Market Microstructure and Algorithmic Trading

University of Oxford

Fayçal Drissi

faycal.drissi@eng.ox.ac.uk
http://www.faycaldrissi.com

March 2, 2026

Optimal control of diffusion processes

Let $(X_t^u)_{t\in[0,T]}$ denote a controlled system with dynamics

$$dX_t^u = \mu\left(t, X_t^u, u_t\right)\, dt + \sigma\left(t, X_t^u, u_t\right) dW_t, \qquad X_0^u = X_0,$$

The agent has a **performance criterion** they wish to maximise

$$\boxed{\sup_{u\in\mathcal{A}} \mathbb{E}\left[G(X_T^u) + \int_0^T F(s, X_s^u, u_s)\, ds\right]}.$$

We define a class of problems indexed by time:

$$\boxed{H(t,x) = \sup_{u \in \mathcal{A}_t} \mathbb{E}_{t,x}\left[G(X_T^u) + \int_t^T F(s, X_s^u, u_s)\, ds \right],}$$

where $\left(X_s^{x,u}\right)_{s \in [t,T]}$ follows the dynamics

$$dX_s^u = \mu\left(t, X_s^u, u_s\right)\, ds + \sigma\left(s, X_s^u, u_s\right) dW_s, \qquad X_t^u = x,$$

Dynamic Programming Principle: The value function *H* satisfies

$$\boxed{H(t,x) = \sup_{u \in \mathcal{A}_t} \mathbb{E}_{t,x}\left[H(T, X_T^u) + \int_t^T F(s, X_s^u, u_s)\, ds \right],}$$

for all $t \in [0, T]$ and $x \in \mathbb{R}$.

The DPP connects the value function to its future expected value, regularised by the expected penalty *F*.

The infinitesimal version of the DPP gives the **Dynamic Programming Equation** (DPE) or the **Hamilton-Jacobi-Bellman** equation (HJB):

$$\boxed{\begin{array}{c} \partial_t H(t,x) + \sup_{u\in\mathcal{A}}(\mathcal{L}_t^u H(t,x) + F(t,x,u)) = 0 \\ \\ \text{subject to the terminal condition } \; H(T,x) = G(x), \end{array}}$$

where $\mathcal{L}_t^u$ is the infinitesimal generator of the process $X_t^{x,u}$.

For the diffusion process

$$dX_t^u = \mu\left(t, X_t^u, u_t\right)\, dt + \sigma\left(t, X_t^u, u_t\right) dW_t, \qquad X_0^u = X_0,$$

the infinitesimal generator acts on functions *H* as follows:

$$\boxed{\mathcal{L}_t^u H(t,x) = \mu(t,x,u)\,\partial_x H(t,x) + \frac{1}{2}\sigma(t,x,u)^2\,\partial_{xx}^2 H(t,x)}$$

Sketch of the proof The dynamic programming principle gives

$$H(t,x) = \sup_u \left\{ \mathbb{E}\left[H(t+h, X^u_{t+h}) + \int_t^{t+h} F(s, X^u_s, u_s)\, ds \,\middle|\, X_t = x \right] \right\} ,$$

for any $h \in (0, T-t)$.

By Itô's formula we have

$$\begin{aligned} & H\left(t+h, X^u_{t+h}\right) - H(t,x) \\ &= \int_t^{t+h} \left(\partial_t + \mathcal{L}^u\right)[H](s, X^u_s) ds + \int_t^{t+h} \sigma(s, X^u_s, u_s) \partial_x H(s, X^u_s) dW_s \,, \end{aligned}$$

where $\mathcal{L}$ is the infinitesimal generator of X^u.

Divide by h on both sides and sending to 0, to obtain the HJB

$$\partial_t H(t,x) + \sup_u (\mathcal{L}^u_t H(t,x) + F(t,x,u)) = 0.$$

- An agent holds an **initial position** Q_0 at time $t = 0$ that they wish to **unwind** over a time window $[0, T]$.

- **Inventory**:

$$\boxed{dQ^\nu_t = \nu_t \, dt}$$

- **Price**:

$$\boxed{dS^\nu_t = \sigma \, dW_t + k \, \nu_t \, dt, \quad S_0 \in \mathbb{R}^\star_+ \text{ is known}}$$

- **Cash**:

$$\boxed{dX^\nu_t = -\nu_t \, \tilde{S}^\nu_t \, dt = -\nu_t \, (S^\nu_t + \eta \, \nu_t) \, dt, \quad X_0 \in \mathbb{R} \text{ is known.}}$$

- **Admissible** strategies: **no unwind** constraint.

Performance criterion

- The agent's objective is to complete the liquidation, but we allow to fall short of this target (no fuel constraint) so $Q_T \neq 0$.
- We add a **terminal penalty** term to include any costs incurred when trading the terminal inventory Q_T.
- To encode risk aversion or urgency, the CJ framework includes a **running inventory penalty** that does not impact the wealth.

$$\mathbb{E}\left[\underbrace{X_T^\nu}_{\text{Terminal Cash}} + \underbrace{Q_T^\nu\left(S_T^\nu - \alpha Q_T^\nu\right)}_{\text{Terminal Execution}} - \underbrace{\phi\int_0^T (Q_u^\nu)^2}_{\text{Inventory Penalty}}\right]$$

$$= \mathbb{E}\left[\underbrace{X_T^\nu + Q_T^\nu\, S_T^\nu}_{\text{Terminal Wealth}} - \underbrace{\left(\alpha Q_T^\nu\right)^2}_{\text{Terminal Penalty}} - \underbrace{\phi\int_0^T (Q_u^\nu)^2}_{\text{Inventory Penalty}}\right],$$

1 Stochastic Optimal control

2 Cartea-Jaimungal framework
- The model
- The optimization problem
- The solution
- Discussion

3 Optimal execution with nonlinear costs
- The model
- The optimization problem
- The solution
- Discussion

4 References

Value function

$$H(t,x,S,q) = \sup_{\nu \in \mathcal{A}} \mathbb{E}_{t,x,S,q} \left[X_T^\nu + Q_T^\nu S_T^\nu - (\alpha Q_T^\nu)^2 - \phi \int_t^T (Q_u^\nu)^2 \right],.$$

where $\mathbb{E}_{t,x,S,q}$ is the expectation conditional on $X_t^\nu = x$, $S_t^\nu = S$, and $Q_t^\nu = q$.

Writing the HJB

$$dQ_t = \nu_t \, dt, \qquad dS_t^\nu = \sigma \, dW_t + k \, \nu_t \, dt, \quad dX_t = -\nu_t \, (S_t + \eta \, \nu_t) \, dt$$

The HJB is

$$0 = \partial_t H + \sup_\nu \left\{ \nu \, \partial_q H + k \, \nu \, \partial_S H + \frac{1}{2} \sigma^2 \, \partial_{SS} - \nu \, (S + \eta \, \nu) \, \partial_x H - \phi \, q^2 \right\}$$

with terminal condition

$$H(T, x, S, q) = x + S \, q - \alpha \, q^2, \quad \forall \, (x, S, q) \in \mathbb{R}^3.$$

Rearrange to obtain

$$0 = \left(\partial_t + \frac{1}{2}\sigma^2\,\partial_{SS}\right) H - \phi\, q^2 + \sup_{\nu} \left\{ -\nu\,(S + \eta\,\nu)\,\partial_x H + k\,\nu\,\partial_S H + \nu\,\partial_q H \right\},$$

with terminal condition

$$H(T, x, S, q) = x + S\,q - \alpha\, q^2, \quad \forall\,(x, S, q) \in \mathbb{R}^3.$$

The first-order condition allows us to obtain the optimal trading speed as a function of the value function (feedback form):

$$\boxed{\nu^\star = \frac{1}{2\,\eta}\;\frac{-S\,\partial_x H + k\,\partial_S H + \partial_q H}{\partial_x H}.}$$

Substitute the optimal feedback control into the HJB to find

$$0 = \left(\partial_t + \frac{1}{2}\sigma^2\,\partial_{SS}\right)H - \phi\,q^2 + \frac{1}{4\,\eta}\;\frac{\left(-S\,\partial_x H + k\,\partial_S H + \partial_q H\right)^2}{\partial_x H},$$

with terminal condition

$$H(T, x, S, q) = x + S\,q - \alpha\,q^2.$$

We must propose an **ansatz**. The terminal condition suggests

$$\boxed{H(t,x,S,q) = \underbrace{x}_{\text{accumulated cash}} + \underbrace{q\,S}_{\text{value of shares}} + \underbrace{h(t,S,q)}_{\text{added value}},}$$

where $h(t,S,q)$ is still to be determined.

Use

$$\partial_x H = 1, \quad \partial_S H = q + \partial_S h, \quad \partial_q H = S + \partial_q h, \quad \partial_{SS} H = \partial_{SS} h$$

to simplify the HJB to

$$\begin{aligned} 0 &= \left(\partial_t + \frac{1}{2}\sigma^2\,\partial_{SS}\right) H - \phi\,q^2 + \frac{1}{4\,\eta}\,\frac{\left(-S\,\partial_x H + k\,\partial_S H + \partial_q H\right)^2}{\partial_x H}, \\ \Longrightarrow 0 &= \left(\partial_t + \frac{1}{2}\sigma^2\,\partial_{SS}\right) h - \phi\,q^2 + \frac{1}{4\,\eta}\,\frac{\left(k\,\left(q + \partial_S h\right) + \partial_q h\right)^2}{\partial_x H} \end{aligned}$$

subject to the terminal condition

$$h(T,S,q) = -\alpha\,q^2.$$

The PDE contains no explicit dependence on the state variable S and the terminal condition is independent of S
$\implies$ h does not depend on S.

We write $h(t, S, q) = h(t, q)$ and $\partial_S h(t, S, q) = 0$

The PDE for h becomes

$$\boxed{0 = \partial_t h - \phi\, q^2 + \frac{1}{4\,\eta}\,\left(k\,q + \partial_q h\right)^2\,,}$$

and the optimal feedback control simplifies to

$$\nu^\star = \frac{1}{2\,\eta}\,\frac{-S\,\partial_x H + k\,\partial_S H + \partial_q H}{\partial_x H} \implies \boxed{\nu^\star = \frac{1}{2\,\eta}\,\left(k\,q + \partial_q h\right).}$$

The solution admits a separation of variables and takes the form of second-degree polynomial in q.

$$\boxed{h(t,q) = h_0(t) + h_1(t)\,q + h_2(t)\,q^2\,.}$$

Use

$$\begin{cases} \partial_t h = & \partial_t h_0(t) + h_1(t)\,q + h_2(t)\,q^2 \\ \partial_q h = & h_1(t) + 2\,h_2(t)\,q \end{cases}$$

to write

$$\begin{aligned} & 0 = \partial_t h - \phi\,q^2 + \frac{1}{4\,\eta}\,(k\,q + \partial_q h)^2 \\ \Longrightarrow\ & 0 = \left(\partial_t h_0(t) + h_1(t)\,q + h_2(t)\,q^2\right) - \phi\,q^2 + \frac{1}{4\,\eta}\,(k\,q + h_1(t) + 2\,h_2(t)\,q)^2 \\ \Longrightarrow\ & 0 = \left\{\partial_t h_2\,(t) - \phi + \frac{1}{4\,\eta}\,(k\ + 2\,h_2\,(t))^2\right\} q^2 \\ & \quad + \left\{\partial_t h_1\,(t) + \frac{1}{2\,\eta} h_1\,(t)\,(k\ + 2\,h_2\,(t))\right\} q + \left\{\partial_t h_0 - \frac{1}{4\,\eta}\,h_1\,(t)^2\right\} \end{aligned}$$

If the equality above is verified for all q, then

$$\boxed{\begin{cases} 0 = & \partial_t h_2(t) - \phi + \frac{1}{4\,\eta}\,\left(k\,+2\,h_2(t)\right)^2 \\ 0 = & \partial_t h_1(t) + \frac{1}{2\,\eta} h_1(t)\left(k\,+2\,h_2(t)\right) \\ 0 = & \partial_t h_0 - \frac{1}{4\,\eta}\,h_1(t)^2\,, \end{cases}}$$

subject to the terminal conditions $h_2(T) = -\alpha,\ h_1(T) = 0,$ and $h_0(T) = 0$.

The solution to the ODE in h_1 is $h_1(t) = 0$.

The ODE in h_2 is a **Riccati** ODE. It can be solved explicitly:

$$\boxed{\mathfrak{h}_2(t) = \sqrt{\eta\phi}\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}}\,,}$$

where

$$\gamma = \sqrt{\frac{\phi}{\eta}} \quad \text{and} \quad \zeta = \frac{\alpha - \frac{1}{2}k + \sqrt{\eta\phi}}{\alpha - \frac{1}{2}k - \sqrt{\eta\phi}}.$$

We have fully determined the solution to the HJB.

The optimal trading strategy $\nu^\star$ can be obtained from the feedback form:

$$\boxed{\nu_t^* = -\gamma \frac{\zeta\ e^{\gamma(T-t)} + e^{-\gamma(T-t)}}{\zeta\ e^{\gamma(T-t)} - e^{-\gamma(T-t)}} Q_t^{\nu^*}.}$$

And

$$\boxed{Q_t^{\nu^*} = \frac{\zeta\ e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta\ e^{\gamma T} - e^{-\gamma T}} Q_0.}$$

1 Stochastic Optimal control

2 Cartea-Jaimungal framework
- The model
- The optimization problem
- The solution
- Discussion

3 Optimal execution with nonlinear costs
- The model
- The optimization problem
- The solution
- Discussion

4 References

Deterministic strategy. the optimal strategy is deterministic even though we did not restrict the space of admissible strategies.

Optimal trading curves and model parameters can be estimated prior to the trading window.

Fuel constraint. In the limit when the quadratic liquidation penalty goes to infinity, i.e., $\alpha \to \infty$,

$$\zeta \underset{\alpha\to\infty}{\longrightarrow} 1 \implies Q_t^{\nu^*} \xrightarrow[\alpha\to+\infty]{} \frac{\sinh\left(\gamma(T-t)\right)}{\sinh\left(\gamma T\right)} Q_0,$$

which <u>does not depend on the permanent impact *k*.</u>

Urgency of trading. As the running penalty increases, the optimal strategy aims to sell more assets sooner.

ϕ is the agent's urgency and penalises holding inventory.

When ϕ approaches zero (no aversion to risk), the optimal curve resembles a straight line

$$Q_t^{\nu^\star} \underset{\phi\to 0}{\longrightarrow} \frac{t}{T+k/\alpha}.$$

Fuel constraint. In the limit when the quadratic liquidation penalty goes to infinity, i.e., $\alpha \to \infty$,

$$\zeta = \frac{\alpha - \frac{1}{2}k + \sqrt{\eta\phi}}{\alpha - \frac{1}{2}k - \sqrt{\eta\phi}} \underset{\alpha\to\infty}{\longrightarrow} 1 \implies Q_t^{\nu^*} \xrightarrow[\alpha\to+\infty]{} \frac{\sinh(\gamma(T-t))}{\sinh(\gamma T)} Q_0,$$

which **does not depend on the permanent impact k.**

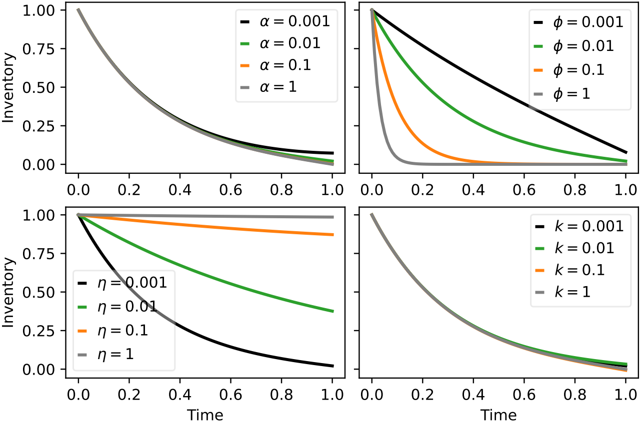
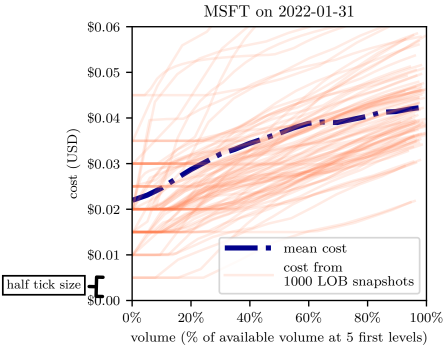


Figure: Optimal trading curves for multiple values of the parameters: terminal penalty α, running inventory penalty ϕ, execution costs η, and permanent impact k. The default parameter values are $\alpha = 0.01$, $\phi = 0.01$, $k = 0.001$, and $\eta = 0.001$. The other parameters are $T = 1$ and $Q_0 = 1$.

Recent literature argues that a power law function or other nonlinear functions fit the observed data better.



We assume that the execution price per share received by the agent is

$$\tilde{S}_t^\nu = S_t^\nu - f(\nu_t).$$

- An agent holds an **initial position** Q_0 at time $t = 0$ that they wish to **unwind** over a time window $[0, T]$.

- **Inventory**:

$$\boxed{dQ_t^\nu = \nu_t \, dt}$$

where ν is restricted to be negative to enforce liquidation.

- **Price**:

$$\boxed{dS_t^\nu = \sigma \, dW_t + k \, \nu_t \, dt, \quad S_0 \in \mathbb{R}_+^\star \text{ is known}}$$

- **Cash**:

$$\boxed{dX_t^\nu = -\nu_t \, \tilde{S}_t^\nu \, dt = -\nu_t \, (S_t^\nu + f(\nu_t)) \, dt, \quad X_0 \in \mathbb{R} \text{ is known.}}$$

- **Admissible** strategies: **no unwind** constraint.

Performance criterion

- The agent's objective is to complete the liquidation, but we allow to fall short of this target (no fuel constraint) so $Q_T \neq 0$.
- We add a **terminal penalty** term to include any costs incurred when trading the terminal inventory Q_T.
- We include a **running inventory penalty**.
- Performance criterion:

$$\mathbb{E}_{t,x,S,q}\left[X_T^\nu + Q_T^\nu S_T^\nu - (\alpha Q_T^\nu)^2 - \phi \int_0^T (Q_u^\nu)^2\right],$$

Value function

$$H^{\nu}(t,x,S,q) = \mathbb{E}_{t,x,S,q}\left[X_T^{\nu} + Q_T^{\nu}\,S_T^{\nu} - (\alpha Q_T^{\nu})^2 - \phi\int_t^T (Q_u^{\nu})^2\right].$$

where $\mathbb{E}_{t,x,S,q}$ is the expectation conditional on $X_t^{\nu} = x$, $S_t^{\nu} = S$, and $Q_t^{\nu} = q$.

Writing the HJB

$$dQ_t = \nu_t\,dt, \qquad dS_t^\nu = \sigma\,dW_t + k\,\nu_t\,dt, \quad dX_t = -\nu_t\,(S_t + f(\nu_t))\,dt$$

The HJB is

$$0 = \partial_t H + \sup_\nu \left\{ \nu\,\partial_q H + k\,\nu\,\partial_S H + \frac{1}{2}\sigma^2\,\partial_{SS} - \nu\,\left(S + f(\nu)\right)\partial_x H - \phi\,q^2 \right\}$$

with terminal condition

$$H(T, x, S, q) = x + S\,q - \alpha\,q^2, \quad \forall\,(x, S, q) \in \mathbb{R}^3.$$

Rearrange to obtain

$$0 = \left(\partial_t + \frac{1}{2}\sigma^2 \partial_{SS}\right) H - \phi q^2 + \sup_{\nu} \{(-\nu\,(S + f(\nu))\,\partial_x - k\,\nu\,\partial_S + \nu \partial_q)\, H\}$$

with terminal condition

$$H(T, x, S, q) = x + S\,q - \alpha\, q^2, \quad \forall\,(x, S, q) \in \mathbb{R}^3.$$

Use the usual ansatz

$$\boxed{H(t,x,S,q) = x + q\,S + h(t,q)\,,}$$

So

$$\begin{aligned} 0 &= \left(\partial_t + \frac{1}{2}\sigma^2\partial_{SS}\right) H - \phi q^2 + \sup_\nu \left\{\left(-\nu\,(S + f(\nu))\,\partial_x - k\,\nu\,\partial_S + \nu\,\partial_q\right) H\right\} \\ \implies 0 &= \partial_t h - \phi q^2 + \sup_\nu \left\{-\nu\, f(\nu) - (k\,q - \partial_q h)\,\nu\right\}, \end{aligned}$$

subject to the terminal condition $h(T,q) = -\alpha\, q^2$.

We need **conditions** on f for the problem to be well-posed:

$F : \nu \mapsto \nu\, f(\nu)$ must be **convex**. It holds for the linear execution costs.

$v\, f(\nu)$ is convex so the supremum term can be solved and we write

$$\sup_{\nu} \left\{ -\nu\, f(\nu) - (k\, q - \partial_q h)\, \nu \right\} = \tilde{F}\left(-\, (k\, q - \partial_q h) \right).$$

The case of a power law

Assume $Q_0 < 0$ and that we *restrict the trading speed to be positive to enforce acquisition only*.

Assume that the execution costs function is, for $0 < a < 1$,

$$f : \nu \mapsto \eta\, \nu^a$$

So $F : x \mapsto \eta\,\nu^{1+a}$, and $\tilde{F}(p) = \sup_x \left\{x\,p - \eta\,x^{1+a}\right\}$. The supremum can be computed with first order condition

$$p - \eta\,(1+a)\,(x^*)^a = 0 \quad \Rightarrow \quad x^* = \left(\frac{p}{(1+a)\,\eta}\right)^{\frac{1}{a}},$$

so the Legendre-Fenchel transform of F is

$$\tilde{F}(p) = \xi p^{1+\frac{1}{a}}, \quad \xi = \frac{a\,\eta}{((1+a)\,\eta)^{1+\frac{1}{a}}}.$$

The HJB becomes

$$\boxed{\partial_t h - \phi q^2 + F^*\left(-(k\,q - \partial_q h)\right) = 0, \quad \text{and} \quad h(T,q) = -\alpha q^2.} \qquad (1)$$

Effect of non-linearity

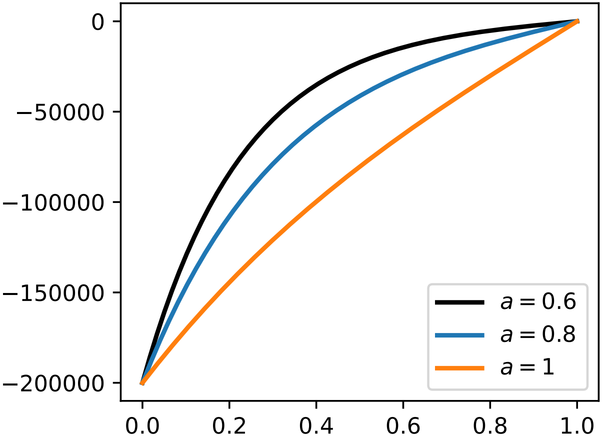


Figure: The effect of non-linear impact on the optimal strate’gy in the case of a power law temporary impact function with power parameter *a*.

- The PDE can rarely be solved analytically and one usually uses numerical approximation techniques.
- When there are too many variables, curse-of-dimensionality.
- Nonlinear execution costs marginally improve the predictive accuracy of market impact models (typically less than 5% R-squared).
- The cost of increased complexity outweighs any gains from better describing market impact.