

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI  
Corso di Laurea Magistrale in Matematica

# Equazione Stocastica di McKean e Particle Method

Tesi di Laurea in Equazioni Differenziali Stocastiche

Relatore:  
Chiar.mo Prof.  
ANDREA PASCUCCI

Presentata da:  
ANDREA RINALDI

Sessione II  
Anno Accademico 2017/2018

*All'Oltre  
che a volte non risponde,  
che della curiosità è la fonte  
e che apra tante porte.*



# Indice

Introduzione	
<b>1 SDE di McKean</b>	<b>7</b>
1.1 Wasserstein Metric . . . . .	8
1.2 Esempio: SDE di McKean-Vlasov . . . . .	9
1.3 Soluzione: Esistenza e Unicità . . . . .	10
<b>2 Particle Method</b>	<b>15</b>
2.1 Particle Method nella SDE di McKean-Vlasov . . . . .	16
2.2 Propagazione del Caos . . . . .	16
2.3 Propagazione del Caos nella SDE di McKean-Vlasov . . . . .	20
<b>3 Modello LSV e sua Calibrazione</b>	<b>25</b>
3.0.1 Opzioni Double No Touch . . . . .	26
3.1 Descrizione del Modello LSV . . . . .	26
3.2 Esempi di Modelli LSV . . . . .	27
3.2.1 Modello LSV di Dupire-Heston . . . . .	27
3.2.2 Modello LSV di SABR . . . . .	28
3.3 Calibrazione del Modello a Volatilità Locale Stocastica . . . . .	29
3.3.1 Volatilità Locale di Dupire . . . . .	29
3.3.2 Legame tra Modelli LV e SV . . . . .	32
3.3.3 Leverage Function . . . . .	34
3.4 Particle Method nei Modelli LSV . . . . .	36
3.4.1 Kernel Regolarizzatori . . . . .	37
3.4.2 Algoritmo Particle Method . . . . .	37

2

*INDICE*

**4 Conclusioni**

**41**

**5 Appendice**

**43**

**Bibliografia**

**49**

# Introduzione

Rispetto ad una equazione classica di Ito, in una Equazione Differenziale Stocastica di McKean i coefficienti dipendono anche da un terzo parametro cioè la distribuzione di probabilità del processo stesso. Per approssimare la soluzione di una SDE di McKean si può utilizzare il Particle Method un algoritmo di tipo Monte Carlo derivante dalla Fisica Statistica basato sulla simulazione simultanea di cammini interagenti [13]. In option pricing possiamo trovare una applicazione finanziaria di questi concetti.

*Un derivato finanziario è un contratto il cui valore dipende da uno o più titoli o beni, detti sottostanti [29] (azioni, indici azionari, tassi, valute e merci).* In Finanza risulta essere molto importante riuscire a prezzare un derivato, dove per "prezzare" si intende il dare un prezzo equo al contratto basandosi sui dati di mercato.

Il modello, per tanti anni applicato ed ancora largamente diffuso, per prezzare i derivati è quello di Black&Scholes. Nel 1973 Fisher Black e Myron Scholes pubblicarono un articolo [4] nella quale, basandosi sugli studi di Robert Merton [28] e Paul Samuelson [32], riuscirono a fornire una formula chiusa per la valutazione delle opzioni europee, i derivati più semplici e utilizzati. La formula di Black&Scholes è inoltre coerente con il Principio di non arbitraggio che afferma l'impossibilità di ottenere un profitto senza esporsi ad un rischio.

In B&S il processo stocastico che descrive la dinamica del sottostante si suppone evolva come un moto browniano geometrico con coefficiente di diffusione (*volatilità*) costante. La volatilità risulta essere l'unico parametro del modello che non è osservabile dal mercato e la stima di questa costante avviene o attraverso l'analisi delle serie storiche dei prezzi (volatilità storica), o osservando i prezzi delle opzioni quotate sul mercato

(approccio più diffuso). Osservando dunque il prezzo di una Call si inverte la formula di Black&Scholes (il prezzo di una opzione è una funzione strettamente crescente e quindi invertibile) e si ottiene il valore della volatilità. Compiendo questa operazione per diverse scadenze (maturity) e strike, si può però notare come la volatilità implicita non sia affatto costante. In questi casi si parla di *effetto smile* della volatilità implicita e ciò comporta enormi difficoltà nella valutazione di contratti esotici. Gli studi sulla volatilità implicita e l'effetto smile evidenziano l'inadeguatezza del modello Black&Scholes che risulta non essere realistico.

I difetti di cui sopra hanno spinto la ricerca verso lo studio di nuovi modelli che ovviassero ai problemi riscontrati e che riuscissero a descrivere lo smile della volatilità. Si è tentato ad esempio di esprimere la volatilità come una funzione deterministica con parametri il tempo e il sottostante. Questi modelli vengono chiamati "a volatilità locale" e Dupire dimostrò sia l'unicità della funzione di volatilità, sia la coerenza con l'effetto smile osservabile dai dati di mercato [9] [10]. Il problema di questi modelli è la mancata robustezza della funzione volatilità che tende a variare molto in risposta a piccole perturbazioni dei dati di mercato. Inoltre i prezzi di alcuni contratti esotici forniti dal modello (ad esempio le opzioni "forward start"), non sono coerenti coi dati di mercato e ciò implica una errata descrizione della dinamica del mercato.

I modelli a volatilità stocastica prevedono invece che la volatilità sia guidata da un processo stocastico costruito in modo tale che sia concorde con alcune osservazioni effettuate sulle serie storiche della volatilità. Nuovamente anche questi modelli riescono a descrivere bene gli smiles di volatilità implicita, risolvono inoltre i problemi dei modelli a volatilità locale sulle opzioni forward start e sulla stabilità dei parametri, ma, poiché i derivati esotici sono così vari, anche in questo caso vi è una categoria che non viene ben apprezzata dai modelli di cui sopra. Le opzioni barriera sui tassi di cambio (nella quale il pagamento avviene solamente se le due valute non hanno mai superato un determinato range di scambio) sono contratti molto importanti per via della loro enorme liquidità ed appartengono proprio a tale categoria. Non si può dunque ignorare tale problematica. Il modello a volatilità stocastica sovrastima il prezzo di tali derivati, mentre il modello a volatilità locale lo sottostima. Si è dunque iniziato a pensare che la dinamica intrinseca del mercato sia una via di mezzo fra i due modelli.

Nei modelli a *volatilità locale stocastica* vi è un parametro ulteriore che varia in modo tale da avvicinarsi maggiormente ai modelli con volatilità locale o a quelli con volatilità stocastica a seconda dei casi e riuscire così a mantenere i lati positivi di entrambi gli approcci eliminando quelli negativi. Il problema cruciale di questa nuova classe di modelli è la loro calibrazione che porta al dover risolvere una Equazione Differenziale Stocastica di McKean. Soluzioni analitiche non sono possibili e dunque si cerca la soluzione con metodi numerici.

I metodi classici come quello delle differenze finite non sono utili in caso di grandi dimensioni molto frequenti ad esempio nelle applicazioni finanziarie. In questo ambito si inserisce il *Particle Method* che viene utilizzato per approssimare la soluzione di una SDE di McKean e di conseguenza per calibrare il modello.

Nel primo capitolo si introducono le equazioni differenziali stocastiche di McKean e il caso particolare di McKean-Vlasov. Le SDE di McKean differiscono da quelle classiche di Ito per via del fatto che i coefficienti dipendono anche da un terzo parametro cioè la distribuzione di probabilità del processo stesso. Nel capitolo iniziale viene inoltre introdotta la metrica di Wasserstein studiata in numerosi ambiti scientifici, a causa delle innumerevoli applicazioni, e che in questo caso è utile per la dimostrazione dell'esistenza e dell'unicità della soluzione sotto opportune ipotesi.

Nel secondo capitolo viene invece descritto il Particle Method, metodo computazionale derivante dalla fisica statistica, che, in questo ambito, consiste nell'approssimazione della distribuzione di probabilità incognita del processo con una distribuzione empirica di  $N$  particelle. Viene poi introdotta la proprietà di propagazione del caos che vale per la SDE di McKean e che risulta cruciale per la dimostrazione della convergenza del Particle Method in una SDE di McKean-Vlasov.

Nel terzo ed ultimo capitolo viene descritta la classe di modelli a volatilità locale-stocastica portando anche due esempi molto conosciuti in letteratura: il modello di Dupire-Heston e quello di SABR. Segue il problema della calibrazione di tali modelli che implica la trattazione dei risultati ottenuti da Dupire per la volatilità locale [9][10] e il collegamento tra i modelli LV e quelli SV. Il fulcro però di tale problema sta nella calibrazione della leverage function che induce ad utilizzare il Particle Method. Infine



si parla di un possibile algoritmo per la calibrazione di un modello LSV utilizzando il Particle Method, dei kernel regolarizzatori per l'approssimazione della delta di Dirac e di alcune tecniche di accelerazione per ridurre il costo computazionale dell'algoritmo.

In appendice si trovano alcune nozioni matematiche basilari che vengono riprese nel testo.

# Capitolo 1

## SDE di McKean

L'equazione differenziale stocastica non lineare di McKean è stata introdotta nel 1966 da Henry McKean [26]. Rispetto alla SDE di Ito (5.13), i coefficienti del drift e della volatilità dipendono da un terzo parametro: la distribuzione di probabilità  $\mathbb{P}_t$  di  $X_t$ .

Sia infatti  $X_t \in \mathbb{R}^n$ :

$$dX_t = b(t, X_t, \mathbb{P}_t)dt + \sigma(t, X_t, \mathbb{P}_t)dW_t \quad (1.1)$$

con:

$$\mathbb{P}_t = Law(X_t) \quad X_0 \in \mathbb{R}^n \quad (1.2)$$

Dove il drift  $b(t, X_t, \mathbb{P}_t)$  risulta essere un vettore  $n$ -dimensionale, mentre il coefficiente di diffusione  $\sigma(t, X_t, \mathbb{P}_t)$  è una matrice  $N \times d$ . I due coefficienti dipendono entrambi dal valore  $X_t$  del processo e anche dalla distribuzione di probabilità  $\mathbb{P}_t$  di  $X_t$  (5.5).  $W_t$  è invece un Moto Browniano  $d$ -dimensionale.

La funzione densità di probabilità  $p(t, \cdot) \equiv \mathbb{P}_t$  di  $X_t$  è soluzione della PDE di Fokker Planck:

$$-\partial_t p - \sum_{i=1}^n \partial_i (b^i(t, x, \mathbb{P}_t) p(t, x)) + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} \left( \sum_{k=1}^d \sigma_k^i(t, x, \mathbb{P}_t) \sigma_k^j(t, x, \mathbb{P}_t) p(t, x) \right) = 0 \quad (1.3)$$

Con le condizioni iniziali:

$$\lim_{t \rightarrow 0} p(t, x) = \delta(x - X_0) \quad (1.4)$$

dove  $\delta$  è la delta di Dirac.

L'equazione è non lineare in quanto  $b^i(t, x, \mathbb{P}_t)$  e  $\sigma_k^i(t, x, \mathbb{P}_t)$  dipendono dall'incognita  $\mathbb{P}_t$ . L'unicità e l'esistenza sono provate se i coefficienti di Drift e di Volatilità sono funzioni lipschitz-continue di  $\mathbb{P}_t$  rispetto alla metrica di Wasserstein [14][27]. L'introduzione del Particle Method è dovuto proprio al tentativo di risolvere una equazione differenziale stocastica non lineare come la SDE di McKean, in quanto la calibrazione del modello LSV (Local Stochastic Volatility) equivale a risolvere una equazione di questo tipo. Uno dei punti cruciali sarà infatti l'approssimazione della legge  $\mathbb{P}_t$  con la distribuzione empirica (2.1) di un numero  $N$  fissato, molto grande, di particelle.

## 1.1 Wasserstein Metric

La distanza di Monge-Kantorovich, o metrica di Wasserstein, fra due misure di probabilità  $\mathbb{P}_1$  e  $\mathbb{P}_2$  su  $\mathbb{R}^n$  è definita come:

$$d_{MK}(\mathbb{P}_1, \mathbb{P}_2)^p = \inf_{\tau \in \mathcal{P}(\mathbb{R}^n, \mathbb{R}^n)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d(x, y)^p \tau(dx, dy)$$

con marginali  $\mathbb{P}_1$  e  $\mathbb{P}_2$

con  $p \geq 1$  e  $(\mathbb{R}^n, d)$  spazio metrico completo.  $\tau$  è una misura di probabilità su  $\mathbb{R}^n \times \mathbb{R}^n$  con marginali sulla prima e sulla seconda coordinata, rispettivamente  $\mathbb{P}_1$  e  $\mathbb{P}_2$ , che realizzano il limite inferiore dell'attesa di una qualche distanza  $d(\cdot, \cdot)$  chiamata spesso in letteratura *funzione di costo*. Numerosi sono i risultati, riassunti in [11], ottenuti sull'esistenza e l'unicità dell'estremo inferiore. Sudakov nel 1976, per il più generale *problema di Monge*, mostrò che se le misure  $\mathbb{P}_1$  e  $\mathbb{P}_2$  sono date su un sottoinsieme limitato di uno spazio di Banach finito e inoltre sono assolutamente continue rispetto alla misura di Lebesgue, allora esiste una misura ottimale, cioè una misura che realizza l'estremo inferiore [34]. Nel caso particolare della formulazione descritta all'inizio di questa sezione e usualmente denominata *problema di Kantorovich*, l'esistenza è stata provata dall'omonimo studioso in [22][23] ed è garantita da una condizione di continuità e compattezza sulla distanza  $d$ . Supponendo invece che le  $\mathbb{P}_i$  siano assolutamente continue rispetto alla misura di Lebesgue e che la funzione di costo sia strettamente convessa, allora la soluzione ottimale è anche unica.

$d_{MK}$  definisce una distanza sull'insieme delle misure di probabilità  $\mathcal{P}_p(\mathbb{R}^n)$  con momenti di ordine  $p$  finiti, cioè:

$$\int d(x_0, x)^p \mathbb{P}(dx) < \infty$$

con  $x_0 \in \mathbb{R}^n$  e  $\mathbb{P} \in \mathcal{P}_p(\mathbb{R}^n)$ .

Tale distanza induce la convergenza debole (5.15) come dimostrato in [36], pertanto:

**Proposizione 1.1.** *Una sequenza di misure di probabilità  $(\mathbb{P}_k)_{k \in \mathbb{N}}$  in  $\mathcal{P}_p(\mathbb{R}^n)$ , soddisfa:*

$$d_{MK}(\mathbb{P}_k, \mathbb{P}) \xrightarrow[k \rightarrow \infty]{} 0 \quad \mathbb{P} \in \mathcal{P}(\mathbb{R}^n)$$

se e solo se  $\mathbb{P}_k \rightarrow \mathbb{P}$  in senso debole:

$$\forall \varphi \in C_b(\mathbb{R}^n), \quad \int \varphi(x) \mathbb{P}_k(dx) \xrightarrow[k \rightarrow \infty]{} \int \varphi(x) \mathbb{P}(dx)$$

Nel lavoro di Villani [36] vi è anche la dimostrazione del fatto che  $d_{MK}$  sia effettivamente una distanza.

**Esempio 1.2.** Siano  $\mu_1 = \delta_a$  e  $\mu_2 = \delta_b$  due delta di Dirac centrate rispettivamente su  $a, b \in \mathbb{R}$ . Sia presa in considerazione la distanza su  $\mathbb{R}$  data dal valore assoluto. L'unica  $\tau \in \mathcal{P}(\mathbb{R}, \mathbb{R})$ , che ha per marginali rispettivamente  $\mu_1$  e  $\mu_2$ , è  $\delta_{(a,b)}$ :

$$W_p(\mu_1, \mu_2) = |a - b|, \quad \forall p \geq 1$$

## 1.2 Esempio: SDE di McKean-Vlasov

Nel caso di una SDE di McKean-Vlasov si ha che:

$$b^i(t, x, \mathbb{P}_t) = \int B^i(t, x, y) \mathbb{P}_t(dy) = \mathbb{E} [B^i(t, x, X_t)]$$

$$\sigma_j^i(t, x, \mathbb{P}_t) = \int \Sigma_j^i(t, x, y) \mathbb{P}_t(dy) = \mathbb{E} [\Sigma_j^i(t, x, X_t)] \quad (1.5)$$

Dunque in questo caso i coefficienti non sono altro che il valore medio di una qualche funzione  $B^i(t, X_t, \cdot)$ , o  $\Sigma_j^i(t, X_t, \cdot)$  con  $1 \leq i \leq n$  e  $1 \leq j \leq d$ , rispetto alla distribuzione  $\mathbb{P}_t$  di  $X_t$ . Infine i due coefficienti  $B$  e  $\Sigma$  si suppone siano Lipschitz-continui in  $x$  e  $y$ .

### 1.3 Soluzione: Esistenza e Unicità

L'esistenza e l'unicità della soluzione della SDE di McKean sono provate nel caso in cui il drift e la volatilità siano lipschitziani in  $x$  e  $\mathbb{P}_t$  rispetto alla distanza di Wasserstein, e con l'aggiunta di una condizione di crescita lineare in  $x$ . Troviamo dunque un risultato analogo a quello che si ottiene per la SDE classica (5.13).

**Teorema 1.3.** *Siano*

- $b : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}^n$
- $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}^{n \times d}$

*due funzioni lipschitziane che soddisfano le seguenti condizioni di crescita lineare:*

$$|b(t, X, \mathbb{P}) - b(t, Y, \mathbb{Q})| + |\sigma(t, X, \mathbb{P}) - \sigma(t, Y, \mathbb{Q})| \leq C(|X - Y| + d_{MK}(\mathbb{P}, \mathbb{Q}))$$

$$|b(t, X, \mathbb{P})| + |\sigma(t, X, \mathbb{P})| \leq C(1 + |X|)$$

*rispetto alla somma della metrica canonica di  $\mathbb{R}^n$  e la distanza di Wasserstein  $d_{MK}$  sull'insieme  $\mathcal{P}_2$  delle misure di probabilità con momento del secondo ordine finito e con  $C$  costante positiva.*

*La SDE non lineare:*

$$dX_t = b(t, X_t, \mathbb{P}_t)dt + \sigma(t, X_t, \mathbb{P}_t)dW_t, \quad X_0 \in \mathbb{R}^n \quad (1.6)$$

*dove  $\mathbb{P}_t$  indica la distribuzione di probabilità di  $X_t$ , ammette un'unica soluzione tale che  $\mathbb{E}(\sup_{0 \leq t \leq T} |X_t|^2) < \infty$ .*

#### **Traccia della Dimostrazione [14]**

Per una dimostrazione più completa nel caso di  $\sigma(\cdot) = 1$  e con  $b$  limitato, si rimanda alla dimostrazione del *Theorem 1.1* in [35], mentre questa versione è tratta dall'idea di Méléard [27].

Sia  $\mathcal{C} = C([0, T], \mathbb{R}^n)$  lo spazio delle funzioni continue da  $[0, T]$  a  $\mathbb{R}^n$ . Sia invece  $\mathcal{P}_2(\mathcal{C})$  lo spazio delle misure di probabilità  $\mathbb{P}$  su  $\mathcal{C}$  tali che  $\mathbb{E}^{\mathbb{P}}[\|Y\|_{\infty}^2] \equiv \mathbb{E}^{\mathbb{P}}[\sup_{0 \leq t \leq T} |Y_t|^2] < \infty$ ,  $\forall Y \in \mathcal{C}$ . Lo spazio  $\mathcal{P}_2(\mathcal{C})$  dotato della metrica di Wasserstein è completo (per la dimostrazione della completezza si rimanda a [8]). Questa metrica induce su  $\mathcal{P}_2(\mathbb{R}^n)$  la

topologia della convergenza debole, come detto nella sezione precedente.

$$D_t(\mathbb{P}, \mathbb{Q}) \equiv \inf \int_{\mathcal{C} \times \mathcal{C}} \sup_{0 \leq s \leq t} |Y_s - X_s|^2 R(dX, dY)$$

dove  $R \in \mathcal{P}_2(\mathcal{C} \times \mathcal{C})$  con distribuzioni marginali  $\mathbb{P}$  e  $\mathbb{Q}$  e dove all'interno dell'integrale è stata scelta la metrica uniforme sullo spazio  $\mathcal{P}_2(\mathcal{C})$ . Consideriamo ora la mappa  $\Phi : \mathcal{P}_2(\mathcal{C}) \rightarrow \mathcal{P}_2(\mathcal{C})$ , con associata a  $\mathbb{P} \in \mathcal{P}_2(\mathcal{C})$  la legge di  $X_t^{\mathbb{P}}$  definita da:

$$dX_t^{\mathbb{P}} = b(t, X_t^{\mathbb{P}}, \mathbb{P}_t)dt + \sigma(t, X_t^{\mathbb{P}}, \mathbb{P}_t)dW_t$$

con  $\mathbb{P}_t$  la distribuzione marginale in  $t$  di  $\mathbb{P}$ . Equivalentemente possiamo scrivere:

$$X_t^{\mathbb{P}} = X_0 + \int_0^t b(s, X_s^{\mathbb{P}}, \mathbb{P}_s) ds + \int_0^t \sigma(s, X_s^{\mathbb{P}}, \mathbb{P}_s) dW_s \quad , t \leq T$$

con  $X_0$  indipendente da  $W$ .

I coefficienti  $b$  e  $\sigma$  sono per ipotesi Lipschitz-continue in  $x$  e soddisfano le condizioni di crescita lineare sopracitate e dunque la SDE precedente ammette un'unica soluzione forte (5.14) e  $\Phi$  prende valori su  $\mathcal{P}_2(\mathcal{C})$ . Possiamo osservare che il processo  $X_t^{\mathbb{P}}$  risolve l'equazione (1.6) se e solo se la sua legge è un punto fisso di  $\Phi$ . Infine bisogna verificare che  $\Phi^N$  sia effettivamente una contrazione (stretta), in quanto questo garantisce che  $\Phi$  ammetta un unico punto fisso.

Siano  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathcal{C})$ , poniamo che:

- $\delta X_s \equiv X_s^{\mathbb{P}} - X_s^{\mathbb{Q}}$
- $\delta b_s = b(s, X_s^{\mathbb{P}}, \mathbb{P}_s) - b(s, X_s^{\mathbb{Q}}, \mathbb{Q}_s)$
- $\delta \sigma_s = \sigma(s, X_s^{\mathbb{P}}, \mathbb{P}_s) - \sigma(s, X_s^{\mathbb{Q}}, \mathbb{Q}_s)$

Usando la relazione:

$$(a_1 + \dots + a_n)^2 \leq n(a_1^2 + \dots + a_n^2)$$

Per prima cosa scomponiamo:

$$|\delta X_s|^2 \leq 2 \left( \left| \int_0^s \delta b_u du \right|^2 + \left| \int_0^s \delta \sigma_u dW_u \right|^2 \right)$$

e applicando la disuguaglianza di Hölder si ottiene:

$$\leq 2 \left( s \int_0^s |\delta b_u|^2 du + \left| \int_0^s \delta \sigma_u dW_u \right|^2 \right)$$

Dalla Disuguaglianza Massimale di Doob (5.11) prima e dalla proprietà di Lipschitz dei coefficienti poi, segue che:

$$\begin{aligned} h(t) &\equiv \mathbb{E} \left[ \sup_{s \leq t} |\delta X_s|^2 \right] \leq 2 \left( t \int_0^t \mathbb{E}[|\delta b_u|^2] du + 4 \int_0^t \mathbb{E}[|\delta \sigma_u|^2] du \right) \\ &\leq C \left( \int_0^t \mathbb{E}[|\delta X_u|^2] du + \int_0^t d^2(\mathbb{P}_u, \mathbb{Q}_u) du \right) \\ &\leq C \left( \int_0^t h(u) du + \int_0^t d^2(\mathbb{P}_u, \mathbb{Q}_u) du \right) \end{aligned}$$

Dove  $C$  è una costante positiva che dipende da  $T$ . Grazie al Lemma di Gronwall si può dedurre inoltre che esiste una costante positiva  $K$  che dipende solo da  $T$ , tale che se  $t \leq T$ :

$$h(t) \leq K \int_0^t d^2(\mathbb{P}_u, \mathbb{Q}_u) du$$

Siccome poi  $D_t^2(\Phi(\mathbb{P}), \Phi(\mathbb{Q})) \leq \mathbb{E}[\sup_{s \leq t} |\delta X_s|^2]$  e  $d(\mathbb{P}_s, \mathbb{Q}_s) \leq D_s(\mathbb{P}, \mathbb{Q})$ , si ha che per  $t \leq T$ :

$$D_t^2(\Phi(\mathbb{P}), \Phi(\mathbb{Q})) \leq K \int_0^t D_s^2(\mathbb{P}, \mathbb{Q}) ds$$

e iterando questa disuguaglianza otteniamo che:

$$D_T^2(\Phi^N(\mathbb{P}), \Phi^N(\mathbb{Q})) \leq K^N \int_0^T \frac{(T-s)^{N-1}}{(N-1)!} D_s^2(\mathbb{P}, \mathbb{Q}) ds \leq \frac{K^N T^N}{N!} D_T^2(\mathbb{P}, \mathbb{Q})$$

per  $N \in \mathbb{N}^*$ , con  $N$  molto grande,  $\Phi^N$  è una contrazione, quindi  $\Phi$  ammette un unico punto fisso.  $\square$

Si può dimostrare che la funzione densità di probabilità

$$p(t, y)dy \equiv \mathbb{P}_t(dy)$$

di  $X_t$  è una soluzione della PDE di Fokker Planck:

$$-\partial_t p(t, \cdot) - \sum_{i=1}^n \partial_i (b^i(t, x, \mathbb{P}_t) p(t, \cdot)) + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} \left( \sum_{k=1}^d \sigma_k^i(t, x, \mathbb{P}_t) \sigma_k^j(t, x, \mathbb{P}_t) p(t, \cdot) \right) = 0 \quad (1.7)$$

con le condizioni iniziali:

$$\lim_{t \rightarrow 0} p(t, x) = \delta(x - X_0)$$

Questa è una PDE non lineare in quanto  $b^i$  e  $\sigma_k^i$  dipendono dall'incognita  $p$ . Si prova che il processo  $X_t$  ha distribuzione che soddisfa in senso debole la PDE non lineare (1.7). Sia  $f \in C_b^2(\mathbb{R}^n)$  (dove  $C_b^2$  indica l'insieme delle funzioni continue limitate e due volte differenziabili), applicando il lemma di Ito (5.10) troviamo:

$$f(X_t) = f(X_0) + \int_0^t \sum_{i=1}^n \sigma^i(s, X_s, \mathbb{P}_s) \partial_i f(X_s) dW_s + \int_0^t \left( \frac{1}{2} \sum_{i,j=1}^n \sum_{k=1}^d \sigma_k^i(s, X_s, \mathbb{P}_s) \sigma_k^j(s, X_s, \mathbb{P}_s) \partial_{ij} f(X_s) + \sum_{i=1}^n b^i(s, X_s, \mathbb{P}_s) \partial_i f(X_s) \right) ds \quad (1.8)$$

Facendo l'attesa e integrando per parti si ottiene una versione debole della (1.7).





# Capitolo 2

## Particle Method

Il Particle Method è un algoritmo di tipo Monte Carlo basato sulla simulazione simultanea di cammini interagenti. Nel caso della SDE di McKean la distribuzione di probabilità  $\mathbb{P}_t$  di  $X_t$  è una incognita. Per poter quindi simulare l'andamento di un processo  $X_t$ , soluzione della SDE di McKean, si sostituisce  $\mathbb{P}_t$  con una distribuzione empirica, di un numero grande fissato  $N$  di particelle, ottenuta da:

$$\mathbb{P} \approx \mathbb{P}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}$$

dove le  $(X_t^{i,N})_{1 \leq i \leq N}$  sono le soluzioni della seguente SDE  $\mathbb{R}^N$ -dimensionale classica:

$$dX_t^{i,N} = b\left(t, X_t^{i,N}, \mathbb{P}_t^N\right) dt + \sigma\left(t, X_t^{i,N}, \mathbb{P}_t^N\right) dW_t^i$$

$$Law\left(X_0^{i,N}\right) = \mathbb{P}_0$$

nella quale:

- $\mathbb{P}_t^N$  è una misura su  $\mathbb{R}^N$ .
- $\{W_t^i\}_{1 \leq i \leq N}$  sono  $N$  moti browniani  $d$ -dimensionali;
- Le  $X_t^{i,N}$  sono soluzioni di una SDE classica in quanto i coefficienti  $b$  e  $\sigma$  dipendono da  $(t, X_t^{i,N}, \mathbb{P}_t^N)$ , ma la distribuzione empirica non è altro che una funzione delle  $X^i$ .

Si parla di "Particle" Method e i cammini  $X^{i,N}$  vengono considerati particelle a seguito di analogie con la fisica statistica. Rispetto ai classici metodi di Monte Carlo, nel Particle Method le particelle sono bosoniche, cioè indistinguibili, e interagenti.

## 2.1 Particle Method nella SDE di McKean-Vlasov

Nel caso della SDE di McKean-Vlasov le  $\{X_t^{i,N}\}_{1 \leq i \leq N}$  risultano essere dei processi di Ito  $n$ -dimensionali ottenuti da:

$$dX_t^{i,N} = \left( \int B(t, X_t^{i,N}, y) d\mathbb{P}_t^N(y) \right) dt + \left( \int \Sigma(t, X_t^{i,N}, y) d\mathbb{P}_t^N(y) \right) dW_t^i \quad (2.1)$$

che equivale a:

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^N B(t, X_t^{i,N}, X_t^{j,N}) dt + \frac{1}{N} \sum_{j=1}^N \Sigma(t, X_t^{i,N}, X_t^{j,N}) dW_t^i \quad (2.2)$$

Si può notare come i cammini  $X^{i,N}$  interagiscano fra loro e dunque non è sufficiente conoscere  $t$  e  $X_t^{i,N}$  per determinare i coefficienti  $B$  e  $\Sigma$  del cammino  $i$  al tempo  $t$  e non è possibile simulare  $X^{i,N}$  al tempo  $t + \delta t$ . Per farlo bisogna conoscere anche gli altri cammini  $X^{j,N}$  con  $j \neq i$ .

## 2.2 Propagazione del Caos

Se al tempo  $t = 0$  le  $X_t^{i,N}$  sono particelle indipendenti allora, per  $N$  che tende ad infinito e per ogni  $t > 0$  fissato, le  $X_t^{i,N}$  sono asintoticamente indipendenti e la loro misura di probabilità empirica  $\mathbb{P}_t^N$  converge in distribuzione (5.16) alla distribuzione reale, nel senso di vera,  $\mathbb{P}_t$ . La distribuzione delle misure di probabilità empiriche  $\mathbb{P}_t^N$  converge quindi verso la delta di Dirac centrata nella misura  $\mathbb{P}_t$ :

$$\frac{1}{N} \sum_{i=1}^N \varphi(X_t^{i,N}) \xrightarrow[N \rightarrow \infty]{L^1} \int_{\mathbb{R}^N} \varphi(x) p(t, x) dx$$

Dove  $p(t, \cdot)$  è la soluzione fondamentale della PDE non lineare di Fokker-Planck. La proprietà di cui sopra si chiama *Proprietà di Propagazione del Caos* ed è stata introdotta da Sznitman [35]. Godere di tale proprietà implica dunque la convergenza del

Particle Method. Possiamo anche dire che se  $N$  è abbastanza grande, la PDE di Fokker-Planck  $(\mathbb{R}^n)^N$ -dimensionale lineare approssima la PDE di Fokker Planck non lineare di dimensione  $n$ .

**Definizione 2.1 (Misura Empirica).** *Siano  $X_1, \dots, X_N$  variabili aleatorie indipendenti e identicamente distribuite (i.i.d.) con legge  $\mu$ . La misura empirica associata alla configurazione  $(X_1, \dots, X_N)$  è:*

$$\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i} \quad (2.3)$$

Possiamo notare che:

$$\mathbb{E}_\mu [\hat{\mu}^N] = \mu$$

cioè per ogni evento  $A$ :

$$\mathbb{E}_\mu [\hat{\mu}^N(A)] = \mu(A)$$

**Definizione 2.2 (Distribuzione  $\mu$ -Caotica).** *Sia  $\{\mu^N\}_{N \in \mathbb{N}}$  una sequenza di distribuzioni di probabilità simmetriche su  $(\mathbb{R}^n)^N$ . Sia  $\mu$  una misura di probabilità su  $\mathbb{R}^n$ . Diciamo che  $\mu^N$  è  $\mu$ -caotica se per ogni intero  $k \geq 1$  e per ogni insieme di funzioni test  $\varphi_1, \dots, \varphi_k \in C_b(\mathbb{R}^n)$ , cioè continue e limitate, abbiamo:*

$$\int \varphi_1(x_1) \cdots \varphi_k(x_k) \mu^N(dx_1, \dots, dx_N) \xrightarrow{N \rightarrow \infty} \int \varphi_1 d\mu \cdots \int \varphi_k d\mu$$

*cioè con il tendere di  $N$  verso infinito, le  $k$  fissate particelle su  $N$  sono asintoticamente indipendenti e identicamente distribuite. Diciamo che  $\mu^N$  è caotica se esiste  $\mu$  tale che  $\mu^N$  è  $\mu$ -caotica.*

L'indipendenza stocastica di  $k$  fissate particelle, di un sistema che ne contiene molte altre, persiste nel tempo al tendere ad infinito del numero di particelle che compongono il sistema.

**Definizione 2.3 (Propagazione del Caos).** *Consideriamo una successione di distribuzioni di probabilità simmetriche che evolvono come una SDE  $N$ -dimensionale che associa ad una misura di probabilità iniziale  $\mu_0^N$  una misura  $\mu_t^N$  al tempo  $t$ . Diciamo che questa successione propaga il caos se, per ogni misura caotica iniziale  $\mu_0^N$  e ogni  $t > 0$ ,  $\mu_t^N$  è caotico.*

**Teorema 2.4.** Sia  $\{\mu^N\}_{N \in \mathbb{N}}$  una sequenza di distribuzioni di probabilità simmetriche su  $(\mathbb{R}^n)^N$  e sia  $\mu$  una misura di probabilità su  $\mathbb{R}^n$ . Le seguenti quattro proprietà sono equivalenti:

(i)  $\{\mu^N\}_{N \in \mathbb{N}}$  è  $\mu$ -caotica.

(ii)  $\forall \varphi_1, \varphi_2 \in C_b(\mathbb{R}^n)$ :

$$\int \varphi_1(x_1)\varphi_2(x_2)\mu^N(dx_1, \dots, dx_N) \xrightarrow{N \rightarrow \infty} \int \varphi_1 d\mu \int \varphi_2 d\mu$$

(iii) Siano  $X_1, \dots, X_N$  variabili aleatorie tali che  $\text{Law}(X_1, \dots, X_N) = \mu^N$ . Allora,  $\forall \varphi \in C_b(\mathbb{R}^n)$ :

$$\frac{1}{N} \sum_{i=1}^N \varphi(X_i) \xrightarrow{N \rightarrow \infty} \int \varphi d\mu$$

(iv) Sia  $\hat{\mu}^N$  la misura empirica associata a  $\mu^N$ . Allora:

$$\mathbb{E}_{\mu^N} \left[ \left| \int \varphi d\hat{\mu}^N - \int \varphi d\mu \right| \right] \xrightarrow{N \rightarrow \infty} 0$$

$\forall \varphi \in C_b(\mathbb{R}^n)$

La convergenza del Particle Method cioè:

$$\frac{1}{N} \sum_{i=1}^N \varphi(X_t^{i,N}) \xrightarrow{N \rightarrow \infty} \int \varphi(x)p(t, x)dx$$

$\forall t$  e  $\forall \varphi \in C_b(\mathbb{R}^n)$

equivale dunque alla propagazione del caos per la SDE  $(X_t^{1,N}, \dots, X_t^{N,N})$ .

**Dimostrazione**

• (i)  $\rightarrow$  (ii): Applico la definizione di  $\mu$ -caotica con  $k = 2$ .

• (ii)  $\rightarrow$  (iii): Si ha:

$$\mathbb{E}_{\mu^N} \left[ \left( \frac{1}{N} \sum_{i=1}^N \varphi(X_i) - \int \varphi d\mu \right)^2 \right]$$

$$\begin{aligned}
&= \mathbb{E}_{\mu^N} \left[ \frac{1}{N^2} \sum_{i,j=1}^N \varphi(X_i)\varphi(X_j) - 2 \int \varphi d\mu \frac{1}{N} \sum_{i=1}^N \varphi(X_i) + \left( \int \varphi d\mu \right)^2 \right] \\
&= \frac{1}{N} \mathbb{E}_{\mu^N} [\varphi(X_1)^2] + \frac{N-1}{N} \mathbb{E}_{\mu^N} [\varphi(X_1)\varphi(X_2)] - 2 \int \varphi d\mu \mathbb{E}_{\mu^N} [\varphi(X_1)] + \left( \int \varphi d\mu \right)^2
\end{aligned}$$

Ma dalla (ii) sappiamo che:

$$\begin{aligned}
&- \mathbb{E}_{\mu^N} [\varphi(X_1)\varphi(X_2)] \xrightarrow{N \rightarrow \infty} \left( \int \varphi d\mu \right)^2 \\
&- \mathbb{E}_{\mu^N} [\varphi(X_1)] \xrightarrow{N \rightarrow \infty} \int \varphi d\mu
\end{aligned}$$

e dunque:

$$\xrightarrow{N \rightarrow \infty} 0 + \left( \int \varphi d\mu \right)^2 - 2 \left( \int \varphi d\mu \right)^2 + \left( \int \varphi d\mu \right)^2 = 0$$

- (iii)→(iv): Segue dalla definizione di convergenza in media.
- (iv)→(i): Sia  $k \geq 1$  e siano  $\varphi_1, \dots, \varphi_k \in C_b(\mathbb{R}^n)$ . Abbiamo che:

$$\begin{aligned}
&\left| \mathbb{E}_{\mu^N} [\varphi_1(X_1) \cdots \varphi_k(X_k)] - \int \varphi_1 d\mu \cdots \int \varphi_k d\mu \right| \\
&\leq \left| \mathbb{E}_{\mu^N} [\varphi_1(X_1) \cdots \varphi_k(X_k)] - \mathbb{E}_{\mu^N} \left[ \int \varphi_1 d\hat{\mu}^N \cdots \int \varphi_k d\hat{\mu}^N \right] \right| \\
&\quad + \left| \mathbb{E}_{\mu^N} \left[ \int \varphi_1 d\hat{\mu}^N \cdots \int \varphi_k d\hat{\mu}^N \right] - \int \varphi_1 d\mu \cdots \int \varphi_k d\mu \right|
\end{aligned}$$

Il secondo termine alla destra del segno di disuguaglianza converge a zero per l'ipotesi (iv), mentre il primo termine si può vedere in questo modo:

$$\left| \mathbb{E}_{\mu^N} [\varphi_1(X_1) \cdots \varphi_k(X_k)] - \frac{1}{N^k} \sum_{i_1, \dots, i_k=1}^N \mathbb{E}_{\mu^N} [\varphi_1(X_{i_1}) \cdots \varphi_k(X_{i_k})] \right|$$

In questa sommatoria ci sono  $\frac{N!}{(N-k)!}$  termini con indici  $i_1, \dots, i_k$  tutti diversi. Per la simmetria questi sono uguali a  $\mathbb{E}_{\mu^N} [\varphi_1(X_1) \cdots \varphi_k(X_k)]$  e possono essere aggiunti al primo termine. Gli altri termini sono limitati da  $M^k$  con  $M = \sum_j \|\varphi_j\|_\infty$ .

Otteniamo dunque che:

$$\begin{aligned}
&\left( 1 - \frac{1}{N^k} \frac{N!}{(N-k)!} \right) \mathbb{E}_{\mu^N} [\varphi_1(X_1) \cdots \varphi_k(X_k)] + \left( 1 - \frac{1}{N^k} \frac{N!}{(N-k)!} \right) M^k \\
&\leq 2M^k \left( 1 - \frac{1}{N^k} \frac{N!}{(N-k)!} \right) \xrightarrow{N \rightarrow \infty} 0
\end{aligned}$$

## 2.3 Propagazione del Caos nella SDE di McKean-Vlasov

La SDE di McKean-Vlasov propaga il caos, cioè gode della proprietà illustrata nella sezione precedente. Questa è l'asserzione decisiva per la convergenza del Particle Method applicato alla SDE di McKean-Vlasov. Per maggiore chiarezza supponiamo che i coefficienti  $b$  e  $\sigma$  non dipendano da  $t$ .

**Teorema 2.5.** *La SDE di McKean-Vlasov gode della proprietà di propagazione del caos.*

Grazie al teorema (1.3) sappiamo che il processo (1.5) esiste e la sua legge è una soluzione debole della PDE di Fokker-Planck (1.3) se  $b$  e  $\sigma$  sono funzioni Lipschitziane in  $(x, y)$ . La dimostrazione del teorema si basa sul metodo di accoppiamento (coupling method) descritto in [35] che prevede l'introduzione di  $N$  processi  $\{Y_t^i\}_{1 \leq i \leq N}$  definiti da:

$$dY_t^i = b(Y_t^i, \mathbb{P}_t)dt + \sigma(Y_t^i, \mathbb{P}_t)dW_t^i$$

con  $Y_0^i = X_0$  e inoltre:

- $b(y, \mathbb{P}_t) \equiv \int b(y, z)\mathbb{P}_t(dz)$
- $\sigma(y, \mathbb{P}_t) \equiv \int \sigma(y, z)\mathbb{P}_t(dz)$
- $\mathbb{P}_t = Law(X_t)$

Queste sono SDE standard che ammettono soluzione forte se  $b$  e  $\sigma$  sono funzioni di Lipschitz. La densità  $q_i(t, x)$  di  $Y_t^i$  soddisfa l'equazione lineare di Fokker-Planck:

$$-\partial_t q_i(t, x) - \sum_{i=1}^n \partial_i (b^i(t, x, \mathbb{P}_t) q_i(t, x)) + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} \left( \sum_{k=1}^d \sigma_k^i(t, x, \mathbb{P}_t) \sigma_k^j(t, x, \mathbb{P}_t) q_i(t, x) \right) = 0$$

con condizione iniziale la Delta di Dirac. Dalla PDE di Fokker-Planck (1.3) sappiamo che la densità  $p(t, x)$  di  $X_t$  è una soluzione. Per unicità:  $\mathbb{P}_t = Law(Y_t^i)$

**Proposizione 2.6.** *Siano  $(X^{i,N})_{1 \leq i \leq N}$  definiti da (2.2). Allora vale che:*

$$\mathbb{E}[|X_t^{1,N} - Y_t^1|] \leq \frac{C(t)}{\sqrt{N}}$$

dove  $C(t)$  è una funzione regolare del tempo indipendente da  $N$ .

**Dimostrazione** Per semplicità abbreviamo  $X^{i,N}$  con  $X^i$ . Dal lemma di Ito (5.10) otteniamo:

$$\begin{aligned} \mathbb{E}[(X_t^1 - Y_t^1)^2] &= 2\mathbb{E} \left[ \int_0^t (X_s^1 - Y_s^1) \left( \frac{1}{N} \sum_{j=1}^N \sigma(X_s^1, X_s^j) - \sigma(Y_s^1, \mathbb{P}_s) \right) dW_s^1 \right] \\ &\quad + \int_0^t \mathbb{E} \left[ \left( \frac{1}{N} \sum_{j=1}^N \sigma(X_s^1, X_s^j) - \sigma(Y_s^1, \mathbb{P}_s) \right)^2 \right] ds \\ &\quad + 2 \int_0^t \mathbb{E} \left[ (X_s^1 - Y_s^1) \left( \frac{1}{N} \sum_{j=1}^N b(X_s^1, X_s^j) - b(Y_s^1, \mathbb{P}_s) \right) \right] ds \end{aligned}$$

Il primo termine si annulla in quanto l'integrale stocastico è una martingala (5.3).

Nel secondo termine possiamo osservare che, applicando la relazione  $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$  e aggiungendo e togliendo due grandezze, abbiamo:

$$\begin{aligned} \left( \frac{1}{N} \sum_{j=1}^N \sigma(X_s^1, X_s^j) - \sigma(Y_s^1, \mathbb{P}_s) \right)^2 &\leq 3 \left( \frac{1}{N} \sum_{j=1}^N (\sigma(X_s^1, X_s^j) - \sigma(X_s^1, Y_s^j)) \right)^2 \\ &\quad + 3 \left( \frac{1}{N} \sum_{j=1}^N (\sigma(X_s^1, Y_s^j) - \sigma(Y_s^1, Y_s^j)) \right)^2 \\ &\quad + 3 \left( \frac{1}{N} \sum_{j=1}^N (\sigma(Y_s^1, Y_s^j) - \sigma(Y_s^1, \mathbb{P}_s)) \right)^2 \end{aligned}$$

Con l'ultimo quadrato denominato  $A_s(\sigma)$ .

Usando poi la disuguaglianza di Cauchy-Schwartz e la proprietà di Lipschitz di  $\sigma$  su  $(x, y)$ , deduciamo:

$$\mathbb{E} \left[ \left( \sum_{j=1}^N \frac{1}{N} \sigma(X_s^1, X_s^j) - \sigma(Y_s^1, \mathbb{P}_s) \right)^2 \right] \leq 6 \|\sigma\|_{Lip}^2 \mathbb{E}[|X_s^1 - Y_s^1|^2] + 3\mathbb{E}[A_s(\sigma)]$$

Nel terzo termine abbiamo invece che:

$$2(X_s^1 - Y_s^1) \left( \frac{1}{N} \sum_{j=1}^N b(X_s^1, X_s^j) - b(Y_s^1, \mathbb{P}_s) \right) \leq (X_s^1 - Y_s^1)^2 + \left( \frac{1}{N} \sum_{j=1}^N b(X_s^1, X_s^j) - b(Y_s^1, \mathbb{P}_s) \right)^2$$



e dunque:

$$\mathbb{E} \left[ 2(X_s^1 - Y_s^1) \left( \frac{1}{N} \sum_{j=1}^N b(X_s^1, X_s^j) - b(Y_s^1, \mathbb{P}_s) \right) \right] \leq \mathbb{E}[(X_s^1 - Y_s^1)^2] + 6\|b\|_{Lip}^2 \mathbb{E}[|X_s^1 - Y_s^1|^2] + 3\mathbb{E}[A_s(b)]$$

Vale dunque la seguente disuguaglianza:

$$\mathbb{E}[(X_t^1 - Y_t^1)^2] \leq (6\|\sigma\|_{Lip}^2 + 6\|b\|_{Lip}^2 + 1) \int_0^t \mathbb{E}[(X_s^1 - Y_s^1)^2] ds + 3 \int_0^t \mathbb{E}[A_s(\sigma) + A_s(b)] ds$$

e grazie al lemma di Gronwall:

$$\mathbb{E}[(X_t^1 - Y_t^1)^2] \leq 3e^{(6\|\sigma\|_{Lip}^2 + 6\|b\|_{Lip}^2 + 1)t} \int_0^t \mathbb{E}[A_s(\sigma) + A_s(b)] ds$$

Mostrare che  $\mathbb{E}[A_s(h)] = O(\frac{1}{N})$  per  $h \in \{b, \sigma\}$  è l'ultimo passaggio della dimostrazione.

Quindi:

$$\mathbb{E}[A_s(h)] = \frac{1}{N^2} \sum_{i,j=1}^N \mathbb{E}[h(Y_s^1, Y_s^i)h(Y_s^1, Y_s^j)] + \mathbb{E}[h(Y_s^1, \mathbb{P}_s)^2] - \frac{2}{N} \sum_{i=1}^N \mathbb{E}[h(Y_s^1, Y_s^i)h(Y_s^1, \mathbb{P}_s)]$$

Siccome  $Law(Y_s^i) = \mathbb{P}_s$  e i processi  $\{Y_s^i\}_{1 \leq i \leq N}$  sono indipendenti, abbiamo:

$$\begin{aligned} \mathbb{E}[A_s(h)] &= \frac{2-N}{N^2} \mathbb{E}[h(Y_s^1, \mathbb{P}_s)^2] + \frac{1}{N^2} \mathbb{E}[h(Y_s^1, Y_s^1)^2] \\ &+ \frac{N-1}{N} \mathbb{E}[h(Y_s^1, Y_s^2)^2] - \frac{2}{N^2} \mathbb{E}[h(Y_s^1, \mathbb{P}_s)h(Y_s^1, Y_s^1)] \leq \frac{K(t)}{N} \end{aligned}$$

Infatti i quattro valori attesi sono finiti per la condizione di Lipschitz su  $b$  e  $\sigma$ , e il fatto che  $\mathbb{E}[\sup_{0 \leq s \leq t} |Y_s|^2] < \infty$ .  $\square$

### Dimostrazione del Teorema 2.5

Denotiamo  $\mu_t$  come la legge della soluzione  $X_t$  della SDE di McKean 1.5. Per quanto visto nel 2.4 al punto (iii) è sufficiente dimostrare che  $\forall \varphi \in C_b(\mathbb{R}^n)$ :

$$\frac{1}{N} \sum_{i=1}^N \varphi(X_t^{i,N}) \xrightarrow[N \rightarrow \infty]{L^1} \int \varphi d\mu_t$$

dove gli  $X_t^{i,N}$  sono definiti da 2.2.

$$\mathbb{E} \left[ \left| \frac{1}{N} \sum_{i=1}^N \varphi(X_t^{i,N}) - \int \varphi d\mu_t \right| \right] \leq$$

$$\mathbb{E}_{\mu^N} \left[ \left| \frac{1}{N} \sum_{i=1}^N \left( \varphi(X_t^{i,N}) - \varphi(Y_t^i) \right) \right| \right] + \mathbb{E}_{\mu^N} \left[ \left| \frac{1}{N} \sum_{i=1}^N \varphi(Y_t^i) - \int \varphi d\mu_t \right| \right]$$

Per costruzione i processi  $\{Y_t^i\}_{1 \leq i \leq N}$  sono indipendenti e identicamente distribuiti con legge  $\mu_t$  e dunque il secondo termine va a zero con il tendere di  $N$  ad infinito per la legge dei grandi numeri.  $\forall \epsilon > 0$  esiste una funzione di Lipschitz  $\varphi_\epsilon$  tale che  $|\varphi - \varphi_\epsilon| \leq \epsilon$  e così il primo termine è limitato da:

$$\begin{aligned} & 2\epsilon + \mathbb{E}_{\mu^N} \left[ \left| \frac{1}{N} \sum_{i=1}^N \left( \varphi_\epsilon(X_t^{i,N}) - \varphi_\epsilon(Y_t^i) \right) \right| \right] \\ & \leq 2\epsilon + \|\varphi_\epsilon\|_{Lip} \mathbb{E}[|X_t^{1,N}|] \leq 2\epsilon + \|\varphi_\epsilon\|_{Lip} \frac{C(t)}{\sqrt{N}} \end{aligned}$$

con l'ultima relazione che deduciamo dalla proposizione 2.6. □



# Capitolo 3

## Modello LSV e sua Calibrazione

Il modello di Black-Scholes afferma che per ogni strike e scadenza il parametro di volatilità è costante e deterministico. Questo è il maggior problema del modello tanto conosciuto quanto utilizzato. L'osservazione empirica dei prezzi di mercato delle opzioni mostra infatti un comportamento non costante del coefficiente di volatilità (comportamento chiamato "effetto smile"). Il valore della volatilità implicita at-the-money (ATM) è infatti inferiore rispetto al valore *in* e *out* the-money (ITM, OTM). Per riuscire ad approssimare questa curva sono stati dunque proposti modelli a volatilità locale (LV), modelli a volatilità stocastica (SV) ed infine modelli a volatilità locale-stocastica (LSV) cioè un mix delle due tipologie precedenti.

I modelli SV e LV riescono a replicare in modo soddisfacente i prezzi delle plain vanilla, ma ciò non basta per approssimare bene la dinamica della volatilità intrinseca nel mercato, infatti per farlo è necessario riuscire a replicare anche i prezzi delle opzioni esotiche [3]. I modelli LV e SV sono interamente determinati dai prezzi delle plain vanilla, cioè la loro calibrazione non lascia alcun parametro libero per poter descrivere anche il comportamento di opzioni path-dependent come ad esempio le barriera dove il prezzo non dipende solamente dal valore finale, ma anche dall'evoluzione del risky asset. Queste ultime tipologie di opzioni sono numerose e molto liquide nel mercato dei *Forex* ed è dunque necessario trovare modelli che riescano a replicarne i prezzi [18].

I modelli a volatilità locale-stocastica sono un mix degli approcci tentati in precedenza. L'idea è di costruire il modello in modo tale da mantenere tutti i lati positivi dei modelli

SV e LV, minimizzando quelli negativi.

### 3.0.1 Opzioni Double No Touch

Le opzioni *Double No Touch* (DNT) sono un esempio di opzioni "path-dependent" di tipo barriera molto frequenti nel mercato FX. Viene fissato un sottostante, una scadenza e un intervallo di prezzo che contiene al proprio interno quello corrente. Il payoff  $\Phi_{DNT}$  è positivo se dall'istante iniziale fino alla scadenza  $T$ , il sottostante non ha toccato una delle due barriere scelte  $(L, U)$  o entrambe.

$$\Phi_{DNT}(S_t) = \lambda I_{(L < \min_{t \in [0, T]} S_t \leq \max_{t \in [0, T]} S_t < U)}$$

$$L < S_0 < U$$

Dove  $I$  indica la funzione indicatrice e  $\lambda$  una qualche funzione positiva che dipende da tutti i parametri in gioco (scadenza, barriere, sottostante). Si può osservare come un modello LV, calibrato ai prezzi di mercato delle opzioni vanilla, sottoprezza le opzioni DNT rispetto al prezzo di mercato. Invece un modello SV, calibrato ai prezzi di mercato delle opzioni vanilla, tende a sovrapprezzare tali opzioni barriera [7] [24]. Intuitivamente si vogliono mixare i due modelli aggiungendo un parametro opportunamente calibrato in modo da poter replicare i prezzi di mercato delle opzioni DNT.

## 3.1 Descrizione del Modello LSV

Questo modello è un processo stocastico in due dimensioni  $X_t = (S_t, V_t)$  cioè lo *spot price* (prezzo di mercato corrente)  $S_t$  e la varianza stocastica  $V_t$ , ed è definito dal seguente sistema:

$$dS_t = \mu_1(S_t, t)dt + L(S_t, t)\sigma_1(S_t, V_t, t)dW_t^1 \quad (3.1)$$

$$dV_t = \mu_2(V_t, t)dt + \sigma_2(V_t, t)dW_t^2 \quad (3.2)$$

$$dW_t^1 \cdot dW_t^2 = \rho dt \quad (3.3)$$

Dove  $W_t = (W_t^1, W_t^2)$  è dunque un moto Browniano 2-dimensionale, mentre  $\rho \in ]-1, 1[$  è la correlazione fra i due moti Browniani. Il coefficiente di diffusione di  $\sigma_t$  è controllato

da una funzione  $L(S_t, t)$ , detta *leverage function* ( $L(S, t)$  invece viene chiamata *leverage surface* o volatilità locale), determinata dalle informazioni del mercato e che svolge la funzione di pesare la parte locale del modello. La leverage function è l'aggiunta che non si trova nei modelli LV e SV. A seconda del valore assunto da  $\sigma_2$  ci si avvicina di più ad un modello locale o ad uno stocastico. Con il suo annullamento infatti si ricadrebbe nel caso di un modello a volatilità locale.

## 3.2 Esempi di Modelli LSV

In letteratura si trovano numerosi esempi di modelli LSV che si differenziano sia per la scelta dei modelli LV e SV da combinare, sia per la scelta del metodo di calibrazione. Di solito viene considerato l'articolo [19] come la fonte originale di questi modelli. Nelle seguenti pagine illustreremo due esempi molto famosi: quello di Heston e quello di SABR. Per ulteriori modelli si consiglia la lettura di [37].

### 3.2.1 Modello LSV di Dupire-Heston

Consideriamo un modello SLV analogo al modello SV di Heston [17] con l'aggiunta della leverage function. Il modello di Heston è probabilmente il più usato fra quelli a volatilità stocastica e per questa ragione funge spesso da base per la costruzione dei modelli LSV. La dinamica dello spot  $S_t$  e della varianza stocastica  $V_t$  sotto la misura neutrale al rischio nel modello LSV di Dupire-Heston è la seguente:

$$dS_t = (r_d - r_f)S_t dt + L(S_t, t)\sqrt{V_t}S_t dW_t^1, \quad S_0 = s \quad (3.4)$$

$$dV_t = k(\theta - V_t)dt + \lambda\sqrt{V_t}dW_t^2, \quad V_0 = v \quad (3.5)$$

$$dW_t^1 \cdot dW_t^2 = \rho dt \quad (3.6)$$

dove  $(k, \theta, \lambda, \rho)$  vengono chiamati parametri di Heston.  $\lambda$  è la *vol-of-vol*: alti valori di  $\lambda$  danno maggior peso alla parte stocastica della volatilità, mentre al contrario valori vicini allo zero riportano il modello a volatilità locale.  $\rho$  è la correlazione fra i due moti Browniani.  $\theta$  è la *long run variance* mentre  $k$  è la *mean reversion*.  $r_d$  e  $r_f$  sono

rispettivamente tassi di interesse domestici e stranieri, ma considereremo in futuro  $r = r_d - r_f$ . La varianza  $V_t$  è un processo di Cox-Ingersoll-Ross (CIR) [6].

Il fatto che ci sia  $\sqrt{V_t}$  implica che il termine di diffusione non soddisfa la condizione di Lipschitz che garantisce l'esistenza e l'unicità della soluzione di (3.4). Si possono però fare le seguenti ipotesi per giungere ad un risultato analogo:

- (i) Supponiamo valori iniziali  $S_0$  e  $V_0$  entrambi positivi.  $\rho \in ]-1, 1[$ , mentre  $k, \theta, \lambda > 0$ . Consideriamo infine  $r_d, r_f \in \mathbb{R}$  e leverage function positiva e limitata.
- (ii) Deve essere soddisfatta la condizione di Feller

$$\frac{2k\theta}{\lambda^2} \geq 1$$

che garantisce la positività della varianza del processo. Se la condizione è soddisfatta e  $V_0 > 0$  come nella prima ipotesi, allora  $P[V_t > 0] = 1 \quad \forall t > 0$ .

Possiamo dunque giungere a questa proposizione:

**Proposizione 3.1 (Unicità ed Esistenza per SLV di Heston).** *Sotto le ipotesi (i), (ii) esiste una sola soluzione  $(S_t, V_t)$  del modello SLV (3.4). Inoltre esiste una funzione  $p(S_t, V_t)$  che è la densità di transizione del modello ed è l'unica soluzione della equazione di Fokker-Planck data da:*

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial S}[rSp] + \frac{\partial}{\partial V}[k(\theta - V)p] - \frac{1}{2} \frac{\partial^2}{\partial S^2}[S^2VL(S, t)^2p] - \frac{1}{2} \lambda^2 \frac{\partial^2}{\partial V^2}[Vp] - \lambda \rho \frac{\partial^2}{\partial S \partial V}[SVL(S, t)p] = 0$$

Una generalizzazione di questo modello che comprende anche salti è il modello di Lipton [25] elaborato sulla spinta delle osservazioni fatte sulle opzioni DNT precedentemente citate.

### 3.2.2 Modello LSV di SABR

Il modello di SABR è un famoso modello a volatilità stocastica [16]. Volendo costruire un modello LSV attenendoci alla dinamica descritta nel modello di SABR otteniamo [30][18]:

$$dS_t = S_t L(S_t, t) f(S_t, \sigma_t, t) dW_t^1 \quad (3.7)$$

$$d\sigma_t = \gamma\xi dW_t^2 \quad (3.8)$$

$$dW_t^1 \cdot dW_t^2 = \gamma\rho dt \quad (3.9)$$

dove  $\gamma \in [0, 1]$  è il parametro di mixing, spesso rappresentato come una percentuale, che moltiplica sia la vol-of-vol  $\xi$  sia il termine di correlazione fra i due moti Browniani. Se il market smile è già ottimamente replicato dal modello SV ottenuto da (3.7) eliminando la leverage function, allora ci si può aspettare che quest'ultima sia sufficientemente vicina ad 1 in quanto buona parte del lavoro è già stato svolto dalla volatilità stocastica.

### 3.3 Calibrazione del Modello a Volatilità Locale Stocastica

I modelli LV e SV possono essere calibrati in modo indipendente e simultaneo ai dati di mercato per ottenere i coefficienti dei rispettivi modelli. Dopo averli ottenuti si può ricavare la superficie di volatilità implicita del modello LSV. Qualsiasi sia il processo  $V_t$  considerato e la volatilità locale, vi sono alcuni passi da seguire per calibrare il modello LSV:

- Si calibra la volatilità locale  $\sigma_{loc}$ ;
- Si calibrano i parametri del processo a volatilità stocastica pura  $V_t$ ;
- Si applica il fattore di mixing (la  $\lambda$  nel modello di Heston, la  $\gamma$  nel modello SABR);
- Si calibra la leverage function  $L(S_t, t)$ .

Tutta la difficoltà nel calibrare il modello SLV si concentra sulla leverage function. Infatti avvalendosi della volatilità locale di Dupire, che differisce dalla leverage function, ci si assicura la calibrazione della parte locale del modello.

#### 3.3.1 Volatilità Locale di Dupire

La volatilità locale di Dupire è un approccio molto utilizzato per la calibrazione di modelli a volatilità locale [9]. Esso consiste nel calcolo della  $\sigma_{loc}$  direttamente dai prezzi



di mercato usufruendo di una equazione alle derivate parziali in avanti che governa la loro dinamica. Nella forma più basilare si ha:

$$\sigma_{Dup}^2(K, t) = \frac{\frac{\partial C}{\partial T}}{\frac{1}{2}K^2 \frac{\partial^2 C}{\partial K^2}} \quad (3.10)$$

Si può dimostrare che in un mercato privo di arbitraggi la volatilità di Dupire è consistente col mercato ed esatta per costruzione. Consideriamo dunque un modello a volatilità locale:

$$dS_t = (r_t - q_t)S_t dt + \sigma_{LV}(S_t, t)S_t dW_t \quad (3.11)$$

Il prezzo neutrale al rischio di una call con strike  $K$  e scadenza  $T$  è:

$$C(S_0, K, T) = P(t, T) \int_0^\infty (S_T - K)^+ f(S_T, T) dS_T \quad (3.12)$$

$$= P(t, T) \int_K^\infty (S_T - K) f(S_T, T) dS_T \quad (3.13)$$

dove  $P(t, T)$  è il fattore di sconto, mentre  $f(S_T, T)$  è la densità marginale dello spot alla scadenza  $T$ .

Differenziando due volte su  $K$  si ha:

$$\frac{\partial C}{\partial K} = P(t, T) \int_K^\infty \frac{\partial}{\partial K} (S_T - K) f(S_T, T) dS_T \quad (3.14)$$

$$\frac{\partial C}{\partial K} = -P(t, T) \int_K^\infty f(S_T, T) dS_T \quad (3.15)$$

$$\frac{\partial^2 C}{\partial K^2} = P(t, T) f(K, T) \quad (3.16)$$

con  $\lim_{S \rightarrow \infty} f(S, t) = 0$ .

Derivando invece rispetto alla scadenza  $T$  si ha:

$$\frac{\partial C}{\partial T} = \frac{\partial}{\partial T} P(t, T) \int_K^\infty (S_T - K) f(S_T, T) dS_T + P(t, T) \int_K^\infty (S_T - K) \frac{\partial}{\partial T} [f(S_T, T)] dS_T \quad (3.17)$$

Siccome  $P(t, T) = \exp(-\int_t^T r_s ds)$  allora  $\frac{\partial P}{\partial T} = -r_T P(t, T)$ :

$$= -r_T C(S_0, K, T) + P(t, T) \int_K^\infty (S_T - K) \frac{\partial}{\partial T} [f(S_T, T)] dS_T \quad (3.18)$$

La densità  $f(S_t, t)$  segue una PDE di Fokker-Planck e quindi soddisfa:

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial S}[\mu S f(S, t)] + \frac{1}{2} \frac{\partial^2}{\partial S^2}[\sigma^2 S^2 f(S, t)] \quad (3.19)$$

dunque andando a sostituire la 3.19 con  $t = T$  nell'equazione (3.18) otteniamo:

$$\frac{\partial C}{\partial T} + r_T C(S_0, K, T) = P(t, T) \int_K^\infty (S_T - K) \times \quad (3.20)$$

$$\left\{ -\frac{\partial}{\partial S}[(r_T - q_T) S f(S, T)] + \frac{1}{2} \frac{\partial^2}{\partial S^2}[\sigma(S, T)^2 S^2 f(S, T)] \right\} dS \quad (3.21)$$

Partendo da quest'ultima equazione, usando le relazioni (3.15) e (3.16), infine integrando per parti [31], si ottiene quanto segue:

**Definizione 3.2 (Equazione di Dupire).** *Sia  $C(T, K)$  il prezzo di una call europea con strike  $K$  e scadenza  $T$ . Allora vale che:*

$$\frac{\partial C}{\partial T} = \frac{1}{2} \sigma^2(K, T) K^2 \frac{\partial^2 C}{\partial K^2} + (r_T - q_T)(C - K \frac{\partial C}{\partial K}) - r_T C \quad (3.22)$$

**Definizione 3.3 (Volatilità Locale di Dupire).** *Risolviendo l'equazione (3.22) per  $\sigma^2 = \sigma(K, T)^2$  si ottiene:*

$$\sigma_{Dup}(K, T)^2 = \frac{\frac{\partial C}{\partial T} + q_T C + (r_T - q_T) K \frac{\partial C}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}} \quad (3.23)$$

Non mancano però i problemi nell'applicare tale formula. Ad esempio per strikes molto *in o out the money* si possono avere delle approssimazioni numeriche poco accurate in quanto numeratore e denominatore potrebbero essere molto piccoli. Il calcolo numerico delle derivate di  $C$  non è stabile in quanto piccole variazioni di  $C$  possono causare enormi variazioni delle derivate. Inoltre i dati di mercato forniscono prezzi solamente per un insieme limitato di strike e scadenze, mentre per poter calcolare le derivate nell'equazione di Dupire bisogna conoscere il prezzo dell'opzione per ogni scadenza  $T$  e per ogni strike  $K$ . Il valore di  $C$  in un qualche punto non considerato dai dati di mercato si può ottenere attraverso tecniche di interpolazione e estrapolazione soggette però ad errore (anche per via della mancata robustezza del calcolo delle derivate) che va ad inficiare il risultato finale [20].

### 3.3.2 Legame tra Modelli LV e SV

I modelli a volatilità locale possono essere pensati come un caso particolare dei modelli a volatilità stocastica nella quale la volatilità locale è vista come una media delle possibili realizzazioni della volatilità istantanea [12].

**Proposizione 3.4.** *Definiamo il seguente modello sotto la misura di probabilità neutrale al rischio:*

$$dS_t = f(V_t)S_t dW_t^1 \quad (3.24)$$

$$dV_t = \alpha(V_t, t)dt + \beta(V_t, t)dW_t^2 \quad (3.25)$$

$$dW_t^1 dW_t^2 = \rho(V_t, t)dt \quad (3.26)$$

con i coefficienti che soddisfano le ipotesi che garantiscono esistenza e unicità della soluzione.

Il modello appena descritto è equivalente ad un modello a volatilità locale come in 3.11 nella quale:

$$\sigma^2(K, T, S_0) = \mathbb{E}^Q[f(V_t)^2 | S_T = K] \quad (3.27)$$

cioè la varianza locale è uguale al valore atteso condizionato di quella istantanea.

#### Dimostrazione:

Sia  $K$  lo strike e  $T$  la scadenza di una Call Europea e prendiamo in considerazione il suo valore non scontato:

$$C(S_0, K, T) = \mathbb{E}^Q[(S_T - K)^+]$$

Differenziando una volta rispetto a  $K$  si ottiene:

$$\frac{\partial C}{\partial K} = -\mathbb{E}^Q[\theta(S_T - K)]$$

dove  $\theta$  è la funzione di Heaviside. Derivando una seconda volta si ha:

$$\frac{\partial^2 C}{\partial K^2} = \mathbb{E}^Q[\delta(S_T - K)]$$

con  $\delta$  la funzione di Dirac.

Ora applicando il lemma di Ito al payoff si ha:

$$d(S_T - K)^+ = \theta(S_T - K)dS_T + \frac{1}{2}f(V_T)^2 S_T^2 \delta(S_T - K)dT \quad (3.28)$$

### 3.3. CALIBRAZIONE DEL MODELLO A VOLATILITÀ LOCALE STOCASTICA 33

Calcolando l'attesa si ottiene:

$$dC = d\mathbb{E}^Q[(S_T - K)^+] = \frac{1}{2}\mathbb{E}^Q[f(V_T)^2 S_T^2 \delta(S_T - K)]dT \quad (3.29)$$

Il valore atteso possiamo riscriverlo nel seguente modo:

$$\begin{aligned} \mathbb{E}^Q[f(V_T)^2 S_T^2 \delta(S_T - K)] &= \mathbb{E}^Q[f(V_T)^2 | S_T = K] K^2 \mathbb{E}^Q[\delta(S_T - K)] \\ &= \mathbb{E}^Q[f(V_T)^2 | S_T = K] K^2 \frac{\partial^2 C}{\partial K^2} \end{aligned}$$

e dunque possiamo riscrivere la 3.29 così:

$$\frac{\partial C}{\partial T} = \frac{1}{2}\mathbb{E}^Q[f(V_T)^2 | S_T = K] K^2 \frac{\partial^2 C}{\partial K^2} \quad (3.30)$$

confrontando questa equazione con quella di Dupire (3.10) si giunge alla tesi:

$$\sigma^2(K, T, S_0) = \mathbb{E}^Q[f(V_t)^2 | S_T = K]$$

Questo risultato è il caso particolare di una affermazione più grande: il *Teorema o Lemma di Gyöngy*.

**Teorema 3.5 (Lemma di Gyöngy).** *Sia  $(X_t)_t$  un processo di Ito  $n$ -dimensionale che soddisfa la seguente SDE:*

$$dX_t = \beta_t dt + \sigma_t dW_t, \quad t \in [0, T] \quad (3.31)$$

*Con condizione iniziale  $X_0 = x_0$ ,  $(W_t)_t$  moto Browniano  $d$ -dimensionale su uno spazio di probabilità filtrato  $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$ ,  $\beta_t$  processo stocastico  $n$ -dimensionale e  $\sigma_t$  processo stocastico  $n \times d$ -dimensionale. Sia inoltre  $\sigma_t \sigma_t^*$  uniformemente definita positiva. Definiamo:*

$$\begin{aligned} b(t, x) &= \mathbb{E}[\beta_t | X_t = x] \\ v^2(t, x) &= \mathbb{E}[\sigma_t \sigma_t^* | X_t = x] \end{aligned}$$

*La SDE data dai coefficienti  $b$  e  $v$  non più casuali*

$$dY_t = b(t, Y_t)dt + v(t, Y_t)dW_t, \quad t \in [0, T] \quad (3.32)$$

*e condizione iniziale  $Y_0 = X_0$ , ammette una soluzione debole  $(Y_t)_t$  tale che  $\forall t \in [0, T]$  le variabili aleatorie  $Y_t$  e  $X_t$  hanno la stessa distribuzione.*

**Definizione 3.6 (Soluzione Debole).** *Sia data l'equazione differenziale stocastica:*

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \quad (3.33)$$

$$X_0 = x_0 \quad (3.34)$$

*Si dice che questa SDE ha soluzione debole se esiste uno spazio di probabilità  $(\Omega, \mathcal{F}, \mathbb{P})$  e un moto Browniano  $\mathcal{F}_t$ -adattato tale che esista un processo  $\bar{X}$ ,  $\mathcal{F}_t$ -adattato, che soddisfi la seguente condizione:*

$$\bar{X}_t = x_0 + \int_0^t \mu(s, \bar{X}_s)ds + \int_0^t \sigma(s, \bar{X}_s)d\bar{W}_s, \quad \forall t > 0 \quad (3.35)$$

Per la dimostrazione del teorema si rimanda a [15].

### 3.3.3 Leverage Function

Sfruttando il Teorema di Gyöngy 3.5 nel caso particolare della relazione 3.27 si ha che la volatilità locale di un processo 3.1 soddisfa tale relazione:

$$\sigma_{loc}^2(K, T) = \mathbb{E}^Q[(L(S_T, T)\sigma_1(S_T, V_T, T))^2 | S_T = K] \quad (3.36)$$

La leverage function esce dal valore atteso in quanto deterministica e quindi si ha che:

$$= L(K, T)^2 \mathbb{E}^Q[\sigma_1(S_T, V_T, T)^2 | S_T = K] \quad (3.37)$$

Dagli studi di Dupire precedentemente citati possiamo affermare che il processo è calibrato agli smile di mercato se e solo se 3.37 coincide con la varianza di Dupire, o più in generale ad una volatilità locale calibrata sul mercato):

$$\sigma_{loc}(K, T) = \sigma_{Dup}(K, T) \quad (3.38)$$

e dunque:

$$L(K, T) = \frac{\sigma_{Dup}(K, T)}{\sqrt{\mathbb{E}^Q[\sigma_1(S_T, V_T, T)^2 | S_T = K]}} \quad (3.39)$$

Il valore atteso del denominatore è ciò che manca da calcolare. Possiamo però osservare che prendendo in considerazione la probabilità congiunta  $p(S_t, V_t, t)$  di  $(S_t, V_t)$  si ha:

$$\mathbb{E}^Q[\sigma_1(S_t, V_t, t)^2 | S_t = K] = \int_0^\infty \sigma(S_t, V, t)^2 p_{V_t | S_t=K}(V, t) dV \quad (3.40)$$

$$= \int_0^\infty \sigma_1(S_t, V_t, t)^2 \frac{p_{S_t, V_t}(K, V, t)}{p_{S_t}(K, t)} dV \quad (3.41)$$

$$= \frac{\int_0^\infty \sigma_1(S_t, V_t, t)^2 p_{S_t, V_t}(K, V, t) dV}{\int_0^\infty p_{S_t, V_t}(K, V, t) dV} \quad (3.42)$$

La leverage function si può riscrivere nella seguente formula:

$$L(t, S, p) = \sigma_{Dup}(S, t) \sqrt{\frac{\int p(t, S, V) dV}{\int \sigma_1(S, V, t)^2 p(t, S, V) dV}} \quad (3.43)$$

La leverage function dipende dalla distribuzione congiunta  $p(S, V, t)$  di  $(S_t, V_t)$ . Approssimare questa densità di probabilità equivale a calibrare la leverage function.

Dunque l'evoluzione del sottostante in 3.1 segue la SDE seguente che rientra nella classe delle equazioni non lineari (in quanto la relativa equazione di Fokker-Planck lo è) di McKean:

$$dS_t = \mu_1(S_t, t) dt + \sigma_{Dup}(S_t, t) \sqrt{\frac{\int p(S_t, u, t) du}{\int \sigma_1(S_t, V_t, t)^2 p(S_t, u, t) du}} \sigma_1(S_t, V_t, t) dW_t^1 \quad (3.44)$$

Il risultato di esistenza e unicità esposto nel primo capitolo non può essere applicato all'equazione (3.44) in quanto in essa non è verificata la condizione di Lipschitzianità. L'esistenza e l'unicità della soluzione perciò non sono ovvie.

Nel lavoro di Abergel e Tachet [1] è stato dimostrato che il problema della calibrazione di un modello a volatilità locale-stocastica è ben posto solamente sotto le seguenti condizioni:

- In un dominio limitato  $[S_{min}, S_{max}] \times [V_{min}, V_{max}]$  e per scadenze a breve termine;
- Se la vol-of-vol è sufficientemente piccola;
- Nel caso di condizioni iniziali e al bordo adeguatamente regolari.

Le condizioni iniziali nel caso da noi considerato coincidono con la Delta di Dirac e dunque questo risultato non può essere applicato. In una recente pubblicazione Jourdain e Zhou hanno mostrato ulteriori risultati parziali dalla quale si può dedurre che vi è una soluzione unica solamente fino ad una certa maturità critica [21]. Intuitivamente se la vol-of-vol è piccola, il valore critico è molto lontano e non si raggiungerà mai. Se invece

la vol-of-vol è abbastanza grande, il valore critico risulta sufficientemente vicino e ciò significa che quando si prezza un contratto con scadenza, ad esempio, fra qualche anno, è possibile che il punto critico venga superato e che non esista quindi una soluzione della PDE di Fokker Planck associata.

### 3.4 Particle Method nei Modelli LSV

Il Particle Method è uno dei tanti metodi per la risoluzione del problema di calibrazione della leverage function per i modelli a volatilità locale-stocastica. Inserendo la distribuzione empirica nella (3.39) si ottiene:

$$\mathbb{E}^{Q^i} [\sigma_1(S_t, V_t, t)^2 | S_t = K] = \frac{\int_0^\infty \sigma_1(S_t, V, t)^2 p_N(K, V, t) dV}{\int_0^\infty p_N(K, V, t) dV} \quad (3.45)$$

$$= \frac{\sum_{i=1}^N \left( \sigma_1(S_t^{i,N}, V_t^{i,N}, t) \right)^2 \delta \left( S_t^{i,N} - K \right)}{\sum_{i=1}^N \delta \left( S_t^{i,N} - K \right)} \quad (3.46)$$

E andando a sostituire nella Leverage Function si ha:

$$L_N(S, t) = \sigma_{Dup}(S, t) \sqrt{\frac{\sum_{i=1}^N \delta \left( S_t^{i,N} - S \right)}{\sum_{i=1}^N \left( \sigma_1(S_t^{i,N}, V_t^{i,N}, t) \right)^2 \delta \left( S_t^{i,N} - S \right)}} \quad (3.47)$$

A questo punto il problema diventa l'approssimazione della delta di Dirac che viene effettuata con i kernel regolarizzatori (3.4.1) ottenendo così:

$$L_N(S, t) = \sigma_{Dup}(S, t) \sqrt{\frac{\sum_{i=1}^N \delta_{t,N} \left( S_t^{i,N} - S \right)}{\sum_{i=1}^N \left( \sigma_1(S_t^{i,N}, V_t^{i,N}, t) \right)^2 \delta_{t,N} \left( S_t^{i,N} - S \right)}} \quad (3.48)$$

Avendo il dato iniziale  $L_N(S_0, 0)$  si ha tutto ciò che serve per la calibrazione.

### 3.4.1 Kernel Regolarizzatori

L'approssimazione della Delta di Dirac è un problema di notevole importanza nella calibrazione dei modelli LSV col Particle Method. Per ovviare a questo problema si usano i *kernel regolarizzatori* cioè funzioni lisce, simmetriche e con integrale uguale ad uno. Uno dei più famosi è il kernel gaussiano:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (3.49)$$

Ma il kernel più utilizzato da Guyon e Labordère nei loro lavori (fra cui [14]) è il kernel quartico in quanto richiede un minor costo computazionale:

$$K(x) = \frac{15}{16} (1 - x^2)^2 I_{\{|x| \leq 1\}} \quad (3.50)$$

La delta di Dirac si approssima quindi nel seguente modo sfruttando i kernel regolarizzatori:

$$\delta_{t,N}(x) = \frac{1}{h_{t,N}} K\left(\frac{x}{h_{t,N}}\right) \quad (3.51)$$

dove  $h$  è la larghezza di banda, che tende a zero al crescere del numero di particelle  $N$ , e a seconda del suo valore cambia il supporto del kernel fissato  $K$ . Così facendo il problema si trasforma nella scelta di una buona larghezza di banda. Guyon e Henry-Labordère in [13] hanno proposto un valore di banda che decresca come  $N^{-\frac{1}{5}}$  per valori molto grandi di  $N$ . In particolare:

$$h = k\bar{\sigma}S_0\sqrt{t}N^{-\frac{1}{5}}$$

dove:

- $k$  è un numero adimensionale che viene scelto seguendo la *rule of thumb* di Silverman [33];
- $\bar{\sigma}$  un livello di volatilità;
- $S_0\sqrt{t}$  la deviazione standard di  $S$ .

### 3.4.2 Algoritmo Particle Method

Un possibile algoritmo per l'applicazione del Particle Method alla calibrazione della leverage function di un modello 3.1 è il seguente:



**Proposizione 3.7.** 1. Prendiamo in considerazione una griglia temporale  $t_k$  per l'intervallo  $[0, T]$ .

2. Inizializziamo  $k = 1$  e poniamo il seguente dato iniziale:

$$L_N(t, S_t) = \frac{\sigma_{Dup}(S_0, 0)}{\sigma_1(S_0, V_0, 0)}, \quad \forall t \in [0, t_1]$$

3. Simuliamo il moto di  $N$  particelle tra  $t_{k-1}$  e  $t_k$  usando uno schema di discretizzazione come quello di Eulero:

$$dS_t^{i,N} = \mu_1(S_t^{i,N}, t)dt + L_N(S_t^{i,N}, t)\sigma_1(S_t^{i,N}, V_t^{i,N}, t)dW_t^{1,i} \quad (3.52)$$

$$dV_t^{i,N} = \mu_2(V_t^{i,N}, t)dt + \sigma_2(V_t^{i,N}, t)dW_t^{2,i} \quad (3.53)$$

4. Calcoliamo la  $L_N(S_{t_k}^{i,N}, t_k)$  con la relazione 3.48 e poniamo

$$L_N(S_t, t) = L_N(S_t, t_k), \quad \forall t \in [t_k, t_{k+1}]$$

5. Infine si pone  $k = k + 1$  e si torna al punto 3 fino alla maturità  $T$ .

*Costo Computazionale:*  $\mathcal{O}(N^2)$  ad ogni passo. Una  $N$  dovuta al numero di stime della leverage function ed un'altra derivante dalle sommatorie.

Il costo computazionale renderebbe alquanto complicato applicare tale metodo se non fosse per alcune tecniche di accelerazione. Si può ad esempio evitare di calcolare la  $L_N(S_{t_k}^{i,N}, t_k)$  per ogni  $i$  e farlo solamente per un numero fissato di particelle molto minore  $N_{S,t} \ll N$  poste su una griglia  $G_{S,t}$  di valori di  $S$ . Inoltre nella 3.48 vi sono tantissimi elementi dal contributo trascurabile nelle sommatorie e pertanto nella valutazione della leverage function si possono considerare le particelle che soddisfano la seguente relazione ignorando le altre:

$$\delta_{t,N} \left( S_t^{i,N} - S \right) > \eta$$

con  $\eta = 10^{-3}$ . Preso il vettore delle particelle, si possono ordinare i suoi elementi in maniera efficiente con un costo pari a  $\mathcal{O}(N \log N)$  e trovare quindi i valori limite  $i_{min}(S)$  e  $i_{max}(S)$ , che soddisfano la condizione precedentemente illustrata, e da questi i valori  $S_t^{i_{min}(S),N}$  e  $S_t^{i_{max}(S),N}$ :

$$S_t^{i_{min}(S),N} = S - h_{t,N} K^{-1}(h_{t,N} \eta) \quad (3.54)$$

$$S_t^{i_{max}(S),N} = S + h_{t,N} K^{-1}(h_{t,N} \eta) \quad (3.55)$$

e questo porterebbe ad un notevole risparmio di operazioni.

Guyon e Labordère hanno potuto constatare che il Particle Method converge con un numero relativamente piccolo di particelle, qualsiasi sia la volatilità stocastica scelta, ed inoltre, grazie alle tecniche di velocizzazione che si possono adottare, hanno potuto apprezzare che il Particle Method è veloce quanto l'algoritmo di Monte Carlo standard [14] [13]. Hanno potuto infine osservare che si ottengono risultati migliori, soprattutto per le opzioni a breve scadenza, adottando una griglia temporale non uniforme.

I due moti Browniani possono essere simulati da  $\Delta W_{t_k}^{i,j} = \sqrt{t_{k+1} - t_k} \cdot Z^{i,j}$  con  $j = 1, 2$  e dove  $(Z^{i,1}, Z^{i,2}) \sim \mathcal{N}_{0,1}$  sono due numeri casuali generati da una distribuzione normale standard e associati alla particella  $i$  nel tempo  $t_k$ .



# Capitolo 4

## Conclusioni

In questo elaborato è stata descritta una SDE di McKean e si è dimostrata l'unicità e l'esistenza di una sua soluzione nel caso particolare in cui il drift e la volatilità siano lipshitziani rispetto alla metrica di Wasserstein. È stato poi descritto il Particle Method, uno dei metodi per trovare la soluzione di una SDE di McKean, nella quale la distribuzione incognita è approssimata con una distribuzione empirica mediante una simulazione di tipo Monte Carlo di cammini simultanei interagenti. La convergenza del Particle Method in una SDE di McKean-Vlasov è stata dimostrata sfruttando la proprietà di propagazione del caos introdotta da Sznitman [35]. I concetti matematici descritti trovano una applicazione in matematica finanza nella calibrazione di un modello di option pricing a volatilità locale-stocastica.

Sono stati introdotti i modelli LSV con accenni alle motivazioni che hanno portato all'introduzione di questa nuova classe. Sono stati descritti due modelli molto conosciuti in letteratura: Modello di Dupire-Heston e di SABR. La calibrazione è il problema più grande dei modelli LSV. Essa richiede innanzitutto la calibrazione della parte locale. Sono stati quindi esposti i risultati ottenuti da Dupire [9][10] ed è stata ricavata l'equazione omonima che garantisce la calibrazione della parte locale del modello. Dopo aver mostrato il legame fra i modelli LV e i modelli SV, si è affrontato il problema della calibrazione della *leverage function*. Abbiamo illustrato come il Particle Method può essere applicato nella calibrazione della Leverage Function sfruttando il fatto che la dinamica di un asset nel modello LSV è descritta da una SDE di McKean. Infine è stato affrontato il problema

dell'approssimazione di una delta di Dirac con kernel regolarizzatori e sono stati esposti un possibile algoritmo basato sul Particle Method ed alcune tecniche di accelerazione.

Ulteriori studi potrebbero riguardare l'estensione del modello LSV anche a modelli che prevedono salti. Se i salti vengono inclusi nel processo del sottostante i risultati sono analoghi a quelli che si otterrebbero con la classe di modelli descritti in questo elaborato (vedasi [14]). Se invece la discontinuità viene introdotta nel processo della volatilità stocastica ciò comporterebbe il dover affrontare una notevole estensione delle fondamenta teoriche [2][5]. Si potrebbero poi considerare modelli con tassi di interesse stocastici ed in tal caso l'applicazione del Particle Method si estenderebbe a comprendere anche questa opportunità [14].

Si potrebbe infine testare, con un programma quale ad esempio Matlab, l'applicazione del Particle Method per l'approssimazione di una soluzione di una SDE di McKean oppure per risolvere il problema della calibrazione di un modello LSV confrontando i risultati ottenuti con altre tecniche numeriche e controllando soprattutto l'esistenza effettiva di una soluzione. Non c'è infatti alcuna certezza sull'esistenza e l'unicità in quanto i coefficienti dell'equazione (3.44) non soddisfano le condizioni di Lipschitzianità. In [14], come in altri lavori di Guyon, è stato osservato empiricamente che l'algoritmo fallisce per valori molto grandi della vol-of-vol.

# Capitolo 5

## Appendice

**Definizione 5.1 (Processo Stocastico Discreto).** *Un processo stocastico discreto su  $\mathbb{R}^N$  è una famiglia di variabili aleatorie  $X = (X_n)_{n \in \mathbb{N}_0}$  definite su uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$  a valori su  $\mathbb{R}^N$ :*

$$X_n : \Omega \longrightarrow \mathbb{R}^N, \quad n \in \mathbb{N}_0$$

- *Una filtrazione sullo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  è una famiglia  $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$  crescente di sotto  $\sigma$ -algebre di  $\mathcal{F}$ .*
- *Si chiama filtrazione naturale per  $X$  la famiglia di  $\sigma$ -algebre  $(\mathcal{F}_n^X)_{n \in \mathbb{N}_0}$  definita da:*

$$\mathcal{F}_n^X = \sigma(X_k, 0 \leq k \leq n)$$

- *Il processo  $X$  si dice adattato alla filtrazione  $(\mathcal{F}_n)$  se  $X_n$  è  $\mathcal{F}_n$ -misurabile  $\forall n \in \mathbb{N}_0$ .*
- *$X$  si dice processo integrabile se  $X_n \in L^1(\Omega, P), \quad \forall n \in \mathbb{N}_0$ .*

*La legge di un processo stocastico discreto  $X = (X_0, \dots, X_n)$  è definita come la distribuzione congiunta delle variabili aleatorie  $X_0, \dots, X_n$ .*

**Definizione 5.2 (Processo Stocastico).** *Un processo stocastico (misurabile) in  $\mathbb{R}^N$  è una famiglia  $(X_t)_{t \in I}$  di variabili aleatorie a valori in  $\mathbb{R}^N$  tale che l'applicazione*

$$X : I \times \Omega \longrightarrow \mathbb{R}^N, \quad X(t, \omega) = X_t(\omega)$$

*è misurabile rispetto alla  $\sigma$ -algebra prodotto  $\mathcal{F} \times \mathcal{B}(I)$ . Il processo stocastico  $X$  quindi associa a  $t$  la variabile aleatoria  $X_t$  in  $\mathbb{R}^N$ .*

- Analogamente al caso discreto si definiscono filtrazione, filtrazione naturale e processo adattato.

**Definizione 5.3 (Martingale).** Sia  $M$  un processo stocastico sommabile e adattato nello spazio con filtrazione  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ . Diciamo che  $M$  è:

- Una martingala rispetto a  $(\mathcal{F})_t$  e alla misura  $P$  se:

$$M_s = \mathbb{E}[M_t | \mathcal{F}_s], \quad \text{per ogni } 0 \leq s \leq t$$

- Una super-martingala se:

$$M_s \geq \mathbb{E}[M_t | \mathcal{F}_s], \quad \text{per ogni } 0 \leq s \leq t$$

- Una sub-martingala se:

$$M_s \leq \mathbb{E}[M_t | \mathcal{F}_s], \quad \text{per ogni } 0 \leq s \leq t$$

e vale che:

$$\mathbb{E}[M_t] = \mathbb{E}[\mathbb{E}[M_t | \mathcal{F}_0]] = \mathbb{E}[M_0] \quad \forall t \geq 0$$

**Definizione 5.4 (Moto Browniano).** Sia  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$  uno spazio di probabilità con filtrazione. Un moto browniano reale è un processo stocastico  $W = (W_t)_{t \geq 0}$  in  $\mathbb{R}$  tale che:

- $W_0 = 0$  q.s.;
- $W$  è  $\mathcal{F}_t$ -adattato e continuo;
- per  $t > s \geq 0$ , la variabile aleatoria  $W_t - W_s$  ha distribuzione normale  $\mathcal{N}_{0, t-s}$  e è indipendente da  $\mathcal{F}_s$ .

**Definizione 5.5 (Legge di un Processo Continuo).** • Sia  $C[0, T] := C([0, T], \mathbb{R}^N)$  lo spazio delle funzioni continue sull'intervallo  $[0, T]$  a valori in  $\mathbb{R}^N$ .

- $C[0, T]$  munito della norma  $\|\cdot\|_\infty$  è uno spazio normato completo. La norma definisce gli aperti di  $C[0, T]$  e quindi la sua  $\sigma$ -algebra dei Borelliani indicata con  $\mathcal{B}(C[0, T])$ .

•

$$X_t : C[0, T] \longrightarrow \mathbb{R}^N, \quad X_t(\omega) = \omega(t)$$

definisce un processo stocastico su  $(C[0, T], \mathcal{B}(C[0, T]))$ .

• Vale che  $\sigma(X_t, t \in [0, T]) = \mathcal{B}(C[0, T])$ .• Dato un processo stocastico continuo  $(X_t)_{t \in [0, T]}$  sullo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  vale:

$$\{X \in H\} \in \mathcal{F}, \quad \forall H \in \mathcal{B}(C[0, T])$$

•

• La probabilità  $P^X$  sullo spazio  $(C[0, T], \mathcal{B}(C[0, T]))$  definita da:

$$P^X(H) = P(X \in H), \quad H \in \mathcal{B}(C[0, T])$$

è detta legge del processo  $X$ .

**Definizione 5.6 (Processo di Ito).** Un processo di Ito è un processo stocastico  $X$  della forma:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad t \in [0, T]$$

dove  $X_0$  è una v.a.  $\mathcal{F}_0$ -misurabile,  $\mu \in \mathbb{L}_{loc}^1$  e  $\sigma \in \mathbb{L}_{loc}^2$ . Nella forma differenziale si ha:

$$dX_t = \mu_t dt + \sigma_t dW_t$$

dove il processo  $\mu$  è chiamato drift, mentre il processo  $\sigma$  è detto coefficiente di diffusione.

**Definizione 5.7 (Lemma (Formula) di Ito).** Sia  $X$  un processo di Ito e  $f = f(t, x) \in C^{1,2}(\mathbb{R}^2)$ . Il processo stocastico

$$Y_t = f(t, X_t)$$

è un processo di Ito e si ha

$$df(t, X_t) = \partial_t f(t, X_t) dt + \partial_x f(t, X_t) dX_t + \frac{1}{2} \partial_{xx} f(t, X_t) d\langle X_t \rangle$$

con  $d\langle X_t \rangle = \sigma_t^2 dt$  e dunque la formula precedente si può riscrivere nel seguente modo:

$$df = \left( \partial_t f + \mu_t \partial_x f + \frac{1}{2} \sigma_t^2 \partial_{xx} f \right) dt + \sigma_t \partial_x f dW_t$$

•



**Definizione 5.8 (Moto Browniano  $d$ -dimensionale).** Sia  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$  uno spazio di probabilità con filtrazione. Un moto browniano  $d$ -dimensionale è un processo stocastico  $W = (W_t)_{t \in [0, \infty[}$  in  $\mathbb{R}^d$  tale che:

- $W_0 = 0$   $P$ -q.s.;
- $W$  è un processo stocastico  $\mathcal{F}_t$ -adattato e continuo;
- per  $t > s \geq 0$ , la variabile aleatoria  $W_t - W_s$  ha distribuzione multi-normale  $N_{0, (t-s)I_d}$ , dove  $I_d$  indica la matrice identità  $d \times d$ , ed è indipendente da  $\mathcal{F}_s$ .

Pertanto si può dire come diretta conseguenza che se  $W = (W^1, \dots, W^d)$  è un moto Browniano  $d$ -dimensionale su  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ , allora  $\forall i = 1, \dots, d$  si ha che:

- $W^i$  è un moto Browniano reale su  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$  e quindi una  $\mathcal{F}_t$ -martingala;
- $W_t^i$  e  $W_t^j$  sono variabili aleatorie indipendenti per  $j \neq i$  e  $t \geq 0$ .

**Definizione 5.9 (Lemma (Formula) di Ito multi-dimensionale).** Sia  $W$  un moto browniano  $d$ -dimensionale. Sia  $f = f(t, x_1, \dots, x_d) \in C^{1,2}(\mathbb{R}^{d+1})$ . Allora  $f(t, W_t)$  è un processo di Ito e vale:

$$df(t, W_t) = \partial_t f(t, W_t)dt + \sum_{i=1}^d \partial_{x_i} f(t, W_t)dW_t^i + \frac{1}{2} \sum_{i=1}^d \partial_{x_i x_i} f(t, W_t)dt$$

**Definizione 5.10 (Lemma (Formula) di Ito multi-dimensionale generale).** Sia  $dX_t = u_t dt + v_t dW_t$  processo a valori in  $\mathbb{R}^N$  con  $u, v \in \mathbb{L}_{loc}^2$ ,  $u$  vettore  $N \times 1$  mentre  $\sigma$  matrice  $N \times d$ . Sia  $f = f(t, x) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^N)$  allora vale:

$$df(t, X_t) = \partial_t f(t, X_t)dt + \nabla f(t, X_t) \cdot dX_t + \frac{1}{2} \sum_{i,j=1}^N \partial_{x_i x_j} f(t, X_t) d\langle X^i, X^j \rangle_t$$

**Definizione 5.11 (Disuguaglianza Massimale di Doob).** Sia  $M$  una martingala nello spazio  $(\Omega, \mathcal{F}, P, \mathcal{F}_n)$ .  $\forall N \in \mathbb{N}$  e  $p \in \mathbb{R}$ , con  $p > 1$ , vale:

$$E \left[ \max_{0 \leq n \leq N} |M_n|^p \right] \leq q^p E[|M_N|^p]$$

dove  $q = \frac{p}{p-1}$ .

Possiamo inoltre formulare la seguente estensione al caso continuo:

Sia  $M$  una martingala continua a destra e  $p > 1$ . Allora  $\forall T$  vale:

$$E \left[ \sup_{t \in [0, T]} |M_t|^p \right] \leq q^p E[|M_T|^p]$$

**Definizione 5.12 (Ipotesi Usuali).** Sia  $(\Omega, \mathcal{F}, P)$  uno spazio di probabilità. Sia

$$\mathcal{N} = \{F \in \mathcal{F} | P(F) = 0\}$$

la famiglia di eventi trascurabili. Diciamo che  $(\mathcal{F}_t)$  è una filtrazione standard se soddisfa le "ipotesi usuali":

- $\mathcal{F}_0$  contiene  $\mathcal{N}$ ;
- la filtrazione è continua a destra cioè, per ogni  $t \geq 0$ , vale:

$$\mathcal{F}_t = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$$

**Definizione 5.13 (Equazione Differenziale Stocastica).** Sia  $Z \in \mathbb{R}^N$  e siano  $b$  e  $\sigma$  due funzioni misurabili deterministiche:

$$b = b(t, x) : [0, T] \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$$

$$\sigma = \sigma(t, x) : [0, T] \times \mathbb{R}^N \longrightarrow \mathbb{R}^{N \times d}$$

Sia  $W$  un moto browniano  $d$ -dimensionale sullo spazio di probabilità filtrato  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$  in cui valgono le ipotesi usuali. Una soluzione relativa a  $W$  della SDE di coefficienti  $Z, b, \sigma$  è un processo  $\mathcal{F}_t$ -adattato e continuo  $(X_t)_{t \in [0, T]}$  tale che:

- $b(t, X_t) \in \mathbb{L}_{loc}^1$  e  $\sigma(t, X_t) \in \mathbb{L}_{loc}^2$ ;

- vale:

$$X_t = Z + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad t \in [0, T]$$

che si può anche scrivere nel seguente modo:

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dW_t, \quad X_0 = Z$$

**Definizione 5.14 (Soluzione di una SDE).** • *La SDE di coefficienti  $Z, b, \sigma$  si dice risolvibile in senso debole se esiste un moto Browniano standard relativamente al quale la SDE ha soluzione.*

- *La SDE di coefficienti  $Z, b, \sigma$  si dice risolvibile in senso forte se, per ogni moto Browniano standard  $W$  assegnato, esiste una soluzione relativa a  $W$ .*

**Definizione 5.15 (Convergenza Debole).** *Sia  $(\mu_n)$  una successione di probabilità su  $\mathbb{R}^n$ , e  $\mu$  una probabilità su  $\mathbb{R}^n$ . Diciamo che  $\mu_n$  converge debolmente a  $\mu$  se per ogni  $f \in C_b$*

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu$$

**Definizione 5.16 (Convergenza in Distribuzione).** *Una successione di v.a.  $(X_n)$  converge in distribuzione, o in legge, a una v.a.  $X$  se la corrispondente successione di distribuzioni  $(\mu_n)$  converge debolmente a  $\mu_X$ . In tal caso scriviamo:*

$$X_n \xrightarrow{d} X$$

# Bibliografia

- [1] Abergel, F. and Tachet, R.: *A nonlinear partial integro-differential equation from mathematic finance*, 2010.
- [2] Bentata, A. and Cont, R.: *Mimicking the Marginal Distributions of a Semimartingale*, 2009.
- [3] Bergomi, L.: *Stochastic Volatility Modeling: A Practitioner's Approach*, Third SMAI European School of Financial Mathematics, 2010.
- [4] Black, F. and Scholes, M.: *The Pricing of Option and Corporate Liabilities*, Journal of Political Economy 81, pp. 637-659, 1973.
- [5] Cont, R. and Tankov, P.: *Financial Modelling With Jump Process*, Chapman and Hall/CRC Financial Mathematics Series, 2003.
- [6] Cox, J.C., Ingersoll, J.E. and Ross, S.: *A theory of the term structure of interest rate*, Econometrica, 54: 373-384, 1985.
- [7] Crosby, J.: *Practicalities of pricing exotic derivatives*, [http://www.john-crosby.co.uk/pdfs/JCrosby\\_OxfordJune2003\\_Exotics.pdf](http://www.john-crosby.co.uk/pdfs/JCrosby_OxfordJune2003_Exotics.pdf), 2013.
- [8] Dobrushin, R. L.: *Prescribing a system of random variables by conditional distributions*, Th. Probab. and its Applic. 3, 469 (1970).
- [9] Dupire, B.: *Pricing With a Smile*, Risk magazine, 7, pp.18-20, 1994.
- [10] Dupire, B.: *A Unified Theory of Volatility*, In Derivatives Pricing: The Class Collection, edited by Peter Carr, Risk Publications, 1996.

- [11] Gangbo, W. and McCann, R.J.: *The Geometry Of Optimal Transportation*, Acta Math. 177:113-161, 1996.
- [12] Gatheral, J.: *The Volatility Surface. A Practitioner's Guide*, John Wiley & Sons Inc, 2006.
- [13] Guyon, J. and Henry-Labordère, P.: *The Smile Calibration Problem Solved*, Available at SSRN: <https://ssrn.com/abstract=1885032> or <http://dx.doi.org/10.2139/ssrn.1885032>, 2011.
- [14] Guyon, J. and Henry-Labordère, P.: *Non Linear Option Pricing*, Chapman & Hall/CRC Financial Mathematics Series, 2014.
- [15] Gyöngy, I.: *Mimicking the One-Dimensional Marginal Distribution of Processes Having an Ito Differential*, Probability Theory and Related Fields, 1986.
- [16] Hagan, P.S., Kumar, D., Lesniewski, A.S. and Woodward D.E.: *Managing Smile Risk*, Wilmott Magazine, 2002.
- [17] Heston, S.: *A closed-form solution for options with stochastic volatility with applications to bond and currency options*. Rev. Financ. Stud., 6: 327-343, 1993.
- [18] Homescu, C.: *Local Stochastic Volatility Models: Calibration and Pricing*, 2014.
- [19] Jex, M., Henderson, R. and Wang, D.: *Pricing exotics under the smile*, Risk, (November):72-75, 1999.
- [20] Jiang, L., Chen, Q., Wang, L. and Zhang E, J.: *A new well-posed algorithm to recover implied local volatility*, Institute of Physics Publishing, Quantitative Finance volume 3: 451-457, 2003.
- [21] Jourdain, B. and Zhou, A.: *Existence of a calibrated regime switching local volatility model and new fake Brownian motions*, 2016.
- [22] Kantorovich, L.: *On the translocation of masses*, C.R. (Doklady) Acad. Sci. URSS (N.S.), 37: 199-201, 1942.

- [23] Kantorovich, L.: *On a Problem of Monge*(in Russo), Uspekhi Math. Nauk., 3:225-226, 1948.
- [24] Lipton, A. and McGhee, W.: *Universal Barriers*, Risk Magazine, 2002.
- [25] Lipton, A.: *The Vol Smile Problem*, Risk Magazine, (February):61-65, 2002.
- [26] McKean, H.P.: *Application of Brownian Motion to the equation with nonlinear parabolic equations*, Proc. Natl. Acad. Sci. U.S.A., 56(6):1907-1911, 1966.
- [27] Méléard, S.: *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, in Probabilistic Models for Non-linear Partial Differential Equations, Lecture Notes in Mathematics 1627, Springer-Verlag, 1996.
- [28] Merton, R.: *Theory of Rational Option Pricing*, Bell J. Econ. Manag. Sci 4, 141-183, 1973.
- [29] Pascucci, A.: *Calcolo stocastico per la finanza*, UNITEXT. Springer Milan; 2008.
- [30] Reghai, A., Klaeyle, V. and Boukhaffa, A.: *LSV Models with a Mixing Weight*, 2012.
- [31] Rouah, F.D.: *Derivation of Local Volatility*, <http://www.frouah.com>
- [32] Samuelson, P. A. and Merton, R.: *A Complete Model of Warrant Pricing that Maximizes Utility*, Indus. Management Rev. 10, 17-46, 1969.
- [33] Silverman, B.W.: *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, New York, 1986.
- [34] Sudakov, V.N.: *Geometric Problems in the Theory of infinite-dimensional probability distributions*, Proc. Steklov Inst. Math., 141:1-178, 1979.
- [35] Sznitman, A.S.: *Topics in Propagation of Chaos*, Ecole d'été de probabilités de Saint-Flour XIX - 1989, volume 1464 of Lect. Notes in Math. Springer-Verlag, 1991.
- [36] Villani, C.: *Topics in Optimal Transportation*, Graduate Studies in Mathematics (58), American Mathematical Society, 2003.

- [37] Wystup, U.: *FX Options models trends*. <http://www.slideshare.net/wystup/fxmodels-slidesenglpublic> , 2013.