, the joint density for the interpolator is

$$P(W_{t_0} \in dx)P(\tau(x) \in dt)P(t_0 + \sigma(x) > T)/dxdt$$
(16)

where

$$P(t_0 + \sigma(x) > T) = 1 - \int_0^{T-t_0} \frac{x}{\sqrt{2\pi r^3}} e^{-x^2/2r} dr$$
(17)

and we can obtain the corresponding conditional density for the interpolator.

From Theorem 9 (Appendix), we have

$$P(\gamma(T) \in d(t_0 - t), \ Z(t_0) \in dx, \ |W(T)| > 0)/d(-t)dx$$

$$= 2p(t_0 - t, 0, 0)g(t_0, 0, x) \int_0^\infty q(T - t_0; x, y)dy.$$
(18)

A little calculation shows that the last integral in the above line coincides with (17), hence (18) yields the same conditional density as (16). So, we conclude that the interpolator for hidden Brownian motion coincides with Brownian excursion or meander.

The conditional densities for the filters and the interpolators at fixed times are compared in Fig 1. We used Matlab for the numerical simulations in this article. The densities for the interpolator seem to have smaller variance, hence we can infer that they yield more accurate estimates.

4 A Last Exit Decomposition

For Brownian motion, the joint density (15) has an another decomposition which is called the last exit decomposition (LED) as follows:

$$p(t_0 - t, 0, 0)v_x(t, 0)v_x(s, 0).$$

This principle is proved for more general Markov processes in Getoor et al. [6] and [7], and for scalar Brownian motion with smooth time varying boundaries in Salminen [14]. Here we give a generalized version of the result in [14]. This provides us some convenience since in this decomposition, we need not know $u_x(t, y)$.

Theorem 8 Let W_t be n-dimensional Brownian motion. Then, for W_t and set-valued function $S : [0, \infty) \to \mathcal{P}_{cc}(\mathbb{R}^n)$, the joint density in (14) is decomposed as

$$q(x; t, y; s, z) = p_{t_0-t}(y)v_x(t, y)v_x(s, z).$$

Proof. It is enough for us to prove

$$u(x;t,y) := \frac{P\left[\xi \in dx; \ (\tau(\xi), \chi_{\tau(\xi)}^{t_0}) \in d(t,y)\right]}{dx \, d(t,y)} = p_{t_0-t}(y)v_x(t,y), \quad (t,y) \in \partial \mathbb{S}_{(0,t_0)}$$

Let $C \subset S(t_0)$ be an open set in \mathbb{R}^n , and \tilde{A} be an open set in $\partial \mathbb{S}_{(0,t_0)}$, then

$$\begin{split} P_{0}\Big[W_{t_{0}} \in S, (\tau(W_{t_{0}}), \chi_{\tau(W_{t_{0}})}) \in \tilde{A}\Big] \\ &= \int_{C} P_{0}\Big[W_{t_{0}} \in dx, (\tau(W_{t_{0}}), \chi_{\tau(W_{t_{0}})}) \in \tilde{A}\Big] \\ &= \int_{C} P_{\chi^{x}}\Big[(\tau(x), \chi_{\tau(x)}) \in \tilde{A}\Big]P_{0}(W_{t_{0}} \in dx) \\ &= \int_{C} P_{\chi^{x}}\Big[(\tau(x), \chi_{\tau(x)}) \in \tilde{A}|\bar{W}_{t_{0}} = 0\Big]P_{0}(W_{t_{0}} \in dx) \\ &= \int_{C} \lim_{dh \to 0} \frac{P_{\chi^{x}}\Big[(\tau(x), \chi_{\tau(x)}) \in \tilde{A}, \bar{W}_{t_{0}} \in dh\Big]}{p_{t_{0}}(x, 0) \cdot dh} \cdot p_{t_{0}}(0, x)dx \\ &= \int_{C} \lim_{dh \to 0} P_{\chi^{x}}\Big[\bar{W}_{t_{0}} \in dh\Big|(\tau(x), \chi_{\tau(x)}) \in \tilde{A}\Big]/dh \cdot P_{\chi^{x}}\Big[(\tau(x), \chi_{\tau(x)}) \in \tilde{A}\Big]dx \\ &= \int_{C} p_{t_{0}-\tau(x)}(\chi_{\tau(x)}, 0) \cdot P_{\chi^{x}}\Big[(\tau(x), \chi_{\tau(x)}) \in \tilde{A}\Big]dx. \end{split}$$

By the Lebesgue differentiation w.r.t x, the last line above becomes

$$p_{t_0-\tau(x)}(\chi_{\tau(x)},0)\cdot P_{\chi^x}\big[(\tau(x),\chi_{\tau(x)})\in \tilde{A}\big].$$

Again, taking a sequence of balls \tilde{A}_n on $\partial \mathbb{S}_{(0,t_0)}$ that shrinks to (t, y) nicely, we obtain

$$u(x;t,y) = \lim_{n \to \infty} p_{t_0 - \tau(x)}(\chi_{\tau(x)}, 0) \cdot \frac{P_{\chi^x} \left[(\tau(x), \chi_{\tau(x)}) \in \tilde{A}_n \right]}{\mu_n(\tilde{A}_n)}$$

= $p_{t_0 - t}(y) v_x(t, y).$

Example 2 (Scalar Brownian motion hidden by an interval)

Suppose W_t is hidden by the interval S(t) = [0, 1] and $W_1 \in [0, 1]$. Due to Theorem 8, we can obtain the explicit filter and interpolator at each t_0 for this case. Suppose W_t is lastly observed at (2/3, 0) i.e. $(\tau(W_1), \chi_{\tau(W_1)}) = (1/3, 0)$. We compare the conditional density functions at t = 1 for the two decompositions given above.

Let 0 < x < a, W_t^x be Brownian motion starting at x and τ_0 and τ_a be the first hitting times of W_t^x to y = 0 and y = a. Then for t > 0,

$$f_{W^x}^0(t) := P^x(\tau_0 \in dt, \tau_0 < \tau_a)/dt, \quad f_{W^x}^a(t) := P^x(\tau_a \in dt, \tau_a < \tau_0)/dt,$$

$$f_{W^x}^0(t) = \frac{1}{\sqrt{2\pi t^3}} \sum_{n=-\infty}^{\infty} (2na+x) \exp\left\{-\frac{(2na+x)^2}{2t}\right\},$$

$$f_{W^x}^a(t) = \frac{1}{\sqrt{2\pi t^3}} \sum_{n=-\infty}^{\infty} (2na+a-x) \exp\left\{-\frac{(2na+a-x)^2}{2t}\right\}.$$

By Theorem 8, we have

$$u(x, 1/3, 0) = p_{2/3}(0) \cdot f_{W^x}^0(1/3)$$
(19)

and alternatively, we can decompose (19) as in (15), $u(x, 1/3, 0) = p_1(0) \cdot u_x(1/3, 0)$ where $u_x(1/3, 0)$ is the first hitting density of (reverse) Brownian bridge at (1/3, 0). We can obtain approximate $u_x(1/3, 0)$ by solving the partial differential equation

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} - \frac{x}{1-t}\frac{\partial u}{\partial x} = 0, \qquad (t,x) \in [0,1) \times (0,1)$$

with the boundary conditions

$$u(t,0) = \varphi_n^{(1/3)}(t), \quad t \in [0,1),$$

$$u(1,x) = 0, \quad x \in (0,1), \text{ and}$$

$$u(t,1) = 0, \quad t \in [0,1),$$

where

$$\varphi_n^{(1/3)}(t) = \frac{n}{2} I_{\left[\frac{1}{3} - \frac{1}{n}, \frac{1}{3} + \frac{1}{n}\right]}(t).$$

We use finite difference method and the approximate conditional densities are depicted in Fig 2 along with those by (19) for two different (τ, χ^1_{τ}) 's.

5 Optimal Reconstruction of Lost Paths

Example 3 (Scalar Brownian motion hidden by an interval II)

For the same setting as in *Example 2*, we can obtain the interpolated estimates at each time on the hidden interval using Theorem 7. In Fig 3, the reconstructed paths are depicted using the conditional means at each time for the interpolator in Theorem 7. We considered two cases. One is 'bounded' (Brownian) excursion and the other is 'crossing' (Brownian) excursion, i.e. the signal reappears on the other boundary. The second case shows that the simple linear interpolation is not the optimal reconstruction for the hidden crossing excursion. A big discontinuity occurs for the filtered estimates at the reappearing time.

Example 4 (Planar Brownian motion hidden by a quadrant)

Let $X_t = (W_t^{(1)}, W_t^{(2)})$ be 2-dimensional Brownian motion starting at (0,0)and suppose X_{t_0} is hidden by the quadrant $S(t) = [0, \infty) \times [0, \infty)$. We want to reconstruct lost path on $I = (t_0 - \tau, t_0 + \sigma)$ for the following two cases. We let $t_0 + \sigma < T$.

Case 1 : X_t lastly observed at $(y_1, 0)$ and reappeared at $(0, y_2)$; On I, $(W_t^{(1)}|Y_{[0,T]})$ is scalar Brownian motion starting at $(t_0 - \tau, y_1)$ which firstly hits zero at $t = t_0 + \sigma$ i.e. time reversed Brownian meander with fixed end point. And $(W_t^{(2)}|Y_{[0,T]})$ on I is the Brownian meander starting from $(t_0 - \tau, 0)$ with fixed end point $(t_0 + \sigma, y_2)$. We can evaluate the each mean path using the density function in Corollary 10 (Appendix).

Case 2 : X_t lastly observed at $(y_1, 0)$ and reappeared at $(y_2, 0)$; On I, $(W_t^{(1)}|Y_{[0,T]})$ behaves as Brownian bridge from $(t_0 - \tau, y_1)$ to $(t_0 + \sigma, y_2)$ with lower bound zero and its mean path is evaluated using Corollary 11. $(W_t^{(2)}|Y_{[0,T]})$ behaves as standard Brownian excursion on I, whose mean path is given as

$$E(W_t^{(2)}|Y_{[0,T]}) = \frac{2}{\sqrt{\pi}} \sqrt{\frac{2(t-t_0+\tau)(t_0-t+\sigma)}{\tau+\sigma}}, \qquad t \in I.$$

 $E(X_t|Y_{[0,T]})$ is depicted in Fig 4 for both two cases, where R is the sojourn time in the obstacle.

5.1 Finite Difference Method

We illustrate the algorithm by FDM by an example. Consider the simple obstacle $S(t) = (0, \infty), t > 0$ and the Ornstein-Uhlenbeck process X_t satisfying the following equation:

$$dX_t = X_t dt + dW_t, \quad X_0 = 1,$$
 (20)

Note that the coefficients in (20) satisfies the 4 conditions in Section 2. Suppose X_t is lastly observed at (1,0) and firstly reappeared at (3,0). We want to reconstruct the lost path.

Since (20) has the explicit solution

$$X_t = e^{-t} \int_0^t e^s dW_s, \quad t > 0,$$

 X_t is a zero mean Gaussian process with variance $V(t) = \frac{1}{2}(1 - e^{-2t})$. Hence

$$p(t,x) = \frac{1}{\sqrt{\pi(1-e^{-2t})}} e^{-x^2/(1-e^{-2t})}.$$
(21)

The PDE's for $u_x(t, y)$ and $v_x(s, z)$ is given as

$$\frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} + x \left(1 - \frac{2}{1 - e^{-2(3-t)}}\right) \frac{\partial u}{\partial x} = 0, \text{ and}$$
$$\frac{\partial v}{\partial t} + \frac{1}{2} \frac{\partial^2 v}{\partial x^2} - x \frac{\partial v}{\partial x} = 0, \text{ for } (t, x) \in [1, 3] \times [0, \infty).$$

We take 8 instead of ∞ and equally divide the rectangle $[1,3) \times [0,8)$ by 8,000× 320. We evaluate two matrices U and V of the same size $m \times n$ where n = 8001 and m = 321 for the functions u and v. As boundary conditions, we impose 10,000 on U(1,8001) for the Dirac- δ function and evaluate U by backward finite difference method. Likewise, we evaluate V backward by imposing 10,000 on the other time end point. The elements of U and V are arrayed in an unusual way as follows:

$$U(V) = \begin{bmatrix} u_{m,n}(v_{m,1}) \cdots & u_{m,1}(v_{m,n}) \\ \vdots & \ddots & \vdots \\ u_{1,n}(v_{1,1}) & \cdots & u_{1,1}(v_{1,n}) \end{bmatrix},$$

where $u_{1,n} = v_{1,n} = 10,000$. For each time discretization point $t_i \in [1,3]$ we use $p(t_i, x)$ in (21), *i*th column of V, and (8001 - i)th column of U to evaluate the interpolated estimate in Theorem 7. The estimated values are plotted in Fig 5.

5.2 Monte Carlo Method by Simulation of Conditional Brownian Motions

When we reconstruct the whole hidden path, we evaluated conditional expectations at each time point using complicated conditional densities. For Brownian motion, we can generate various conditioned paths by some transforms of Brownian motion and Brownian bridge. For example, let B_t^i , i = 1, 2, 3 be 3 independent copies of Brownian bridge from (0,0) to (1,0). Then

$$B_t^{me,r} = \sqrt{(rt + B_t^1)^2 + (B_t^2)^2 + (B_t^3)^2}, \quad 0 \le t \le 1$$

is the Brownian meander with fixed end point (1, r) which is needed in *Example* 4. We can also generate Brownian excursion path by transposing the pre-minimum part and post-minimum part of Brownian bridge as can be seen in Vervaat [16]. For other transforms, see [2] and references therein. Hence, we can get approximate estimates by Monte Carlo methods using these transformed paths. Since

we are interested in the conditional mean or expectation of some functionals of the process, we need not evaluate the conditional density at each time in this approach. When one of the boundary in *Example 3* is not a constant, this method is more useful. We can choose out paths which satisfy given boundary conditions.

For example, consider the obstacle S(t) = (0, 4t/5), t > 0 and W_t is lastly observed at (1,0) and reappeared at (2,0). We took a step-size $1/2^9$ and generated 5000 excursion paths. 2842 paths are discarded which do not obey the upper boundary. The approximate estimates are compared with those by finite difference method. See Fig 6.

Appendix : Some f.d.d.'s for conditional Brownian motion

Consider the standard Brownian meander on $(\gamma(t_0), t_0)$ where $\gamma(t_0) = t_0 - \tau$. Let $Z(u) = |W(\gamma(t_0) + u)|$. We quote the finite dimensional distribution of Brownian meander given in Theorem 4 in Chung [4].

Theorem 9 Let $m \ge 1$, $0 < u_1 < u_2 < ... < u_m < t_0 - t < t_0$ and $y_1, ..., y_{m+1}$ be arbitrary positive numbers. We have

$$P(\gamma(t_0) \in dt; Z(u_1) \in dy_1, ..Z(u_m) \in dy_m; |W(t_0)| \in dy_{m+1})$$

 $= 2p(t,0,0)dt \ g(u_1;0,y_1)dy_1 \ q(u_2-u_1;y_1,y_2)dy_2 \ ...q(u_m-u_{m-1};y_{m-1},y_m)dy_m \\ \cdot q(t_0-t-u_m;y_m,y_{m+1})dy_{m+1}.$

where p(t; x, y) is the transition density of W_t , $g(t; 0, y) = \frac{|y|}{\sqrt{2\pi t^3}} e^{-y^2/2t}$, and q(t; x, y) = p(t; x, y) - p(t; x, -y).

We immediately have

Corollary 10 (Brownian meander with fixed end point) For $0 < u < t_0$,

$$P(Z(u) \in dy | \gamma(t_0) \in dt, |W(t_0)| \in dy_1) = \frac{g(u; 0, y) \cdot q(t_0 - t - u; y, y_1)}{g(t_0 - t; 0, y_1)} \, dy.$$

Corollary 11 (Brownian bridge with a lower bound) Let B_t be Brownian bridge from $(0, y_1)$ to (t_0, y_2) .

$$P(B_u \in dy | B_u > 0 \text{ for } u \in [0, 1]) = \frac{q(u; y_1, y) \cdot q(t_0 - t - u; y, y_2)}{q(t_0; y_1, y_2)} \, dy$$

References

- B. D. O. Anderson, Reverse-time diffusion equation models. Stochastic Process. Appl. 12, no. 3, (1982) 313–326.
- [2] J. Bertoin, J. Pitman, and J. R. De Chavez, Constructions of a Brownian path a given minimum, Electron. Comm. Probab., (1999) 31–37.
- [3] D. A. Castanon, Reverse-time diffusion processes, IEEE Trans. Inform. Theory., IT28, no. 3, (1982) 953–956.
- [4] K. L. Chung, Excursions in Brownian motion, Ark. Mat., 14, (1976) 213–313.
- [5] H. Dym, and H. P. McKean, Gaussian Processes, Function Theory, and Inverse spectral problems, AP (1976).
- [6] R. K. Getoor and M. J. Sharpe, Last exit decompositions and distributions, Indiana Univ. Math. J., 23, (1973) 377–404.
- [7] R. K. Getoor, Excursions of a markov process, Ann. Probab., 7, (1979) 244–266.
- [8] A. J. E. M. Jannsen, N. J. V. Raymond and B. V. Lodewijk, Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes, IEEE Tran. Acous. Speech. Signal Processing, vol. ASSP-34, No 2, (1986) 317–329.
- [9] I. Karatzas and S. E. Shreve, Brownian Motion and Stochastic Calculus, Springer (1991).
- [10] H. G. Kim, Filtering of hidden diffusion processes, Stochastics and Stochastics Rep., vol 71, (2001) 217–226.
- [11] H. G. Kim and D. Nam, Optimal estimation of diffusion processes hidden by general obstacles, J. Appl. Probab., 38, (2001) 1067–1073.
- [12] J. L. Lewis and Margaret A. M. Murray, The method of layer potentials for the hit equation in time-varing domains, Mem. Amer. Math. Soc., 114, no. 545 (1995).
- [13] T. Misawa, Stochastic dynamical interpolation formula to time reversible diffusion processes J. Math Phys., 34, (1993) 775–784.
- [14] P. Salminen, On the first-hitting time and the last exit time for a Brownian motion to/from a moving boundary, Adv. in Appl. Probab., 20, (1988) 411–426.
- [15] L. Takacs, Bernoulli excursion and its various applications, Adv. in Appl. Probab., 23, (1991) 557–585.
- [16] W. Vervaat, A relation between Brownian bridge and Brownian excursion, Ann. Probab., 7, No. 1, (1979) 143–149.
- [17] J. C. Zambrini, Euclidean quantum mechanics, Phys. Rev. A (3), 35, no. 9, (1987) 3631–3649.



Fig. 1. Conditional densities for filter and interpolator; Lastly observes at (1,0) & reappeared at (2,0).



Fig. 2. Conditional densities; FDM vs. analytic solutions by LED



Fig. 3. Bounded & crossing Brownian excursion



Fig. 4. Mean of the hidden parts of Brownian motion in a quadrant



Fig. 5. Interpolated estimates by FDM - Mean of B & O-U excursions



Fig. 6. FDM vs. Monte Carlo method – a moving boundary