PRINCETON UNIVERSITY

ORF 474: High Frequency Trading
Spring 2020
Robert Almgren

Lecture 12b

April 29, 2020

Arrival price algorithms

- Arrival price benchmark:
  market price at order start time
- What kind of strategy should you execute
  if this is your benchmark?
- Answer: execute fast at beginning, then slowly

EXECUTION ALGORITHMS

OCTOBER 2014

Nasdaq

WHY EXECUTION ALGORITHMS VIA NASDAQ?

YOUR NEEDS

- Achieve the best possible execution price based on your chosen strategy.
- Minimize market impact and prevent information leakage for your large orders.
- Easy access to liquidity across multiple trading venues via Nasdaq Nordic current smart order routing.

DESCRIPTION OF ALGOS - IMSH

(implementation shortfall)

IMSH

"Be as close to the arrival price as possible and adjust aggression in response to market conditions"

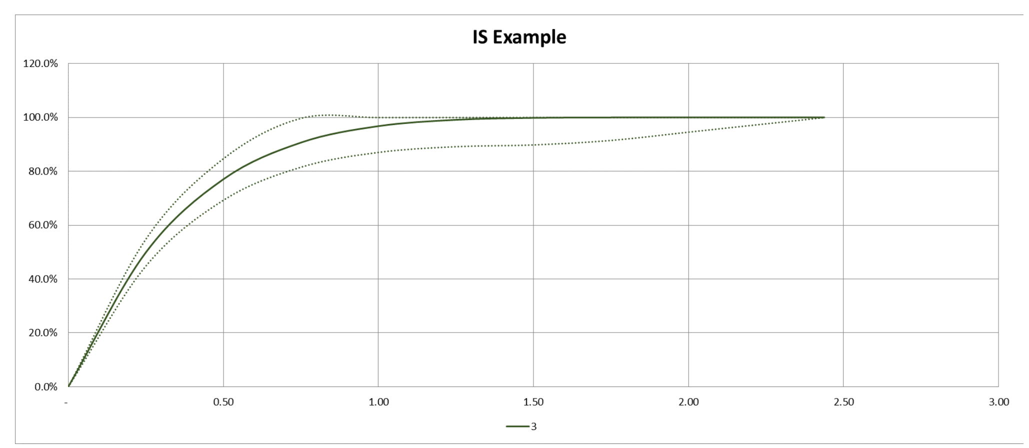
Minimizes the risk-adjusted trading costs relative to the arrival price of the stock and dynamically adjusts aggression as a function of real-time market conditions relative to a chosen benchmark.

This front weighted trajectory tends to result in an average execution price close to the arrival price while minimizing implementation shortfall.

Key parameters:
Min & Max % Vol, I Would Price, Dynamic Benchmark

IMSH

"Try to beat the arrival price and adjust aggression in response to market conditions"



% of Order Finished

Expert opinion

Implementation shortfall

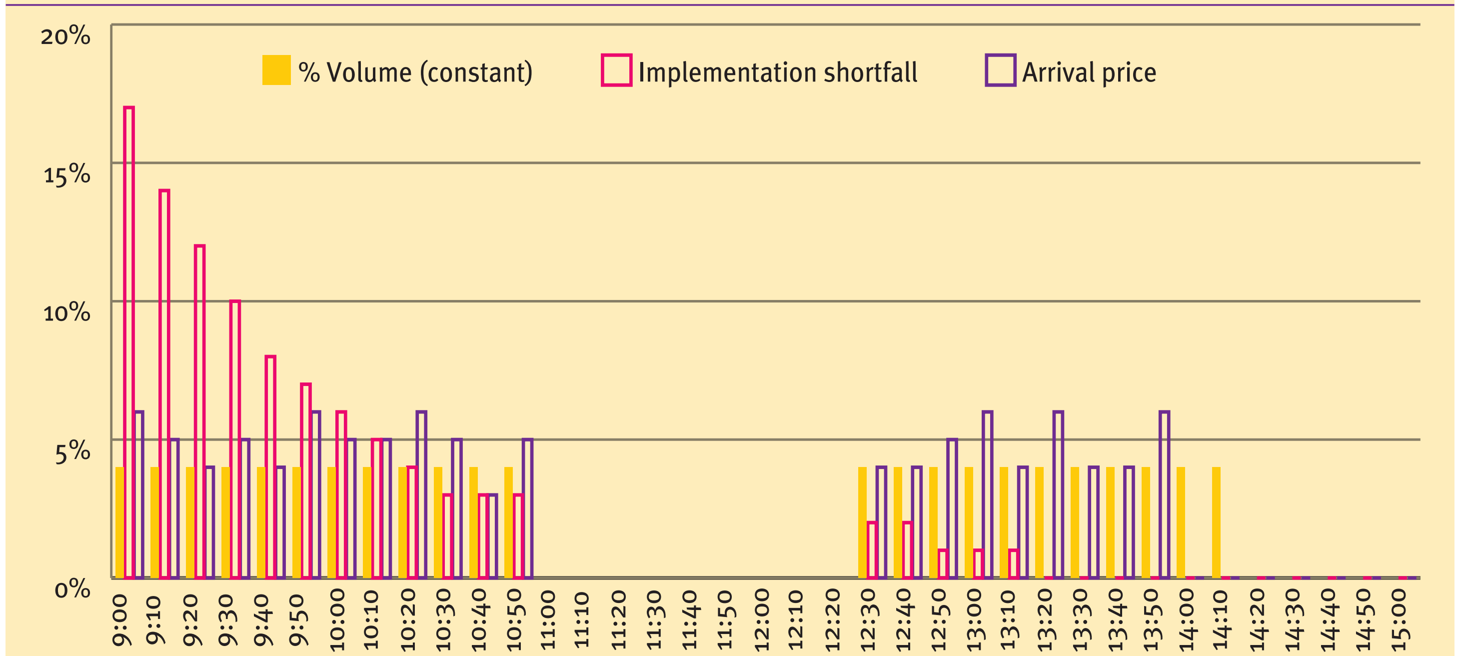
Strategies to fit the benchmark

Mark Maloney

Mark Maloney, director of autobahn Equity® sales at Deutsche Securities in Tokyo, looks at strategies traders can deploy to optimise IS-benchmarked orders.

32 THE TRADE ASIA ISSUE 2 Q1–Q2 2008

FIGURE 1: PARTICIPATION RATE – PARTICIPATION VOLUME, ARRIVAL PRICE AND IMPLEMENTATION SHORTFALL



Trading trajectory – managing your risk

While practitioners often use some variation of a Percent of Volume algorithm for IS-benchmarked orders, research[1] suggests an IS-benchmarked order should be traded differently depending on one's risk profile. In general, the less risk one wants to assume, the more the execution schedule should be front-loaded rather than constant (Figure 1.).

Deutsche empirical analysis reinforces the conclusion that over the long term, a true optimised IS strategy outperforms a percent-participation or VWAP. This is a function of optimising the benefits of trading rapidly (lower exposure to market risk) against the advantages of trading slowly (lower market impact costs.)

Risk

Sebastian Day

03 Jul 2017

Why Robert Almgren no longer trades using Almgren-Chriss

Co-author of popular model says market nuances are critical in optimal execution

He published his influential paper *Optimal execution of portfolio transactions* in the late 90s, co-authoring with Neil Chriss, now managing principal for Hutchin Hill Capital. That paper presented a framework for finding optimal execution strategies by balancing between price impact and volatility.

Its crucial point is that eliminating volatility risk and capturing alpha are reasons to trade quickly, while minimising market impact is a reason to trade slowly. An optimal trading strategy balances the two. The approach was soon adopted across the industry.

The solution was to realise that a standard mathematical approach – calculus of variations, which is used in problems such as finding the shortest path between two points on a curved surface – could provide the answer.

Bank of America

2007

Adaptive Implementation™

The next frontier in algorithmic trading

Robert Almgren, managing director and Head of Quantitative Strategies in Electronic Trading Services at Banc of America Securities LLC, argues that Adaptive Implementation is the next frontier in algorithmic trading, and that doing this well requires a surprising depth of quantitative analysis.

We like to divide the history of algorithmic trading into three generations:

1 ***VWAP.*** The original motivation was the observation that computers could automate routine trading tasks. Rather than have a person sit there and push the button every few minutes through the day, get a computer to do it. This reasoning is still useful today, since it embodies the instruction to "spread it out over time."

2 ***Arrival Price.*** The next development was the realization that once a computer is involved, you may as well program it to do something optimal from the point of view of the investor who is driving the trade. For example, if you are going to do a VWAP execution, should the duration be one hour or one week? How do you choose between executing fast and slow? Answering this question requires measurement of market impact and volatility, and requires the user to specify some balance between the cost of fast trading and the risk of slow trading. Then, within this model, you calculate the optimal strategy; the result is arrival price or implementation shortfall algorithms, the bread and butter of algorithmic trading today. They embody the instructions to "get a big part of the order done quickly to reduce exposure, then reduce cost by trading the rest slowly."

3 ***Adaptive Implementation.*** In their standard form, arrival price strategies precompute an execution trajectory; they do not deviate from this trajectory to accommodate price motions, varying liquidity, or other market conditions and events. Of course in practice, algorithmic providers have highly optimized trade engines that adapt the second-by-second order submission to market events. In addition, a variety of "scaling" strategies vary the overall execution urgency depending on price motion, possibly in relation to index values and other indicators. This note argues that the next frontier in algorithmic trading is building adaptive strategies to the same high standards that we are used to with arrival price.

The first question is what an algorithm should adapt to. The most obvious answer is that it should do "something different" when the price moves away from its initial value. Traditional arrival price ignores this price motion, for good reason: past motions of the price do not necessarily tell you anything about its future motions. If you believe in momentum, then you will want to slow down your execution when the price moves in your favor (say, price decreases on a buy order) to give it time to move further. If you believe in reversion, then you will want to speed up when the price moves in your favor, before it bounces back against you. If you believe in a pure random walk, then you will want to leave your execution strategy unchanged as the price moves.

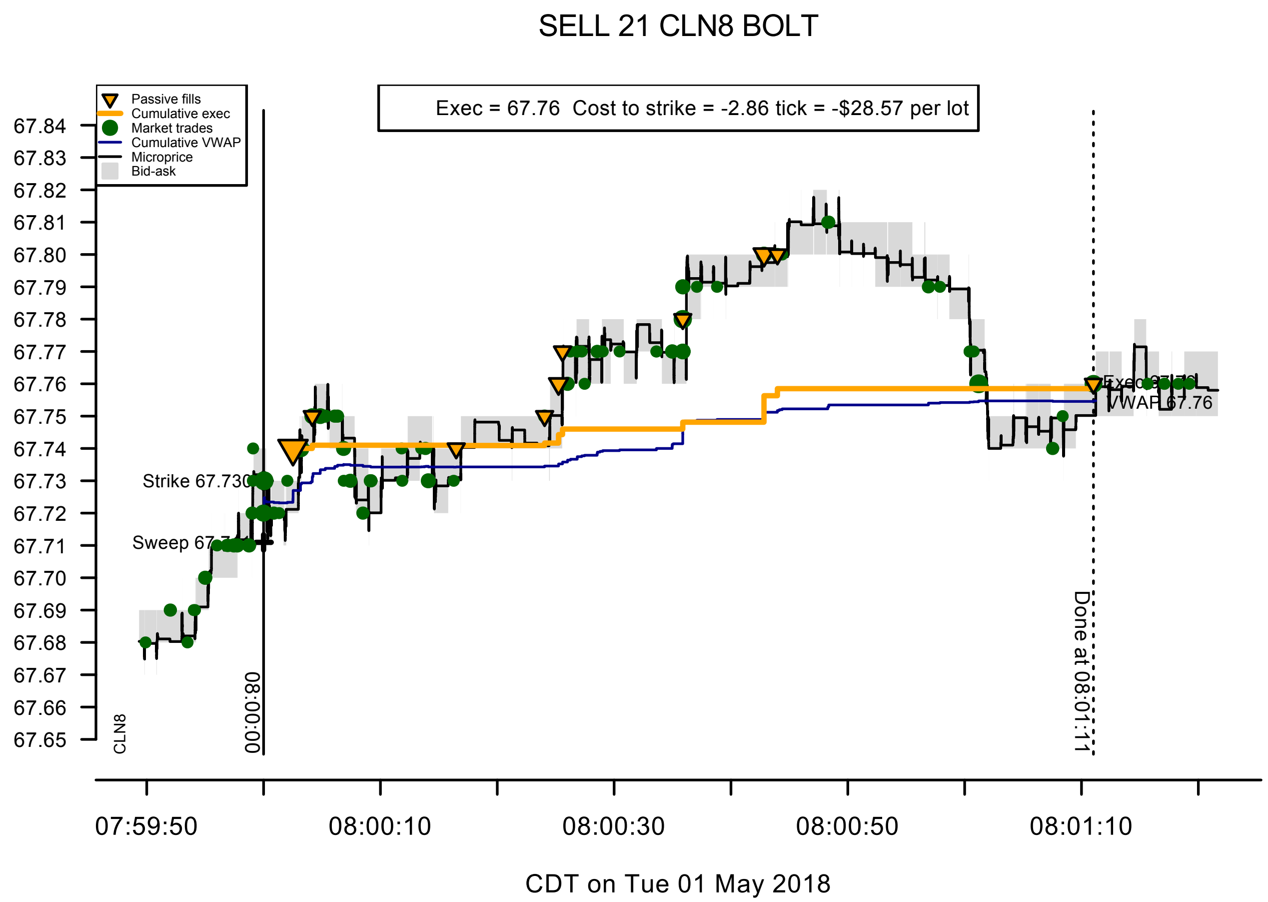
This indifference is not a natural response for most traders, who generally prefer to "get the trade done" when the price has come in their favor, and the existence of this tendency is supported by a large body of literature in behavioral finance. It is difficult to justify building these preferences into an automated trading system, which is expected to achieve quantitatively good performance across a long history of executions. There are various subtle interpretations of the risk-reward trade-off that do lead to strategies having behavior consistent with intuitions, but it is not yet clear what is the right interpretation.

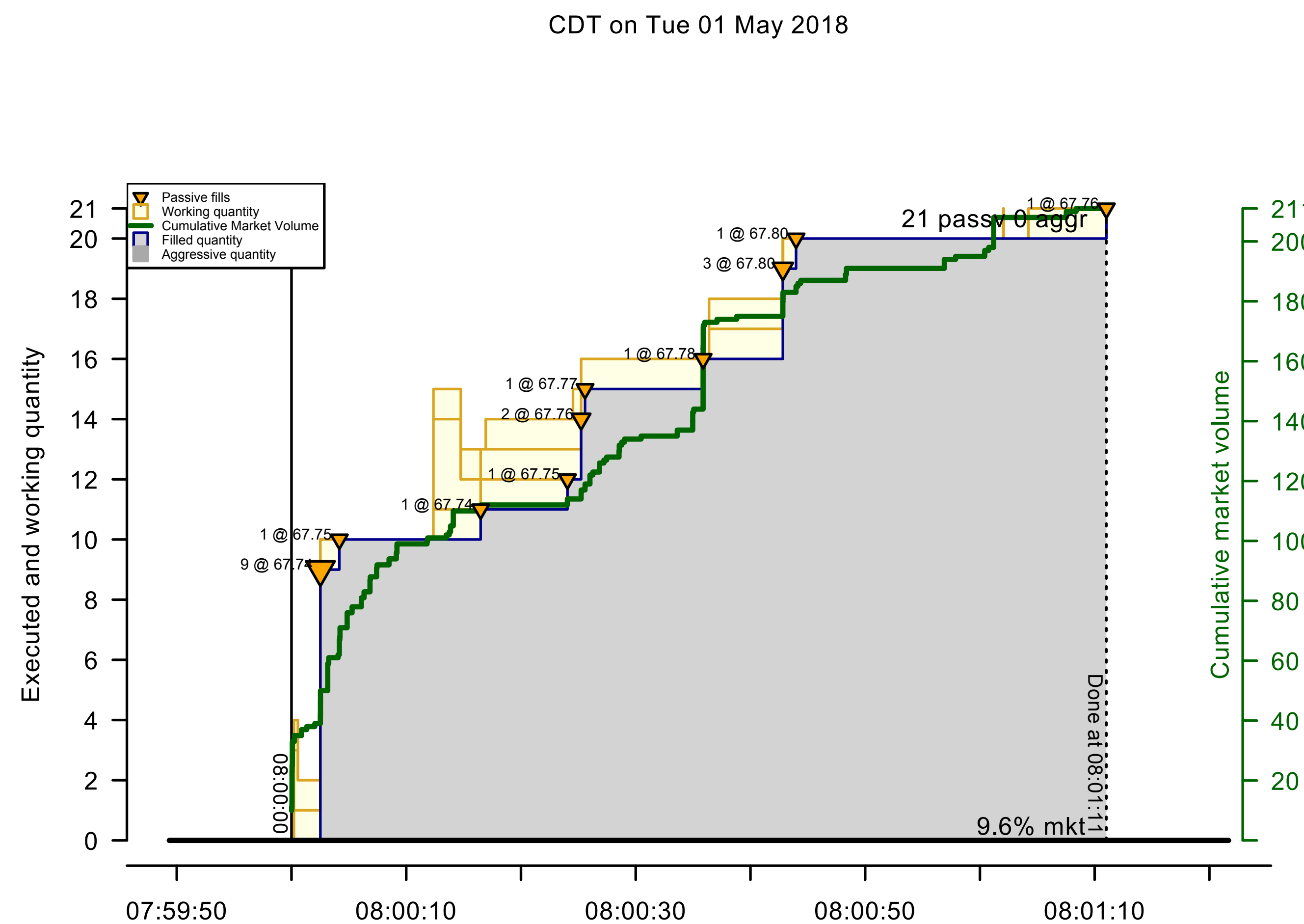
The second, and most important, answer to the question of what an algorithm should adapt to is varying market conditions: whether the market is ready to take or provide the shares you want to trade. The hard thing is getting a computer algorithm to recognize the difference using only the stream of market data. We emphasize that this is not the second-by-second question of when to place a limit order or when to cross the spread, which we assume is already being done excellently. It is the ebb and flow of trading activity and liquidity on a minute-by-minute and hour-by-hour level, as other market participants come and go. Rapidly identifying and optimally adapting to these variations in market activity have the potential to give immense improvements in overall performance.

Identifying instantaneous liquidity depends on detailed analysis of market microstructure, and on having the technology infrastructure to process market data in its full complexity. Adapting the execution to these varying signals is a complex dynamic optimization problem for which a rich set of mathematical tools can be applied. The result of this combination is a highly adaptive optimal strategy that represents the next generation of arrival price. We have these strategies running and the results are excellent.

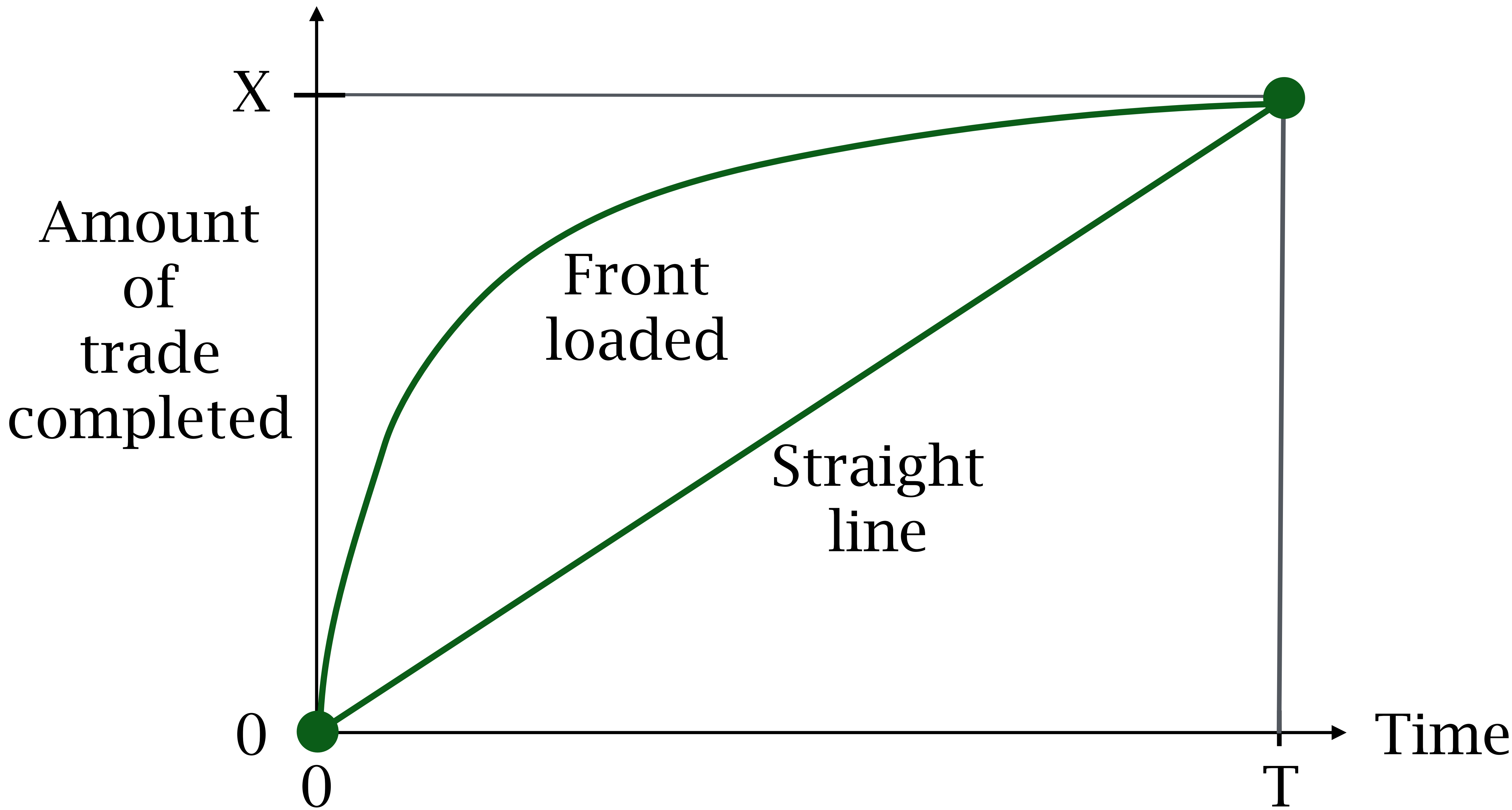
Electronic Trading Services
EASE™ Trading Desk
212.583.8730 • 888.279.1196
ease@bofasecurities.com

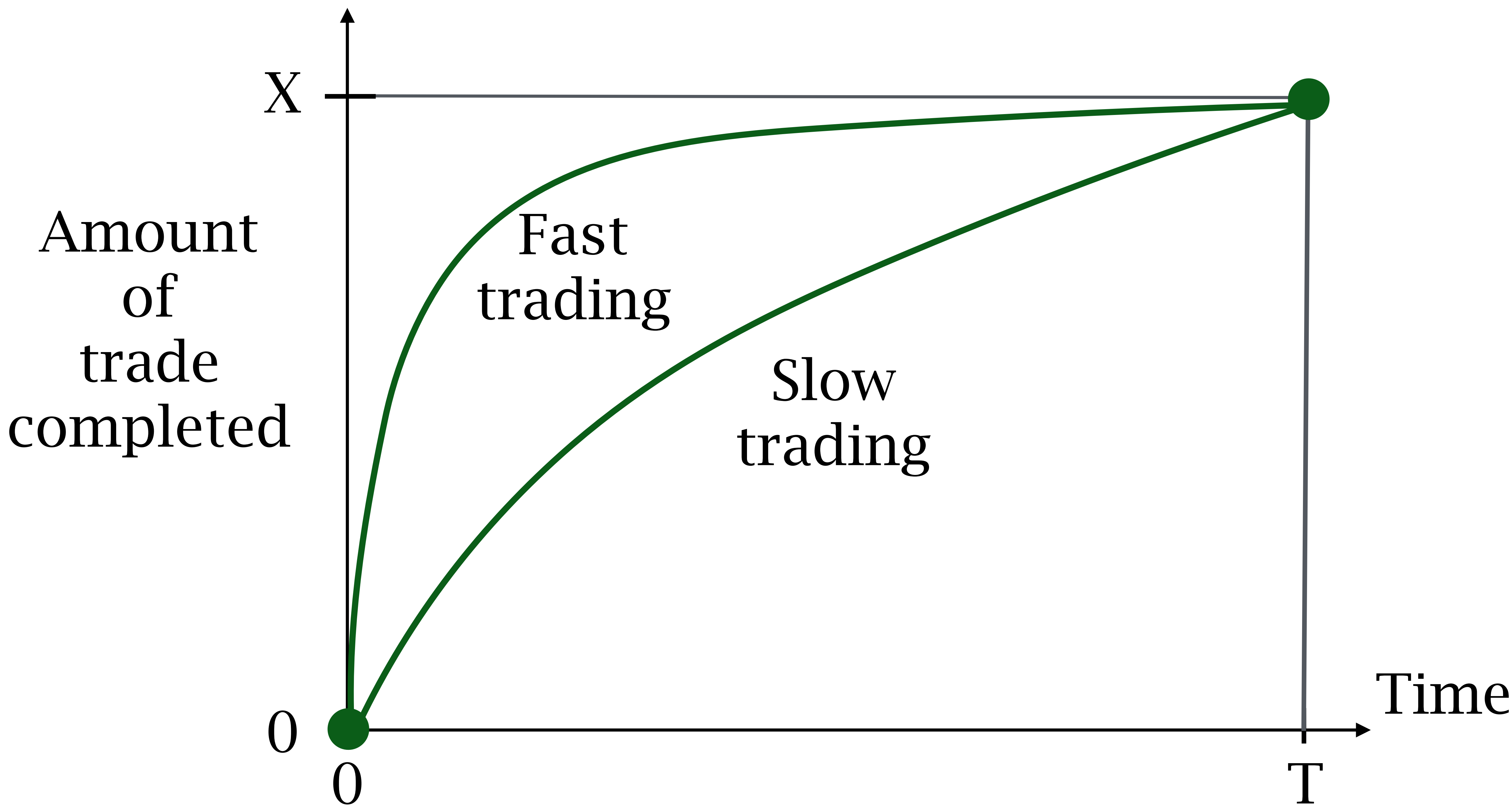
Bank of America


SELL 21 CLN8 BOLT
Exec = 67.76 Cost to strike = -2.86 tick = -$28.57 per lot
Passive fills
Cumulative exec
Market trades
Cumulative VWAP
Microprice
Bid-ask
Strike 67.730
Done at 08:01:11
CDT on Tue 01 May 2018


CDT on Tue 01 May 2018
Passive fills
Working quantity
Cumulative Market Volume
Filled quantity
Aggressive quantity
Executed and working quantity
Cumulative market volume
9.6% mkt
Done at 08:01:11

Trade scheduling problem




X
Amount
of
trade
completed
Fast
trading
Slow
trading
Time
0
0
T

Optimal Execution of Portfolio Transactions*

Robert Almgren† and Neil Chriss‡

December 2000

We consider the execution of portfolio transactions with the aim of minimizing a combination of volatility risk and transaction costs arising from permanent and temporary market impact. For a simple linear cost model, we explicitly construct the *efficient frontier* in the space of time-dependent liquidation strategies, which have minimum expected cost for a given level of uncertainty. We may then select optimal strategies either by minimizing a quadratic utility function, or by minimizing Value at Risk. The latter choice leads to the concept of Liquidity-adjusted VAR, or L-VaR, that explicitly considers the best tradeoff between volatility risk and liquidation costs.

General strategies (not only straight lines)

Purchase X shares, complete by time T

Strategy is $\theta(t)$ for $0 \le t \le T$,
with $\int_0^T \theta(t)\,dt \;=\; X$

$$x(t) = \int_0^t \theta(s)\,ds$$

Assume θ(t) fixed (for 0≤t≤T) at t=0

Public market price

$$P(t) \;=\; P_0 + \sigma W(t) + \nu\, x(t)$$

Private trade price

$$\begin{aligned} \tilde{P}(t) \;&=\; P_t + H\big(\theta(t)\big) \\ &=\; P_0 + \sigma W(t) + \nu\, x(t) + H\big(\theta(t)\big) \end{aligned}$$

Cost to acquire X shares

$$Z \;=\; \int_0^T \tilde{P}(t)\,\theta(t)\,dt$$

Cost to acquire X shares

Temporary impact varies in time because trade rate varies

$$Z = \int_0^T \tilde{P}(t)\,\theta(t)\,dt$$

$$= \int_0^T \Big[\underbrace{P_0}_{\text{Initial price}} + \underbrace{\sigma W(t)}_{\text{Volatility}} + \underbrace{\nu x(t)}_{\text{Permanent impact}} + H\big(\theta(t)\big)\Big]\,\theta(t)\,dt$$

$$= \underbrace{P_0 X + \nu \int_0^T x(t)\,dx(t) + \int_0^T H\big(\theta(t)\big)\,\theta(t)\,dt}_{\text{Nonrandom}} + \underbrace{\sigma \int_0^T W(t)\,dx(t)}_{\text{Normal, mean zero}}$$

Nonrandom (if strategy is non-adaptive)

Normal, mean zero

$$\int_0^T x(t)\,dx(t) = \frac{1}{2} X^2$$

$$\int_0^T W(t)\,dx(t) = \int_0^T \big(X - x(t)\big)\,dW(t) - \underbrace{\Big[\big(X - x(t)\big)W(t)\Big]_{t=0}^{T}}_{=\,0}$$

$$\text{Var} = \int_0^T \big(X - x(t)\big)^2\,dt$$

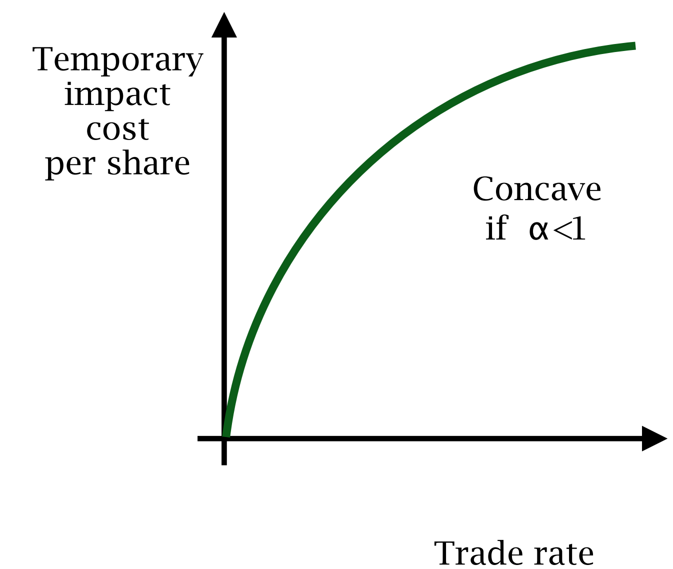
Cost of trading

$$C \;=\; Z \;-\; X\,P_0 \qquad \text{Normal random variable}$$

$$\mathbb{E}(C) \;=\; \frac{1}{2}\nu X^2 \;+\; \int_0^T H(\theta(t))\,\theta(t)\,dt$$

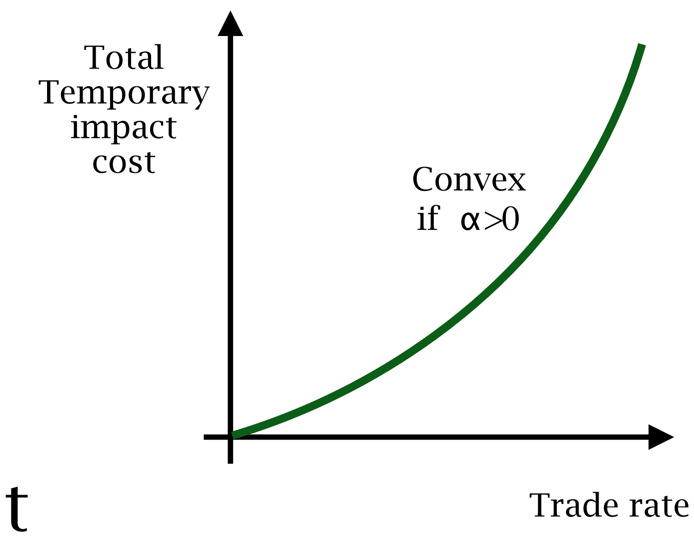
$$\mathrm{Var}(C) \;=\; \sigma^2 \int_0^T (X - x(t))^2\,dt$$

With $H(\theta) = \eta\,\theta^\alpha$



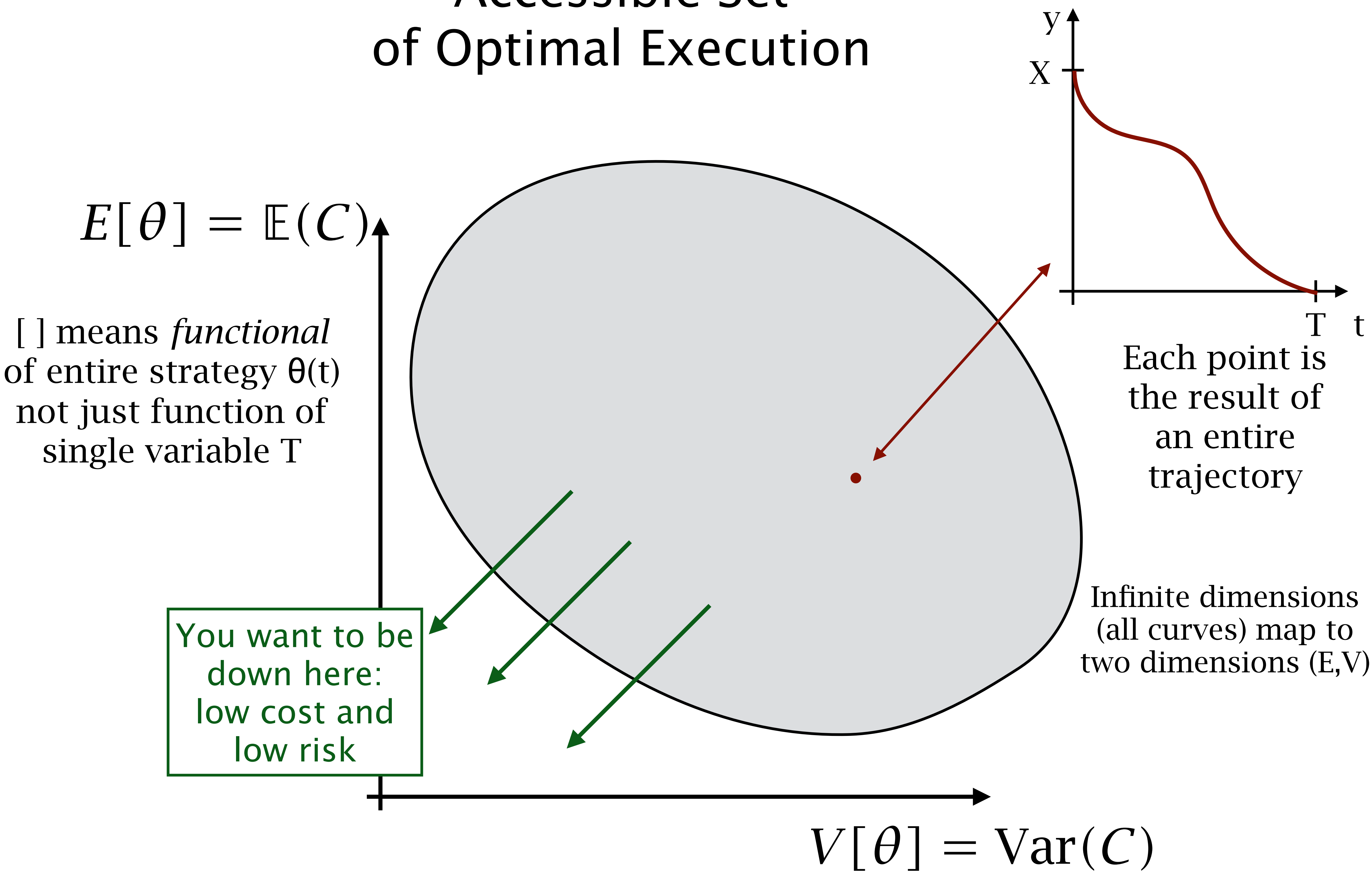
$$E[\theta] \;=\; \mathbb{E}(C) \;=\; \frac{1}{2}\nu X^2 \;+\; \eta \int_0^T \theta(t)^{1+\alpha}\,dt$$

$$V[\theta] \;=\; \mathrm{Var}(C) \;=\; \sigma^2 \int_0^T y(t)^2\,dt$$



$y(t) = X - x(t)$, deviation from target

Accessible Set of Optimal Execution



$$\min_{\{x(t)\,|\,x(0)=0,\ x(T)=X\}} \Big(E[\theta] + \lambda V[\theta] \Big)$$

- How do you find the "best" strategies?
- Boundary of the accessible set on the small-E, small-V side

$$\theta(t) = x'(t)$$

$$E + \lambda V \;=\; \eta \int_0^T \theta(t)^{1+\alpha}\, dt \;+\; \lambda\sigma^2 \int_0^T y(t)^2\, dt$$

$$y(t) = X - x(t)$$

Distance from target

Minimize this by taking θ constant TWAP on whole interval

Minimize this by taking y to 0 (x to X) as rapidly as possible

Linear temporary impact

$$\alpha = 1,\, H(\theta) = \eta\theta$$

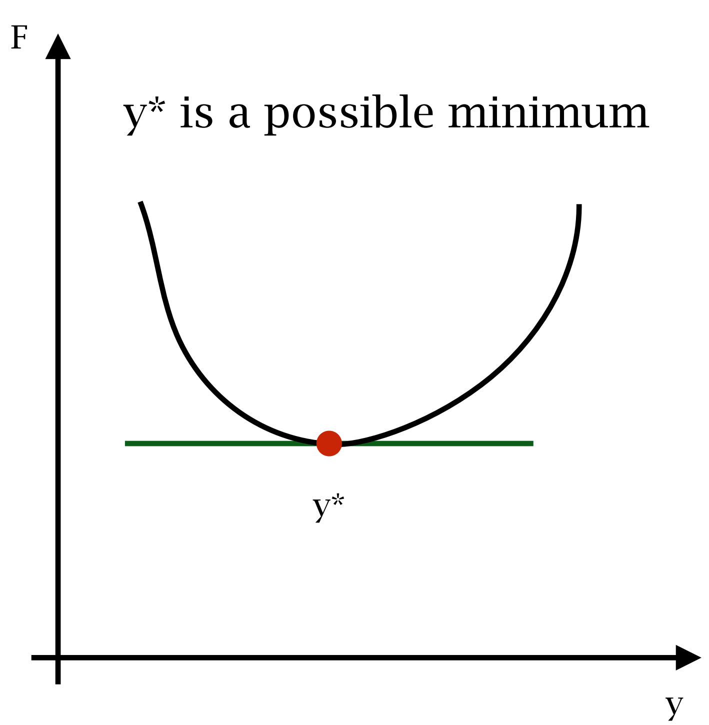
Objective function is quadratic

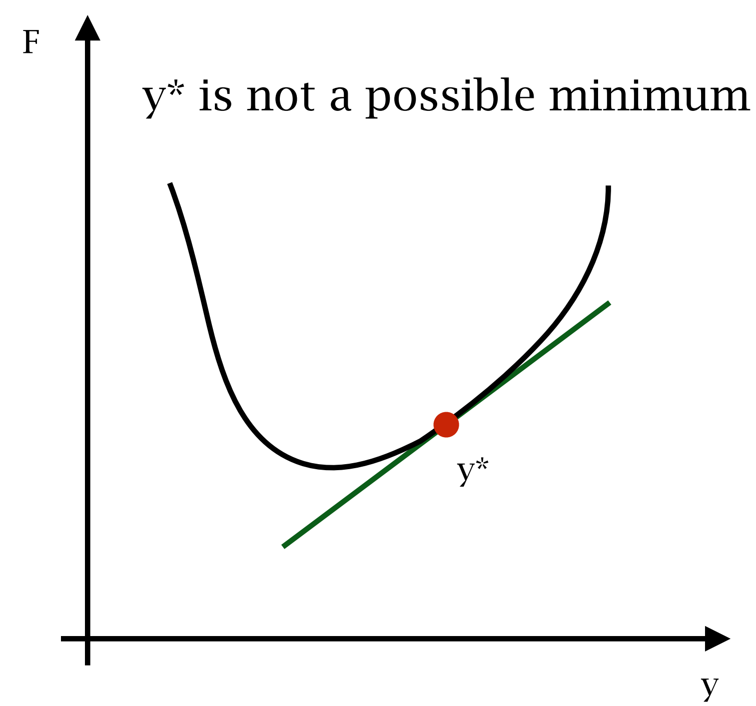
$$F[y] = \eta \int_0^T y'(t)^2\, dt + \lambda\sigma^2 \int_0^T y(t)^2\, dt$$

Minimize over $y(t)$ on $0 \le t \le T$,
with $y(0) = X$ and $y(T) = 0$.

Calculus

Calculus: A necessary condition for a function F(y) to have a minimum at y* is <u>stationarity</u>:
F(y + z) for z small has no linear dependence on z
$F(y+z) = F(y) + O(z^2)$ for <u>all</u> admissible z
($dF/dy = 0$ at $y = y^*$).
If not, then you could find a nearby y** with $F(y^{**})<F(y^*)$





Calculus of variations

$y^*(t)$ is a candidate optimal solution

Perturbation $y^*(t) \mapsto y^*(t) + z(t)$

For $y^* + z$ to be admissible,
need $z(0) = 0$ and $z(T) = 0$

$$F[y+z] = \eta \int_0^T \left(y'(t) + z'(t)\right)^2 dt + \lambda\sigma^2 \int_0^T \left(y(t) + z(t)\right)^2 dt$$

$$\sim F[y] + \underline{2\eta \int_0^T y'(t) z'(t)\, dt + 2\lambda\sigma^2 \int_0^T y(t) z(t)\, dt} + \mathcal{O}(z^2)$$

Linear part must vanish for all z(t)

$$\int_0^T y'(t) z'(t)\, dt = \underbrace{\Big[y'(t) z(t) \Big]_{t=0}^{T}}_{=0 \text{ since } z(0)=z(T)=0} - \int_0^T z(t) y''(t)\, dt$$

$$F[y+z] - F[y] = 2 \int_0^T \underbrace{\Big[\lambda\sigma^2 y(t) - \eta\, y''(t) \Big]}_{\substack{\text{Must be zero if integral} \\ \text{is zero for all } z(t)}} z(t)\, dt \;+\; \mathcal{O}(z^2)$$

Any optimal y(t) must satisfy

$$y''(t) = \kappa^2 y(t), \qquad \kappa^2 = \frac{\lambda\sigma^2}{\eta}$$

κ is *time scale of optimal execution*

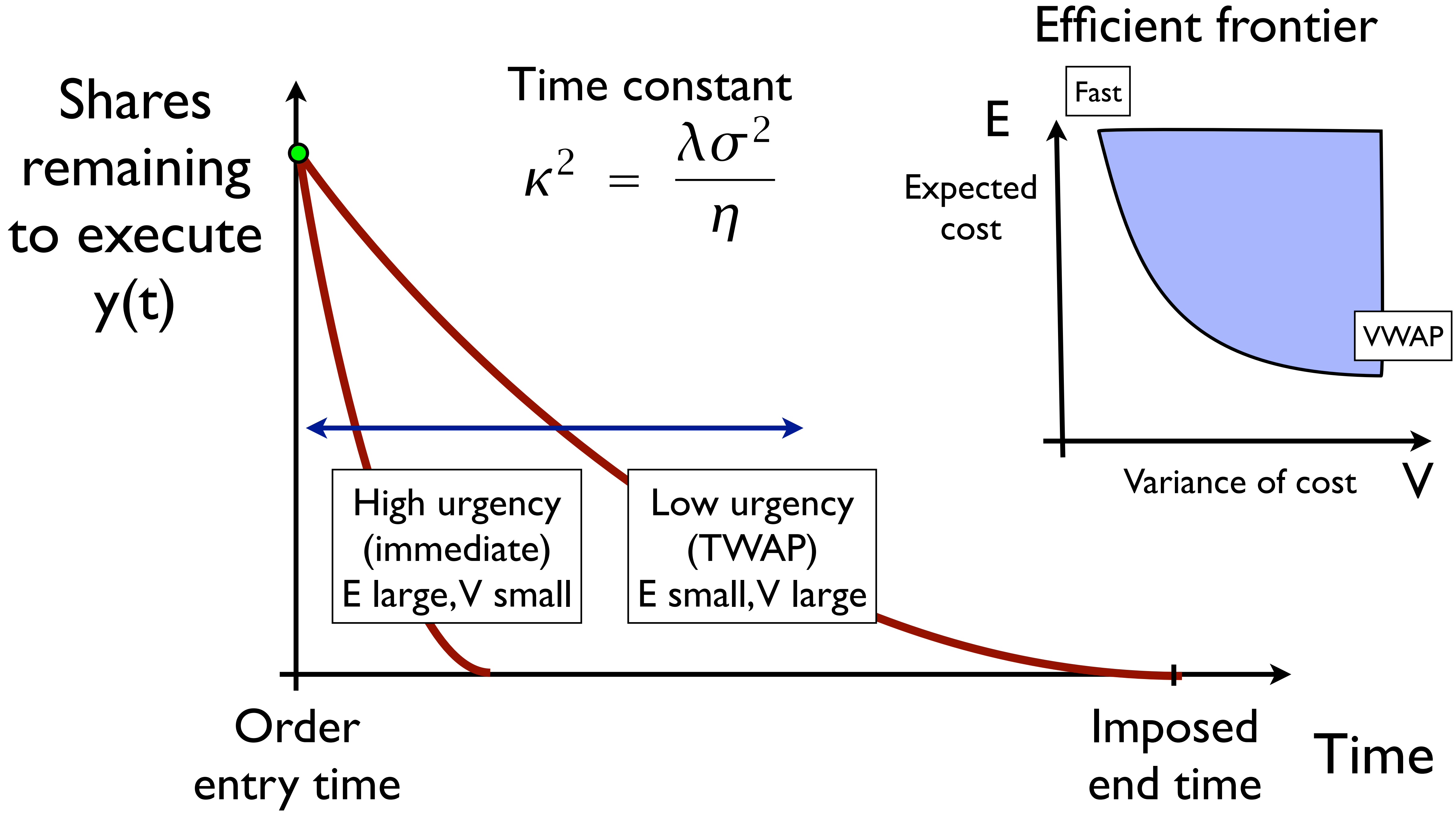
Solution to $y'' = \kappa^2 y$ is

$$y(t) = Ae^{\kappa t} + Be^{-\kappa t}$$

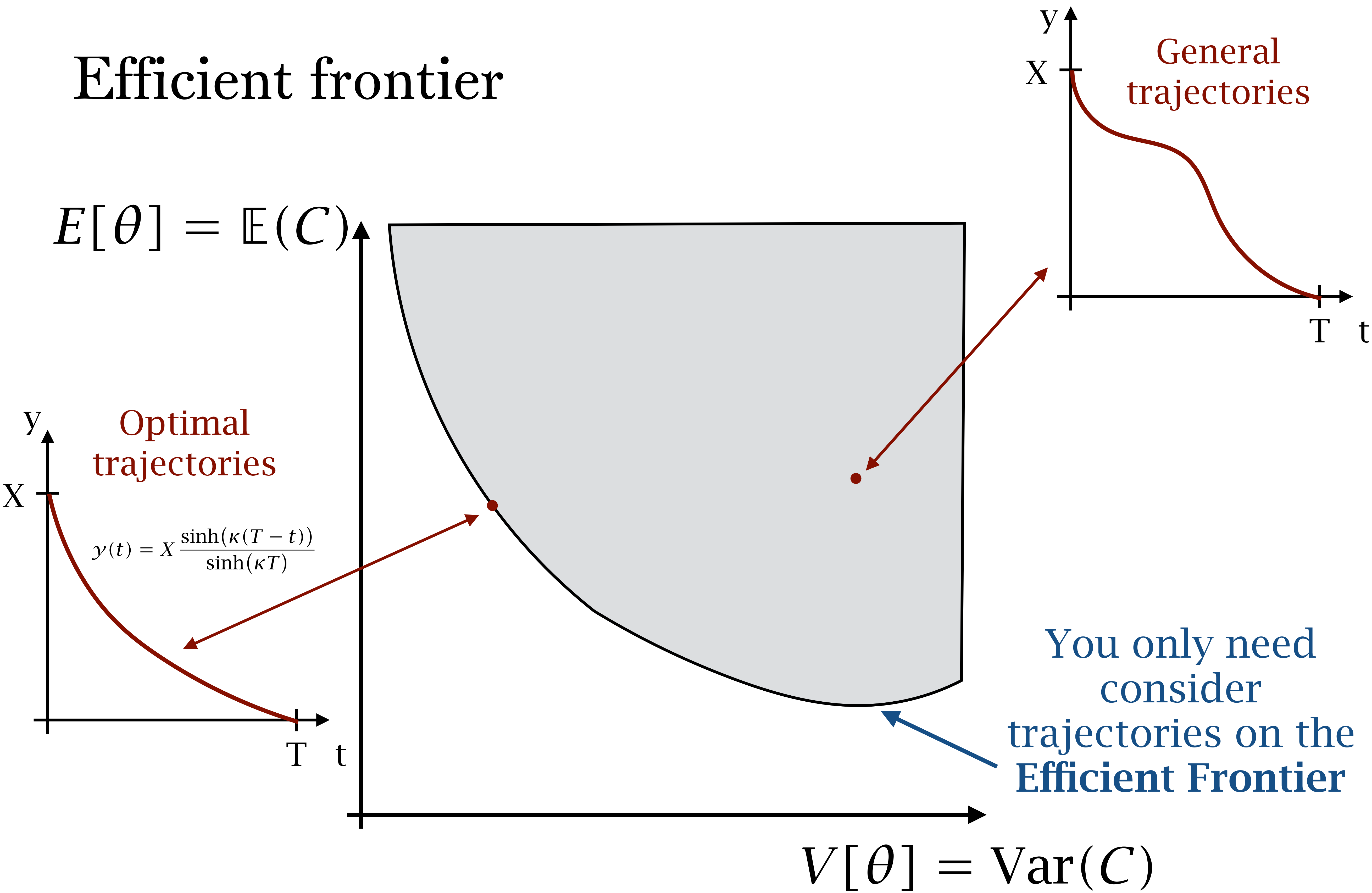
With $y(0) = X$ and $y(T) = 0$, A and B are determined to give

$$y(t) = X \frac{\sinh\big(\kappa(T-t)\big)}{\sinh(\kappa T)} \qquad \sinh x = \frac{1}{2}\left(e^x - e^{-x}\right)$$

Arrival price solution



Efficient frontier



Extensions

- Portfolio trading

  trade more liquid first to reduce risk
- Nonlinear temporary impact function

  more difficult to get analytic solutions
- Drift in price process (homework)
- Short- or long-term pricing signals
- Randomly varying liquidity
- More general impact functions

Nonlinear impact

Optimal execution with nonlinear impact functions and trading-enhanced risk

ROBERT F. ALMGREN

Applied Mathematical Finance **10**, 1–18 (2003)

$$h(v) = \eta v^k$$

$$\frac{x(t)}{X} = \begin{cases} \left(1 + \dfrac{1-k}{1+k}\,\dfrac{t}{T_*}\right)^{-(1+k)/(1-k)} & \text{if } 0 < k < 1 \\ \exp\left(-\dfrac{t}{T_*}\right) & \text{if } \quad k = 1 \\ \left(1 - \dfrac{k-1}{k+1}\,\dfrac{t}{T_*}\right)^{(k+1)/(k-1)} & \text{if } k > 1 \end{cases}$$

$$T_* = \left(\frac{k\eta X^{k-1}}{\lambda\sigma^2}\right)^{1/(k+1)}$$

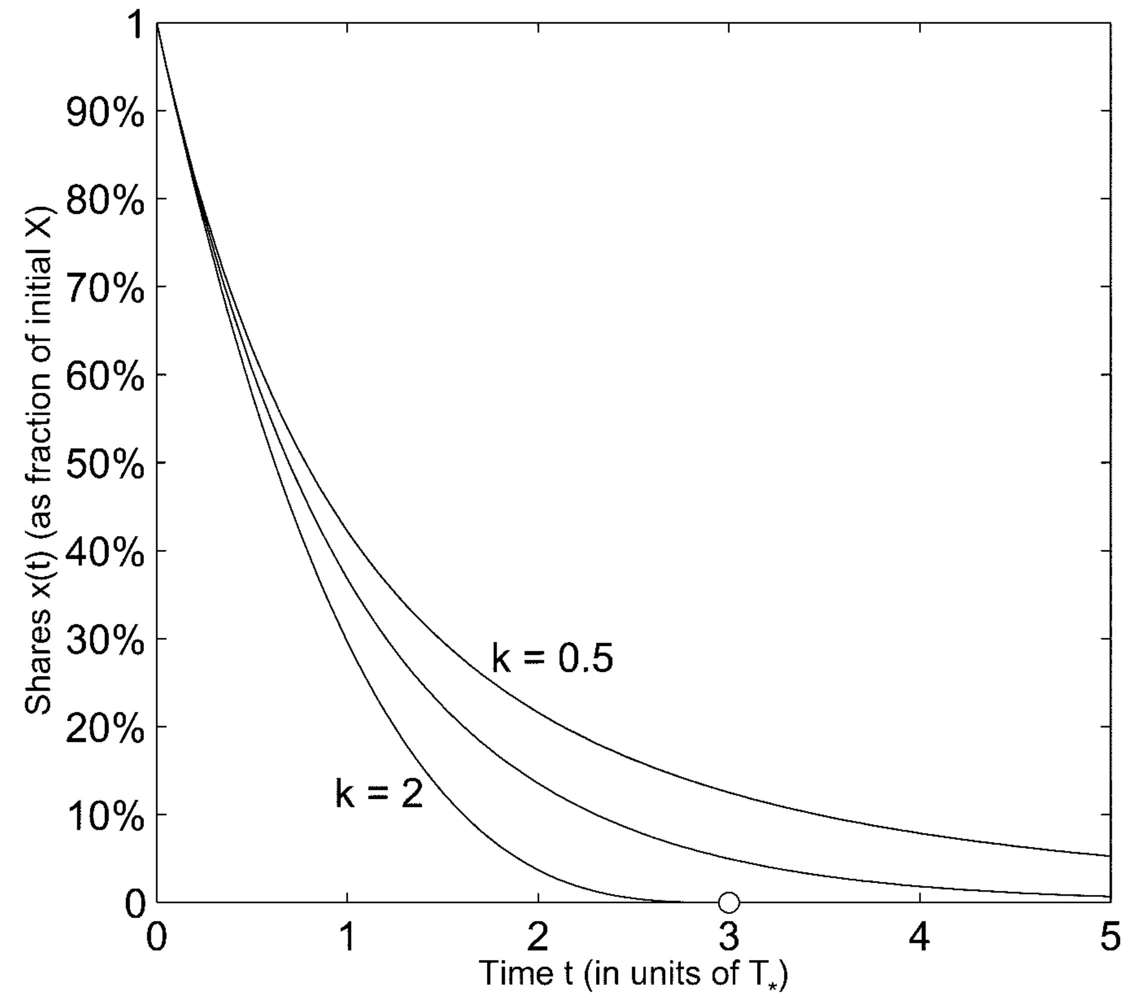


Fig. 2. Optimal solution trajectories $x(t)$, for $k=\frac{1}{2}$, 1, 2. For each value of k, the solution has a 'universal' form, which may be scaled as necessary for different choice of time scales T_* and initial portfolio size X. The disc shows $T_{max}=3T_*$ for $k=2$.

Portfolio trading

Optimal Execution
of Portfolio Transactions*

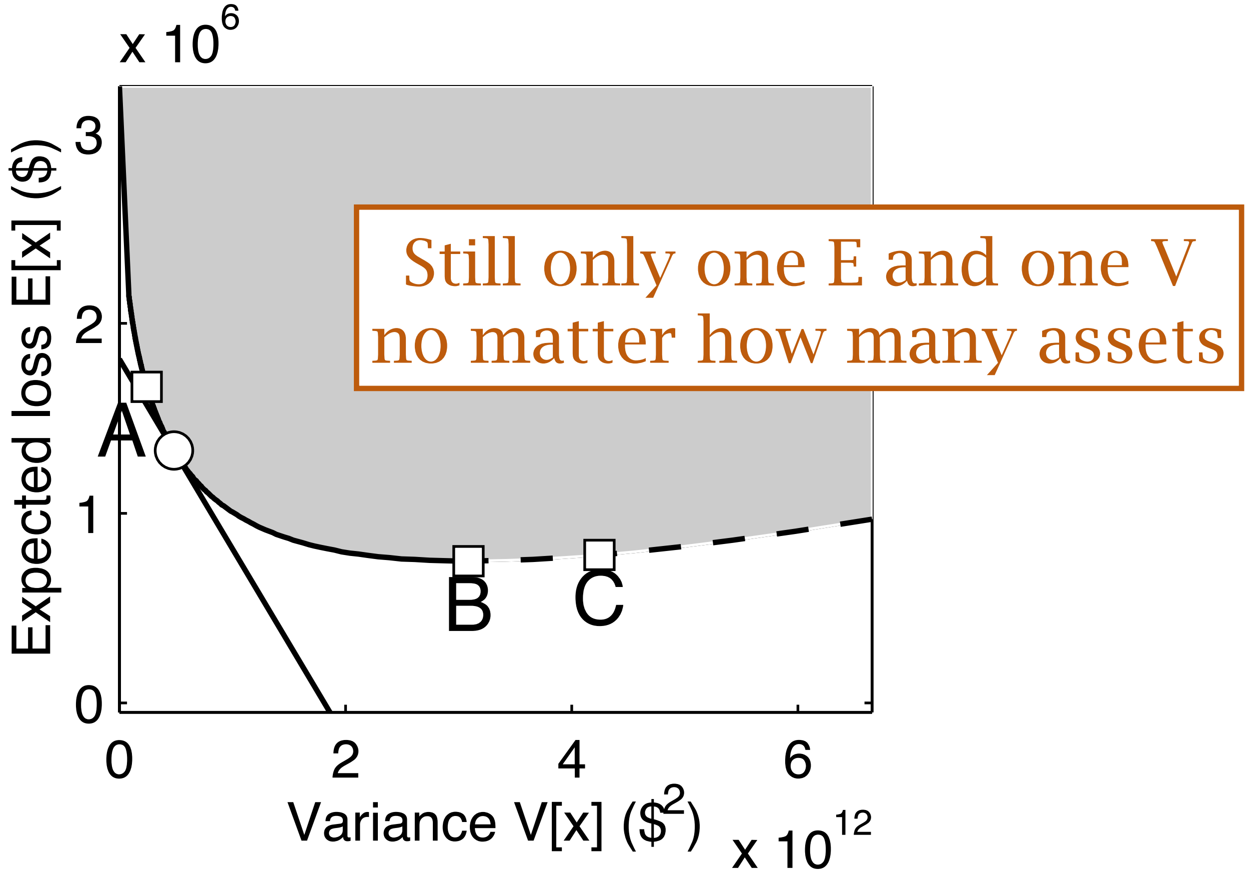
Robert Almgren† and Neil Chriss‡

December 2000

$$g(v) \;=\; \Gamma\, v, \qquad h(v) \;=\; \epsilon\, \mathrm{sgn}(v) \,+\, \mathrm{H}\, v,$$

Γ and H are $m \times m$ matrices,

$$C \;=\; \sigma\sigma^{\mathrm{T}}$$



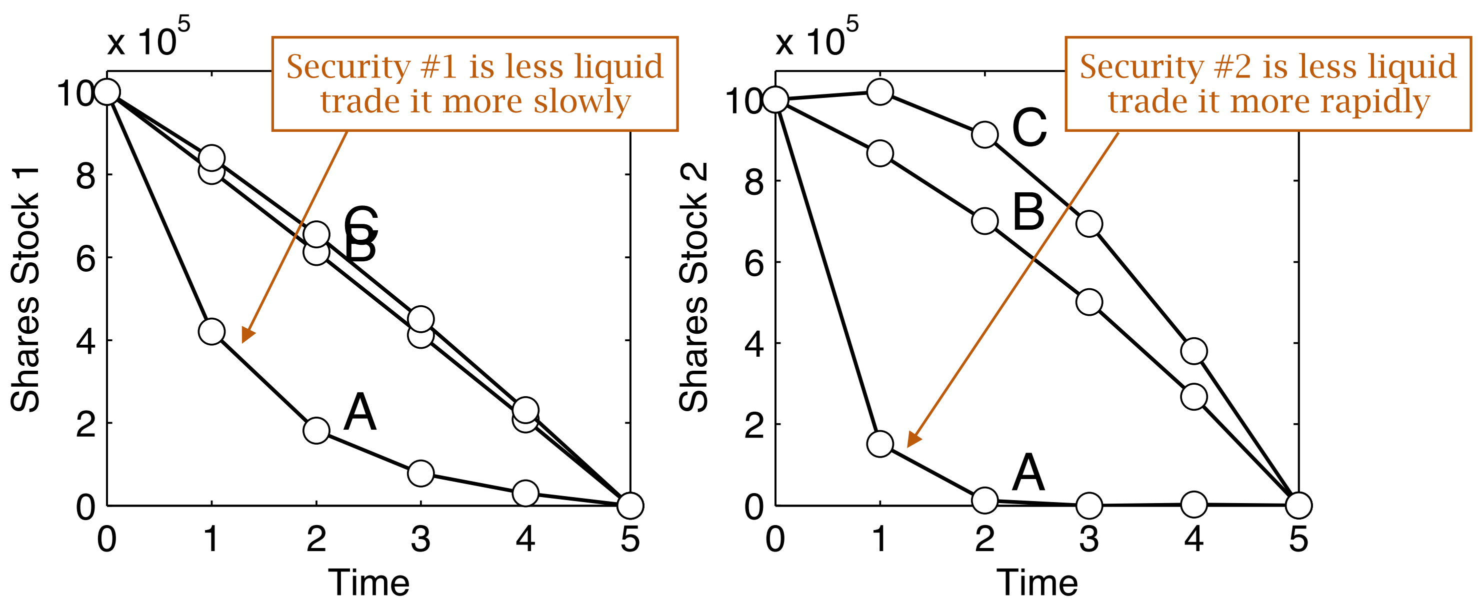


Figure 6: *Optimal trajectories for two securities.* As in Figure 5, for (A) $\lambda = 2 \times 10^{-6}$, (B) the naïve strategy with $\lambda = 0$, (C) $\lambda = -5 \times 10^{-8}$.

Varying liquidity and volatility

Optimal Trading with Stochastic Liquidity and Volatility*

Robert Almgren†

SIAM J. FINANCIAL MATH.
Vol. 3, pp. 163–181

$$dS(t) = \sigma(t)\, dB(t), \qquad \text{Random volatility}$$

$$\tilde{S}(t) = S(t) + \eta(t)\, v(t), \qquad \text{Random liquidity}$$

$$\mathcal{C} = \int_t^T \sigma(s)\, x(s)\, dB(s) + \int_t^T \eta(s)\, v(s)^2\, ds. \qquad \text{Trade cost}$$

$$\eta(t) = \bar{\eta} \exp \xi(t) \qquad \text{and} \qquad \sigma(t) = \bar{\sigma} \exp\left(-\frac{\xi(t)}{2}\right),$$

$$d\xi = a(\xi)\, dt + b(\xi)\, dB_{\mathrm{L}}(t),$$

Mean-reverting process
for liquidity and volatility

Nonlinear partial differential equation
for trade cost and strategy

$$-c_t = \lambda \sigma^2 x^2 - \frac{c_x^2}{4\eta} + a c_\xi + \frac{1}{2} b^2 c_{\xi\xi}.$$

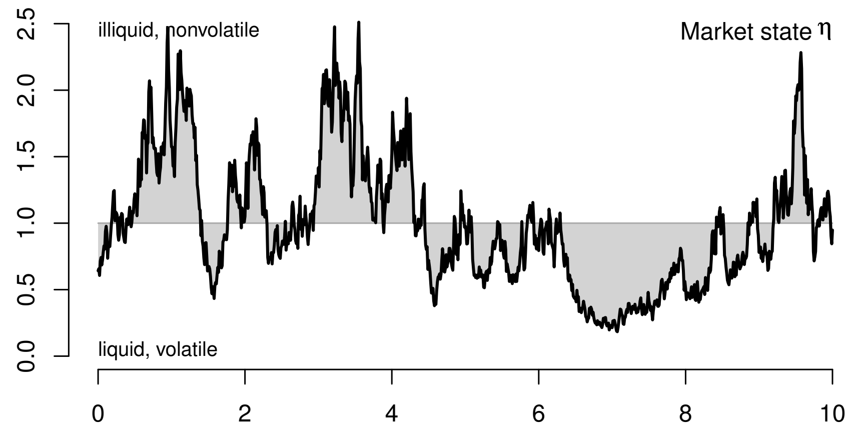
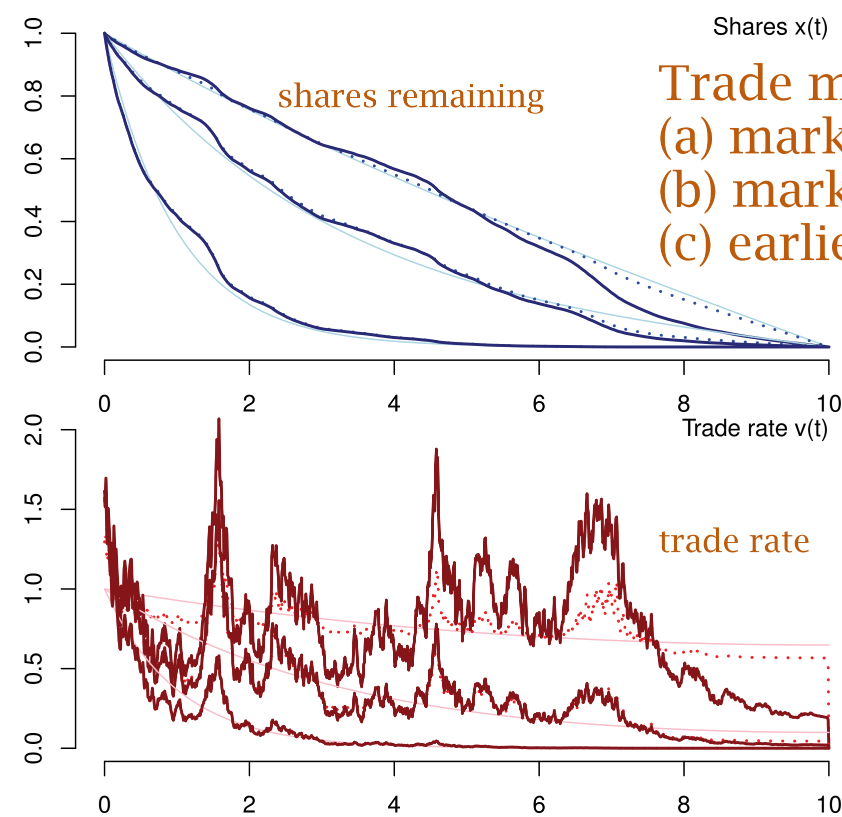


Figure 4. *Realization of the market state trajectory used for the example numerical simulations, with $\beta = 1$. The horizontal axis is time measured in units of the mean-reversion time. The value plotted is the market impact coefficient $\eta(t)$, so large values mean an illiquid market. In our simplified "coordinated variation" market assumption, the volatility $\sigma(t)$ varies inversely with $\eta(t)$: the market is either liquid and volatile, or illiquid and nonvolatile.*



Trade more when
(a) market is more liquid, and
(b) market is more volatile, and
(c) earlier in the interval

Figure 5. *Optimal trajectories for $\lambda = 0.1$ (highest curves), 0.3, and 1 (lowest curves). The horizontal axis is time t forward from the initial time, measured in units of the market mean-reversion time: the program must be completed by $t = T = 10$. The upper panel is shares remaining; the lower panel is the trade rate. Light lines are the nonadaptive solution in the mean market state; dotted lines are the "rolling horizon" approximation. In the lower panel, the trade rate is normalized so that each nonadaptive solution has the same initial trade rate.*

Trading with signals

Finance Stoch
https://doi.org/10.1007/s00780-019-00382-7

CrossMark

Incorporating signals into optimal trading

Charles-Albert Lehalle[1,2] · Eyal Neuman[2,3]

Published online: 14 February 2019

Springer

Abstract We incorporate a Markovian signal in the optimal trading framework which was initially proposed by Gatheral et al. (Math. Finance 22:445–474, 2012) and provide results on the existence and uniqueness of an optimal trading strategy. Moreover, we derive an explicit singular optimal strategy for the special case of an Ornstein–Uhlenbeck signal and an exponentially decaying transient market impact. The combination of a mean-reverting signal along with a market impact decay is of special interest, since they affect the short term price variations in opposite directions. Later, we show that in the asymptotic limit where the transient market impact becomes instantaneous, the optimal strategy becomes continuous. This result is compatible with the optimal trading framework which was proposed by Cartea and Jaimungal (Appl. Math. Finance 20:512–547, 2013). In order to support our models, we analyse nine months of tick-by-tick data on 13 European stocks from the NASDAQ OMX exchange. We show that order book imbalance is a predictor of the future price move and has some mean-reverting properties. From this data, we show that market participants, especially high-frequency traders, use this signal in their trading strategies.

Fig. 2 Simulation of the optimal inventory X^* which corresponds to the trading speed r^* from Proposition 3.2 (b). In the black curve, we present the optimal inventory in the absence of a signal. The red region is a plot of 1000 trajectories of the optimal inventory X^*. The parameters of the model are $\gamma = 0.1$, $\sigma = 0.1$, $I_0 = 0$, $T = 10$, $\kappa = 0.5$, $\phi = 0.1$, $X_0 = 10$ and $\varrho = 10$

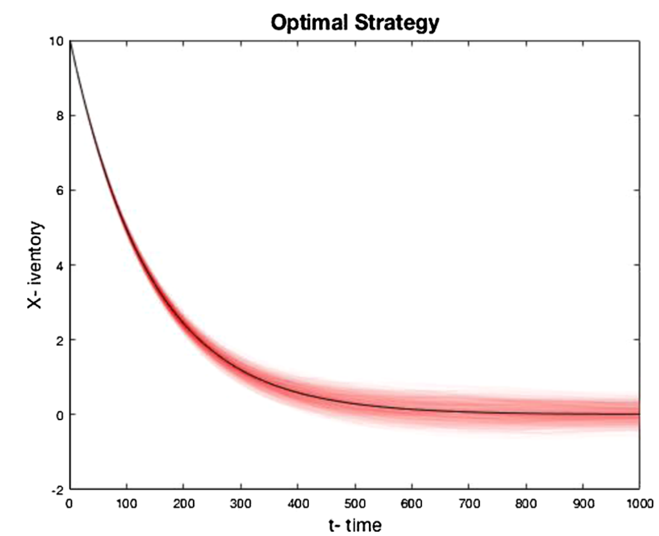
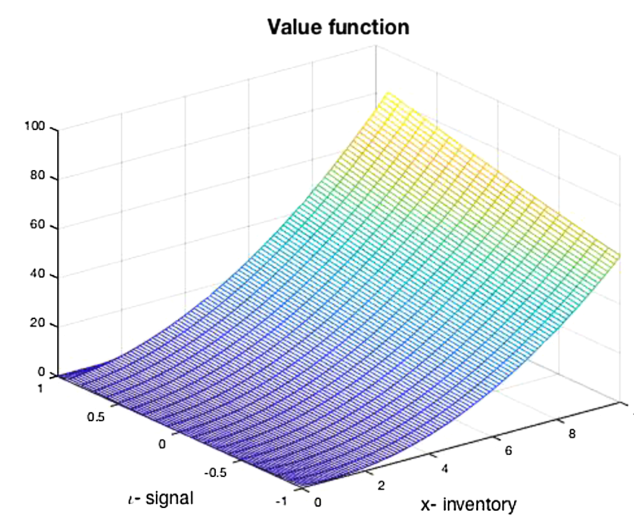
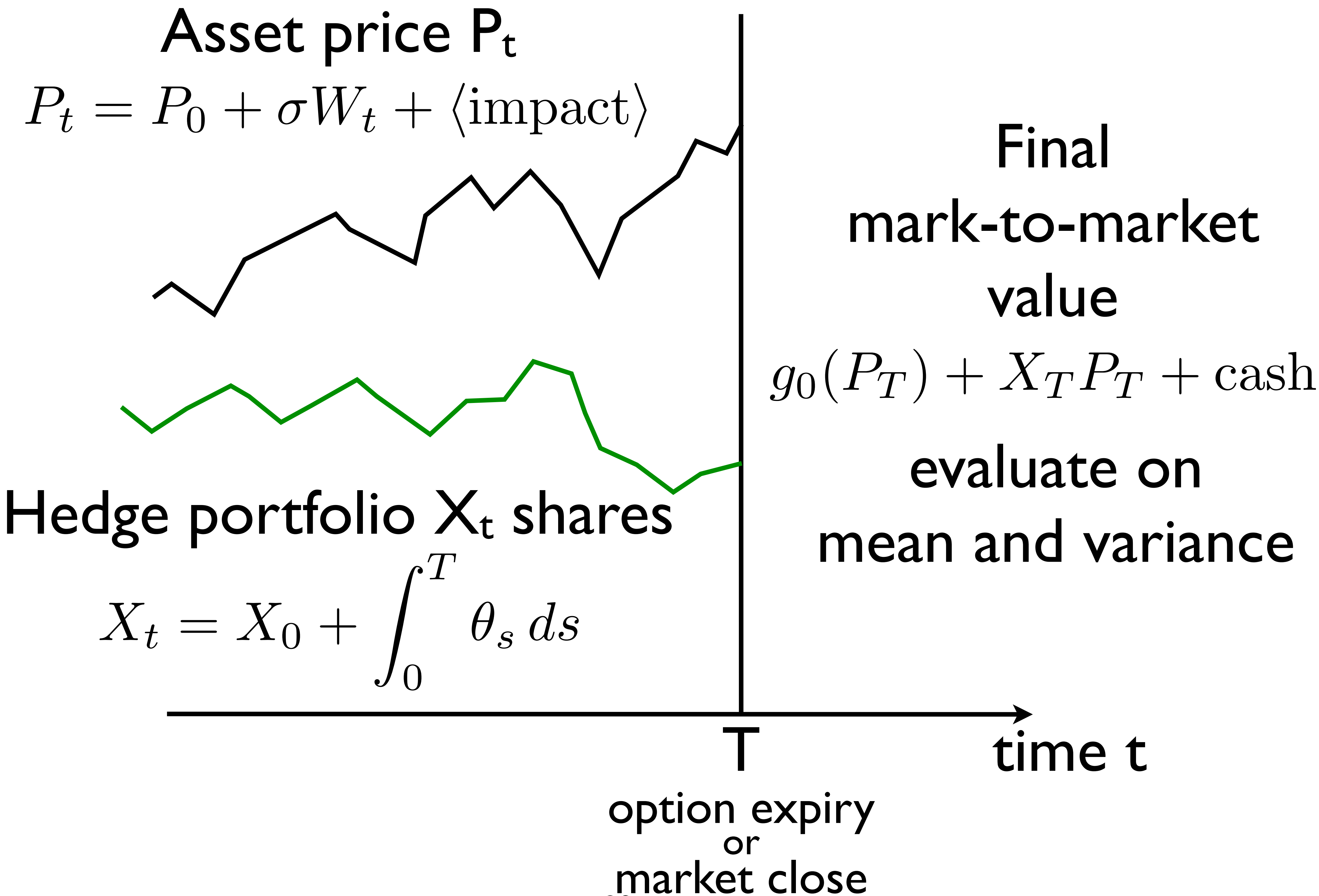


Fig. 3 Plot of the value function $V(0, \iota, c, x, p) - (c - xp)$ from (3.5) when the signal is an OU process. The parameters of the model are $\gamma = 0.1$, $\sigma = 0.1$, $T = 10$, $\kappa = 0.5$, $\phi = 0.1$, $X_0 = 10$ and $\varrho = 10$



Option hedging



Market Microstructure and Liquidity
Vol. 2, No. 1 (2016) 1650002 (26 pages)

DOI: 10.1142/S2382626616500027

Option Hedging with Smooth Market Impact

Robert Almgren*
Quantitative Brokers, New York, NY, USA
ralmgren@quantitativebrokers.com

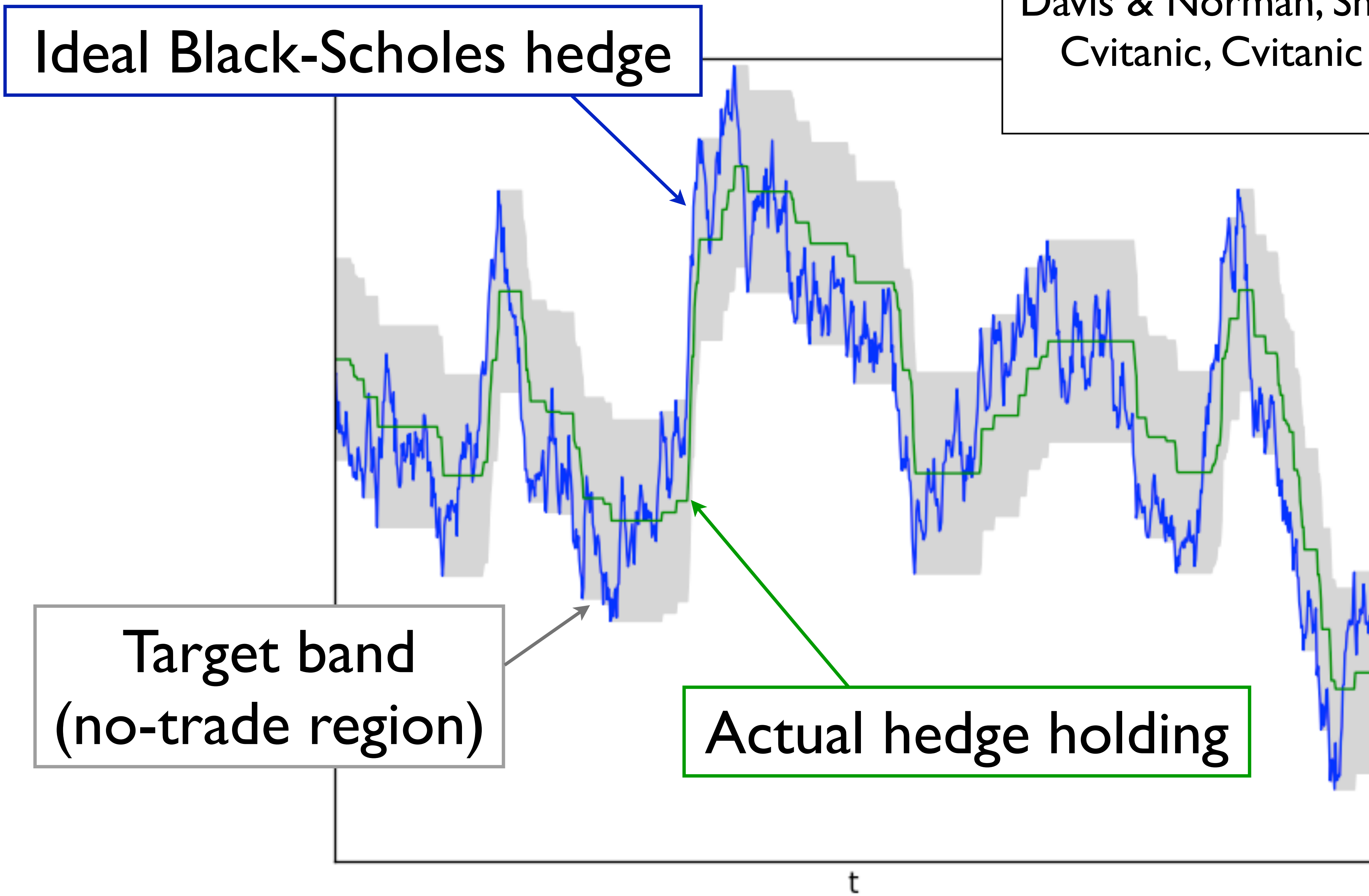
Tianhui Michael Li
The Data Incubator, New York, NY, USA

Accepted 15 April 2016
Published

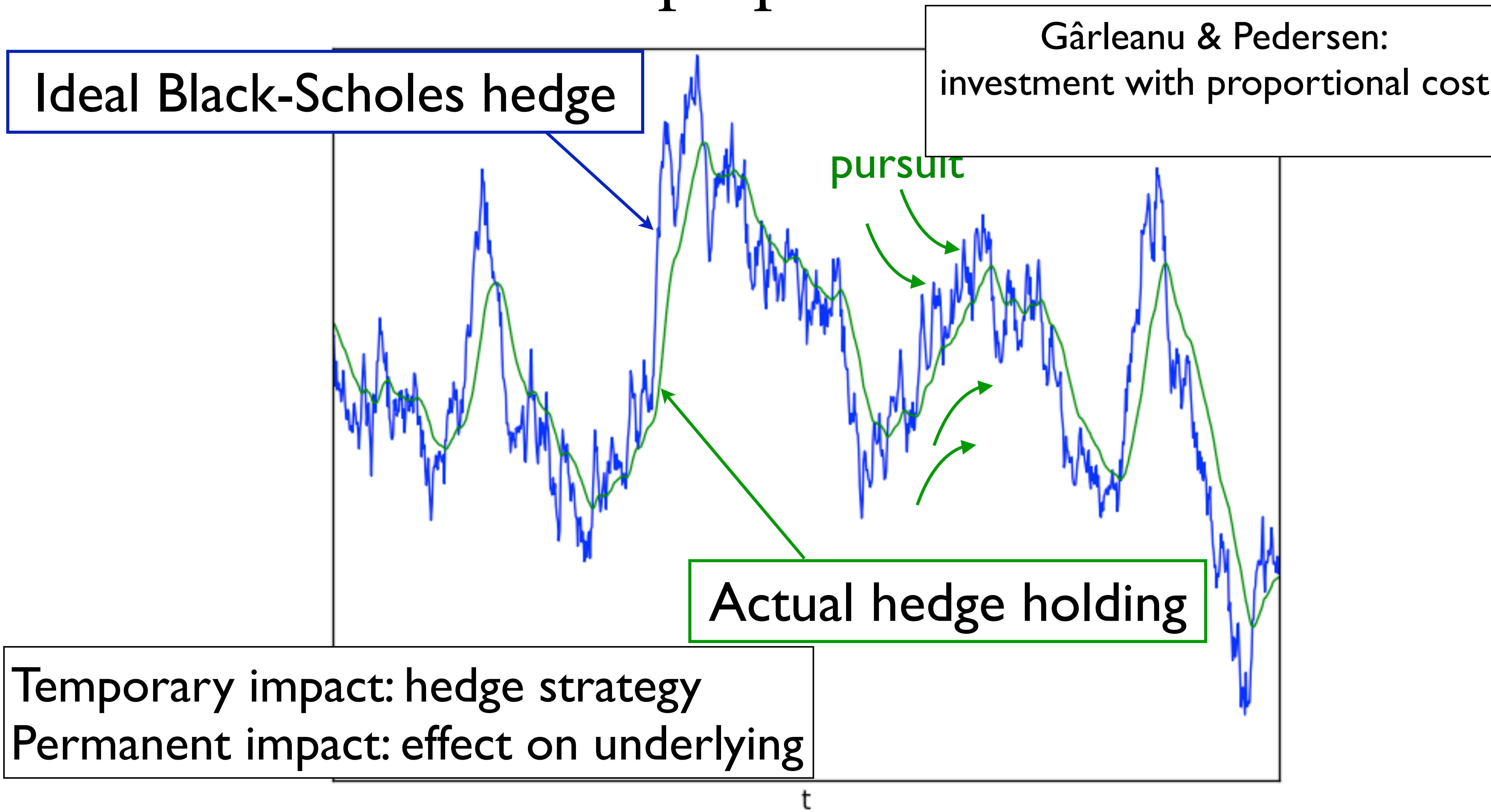
We consider intraday hedging of an option position, for a large trader who experiences temporary and permanent market impact. We formulate the general model including overnight risk, and solve explicitly in two cases which we believe are representative. The first case is an option with approximately constant gamma: the optimal hedge trades smoothly towards the classical Black–Scholes delta, with trading intensity proportional to instantaneous mishedge and inversely proportional to illiquidity. The second case is an arbitrary non-linear option structure but with no permanent impact: the optimal hedge trades toward a value offset from the Black–Scholes delta. We estimate the effects produced on the public markets if a large collection of traders all hedge similar positions. We construct a stable hedge strategy with discrete time steps.

Solutions with bid-ask spread cost

Davis & Norman, Shreve & Soner, Cvitanic, Cvitanic & Karatzas

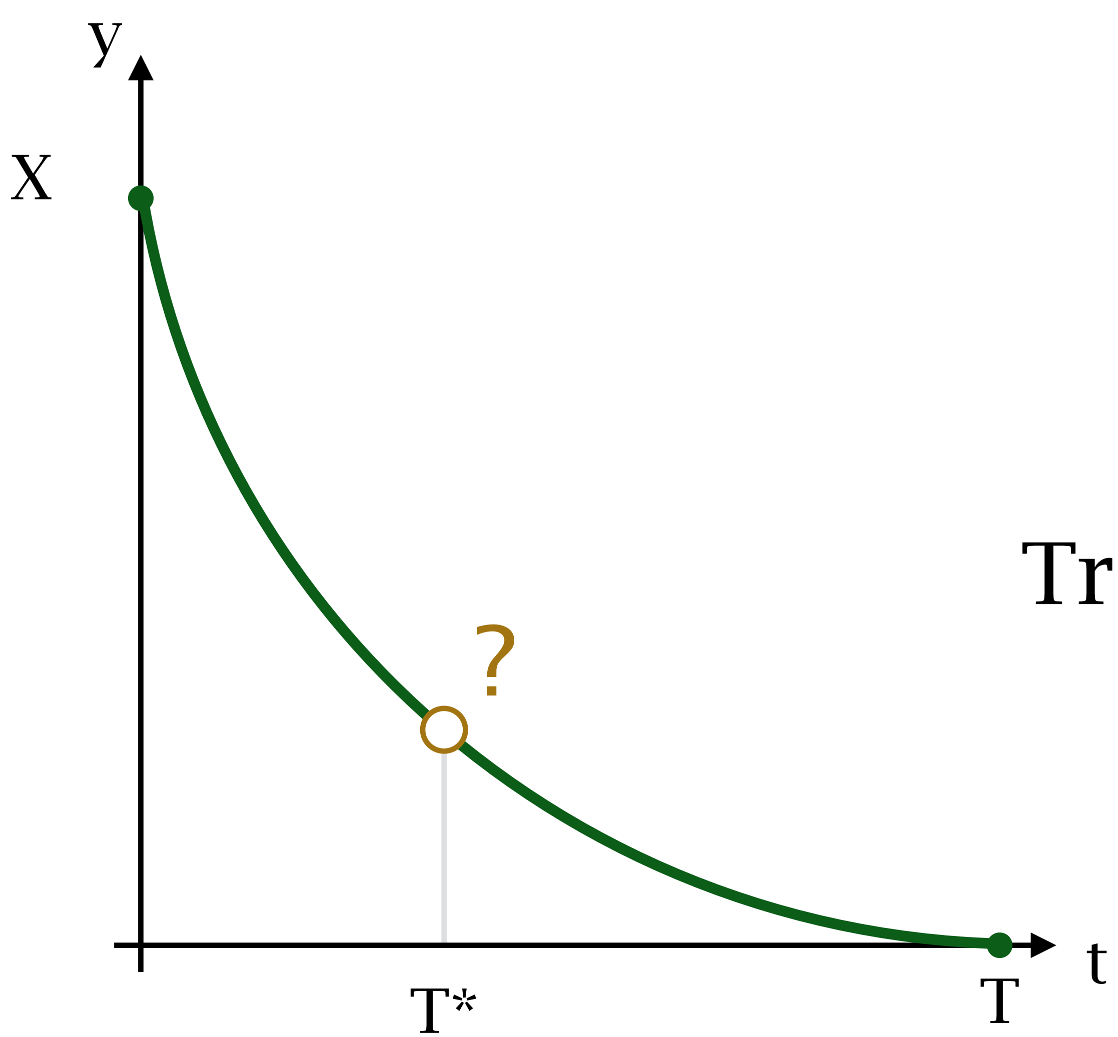


Our solutions with proportional cost



$$\theta_t \;=\; -\kappa\, h\big(\kappa(T-t)\big)\cdot\big(X_t - \text{target}\big)$$

Static vs dynamic



If you recalculate the trajectory at $t^* < T$, using the same λ and your utility function at that time, then you continue the original trajectory. Trajectories are *time-stable.*

There are subtleties.

Trajectories with drift

$$dP_t = \alpha\, dt + \sigma\, dB_t + \gamma\, \theta_t$$

A static position of size x has

- gain per unit time α x
- variance per unit time σ^2 x^2

Markowitz optimal portfolio

$$\min_x \left(\alpha x - \lambda \sigma^2 x^2 \right)$$

$$x_* = \frac{\alpha}{2\lambda\sigma^2}$$