Reinforcement Learning for Optimized Trade Execution

Yuriy Nevmyvaka, Lehman Brothers (& CMU)
Yi Feng, Penn
Michael Kearns, Penn (& Lehman Brothers)

ICML 2006

Background on Market Microstructure

- Consider a typical exchange for some specific stock
- Limit order: specify price (away from the market)
- (Non-executable) Orders are placed in the buy or sell book
  - sorted by price; top prices are the bid and ask
- (Partially) Executable orders are filled immediately
  - prices determined by standing orders in the book
  - one order may execute at multiple prices
  - the "mechanical" component of market impact
- Market order: limit order with an extreme price
- Full order books now visible in real time
- What are they good for?

refresh | island home | disclaimer | help

MSFT

GET STOCK

MSFT go

Symbol Search

LAST MATCH		TODAY'S ACTIVITY	
Price	24.0700	Orders	52,983
Time	14:57:07.72	Volume	10,243,212

BUY ORDERS		SELL ORDERS	
SHARES	PRICE	SHARES	PRICE
500	24.0620	500	24.0690
6,000	24.0610	500	24.0690
5,000	24.0600	500	24.0700
100	24.0600	200	24.0800
1,100	24.0550	1,981	24.0900
100	24.0500	412	24.0900
5,000	24.0500	3,000	24.0980
200	24.0500	500	24.1000
3,294	24.0500	100	24.1200
1,000	24.0500	2,800	24.1400
3,000	24.0430	5,000	24.1400
100	24.0400	1,000	24.1400
5,503	24.0400	5,000	24.1500
2,100	24.0300	400	24.1600
2,800	24.0300	1,000	24.1700
(412 more)		(694 more)	

As of 14:57:16.178

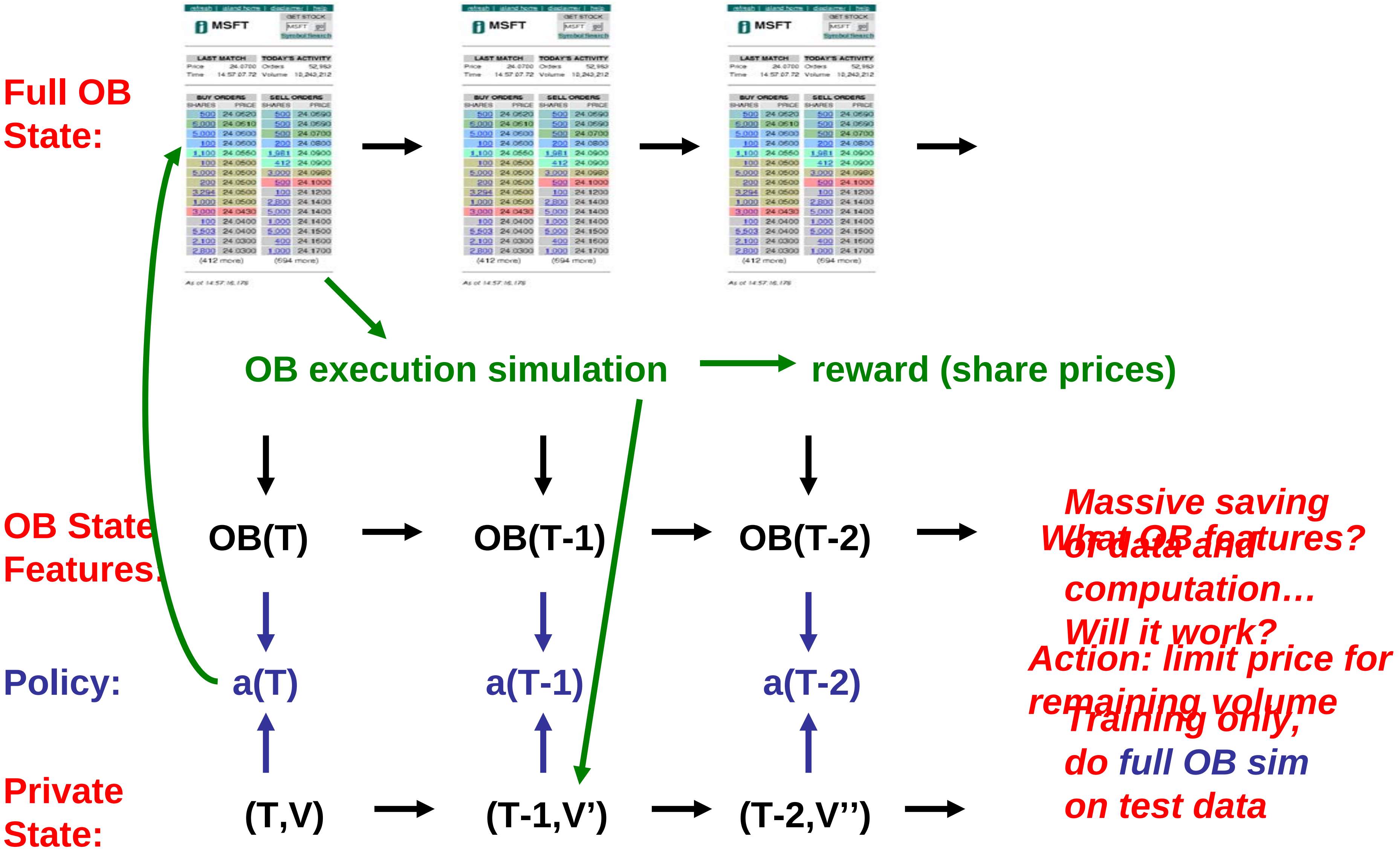
Optimized Trade Execution

- Canonical execution problem: sell V shares in T time steps
  - must place market order for any unexecuted shares at time T
  - trade-off between price, time… and liquidity
  - problem is ubiquitous
- Canonical goal: Volume Weighted Average Price (VWAP)
  - attempt to attain per-share average price of executions
  - widely used on Wall Street; reduces risk sources to execution

RL for Optimized Execution

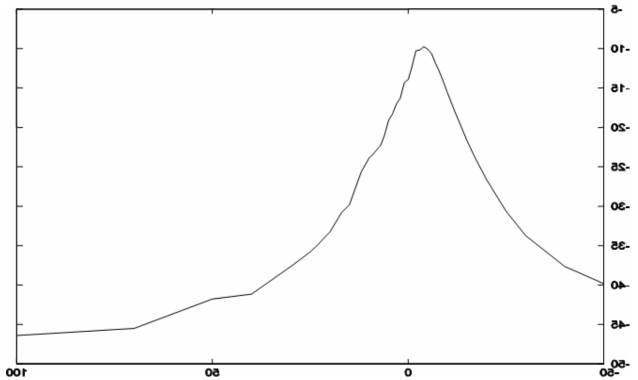
- Basic idea: execution as state-based stochastic optimal control
  - state: time and shares remaining… what else?
  - actions: position(s) of orders within the book
  - rewards: prices received for executions
  - stochastic: because same state may evolve differently in time
- This work: large-scale application of RL to microstructure
- Related work:
  - Bertsimas and Lo
  - Coggins, Blazejewski, Aitken

“No Impact” State Factorization



Experimental Details

- Stocks: AMZN, NVDA, QCOM (varying liquidities)
- V = 5K and 10K shares
  - divided into 1, 4 or 8 levels of observed discretization
- T = 2 and 8 mins
  - divided into 4 or 8 decision points
- Explored a variety of OB state features
- Learned optimal strategy on 1 year of INET training data
- Tested strategy on subsequent 6 months of test data
- Evaluation:
  - compare to optimized submit and leave strategies
    - best single limit order price at start of trading interval
    - simplest form of learning
  - performance criterion: implementation shortfall
    - basis points compared to all shares at initial spread midpoint
    - an unattainable ideal (infinite liquidity assumption)

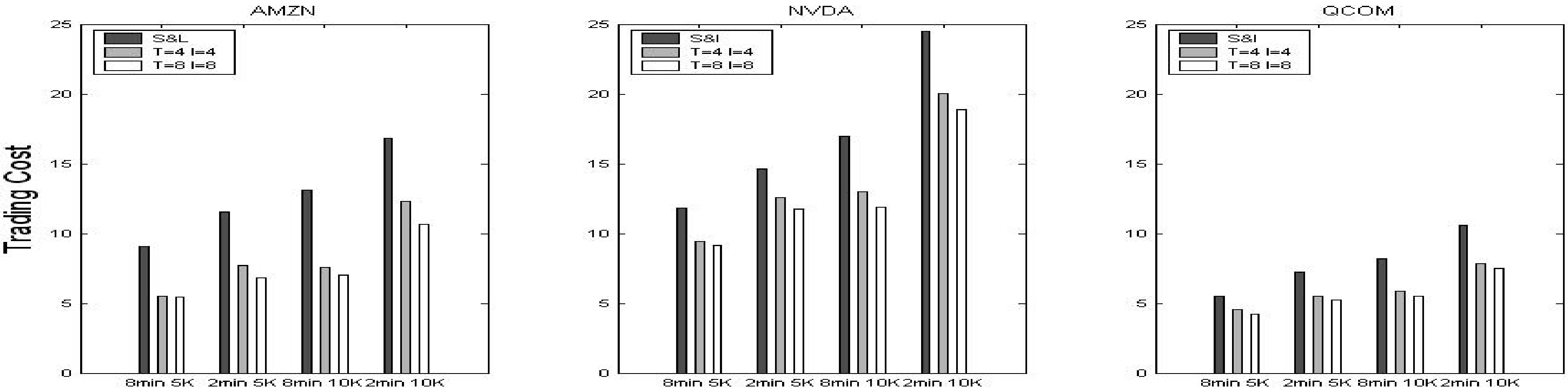


Trading Cost vs. Limit Price

deep in OB M.O. at start

Results

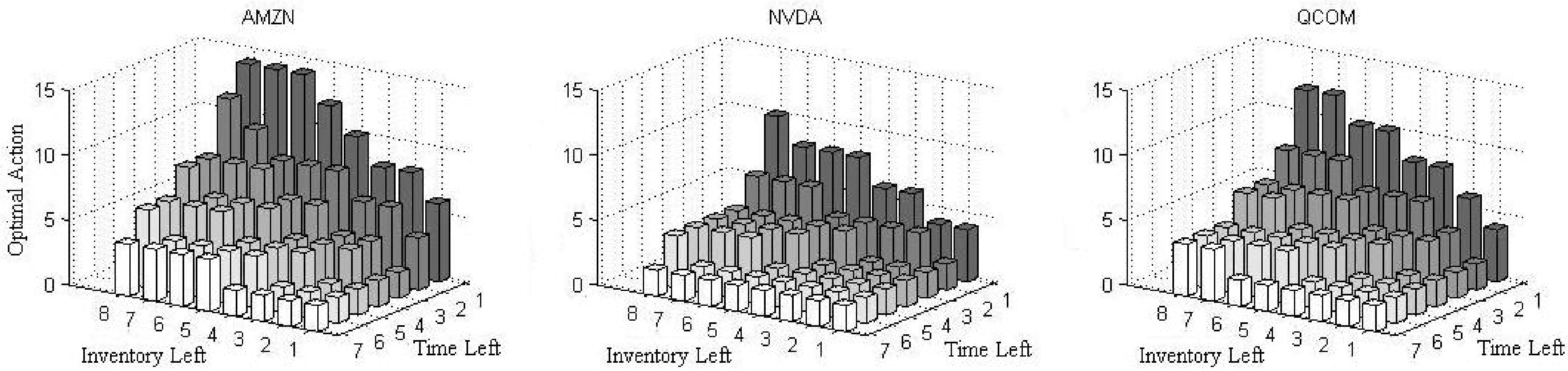
Private State Variables Only: Time and Inventory Remaining



Average Improvement Over Optimized Submit-and-Leave

T=4 I=1	27.16%	T=8 I=1	31.15%
T=4 I=4	30.99%	T=8 I=4	34.90%
T=4 I=8	31.59%	T=8 I=8	35.50%

Strategy Visualization (10K, 2min)

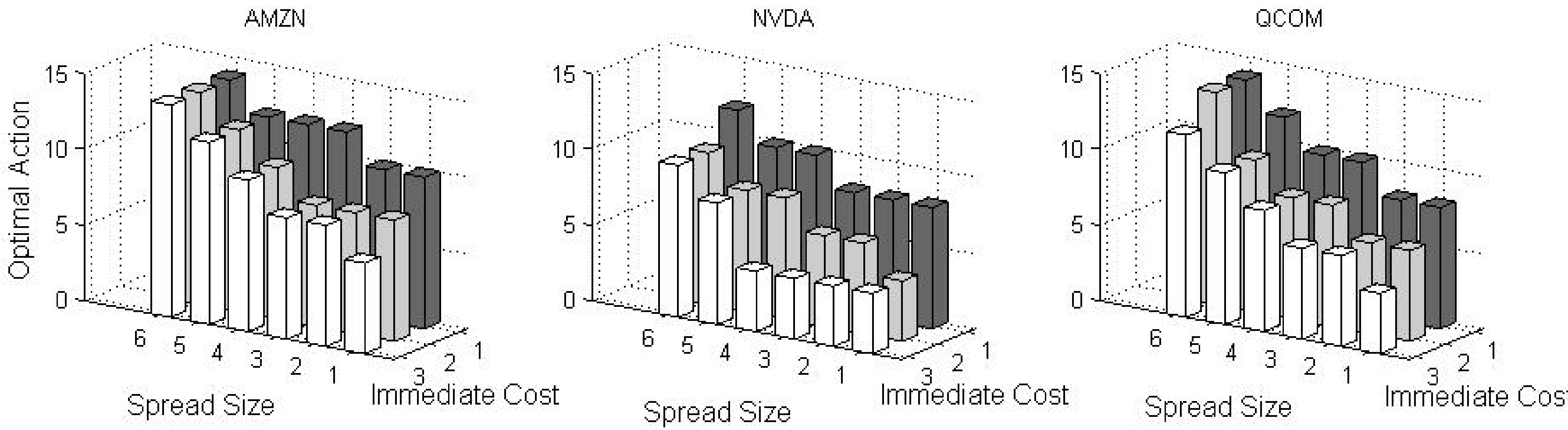


General shape is intuitive, but (stock-specific) numerical optimization matters!

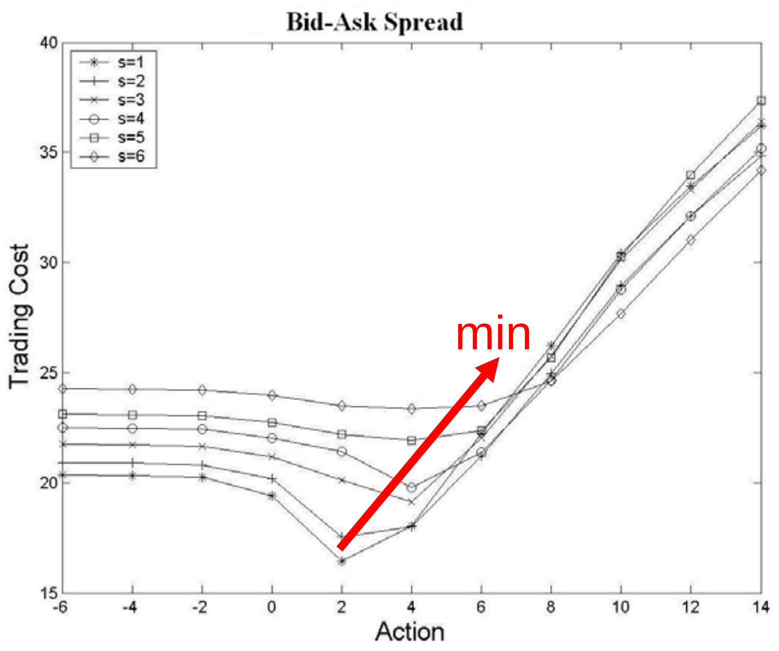
Improvement From Order Book Features

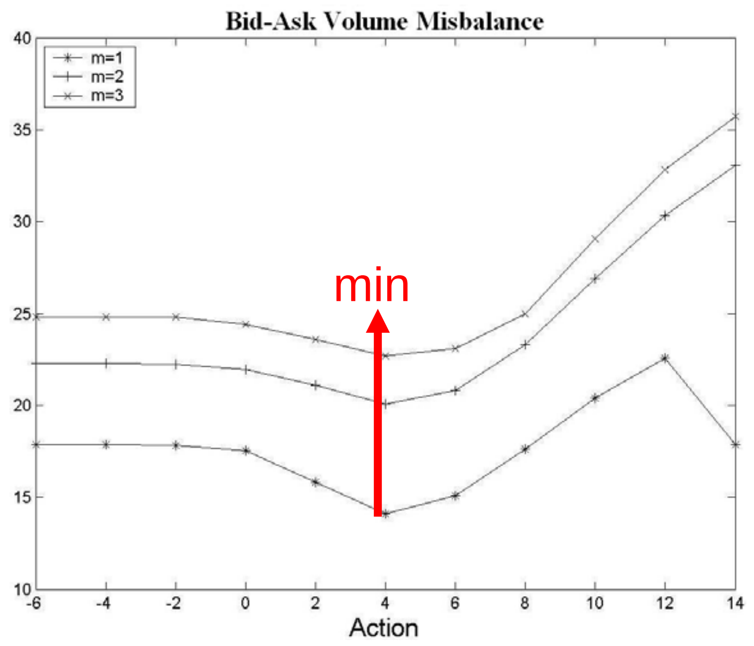
Bid Volume	-0.06%	Ask Volume	-0.28%
Bid-Ask Volume Misbalance	0.13%	Bid-Ask Spread	**7.97%**
Price Level	0.26%	Immediate Market Order Cost	**4.26%**
Signed Transaction Volume	**2.81%**	Price Volatility	-0.55%
Spread Volatility	1.89%	Signed Incoming Volume	0.59%
Spread + Immediate Cost	**8.69%**	Spread+ImmCost+Signed Vol	**12.85%**

Strategy Visualization II



Q-Values: Trading Costs vs. Actions





predictive and actionable

predictive but not actionable

Future Work

- “Fancier” RL
  - function approximation
  - may permit richer feature set, but…
- RL for other stylized trading problems
  - market-making strategies
- Theory: low-impact RL?