Endogenous liquidity crises

Antoine Fosset et al J. Stat. Mech

Maxime Pedron Augustin Boissier James Swomley Leon Kloker

Stanford University, ICME

MATH 237A

May 30, 2023

Presentation Overview

How reasonable are the standard economics theories ?

Agents in **standard economic models** are seen as perfect optimisers :

- Fully rational
- Fully informed
- Homogeneous
- Independent (isolated)

$\Rightarrow$ **Collective behaviour is the reflection of individual behaviour** (representative agents/firms/central banks)

Not hard to see that **real economic agents** are :

- **Irrational**
- **Partially informed**
- **Highly heterogenous**
- **Connected** (they influence one another)

$\Rightarrow$ **Aggregate outcomes do not reflect individual motives**

Excess Volatility

Excess Volatility Puzzle

$\Rightarrow$ The volatility of financial markets, but also of large economies is **much too large** to be explained by *"fundamentals"* [1] [2]

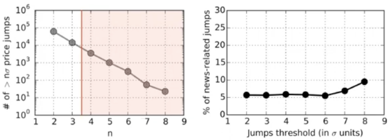


Figure: Stock moves distribution has a fat tail: **most of the market volatility is endogenous in nature**, in contradiction with standard economic theory

[1] DM. Cutler, JM. Porterba *What moves stock prices ?* 1989

[2] Bouchaud et al., *Exogenous and Endogenous Price Jumps Belong to Different Dynamical Classes* 2021

Excess Volatility

SP500 flash crash of 6th May 2010 is an example of an endogenous extreme event.

Other examples:

- May 28, 1962 (US Stock Market)
- October 15, 2014 (Treasury bond flash crash)
- October 7, 2016 (British pound flash crash)
- August 2, 2020 (Bitcoin flash crash)

Even for exogenous crises: while a large number of crashes are triggered by exogenous factors, endogenous mechanisms are at play a reinforcing role.

Continuous double action

The vast majority of modern market use an electronic **limit order book (LOB)** updated in real time and observable by all traders.

Each market participant may

1. provide firm trading opportunities to the rest of the market by posting a **limit order** at a specified price
2. accept such trading opportunities by placing a **market order**, at the current best price

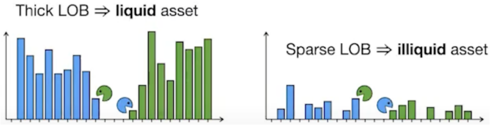
Figure: [a]

[a] OMI 446: Physics of Socioeconomic Systems, Michael Benzaquen

Liquidity and Market impact

When looking at microstructure mechanics, it becomes clear that **trades consume liquidity** and **mechanically impact prices**

Market **liquidity** = capacity of the market to accomodate a large market order
(Liquidity is difficult to define because it is a **dynamical concept**, limit order are continuously deposited, cancelled and executed against incoming market order)

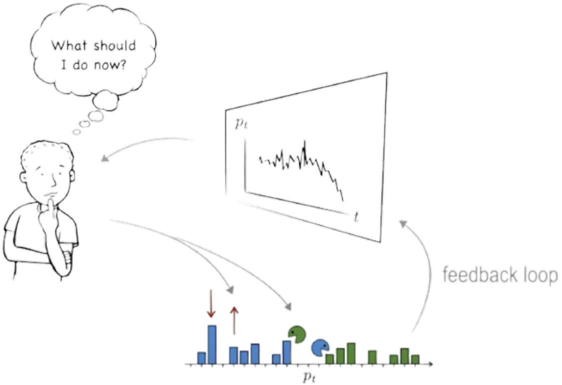


The available volume in the order book at a given instant in time is a very small fraction, typically $< 1\%$, of the total daily traded volume, which is itself also very small compared to the total market capitalisation

Feedback loops and liquidity seizures

Large endogenous price moves seem to be the result of **feedback loops** that lead to liquidity dry outs

Empirical data indeed reveal that the liquidity flow into the order book (limit orders, cancellations and market orders) is influenced by past price changes.

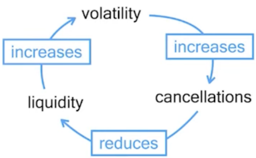


Feedback loops and liquidity seizures

A burst of volatility creates anxiety for liquidity providers, who fear that some information about the future price, unknown to them, is the underlying reason for the recent price changes.

→ Increased reluctance to provide liquidity: likely to cancel their existing limit orders and less likely to refill the order book

→ Less liquidity is likely to amplify the future price moves, thereby creating an unstable feedback loop which might result in a liquidity breakdown



Empirical Evidence - Introduction

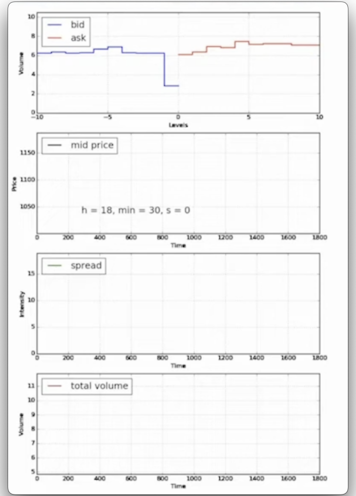
In the following slide, we will be showing the 2010 flash crash's effect on spread dynamics and volume :

- The instruments traded is the **E-mini SP 500 Index Futures** (CME), an instrument to get market exposure to the SP500, which was strongly affected during the flash crash
- The graphs update each ms ie the graphs represent **30-minutes** worth of trading on the 6th May 2010, starting a little before 2:30PM EDT

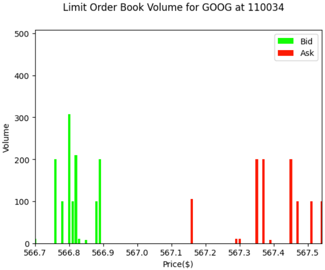
There are 4 graphs in the slide :

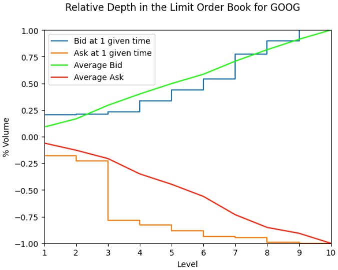
1. Relative depth of market order book
2. Mid-Price of e-mini SP500 futures vs Time
3. Spread of the e-mini SP500 futures vs Time
4. Total Volume vs Time

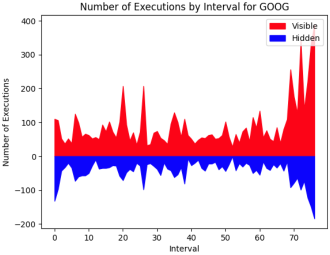
Empirical Evidence



Google Stock Order Book example







Empirical evidence

To account for the joint dynamics of liquidity and price, we introduce the following 6-dimensional process for order book events

$$N_t = \left(N_t^{\mathrm{C,b}}, N_t^{\mathrm{LO,b}}, N_t^{\mathrm{MO,b}}, N_t^{\mathrm{MO,a}}, N_t^{\mathrm{LO,a}}, N_t^{\mathrm{C,a}}\right)$$

$\lambda_t \in \mathbb{R}^6$ intensity of the process

Q-Hawkes process

Even though the paper does not go into detail about the price dynamics, we can make the assumption that the price process P is driven by a Brownian motion dynamic, ie $P_t = f((W_s)_s, t)$

Equation for the Queue-reactive Hawkes process

$$\lambda_t = \alpha_0$$

Base rate

$$+ \int_0^t \phi(t-s)d\boldsymbol{N}_s$$

Standard linear Hawkes contribution: effect of past events on the current event rate

$$+ \int_0^t \boldsymbol{L}(t-s)dP_s$$

Linear contribution of price fluctuations: effect of past local price trend on current rate

$$+ \iint_0^t \boldsymbol{K}(t-s, t-u)dP_s dP_u$$

Quadratic contribution of price fluctuations: effect of volatility and square trend on

Empirical Evidence

$$\lambda_t = \alpha_0 + \boxed{\int_0^t \underline{\phi}(t-s)dN_s}$$

Estimate $\underline{\phi}$ = $[\phi_{i,j}] \in \mathbb{R}^{6x6}$ by

using the EM & MLE algorithm given by **E. Bacry**'s library **tick**.

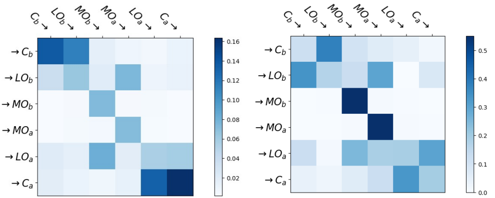


Figure: $|\phi|$ for Apple & Google stock order book calibration using sum of exponential kernels, between 9:30am-10:30am 06/21/2012

The kernel norms $||\phi_{ij}||$ represent the average number of events of type i caused by an event of type j. Antoine Fosset proved that $\phi_{i,i}(t)$ decreases at the speed $\frac{1}{t^{\frac{3}{2}}}$

Empirical Evidence - Fitting Part

To use the favorable characteristics of the exponential kernel in Hawkes processes, we opt for its fitting by employing a set of $(f_{\alpha,\beta})$ large enough, where $f_{\alpha,\beta}(t) = \alpha e^{-\beta t}$. This is due to the validity of the following equation:

$$\lambda_i(t) = \alpha_0^i + \sum_{j=1}^{p} \int_0^t \phi_{ij}(t-s)dN_j(s), \quad i = 1, \cdots, p\,(p = 6 \text{ here})$$

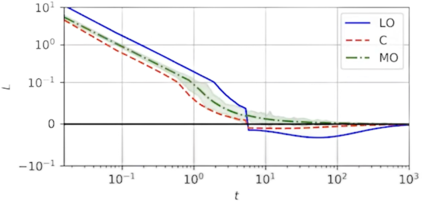
$$\begin{aligned}\|\phi_{ij}\| &= \int_0^\infty \phi_{ij}(t)dt \\ &= \int_0^\infty \sum_{u=1}^{m} \alpha_{ij}^u e^{-\beta^u t} dt \\ &= \sum_{u=1}^{m} \frac{\alpha_{ij}^u}{\beta^u}\end{aligned}$$

E. Bacry advised us to take the family generated by $(f_{\alpha,k\beta})$ for $k < N$.

⚠ On theory it's impossible to have 2 events at the same time, but in reality this can happen because of time discretization

Empirical Evidence

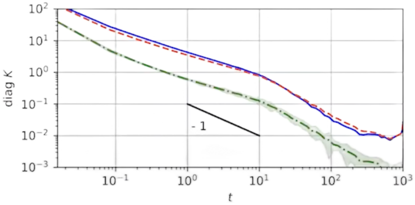
$$\lambda_t = \alpha_0 + \int_0^t \underline{\phi}(t-s)d\boldsymbol{N}_s + \boxed{\int_0^t L(t-s)dP_s} + \iint_0^t \boldsymbol{K}(t-s, t-u)dP_s dP_u$$



Coordinates of $L(t)$ decreases as the speed of $\frac{1}{t}$

Empirical evidence

$$\lambda_t = \alpha_0 + \int_0^t \underline{\phi}(t-s)d\boldsymbol{N}_s + \int_0^t L(t-s)dP_s + \boxed{\iint_0^t \boldsymbol{K}(t-s, t-u)dP_s dP_u}$$



$K(t, t)$ decreases as the same speed as L , $\frac{1}{t}$

Zumbach Effect

Separate the contributions of **trend** and of **volatility** to the quadratic feedback, a meaningful approximation for K is given by the sum of :

- a purely **diagonal matrix**
- a **rank-one** contribution

$$K^i(t-s,t-u) = \underbrace{K_d^i \psi^i(t-s) 1_{\{s=u\}}}_{\text{purely \textbf{diagonal matrix}}} + \underbrace{K_1^i Z^i(t-s) Z^i(t-u)}_{\text{rank-one contribution}} \quad (1)$$

$$\lambda_t = \alpha_0 + \int_0^t \underline{\phi}(t-s) d\mathbf{N}_s + \int_0^t L(t-s) dP_s + \underbrace{\iint_0^t K(t-s,t-u) dP_s dP_u}_{\underbrace{\int_0^t \psi^i(t-s)\,(dP_s)^2}_{\text{effect of past volatility}} + \underbrace{\left[\int_0^t Z^i(t-s) dP_s\right]^2}_{\text{effect of past (unsigned) trends}}}$$

Zumbach effect: past trends, regardless of their signs lead to an increase in activity in the order book

What did we learn from the data?

Past events and past price changes **influence** future events
Past trends, regardless of their sign **increase** future volatility

The rate of *truly exogenous events* is gound to be **much smaller** than the total event rate ($\approx$ 10%). Suggests that the system is close to a critical point because stronger feedback kernels would lead to instabilities

Even worse, the quadratic feedback terms in K is the dominant effect in the Queue Hawkes equation!

$\Rightarrow$ **Can we build a tractable agent-based model that would give further insights?**

Santa Fe Model

The **Santa Fe model** is a purely stochastic agent-based model:

- Zero-intelligence agents place orders at random
- Collection of additive limit and market order arrival Poisson rates
- Constant cancellation rate
- Shown to capture certain empirical properties (i.e. mean bid-ask spread)

Model shortcomings:

- Fails to account for relation between spread and volatility
- Long-range correlation between market order flows is absent (and therefore prices are mean-reverting)

$\Rightarrow$ **Santa Fe model is useful for identifying qualitative characteristics of order book dynamics**

Santa Fe Model

Parameters:

- μ: market order rate (shares/time)
- α: limit order rate (shares/(price*time))
- δ: order cancellation rate (1/time)
- σ: characteristic order size (shares)
- dp: tick size (price)

All order flows are modeled by Poisson processes. Two Poisson processes are simulated: **(1)** $N_{MO}(t)$ with intensity μ for market orders, and **(2)** $N_{LO}(t)$ with intensity α for limit orders.

Market orders arrive in chunks of σ shares at rate μ, with even probability 1/2 of being ask or bid.

Limit orders arrive in chunks of σ shares at rate α, with even probability 1/2 of being ask or bid. Offers are placed with uniform probability at integer multiples of tick size dp within the range $(b(t), \infty)$ for asks and $(-\infty, a(t))$ for bids, where $a(t)$ and $b(t)$ are the best ask and bid prices at time t, respectively.

Santa Fe Model

Simulation process:

1. Initialize Poisson processes $N_{MO}(t)$ with intensity μ and $N_{LO}(t)$ with intensity α
2. Every time $N_{MO}(t)$ or $N_{LO}(t)$ hits, choose bid or ask with equal probability
3. **If market order bid:** remove one order at $a(t)$, then update $a(t)$
   If market order ask: remove one order at $b(t)$, then update $b(t)$
   If limit order bid or ask: choose price in acceptable range with uniform probability and either carry out transaction and update $a(t)/b(t)$, or keep limit order on the book if no transaction is possible
4. Cancel any existing limit orders with probability δ

Santa Fe Model with Feedback

We consider a modified Santa Fe model where the feedback of prior price changes on event rates is considered:

- Market orders can fall at either best bid or best ask with even probability $1/2$ and total rate 2μ
- Bid and ask limit orders fall uniformly and at the same rate λ, and only within one tick of their respective best prices
- Cancellations occur at rate ν_t per outstanding limit order, where ν_t is given by Q-Hawkes process with only the Zumbach term:

$$\nu_t = \nu_0 + \alpha_k \left(\int_0^t \sqrt{2\beta} e^{-\beta(t-s)} dP_s \right)^2$$

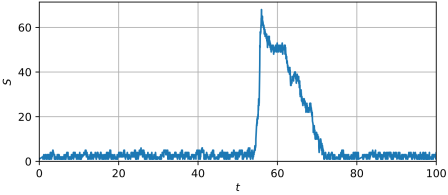
ν_t - cancellation rate
α_k - feedback intensity
β^{-1} - integration timescale
dP_s - price change at time s in tick units, from a simple martingale-based model

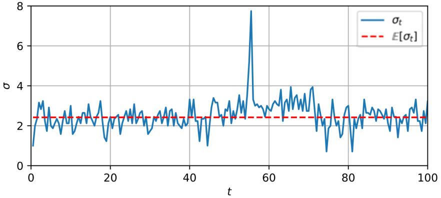
Note: $\alpha_k = 0$ specifies condition for original Santa Fe model

Trajectory Simulation

By defining initial conditions and initializing a price grid, we can then simulate a multidimensional, non-homogeneous Poisson process.

Typical results in a stable phase:



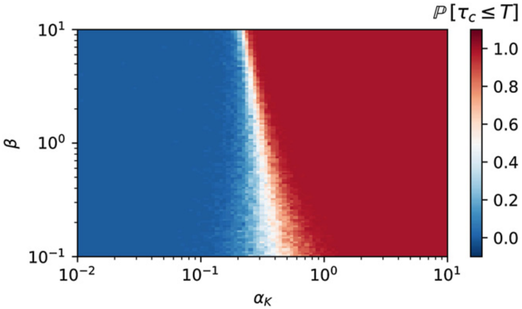


Stability Simulation

We can also use numerical simulation to observe liquidity crisis probability $P[\tau_c \leq T]$, or the probability that a crisis time τ_c occurs during the simulation.

This simulation reveals:

1. Large feedback intensities α_k yield unstable markets
2. Longer integration timescales β^{-1} yield more stable markets
3. Crossover threshold decreases as β^{-1} decreases



Phase Transition

The simulation suggests the existence of a phase transition, but we must confirm mathematically by analyzing the following double limit behavior:

$$\boxed{\lim_{N\to\infty}\lim_{T\to\infty} P[\tau_c \leq T]}$$

For finite N, there is a nonzero probability that the order book eventually empties

$$\lim_{N\to\infty}\lim_{T\to\infty} P[\tau_c \leq T, \alpha_k] = 1, \forall \alpha_k$$

For infinite N, the double limit behavior may depend on the model's parameters

$$\lim_{T\to\infty}\lim_{N\to\infty} P[\tau_c \leq T, \alpha_k] = \begin{cases} 1, & \text{for } \alpha_k > \alpha^* \\ 0, & \text{for } \alpha_k < \alpha^* \end{cases}$$

Numerical simulations can only consider finite N and T, so we turn to finite size scaling to recreate infinite sizes and waiting times

Finite Size Scaling

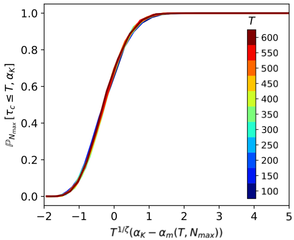
If a true phase transition exists at $\alpha_k = \alpha^*$, we expect the following to hold for sufficiently large N and T:

$$P_N[\tau_c \leq T, \alpha_k] = F(T(\alpha_k - \alpha_m(T,N))^\zeta)$$

where

$$\alpha_m(T,N) = \alpha^* - \frac{1}{T^{\frac{1}{\zeta}}} g(\frac{N^\eta}{T}),$$

$F(\cdot)$ is a regular monotonic function ranging from $F(-\infty) = 0$ to $F(\infty) = 1$, and $g(\cdot)$ ranges from $g(0) = \infty$ to $g(\infty) = g_\infty$ constant.



T vs. N^η

$$P_N[\tau_c \leq T, \alpha_k] = F(T(\alpha_k - \alpha_m(T,N))^\zeta);$$

$$\alpha_m(T,N) = \alpha^* - \frac{1}{T^{\frac{1}{\zeta}}} g(\frac{N^\eta}{T})$$

Observations:

1. For $1 \ll T \ll N^\eta$: $\alpha_m \approx \alpha^*$. Greater α_k pushes liquidity crisis probability from 0 to 1, transition region defined around α^* with width $T^{-\frac{1}{\zeta}}$
2. For $T \gg N^\eta$: α_m becomes negative. With large enough T, liquidity crisis probability is close to 1 regardless of parameters

Intuitions:

1. For much larger N^η, system never notices order book boundaries as the spread has not had enough time to grow $\Rightarrow$ liquidity crisis probability is mostly dependent on parameters
2. For much larger T, spread has likely spanned size of order book $\Rightarrow$ liquidity crisis has likely taken place regardless of parameters

Agent-Based Model Takeaways

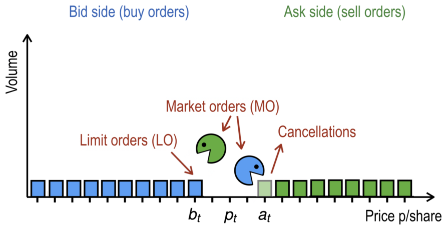
The results of the Santa Fe model are suggestive but not rigorous proof of the following

Critical value existence: There exists a critical value α^* of feedback parameter α_k such that for $\alpha_k < \alpha^*$ an infinitely large order book never empties, and for $\alpha_k > \alpha^*$ an infinitely large order book empties with probability 1.

Second-order nature:

1. With fixed parameters, rescaling is consistent and thus the indication of a second-order phase transition can be trusted
2. A power-law decay is observed near criticality

As Santa Fe model is very complex, introduce a simple, analytically tractable model that sets aside LOB dynamics and instead only focuses on the spread dynamics.



Assumptions are:

1. There is only one limit-order per price slot $\Rightarrow$ shape of LOB is fixed
2. Cancel orders only appear at the best ask or bid price
3. CO and MO are lumped into class of spread increasing events
4. LO can only be placed when the spread is open $S_t \geq 2$
5. There is no gap in the LOB except for the spread

$\Rightarrow$ **The LOB is entirely determined by the spread $\mathbf{S_t}$.**

State-dependent Hawkes model for the spread dynamics

Spread opening and spread closing events are both modeled by inhomogeneous Poisson processes N_t^+, N_t^- with intensity λ_t^+, λ_t^-, respectively:

$$\lambda_t^+ = \lambda_0^+ + (\phi * dS^+)_t = \lambda_0^+ + \int_0^t \phi(t-s) dS_s^+ \quad ,$$

$$\lambda_t^- = \lambda_0^- \, \mathbb{1}(S_t \geq 2) \quad ,$$

$$S_t = S_0 + N_t^+ - N_t^- \quad ,$$

where S_t is the spread at time t and $dS_t^+ = \max(0, dS_t) = dN_t^+$. The Hawkes kernel is given by

$$\phi_t = \alpha\beta e^{-\beta t}.$$

Stability criteria can be derived by finding the expectation of the spread at any given time.

Doob decomposition of the spread

The spread dynamics are given by:

$$
\begin{aligned}
S_t &= S_0 + N_t^+ - N_t^- \\
&= S_0 + \int_0^t \lambda_s^+ ds + M_t^+ - \int_0^t \lambda_s^- ds - M_t^- \\
&= S_0 + \int_0^t \left[\lambda_s^+ - \lambda_0^- \mathbb{1}(S_s \geq 2)\right] ds + M_t^+ - M_t^- \\
&= S_0 + \int_0^t \left[\lambda_0^+ + (\phi * dS^+)_s - \lambda_0^- \mathbb{1}(S_s \geq 2)\right] ds + M_t^+ - M_t^-
\end{aligned}
$$

$\Rightarrow$ Doob decomposition of λ_t^+ is needed, otherwise the spread itself will come into play again.

Doob decomposition of the intensity

$$\begin{aligned} \lambda_t^+ &= \lambda_0^+ + \left(\phi * dS^+\right)_t \\ dS_t^+ &= dN_t^+ = \lambda_t^+ dt + dM_t^+ \end{aligned} \quad \Rightarrow \quad \underbrace{\lambda_t^+ = \lambda_0^+ + \left(\phi * \lambda^+\right)_t}_{\textit{Fredholm integral equation}} + \left(\phi * dM^+\right)_t$$

Idea: Turn convolution into multiplication by applying Laplace transform $\mathcal{L}$.

$$\begin{aligned} \mathcal{L}\lambda_t^+ &= \mathcal{L}\lambda_0^+ + \mathcal{L}\phi_t \mathcal{L}\lambda_t^+ \\ \Leftrightarrow \mathcal{L}\lambda_t^+ &= \frac{\mathcal{L}\lambda_0^+}{1 - \mathcal{L}\phi_t} \\ \Leftrightarrow \lambda_t^+ &= \mathcal{L}^{-1}\left\{\frac{\mathcal{L}\lambda_0^+}{1 - \mathcal{L}\phi_t}\right\} \end{aligned}$$

Calculating the Laplace and inverse Laplace transforms leads to:

$$\lambda_t^+ = \frac{\lambda_0^+}{1-\alpha}\left(1 - \alpha e^{-(1-\alpha)\beta t}\right) + \left(\phi * dM^+\right)_t$$

Doob decomposition of the spread

$$\boxed{S_t = S_0 + \int_0^t \left[\left(1 - \alpha e^{-(1-\alpha)\beta s}\right) \frac{\lambda_0^+}{1-\alpha} - \lambda_0^- \mathbb{1}(S_s \geq 2) \right] ds + M_t}$$

This decomposition of the spread dynamics allows us to look at the expectation of the spread at any given time. Setting the indicator function to 1 for the stability analysis ($\approx \widetilde{S}_t$) still yields exact stability bounds:

$$\mathbb{E}(\widetilde{S}_t) = \underbrace{S_0 - \frac{\lambda_0^+ \alpha}{(1-\alpha)^2 \beta}}_{\text{constant term}} + \underbrace{\left(\frac{\lambda_0^+}{1-\alpha} - \lambda_0^- \right) t}_{\text{linear growth}} + \underbrace{\frac{\lambda_0^+ \alpha}{(1-\alpha)^2 \beta} e^{-(1-\alpha)\beta t}}_{\text{exponential growth}}$$

$\Rightarrow$ The model has three different stability regimes, depending on which term dominates.

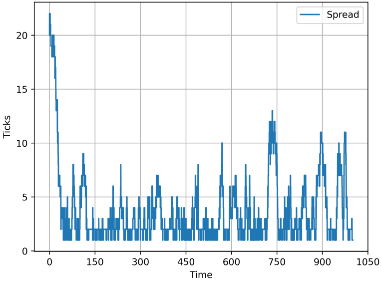
Stable regime: $0 \leq \alpha \leq 1 - \frac{\lambda_0^+}{\lambda_0^-} =: \alpha_c < 1$

In this regime the spread is stable, i.e. a stationary distribution exists, and one can show that:

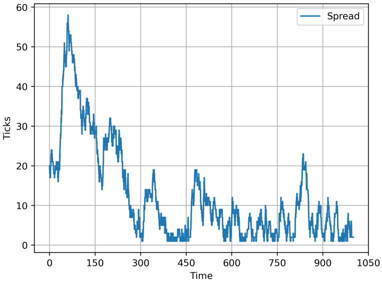
$$\mathbb{P}(S_t = 1) = \frac{\alpha_c - \alpha}{1 - \alpha}$$

$\Rightarrow$ **No liquidity crisis**

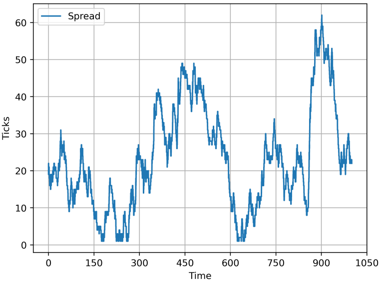
Realization of the spread from a simulation for $\alpha = 0.2$, $\alpha_c = 0.5$, S_0=20:



Realization of the spread from a simulation for $\alpha = 0.4$, $\alpha_c = 0.5$, S_0=20:



Realization of the spread from a simulation for $\alpha = 0.5$, $\alpha_c = 0.5$, S_0=20:



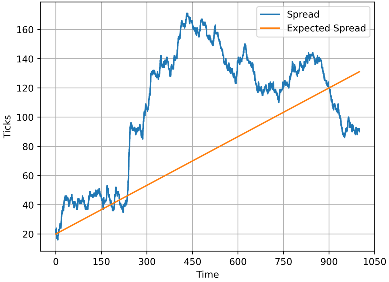
Linear spread growth: $\alpha_c < \alpha < 1$

Here, the spread is linearly unstable and grows with a constant rate:

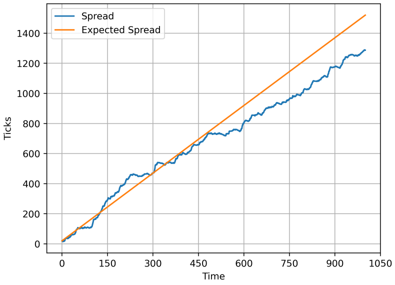
$$\lim_{t\to\infty} \frac{1}{t}\mathbb{E}(S_t) = \frac{\lambda_0^+}{1-\alpha} - \lambda_0^-$$

This is easy to see from the Doob decomposition of the spread.

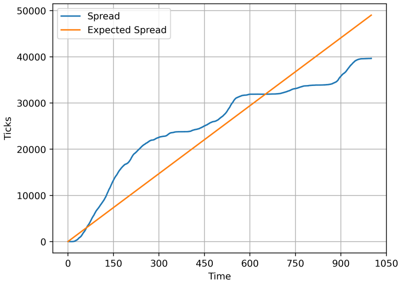
Realization of the spread from a simulation for $\alpha = 0.55$, $\alpha_c = 0.5$, S_0=20:



Realization of the spread from a simulation for $\alpha = 0.8$, $\alpha_c = 0.5$, S_0=20:



Realization of the spread from a simulation for $\alpha = 0.99$, $\alpha_c = 0.5$, S_0=20:



Calculating the diffusion constant of the spread allows approximating the spread as Brownian motion with drift:

$$V = \lim_{t\to\infty} \frac{1}{t}\mathbb{E}[S_t] = \frac{\lambda_0^+}{1-\alpha} - \lambda_0^-, \quad D = \lim_{t\to\infty} \frac{1}{t}\left(\mathbb{E}[S_t^2] - \mathbb{E}[S_t]^2\right) = \frac{\lambda_0^+}{(1-\alpha)^3} + \lambda_0^-,$$

$$\Rightarrow dS_t = V\,dt + D\,dB_t$$

which gives us the first-passage time result:

$$\mathbb{P}_N\left[\tau_c \leq T, \alpha\right] = \int_0^T \frac{N}{\sqrt{2\pi D s^3}} e^{-\frac{(N+Vs)^2}{2Ds}} ds \quad ,$$

for a threshold N of the spread. It is visible now that:

$$\lim_{T\to\infty} \lim_{N\to\infty} P[\tau_c \leq T, \alpha] = 0.$$

Hence, there is no phase transition as in the Santa Fe model and accordingly **no real liquidity crisis** as the order book never empties in finite time.

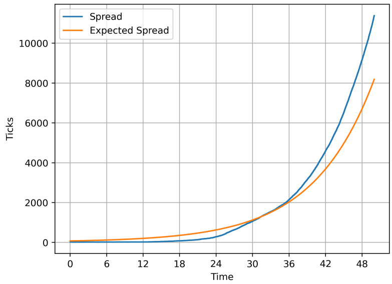
Explosive regime: $\alpha > 1$

For , the exponential term dominates the drift of the spread:

$$S_t \propto e^{(\alpha-1)t}$$

$\Rightarrow$ **A liquidity crisis occurs**, when $T(\alpha - 1) \propto \log N$, where N is the depth of the LOB.

Realization of the spread from a simulation for $\alpha = 1.1$, $\alpha_c = 0.5$, S_0=20:



Conclusion - Endogenous liquidity crises

Using tick-by-tick order book data (on Google & Apple stock), we were able to :

- show that event rates are strongly affected by past price moves
- understand that large price trend and/or volatility tends to increase the rate of market orders and cancellations

$\rightarrow$ subsequently leads to a decrease in liquidity and therefore contributes to increasing volatility, which may lead to a destabilising feedback loop and a liquidity dry-out.