Constructing Markov models for barrier options

Gerard Brunick
joint work with Steven Shreve

Department of Mathematics
University of Texas at Austin

Nov. 14th, 2009
3rd Western Conference on Mathematical Finance
UCSB - Santa Barbara, CA

Outline

Introduction

This is really a talk about "Markovian projection" or constructing Markov mimicking processes.

Main point: It often possible to construction Markov processes which mimick properties of more general non-Markovian processes.

This can be useful for a number of reasons.

1. Difficult and expensive to compute with non-Markovian models or models of large dimension
2. To determine the correct "nonparametric form" for a given application
3. As a tool to understand the general model (calibration) application (which models allow "perfect calibration")

Introduction

Local volatility is a "mimicking result."

Consider a linear pricing model where the risk-neutral dynamics of the stock price are given by

$$\mathrm{d}S_t = \sigma_t \, S_t \, \mathrm{d}W_t,$$

for some process σ.

There is often is local volatility model where the risk neutral dynamics of the stock price are given by:

$$\mathrm{d}\widehat{S}_t = \widehat{\sigma}(t, \widehat{S}_t) \, \widehat{S}_t \, \mathrm{d}W_t$$

with the same European option prices.

Local Volatility

Why are local volatility models attractive?

- simple dynamics
- low dimensional Markov process
- general enough to allow for "perfect calibration" to wide range of option prices
- "Markovian projection" - one can use the local volatility model to characterize the set models consistent with a given set of prices

The local volatility function $\widehat{\sigma}$.

- Dupire (1994) as well as Derman & Kani (1994)

$$\widehat{\sigma}^2(t,x) = \frac{\frac{\partial}{\partial T} C(t,x)}{\frac{1}{2} x^2 \frac{\partial^2}{\partial K^2} C(t,x)}$$

- Gyöngy (1986), Derman & Kani (1998) as well as Britten-Jones & Neuberger (2000). If

$$\widehat{\sigma}^2(t,x) = \mathbb{E}\big[\sigma_t^2 \bigm| S_t = x\big],$$

then $\mathrm{d}\widehat{S}_t = \widehat{\sigma}(t,\widehat{S}_t)\,\widehat{S}_t\,\mathrm{d}W_t$ has the some one-dimensional marginal distributions as $\mathrm{d}S_t = \sigma_t\,S_t\,\mathrm{d}W_t$.

Local Volatility

The relationship between

- European option prices and the
- 1-dimensional risk-neutral marginals of the underlying asset

has been understood since at least Breeden and Litzenberger (1978).

If $C(T,K)$ denotes the price of a European call option with maturity T and strike K and $p(t,x) = \mathbb{P}[S_t \in \mathrm{d}x]$, then

$$\begin{aligned} \frac{\partial^2}{\partial K^2} C(T,K) &= \frac{\partial^2}{\partial K^2} \int (x-K)^+ p(T,x)\,\mathrm{d}x \\ &= \int \delta(x-K)\, p(T,x)\,\mathrm{d}x \\ &= p(T,K) \end{aligned}$$

Krylov (1984) and Gyöngy (1986)

Theorem

Let W be an $\mathbb{R}^r$-valued Brownian motion, and let X solve

$$\mathrm{d}X_t = \mu_s \,\mathrm{d}s + \sigma_s \,\mathrm{d}W_s,$$

where

1. *μ is a bounded, $\mathbb{R}^d$-valued, adapted process, and*
2. *σ is a bounded, $\mathbb{R}^{d\times r}$-valued, adapted process such that $\sigma\sigma^T$ is uniformly positive definite (i.e., there exists $\lambda. > 0$ with $x^T\sigma_t\sigma_t^T x \geq \lambda\|x\|$ for all $t \in \mathbb{R}_+$ and $x \in \mathbb{R}^d$).*

Krylov (1984) and Gyöngy (1986)

Theorem

If the conditions on the last slide are met by

$$\mathrm{d}X_t = \mu_s \,\mathrm{d}s + \sigma_s \,\mathrm{d}W_s,$$

then there exists a weak solution to the SDE:

$$\mathrm{d}\widehat{X}_t = \widehat{\mu}(t, \widehat{X}_t)\,\mathrm{d}t + \widehat{\sigma}(t, \widehat{X}_t)\,\mathrm{d}\widehat{W}_t$$

where

1. $\widehat{\mu}(t, X_t) = \mathbb{E}[\,\mu_t \mid X_t]$ *for Lebesgue-a.e.* t*,*
2. $\widehat{\sigma}\,\widehat{\sigma}^T(t, X_t) = \mathbb{E}[\sigma_t\,\sigma_t^T \mid X_t]$ *for Lebesgue-a.e.* t*, and*
3. $\widehat{X}_t$ *has the same distribution as* X_t *for each fixed* t*.*

General Mimicking Results

1. Given a (non-Markov) Ito process it is possible to find a mimicking process which preserves the distributions of a number of running statistics about the process.

2. If futher technical conditions are met, the mimicking Itô process "drives" a Markov process whose dimension is equal to the number of running statistics.

3. To understand the kinds of running statistics that can be preserved, we need to introduce the notion of an updating function.

Some Notation

We let $C_0(\mathbb{R}_+;\mathbb{R}^d)$ denotes the paths in $C(\mathbb{R}_+;\mathbb{R}^d)$ that start at zero, and we let

$$\Delta : C(\mathbb{R}_+,\mathbb{R}^d)\times\mathbb{R}_+ \to C_0(\mathbb{R}_+,\mathbb{R}^d)$$

denote the map such that

$$\Delta_u(x,t) = x(t+u) - x(t)$$

So $\Delta(x,t)$ is the path in $C_0(\mathbb{R}_+,\mathbb{R}^d)$ that corresponds to the changes x after the time t.

Updating Functions

Definition

Let $\mathcal{E}$ be a Polish space, and let $\Phi : \mathcal{E} \times C_0(\mathbb{R}_+; \mathbb{R}^d) \to C(\mathbb{R}_+; \mathcal{E})$ be a function. We say that Φ is an **updating function** if

1. $x(s) = y(s)$ for all $s \in [0, t]$ implies that $\Phi_s(e, x) = \Phi_s(e, y)$ for all $s \in [0, t]$, and

2. $\Phi_{t+u}(e, x) = \Phi_u\big(\Phi_t(e, x), \Delta(x, t)\big)\ \ \forall t, u \in \mathbb{R}_+$.

If Φ is also continuous as map from $\mathcal{E} \times C_0(\mathbb{R}_+; \mathbb{R}^d)$ to $C(\mathbb{R}_+; \mathcal{E})$, then we say that Φ is a **continuous updating function**.

Example: Process Itself

A trivial updating function: take $\mathcal{E} = \mathbb{R}^d$, and

$$\Phi(e, x) = e + x, \quad e \in \mathbb{R}^d,\ x \in C_0^d,$$

so $X_t = \Phi_t\big(X_0, \Delta(X, t)\big)$.

The updating property reads

$$X_{t+u} = X_t + \Delta_u(X, t)$$

So Φ_{t+u} is function of Φ_t and $\Delta(X, t)$.

Example: Process and Running Max

Let $\mathcal{E} = \{(x, m) \in \mathbb{R}^2 : x \leq m\}$.

x Process position

m Maximum-to-date

Given $x, m \in \mathcal{E}$ and changes $y \in C_0(\mathbb{R}_+; \mathbb{R}^d)$, we update the current location and current maximum-to-date by:

$$\Phi_t(x, m; y) = \big(x + y(t),\ m \vee \max_{0 \leq s \leq t} \big(x + y(s)\big)\big).$$

Example: Process and Running Max

If we take $M_t = \max_{s \leq t} X_t$, then we have

$$\Phi_t\big(X_0, X_0; \Delta(X, 0)\big) = (X_t, M_t)$$

The second property in the definition of updating function amounts to

$$(X_{t+u}, M_{t+u}) = \Big(X_t + \Delta_u(X, t),$$
$$M_t \vee \max_{s \leq u}\big(X_t + \Delta_s(X, t)\big)\Big)$$

So Φ_{t+u} is function of Φ_t and $\Delta(X, t)$.

Example: Entire History

Take

$$\mathcal{E} = \big\{(x,s) \in C(\mathbb{R}_+;\mathbb{R}^d) \times \mathbb{R}_+; x \text{ is constant on } [s,\infty)\big\}.$$

Given an initial path segment $(x,s) \in \mathcal{E}$ and changes $y \in C_0(\mathbb{R}_+;\mathbb{R}^d)$, let $(x,s) \oplus y$ denote the path obtained by appending y to x after time s:

$$\big((x,s) \oplus y\big)(t) = \begin{cases} x(t) & \text{if } t \leq s, \text{ and} \\ x(s) + y(t-s) & \text{if } t > s. \end{cases}$$

Then $\Phi_t(x,s;y) = \big((x,s) \oplus y^t, s+t\big)$ is an updating function, where y^t is the path y stopped at time t.

Example: Entire History

With

$$\mathcal{E} = \big\{(x,s) \in C(\mathbb{R}_+;\mathbb{R}^d) \times \mathbb{R}_+; x \text{ is constant on } [s,\infty)\big\}, \text{ and}$$
$$\Phi_t(x,s;y) = \big((x,s) \oplus y^t, s+t\big),$$

we have $\Phi_t(X^0, 0; \Delta(X,0)) = (X^t, t)$, so Φ tracks the whole path history.

The updating property amounts to

$$(X^{t+u}, t+u) = \big((X^t, t) \oplus \Delta^u(X,t), t+u\big),$$

so again Φ_{t+u} is a function of Φ_t and $\Delta(X,t)$.

General Mimicking Result (B. and Shreve)

Let Y be a $\mathbb{R}^d$-valued process with

$$Y_t \triangleq \int_0^t \mu_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s,$$

where W be an $\mathbb{R}^r$-valued B.M. and μ and σ be an adapted processes with

$$\mathbb{E}\left[\int_0^t \|\mu_s\| + \|\sigma_s \sigma_s^T\| \,\mathrm{d}s\right] < \infty \quad \forall t \in \mathbb{R}_+, \tag{1}$$

Let $\mathcal{E}$ be a Polish space, and let Z be a continuous, $\mathcal{E}$-valued process with $Z = \Phi(Z_0, Y)$ for some continuous updating function Φ.

(Z tracks the running statistics of Y that we care about.)

General Mimicking Result (B. and Shreve)

Then there exists a weak solution to the stochastic system

$$\widehat{Y}_t = \int_0^t \widehat{\mu}(s, \widehat{Z}_s)\,\mathrm{d}t + \int_0^t \widehat{\sigma}(s, \widehat{Z}_s)\,\mathrm{d}\widehat{W}_s, \text{ and}$$
$$\widehat{Z}_t = \Phi(\widehat{Z}_0, \widehat{Y}),$$

where

1. $\widehat{\mu}(t, z) = \mathbb{E}[\mu_t \,|\, Z_t = z]$ a.e. t,
2. $\widehat{\sigma}\widehat{\sigma}^T(t, z) = \mathbb{E}[\sigma_t\,\sigma_t^T \,|\, Z_t = z]$, a.e. t, and
3. $\widehat{Z}_t$ has the same law as Z_t for each t.

Corollary: Process Itself

Suppose X solves

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t$$

and the integrability condition (1) is satisfied.

Then there exists a weak solution to

$$\mathrm{d}\widehat{X}_t = \widehat{\mu}(t, \widehat{X}_t)\mathrm{d}t + \widehat{\sigma}(t, \widehat{X}_t)\mathrm{d}W_t$$

where

1. $\widehat{\mu}(t,x) = \mathbb{E}[\mu_t \mid X_t = x]$ a.e. t,
2. $\widehat{\sigma}\widehat{\sigma}^T(t,x) = \mathbb{E}[\sigma_t \, \sigma_t^T \mid X_t = x]$, a.e. t, and
3. $\widehat{X}_t$ has the same law as X_t for each t.

Corollary: Process and Running Max

Suppose X solves

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t,$$

$M_t = \sup_{s \leq t} X_s$, and the integrability condition (1) is satisfied.

Then there exists a weak solution to

$$\mathrm{d}\widehat{X}_t = \widehat{\mu}(t, \widehat{X}_t, \widehat{M}_t)\mathrm{d}t + \widehat{\sigma}(t, \widehat{X}_t, \widehat{M}_t)\mathrm{d}\widehat{W}_t,$$
$$\widehat{M}_t = \max_{s \leq t} \widehat{X}_t,$$

where

1. $\widehat{\mu}(t, x, m) = \mathbb{E}[\mu_t \,|\, X_t, M_t = x, m]$ a.e. t,
2. $\widehat{\sigma}\widehat{\sigma}^T(t, x, m) = \mathbb{E}[\sigma_t \, \sigma_t^T \,|\, X_t, M_t = x, m]$, a.e. t, and
3. $(\widehat{X}_t, \widehat{M}_t)$ has the same law as (X_t, M_t) for each t.

Main Idea of Proof

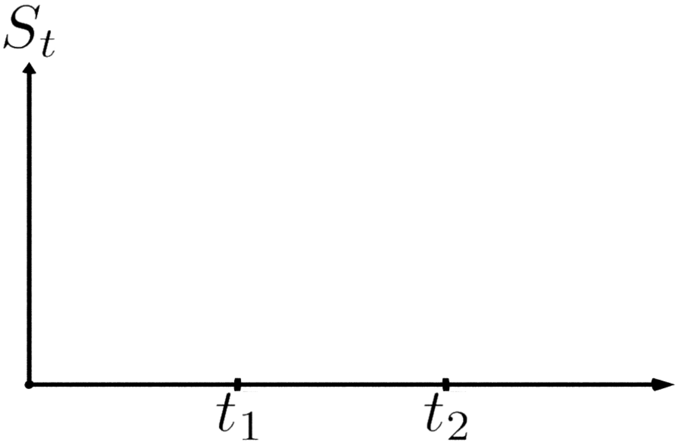
Let S be an Itô process S that solves $\mathrm{d}S_t = \sigma_t\, S_t\, \mathrm{d}W_t$.

We construct processes $\boldsymbol{S^1}$, $\boldsymbol{S^2}$, and $\boldsymbol{S^3}$ on some space with $\mathscr{L}(\boldsymbol{S^1}) = \mathscr{L}(\boldsymbol{S^2}) = \mathscr{L}(\boldsymbol{S^3}) = \mathscr{L}(S)$.

We then piece these processes together to form a process $\widetilde{S}$ with $\mathscr{L}(\widetilde{S}_t) = \mathscr{L}(S_t)$ for all t.

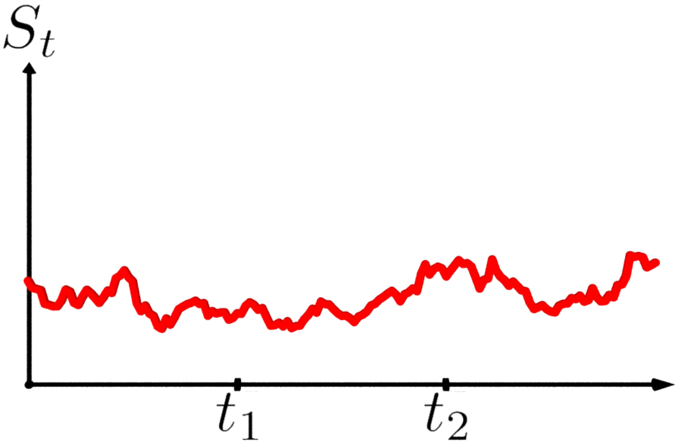
Main Idea of Proof

Suppose S solves $\mathrm{d}S_t = \sigma_t\, S_t\, \mathrm{d}W_t$.



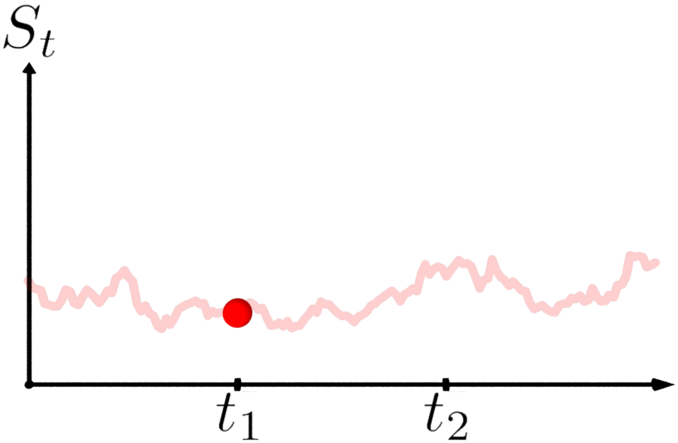
Main Idea of Proof

$$\text{Let } \mathscr{L}(\boldsymbol{S^1}) = \mathscr{L}(S).$$



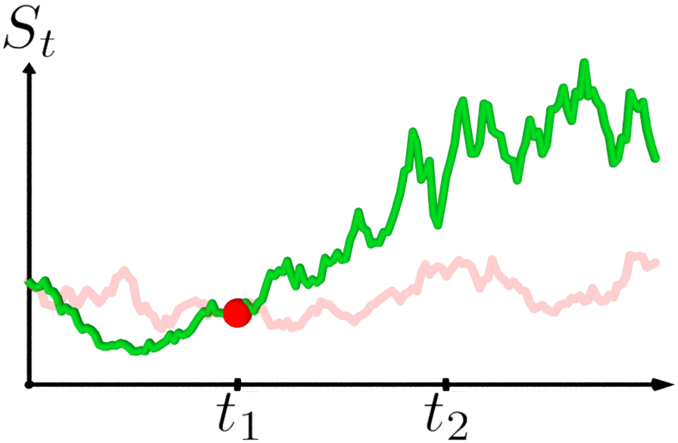
Main Idea of Proof

Forget everything about S^1 except $S^1_{t_1}$.



Main Idea of Proof

$$\text{Let } \mathscr{L}(\boldsymbol{S^2} \mid \boldsymbol{S^1_{t_1}}) = \mathscr{L}(S \mid S_{t_1} = \boldsymbol{S^1_{t_1}}).$$



Main Idea of Proof

Let $\mathscr{L}(\boldsymbol{S^2} \mid \boldsymbol{S^1_{t_1}}) = \mathscr{L}(S \mid S_{t_1} = \boldsymbol{S^1_{t_1}})$.

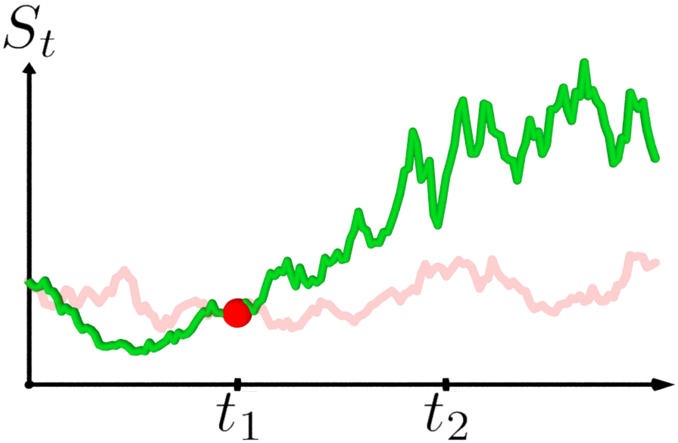
Taking any measurable $A \subset C(\mathbb{R}_+; \mathbb{R})$, notice that

$$\begin{aligned}
\mathbb{P}[\boldsymbol{S^2} \in A] &= \int_{\mathbb{R}} \mathbb{P}[\boldsymbol{S^2} \in A \,|\, \boldsymbol{S^1_{t_1}} = x]\, \mathbb{P}[\boldsymbol{S^1_{t_1}} \in \mathrm{d}x] \\
&= \int_{\mathbb{R}} \mathbb{P}[S \in A \,|\, S_{t_1} = x]\, \mathbb{P}[S_{t_1} \in \mathrm{d}x] \\
&= \mathbb{P}[S \in A].
\end{aligned}$$

In particular, $\boldsymbol{S^2}$ is distributed according to $\mathscr{L}(S)$.

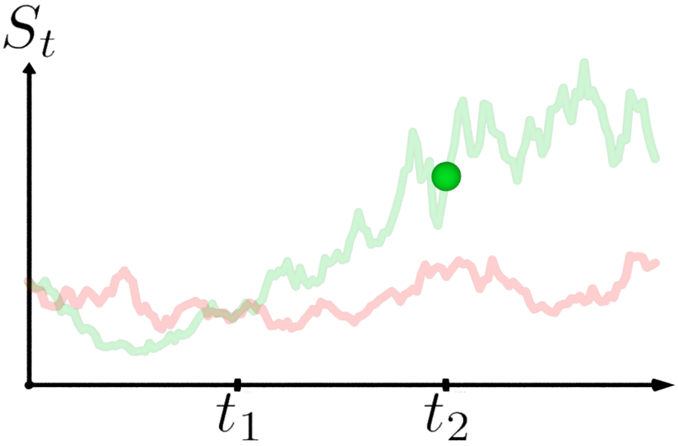
Main Idea of Proof

$$\text{Let } \mathscr{L}(\boldsymbol{S^2} \mid \boldsymbol{S^1_{t_1}}) = \mathscr{L}(S \mid S_{t_1} = \boldsymbol{S^1_{t_1}}).$$



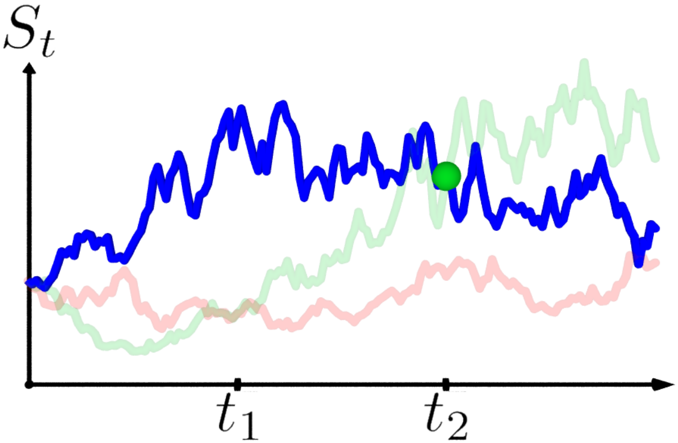
Main Idea of Proof

Forget everything about S^2 except $S^2_{t_2}$.



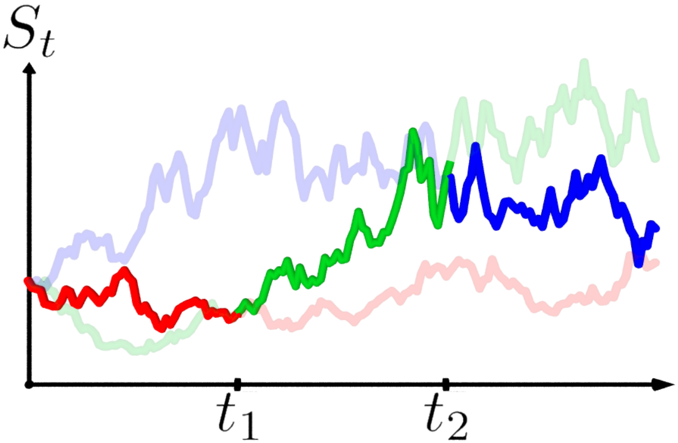
Main Idea of Proof

Let $\mathscr{L}(\boldsymbol{S^3} \mid \boldsymbol{S^1_{t_2}}) = \mathscr{L}(S \mid S_{t_2} = \boldsymbol{S^1_{t_2}})$.



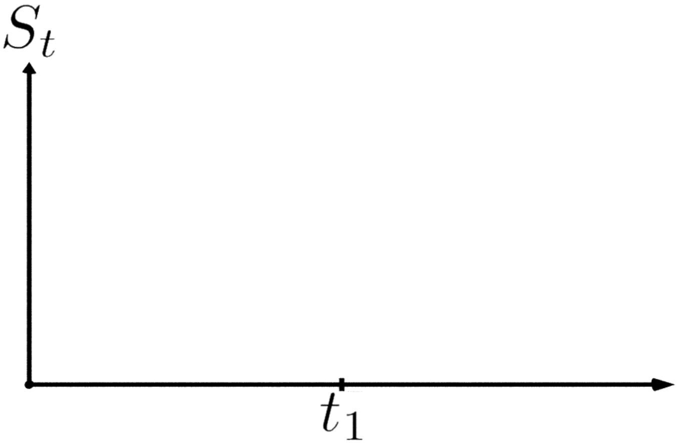
Main Idea of Proof

Set $\widehat{S} \triangleq \boldsymbol{S^1}\,\mathbf{1}_{[0,t_1)} + \boldsymbol{S^2}\,\mathbf{1}_{[t_1,t_2)} + \boldsymbol{S^3}\,\mathbf{1}_{[t_2,\infty)}$.



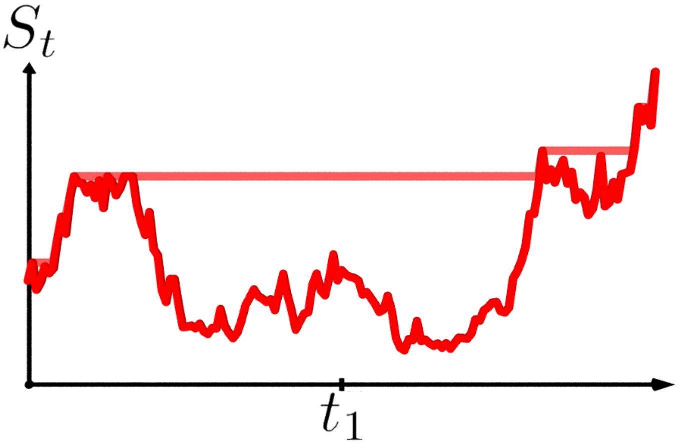
Main Idea of Proof

This still works when we track additional information.



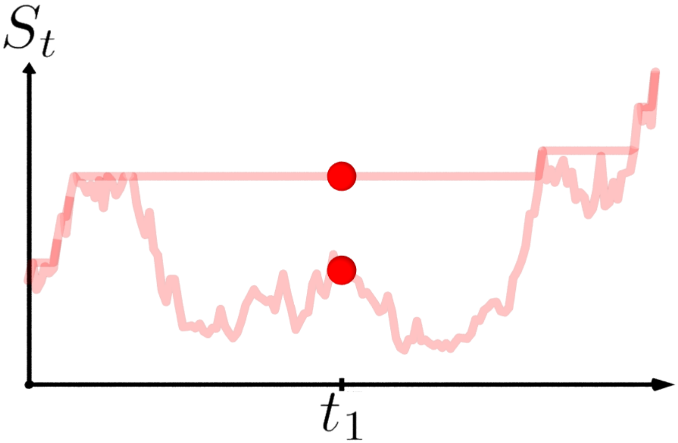
Main Idea of Proof

Let $\mathscr{L}(S^1) = \mathscr{L}(S)$.



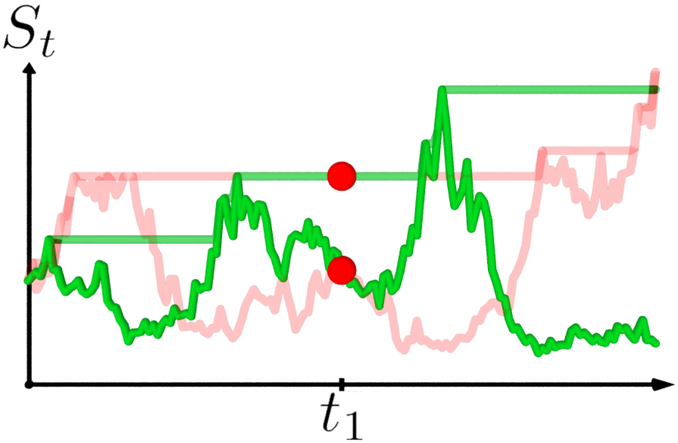
Main Idea of Proof

Forget everything about S^1 except $S^1_{t_1}$ and $M^1_{t_1}$.



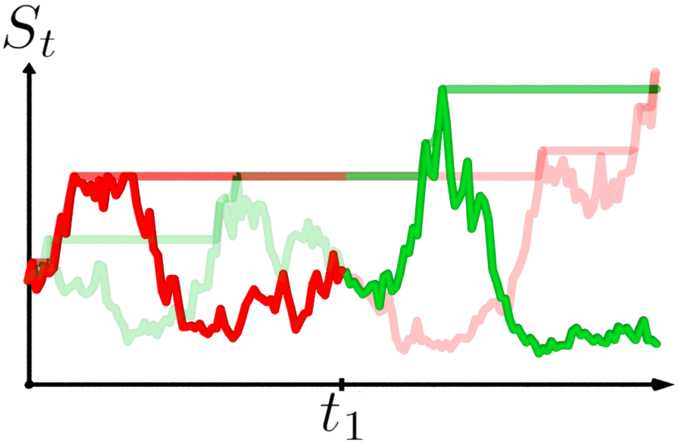
Main Idea of Proof

Let $\mathscr{L}(\boldsymbol{S^2} \mid \boldsymbol{S^1_{t_1}}, \boldsymbol{M^1_{t_1}}) = \mathscr{L}(S \mid S_{t_1} = \boldsymbol{S^1_{t_1}}, M_{t_1} = \boldsymbol{M^1_{t_1}})$.



Main Idea of Proof

$$\text{Set } \widetilde{S} \triangleq \boldsymbol{S^1}\,\mathbf{1}_{[0,t_1)} + \boldsymbol{S^2}\,\mathbf{1}_{[t_1,\infty)}.$$



General Mimicking Result (B. and Shreve)

Then there exists a weak solution to the stochastic system

$$\widehat{Y}_t = \int_0^t \widehat{\mu}(s, \widehat{Z}_s)\,\mathrm{d}t + \int_0^t \widehat{\sigma}(s, \widehat{Z}_s)\,\mathrm{d}\widehat{W}_s, \text{ and}$$
$$\widehat{Z}_t = \Phi(\widehat{Z}_0, \widehat{Y}),$$

where

1. $\widehat{\mu}(t, z) = \mathbb{E}[\mu_t \,|\, Z_t = z]$ a.e. t,
2. $\widehat{\sigma}\widehat{\sigma}^T(t, z) = \mathbb{E}[\sigma_t\, \sigma_t^T \,|\, Z_t = z]$, a.e. t, and
3. $\widehat{Z}_t$ has the same law as Z_t for each t.

Example: Barrier Options

Definition

Given an exercise time, T, an upper barrier, U, and strike, K, the holder of an **up-and-out call option** has the right to exercise a call option at time T with strike K if the stock price has remained below the barrier U. If the stock price crosses the barrier, the option becomes worthless.

Calibration Problem

Given a collection $\{B(T, U, K)\}_{T,U,K}$ of prices for up-and-out call options, we would like to construct a linear pricing model which is consistent with these prices.

Example: Barrier Options

Previous results suggest that we may want to look for a (risk-neutral) model of the form:

$$\begin{aligned} \mathrm{d}S_t &= \sigma(t, S_t, M_t) S_t \, \mathrm{d}W_t \\ M_t &= \max_{s \le t} S_t, \end{aligned}$$

with σ choosen so that

$$\mathbb{E}\left[\mathbf{1}_{\{M_T \le U\}} \left(S_T - K \right)^+ \right] = B(T, L, K).$$

Dupire Formula

Formally, we may recover σ from the prices of corridor options with a Dupire-type formula.

$$\begin{aligned}
\mathrm{B}(T,K,U) &= \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{M_T \le U\}}(S_T - K)^+\right] \\
\frac{\partial\,\mathrm{B}(T,K,U)}{\partial U} &= \mathbb{E}^{\mathbb{Q}}\left[\delta_U(M_T)(S_T - K)^+\right] \\
\frac{\partial^2\,\mathrm{B}(T,K,U)}{\partial T \partial U} &= \mathbb{E}^{\mathbb{Q}}\left[\tfrac{1}{2}\sigma^2(T,K,U)K^2\delta_U(M_T)\delta_K(S_T)^+\right] \\
\frac{\partial^3\,\mathrm{B}(T,K,U)}{\partial K^2 \partial U} &= \mathbb{E}^{\mathbb{Q}}\left[\delta_U(M_T)\delta_K(S_t)\right]
\end{aligned}$$

So

$$\sigma^2(T,K,U) = \frac{2\partial^2 B(T,K,U)/\partial T \partial U}{\partial^3 B(T,K,U)/\partial K^2 \partial U}$$

Markov Property

Theorem

Let $\mathcal{E}$ be a Polish space and let Φ be a continuous updating function Φ.

Consider the stochastic differential equation:

$$\widehat{Y}_t = \int_{t_0}^{t} \widehat{\mu}(s, \widehat{Z}_s)\,\mathrm{d}t + \int_{t_0}^{t} \widehat{\sigma}(s, \widehat{Z}_s)\,\mathrm{d}\widehat{W}_s, \text{ and}$$
$$\widehat{Z}_t = \Phi(\widehat{Z}_{t_0}, \widehat{Y}).$$

If weak uniqueness holds for each initial condition $Z_{t_0} = z_0 \in \mathcal{E}$, then the process Z is strong Markov.

Markov Property

Corollary

Suppose σ is Lipshitz continuous, then weak uniqueness holds for the stochastic differential equation

$$\mathrm{d}S_t = \sigma(t, S_t, M_t)\,\mathrm{d}W_t$$
$$M_t = \max_{s \leq t} S_t,$$

and the process $Z = (S, M)$ is strong Markov.

Conclusions

- It is often possible to construct reduced form models which preserve the prices of path-dependent options.
- Weak uniqueness results allow one to conclude that the reduced form models are Markov.

Open Question?

Let σ be continuous with $1/C \leq \sigma \leq C$ for some constant C. Is this sufficient to ensure weak uniqueness for the stochastic differential equation:

$$\begin{aligned} \mathrm{d}X_t &= \sigma(t, X_t, M_t)\,\mathrm{d}W_t \\ M_t &= \max_{s \leq t} X_t? \end{aligned}$$

References I

M. Britten-Jones and A. Neuberger.
Option prices, implied price processes, and stochastic volatility.
The Journal of Finance, 55(2):839–866, 2000.

E. Derman and I. Kani.
Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility.
International Journal of Theoretical and Applied Finance, 1(1):61–110, 1998.

B. Dupire.
Pricing with a smile.
Risk, 7(1):18–20, 1994.

References II

I. Gyöngy.
Mimicking the one-dimensional marginal distributions of processes having an Itô differential.
Probability Theory and Related Fields, 71(4):501–516, 1986.

N. V. Krylov.
Once more about the connection between elliptic operators and Itôs stochastic equations.
Statistics and Control of Stochastic Processes, Steklov Seminar, pages 214–229, 1984.