Data-Driven Models & Mathematical Finance: Opposition or Apposition?

DPhil in Machine Learning Viva Voce

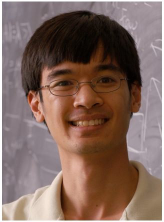
Babak Mahdavi-Damghani[1]

University of Oxford

[1] *babak.mahdavidamghani@oxford-man.ox.ac.uk*

September 19th, 2019

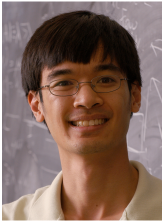
Successful Research Strategies

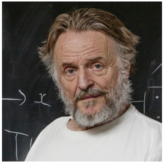


Research Interest vs Publication

- Some dedicate their whole lives to **one complicated problem** with a rare (but often nothing) outstanding outcome (Grigori Perelman).
- Some other prefer adhering to the "**publish or perish**" model at the cost of not producing the same quality research.
- **Terence Tao** suggests to have one big problem to go back to when inspired but adhere to (or at least do not neglect) the "publish or perish" model most of the time.

Successful Research Strategies





Research Interest vs Publication

- Some dedicate their whole lives to **one complicated problem** with a rare (but often nothing) outstanding outcome (Grigori Perelman).
- Some other prefer adhering to the "**publish or perish**" model at the cost of not producing the same quality research.
- **Terence Tao** suggests to have one big problem to go back to when inspired but adhere to (or at least do not neglect) the "publish or perish" model most of the time.

Research Complexity vs Recognition

- You can dedicate lots of energy on **problems you find personally stimulating** but nobody cares about.
- You can dedicate little energy on problems you do not find personally very stimulating but **others** may find useful.
- **John Conway**'s Game of Life happens to be the latter case (surprisingly).

Successful Research Strategies

Research Interest vs Publication

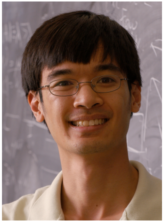
- Some dedicate their whole lives to **one complicated problem** with a rare (but often nothing) outstanding outcome (Grigori Perelman).
- Some other prefer adhering to the "**publish or perish**" model at the cost of not producing the same quality research.
- **Terence Tao** suggests to have one big problem to go back to when inspired but adhere to (or at least do not neglect) the "publish or perish" model most of the time.

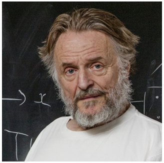
Research Complexity vs Recognition

- You can dedicate lots of energy on **problems you find personally stimulating** but nobody cares about.
- You can dedicate little energy on problems you do not find personally very stimulating but **others** may find useful.
- **John Conway**'s Game of Life happens to be the latter case (surprisingly).

Theory vs Simulation

- **A good theory** should be able to be **simulated**.
- A good **simulation may change/iron out a theory**.
- **Cedric Villani** thinks that the process can go back and forth until the picture becomes clearer.



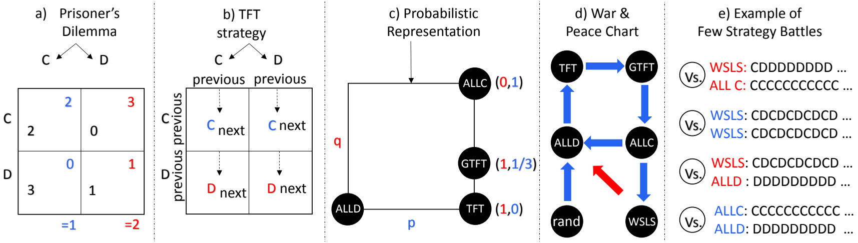


Memorable Events Influencing Research

- How does **Morality** Emerge in human behaviour → Axelrod [3, 4]:

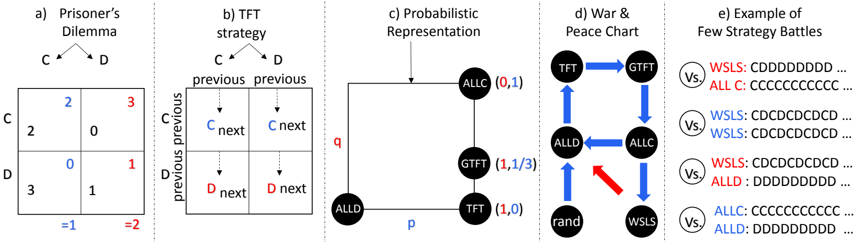
Memorable Events Influencing Research

- How does **Morality** Emerge in human behaviour → Axelrod [3, 4]:



Memorable Events Influencing Research

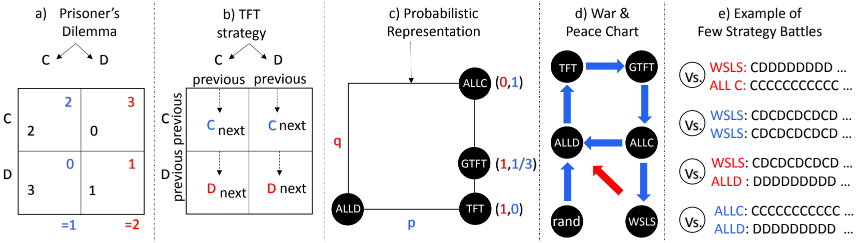
- How does **Morality** Emerge in human behaviour → Axelrod [3, 4]:



- **Shocking work experience** and the necessity to exude that frustration with papers and vulgarization

Memorable Events Influencing Research

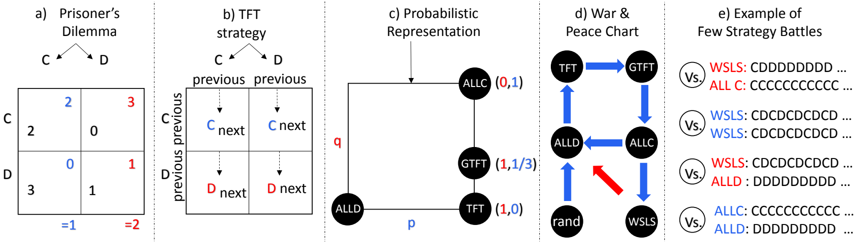
- How does **Morality** Emerge in human behaviour → Axelrod [3, 4]:



- **Shocking work experience** and the necessity to exude that frustration with papers and vulgarization
  - **Lack of Integrity** (Apophenia, Anomaly Detection, Misleading Statement and Actions)

Memorable Events Influencing Research

- How does **Morality** Emerge in human behaviour → Axelrod [3, 4]:

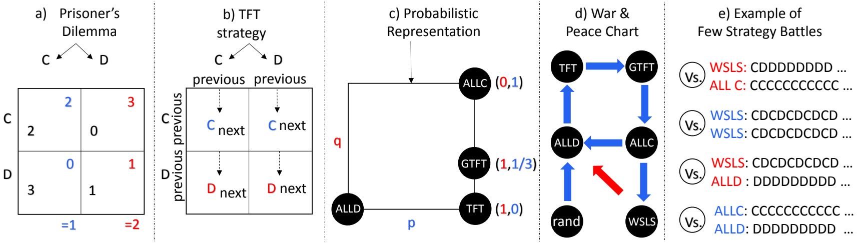


- **Shocking work experience** and the necessity to exude that frustration with papers and vulgarization
  - **Lack of Integrity** (Apophenia, Anomaly Detection, Misleading Statement and Actions)
  - **Models with incredible rigour when unnecessary** but with total **disregard to rigour when it becomes critical** (e.g: systemic risk associated to covariance).

Memorable Events Influencing Research

Part I:

- How does **Morality** Emerge in human behaviour → Axelrod [3, 4]:



Part II:

- **Shocking work experience** and the necessity to exude that frustration with papers and vulgarization
  - **Lack of Integrity** (Apophenia, Anomaly Detection, Misleading Statement and Actions)
  - **Models with incredible rigour when unnecessary** but with total **disregard to rigour when it becomes critical** (e.g: systemic risk associated to covariance).

Context

Market Changing Financial Crisis: 10 year anniversary

Context

Market Changing Financial Crisis: 10 year anniversary

- The **Subprime crisis triggers malaise** [15, 20, 28, 21]: This encourages **Socially Responsible Finance**.

Context

Market Changing Financial Crisis: 10 year anniversary

- The **Subprime crisis triggers malaise** [15, 20, 28, 21]: This encourages **Socially Responsible Finance**.
- **Quantitative Finance was instructed to change** [17]: From **models assuming data to data reassuming the models**,
- There is a wave associated to **tying Classic Financial Mathematics** models towards more **Data Driven** problems: rise of Machine Learning.

Context

Market Changing Financial Crisis: 10 year anniversary

- The **Subprime crisis triggers malaise** [15, 20, 28, 21]: This encourages **Socially Responsible Finance**.
- **Quantitative Finance was instructed to change** [17]: From **models assuming data to data reassuming the models**,
- There is a wave associated to **tying Classic Financial Mathematics** models towards more **Data Driven** problems: rise of Machine Learning.

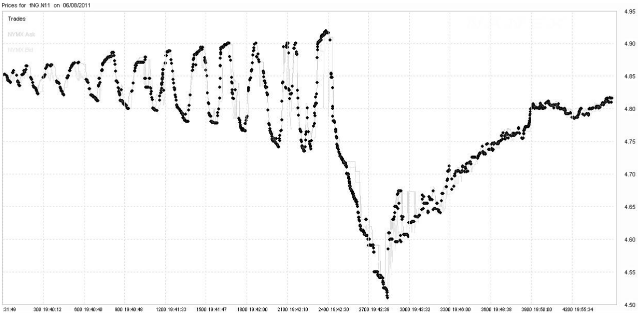
The Rise of Big Data

- **What is Big Data**: lots of anecdotal claims about how big is Big Data [9, 2, 16] but the term refers more to the concept of “**datafication**” (increase in size $\neq$ better confidence interval but rather change in perspective).

Context

Market Changing Financial Crisis: 10 year anniversary

- The **Subprime crisis triggers malaise** [15, 20, 28, 21]: This encourages **Socially Responsible Finance**.
- **Quantitative Finance was instructed to change** [17]: From **models assuming data to data reassuming the models**,
- There is a wave associated to **tying Classic Financial Mathematics** models towards more **Data Driven** problems: rise of Machine Learning.



The Rise of Big Data

- **What is Big Data**: lots of anecdotal claims about how big is Big Data [9, 2, 16] but the term refers more to the concept of “**datafication**” (increase in size $\neq$ better confidence interval but rather change in perspective).
- The **Flash Crashes** (eg: [29]) calls for a **modelling revolution** [5, 12, 6] (**BU vs. TD**): the Brownian motion assumption to model markets is increasingly difficult to defend.

1. Introduction
   - Foreword
   - Historical Context
   - Original Contribution
2. A Bottom-up Approach to the Financial Markets
   - Useful material for Quant Finance from Mathematical Biology
   - Neural Network Architecture and Learning Potential
   - Dynamic of the Financial Market
   - Stability of Financial Systems and Multi-Target Tracking
3. Model Assuming Data vs Data Reassuming the Models
   - Cointelation, Inferred Correlation & Portfolio Optimization
   - Anomaly Detection & Volatility Surface de-Arbitraging
   - Big Data Changing the Vanilla Options Landscape
   - Clustering for Distribution & Regime Change Forecasting
4. Conclusion
   - A Bottom-Up Approach to the Financial Markets
   - Model Assuming Data vs Data Reassuming the Models

Original Contribution: Part I (Opposition)

- Exposing the benefits of bringing **Mathematical Biology and Evolutionary Dynamics into 21st century QF**.

Original Contribution: Part I (Opposition)

- Exposing the benefits of bringing **Mathematical Biology and Evolutionary Dynamics into 21st century QF**.
- Turning classic **QF strategies into NN format** in order to prepare for our strategy tournament,

Original Contribution: Part I (Opposition)

- Exposing the benefits of bringing **Mathematical Biology and Evolutionary Dynamics into 21st century QF**.
- Turning classic **QF strategies into NN format** in order to prepare for our strategy tournament,
- Formalization of the **HFTE rough tournament** and exposing why this is **not a good approach**,

Original Contribution: Part I (Opposition)

- Exposing the benefits of bringing **Mathematical Biology and Evolutionary Dynamics into 21st century QF**.
- Turning classic **QF strategies into NN format** in order to prepare for our strategy tournament,
- Formalization of the **HFTE rough tournament** and exposing why this is **not a good approach**,
- Formalizing the **path of interaction** concept in the context of an **HFTE Game**,

Original Contribution: Part I (Opposition)

- Exposing the benefits of bringing **Mathematical Biology and Evolutionary Dynamics into 21st century QF**.
- Turning classic **QF strategies into NN format** in order to prepare for our strategy tournament,
- Formalization of the **HFTE rough tournament** and exposing why this is **not a good approach**,
- Formalizing the **path of interaction** concept in the context of an **HFTE Game**,
- Performing a **computer tournament with a set of simple fixed classic strategies** as well as proposing some interesting **connection to** other seemingly unconnected concepts (**Morality and Complexity/Diversity** debate),

Original Contribution: Part I (Opposition)

- Exposing the benefits of bringing **Mathematical Biology and Evolutionary Dynamics into 21st century QF**.
- Turning classic **QF strategies into NN format** in order to prepare for our strategy tournament,
- Formalization of the **HFTE rough tournament** and exposing why this is **not a good approach**,
- Formalizing the **path of interaction** concept in the context of an **HFTE Game**,
- Performing a **computer tournament with a set of simple fixed classic strategies** as well as proposing some interesting **connection to** other seemingly unconnected concepts (**Morality and Complexity/Diversity** debate),
- Formalizing the methodology for a **Particle Filter** which aim is to **track ecosystems of strategies** through time by looking at price dynamics only as well as performing few simulations.

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k\sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k \sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k \sigma^2(K, T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),
- Addressing these problems by **constructing the IVP model**,

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k \sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),
- Addressing these problems by **constructing the IVP model**,
- Using **proxying and de-abitraging** (on the **Mid** or leveraging on **Liquidity**) in order to reconstruct implied volatility scenarios,

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k\sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),
- Addressing these problems by **constructing the IVP model**,
- Using **proxying and de-abitraging** (on the **Mid** or leveraging on **Liquidity**) in order to reconstruct implied volatility scenarios,
- Exposing the **Financial Mathematics limiting assumptions** in the context of Markowitz especially when **contrasted with real data**,

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k\sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),
- Addressing these problems by **constructing the IVP model**,
- Using **proxying and de-abitraging** (on the **Mid** or leveraging on **Liquidity**) in order to reconstruct implied volatility scenarios,
- Exposing the **Financial Mathematics limiting assumptions** in the context of Markowitz especially when **contrasted with real data**,
- **Cointelation** model construction and it's mirror **Inferred Correlation** approximation for **data driven applications** and **Proper Market Conduct**,

Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k\sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),
- Addressing these problems by **constructing the IVP model**,
- Using **proxying and de-abitraging** (on the **Mid** or leveraging on **Liquidity**) in order to reconstruct implied volatility scenarios,
- Exposing the **Financial Mathematics limiting assumptions** in the context of Markowitz especially when **contrasted with real data**,
- **Cointelation** model construction and it's mirror **Inferred Correlation** approximation for **data driven applications** and **Proper Market Conduct**,
- Its **Portfolio Optimization** using several the **Bandwise-Gaussian mixture** model as well as **Deep Learning**,

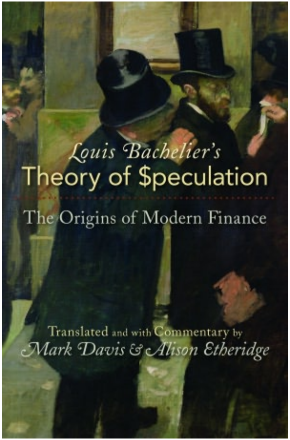
Original Contribution: Part II (Apposition)

- **Detective work**: exposing where $|T\partial_k\sigma^2(K,T)| < 4$ came from and why it is a necessary but **not sufficient condition** because of **wrong data assumptions** (moment explosion vs linearity),
- Exposing the **limitations of the SVI** model (**arbitrage constraint, wings & liquidity**),
- Addressing these problems by **constructing the IVP model**,
- Using **proxying and de-abitraging** (on the **Mid** or leveraging on **Liquidity**) in order to reconstruct implied volatility scenarios,
- Exposing the **Financial Mathematics limiting assumptions** in the context of Markowitz especially when **contrasted with real data**,
- **Cointelation** model construction and it's mirror **Inferred Correlation** approximation for **data driven applications** and **Proper Market Conduct**,
- Its **Portfolio Optimization** using several the **Bandwise-Gaussian mixture** model as well as **Deep Learning**,
- Reconciling discordant Risk: **Anticipative Responsible** VaR.

Two simple questions for two simple definitions

Definition (Top-Down Vs Bottom-Up)

Top-Down (TD): Any **stochastic** quantitative approach which assumes that the **market is random** (or close to random) but for which we can create **interesting dynamic strategies**.

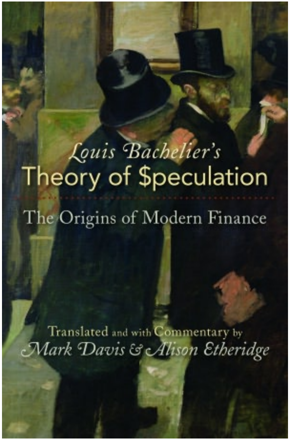


Two simple questions for two simple definitions

Definition (Top-Down Vs Bottom-Up)

Top-Down (TD): Any **stochastic** quantitative approach which assumes that the **market is random** (or close to random) but for which we can create **interesting dynamic strategies**.

Bottom-Up (BU): Any **deterministic** quantitative approach which assumes that **strategies interacting** create the **dynamics of the market** (*new angle*).



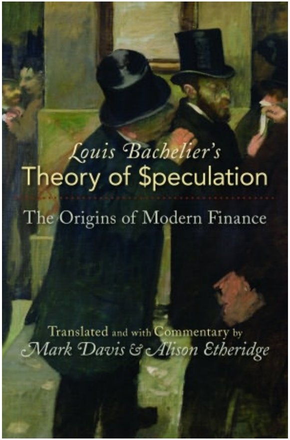
Two simple questions for two simple definitions

We have tried to understand the financial markets using the top down approach for **100 years with (very) partial success**. Can we look at it in a **different angle**?

Definition (Top-Down Vs Bottom-Up)

Top-Down (TD): Any **stochastic** quantitative approach which assumes that the **market is random** (or close to random) but for which we can create **interesting dynamic strategies**.

Bottom-Up (BU): Any **deterministic** quantitative approach which assumes that **strategies interacting** create the **dynamics of the market** (*new angle*).



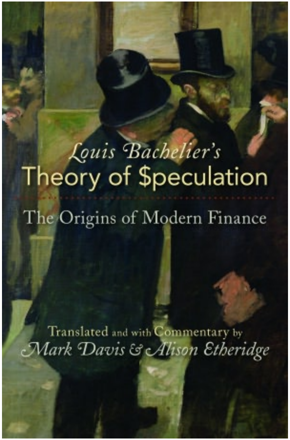
Two simple questions for two simple definitions

We have tried to understand the financial markets using the top down approach for **100 years with (very) partial success**. Can we look at it in a **different angle**?

Definition (Top-Down Vs Bottom-Up)

Top-Down (TD): Any **stochastic** quantitative approach which assumes that the **market is random** (or close to random) but for which we can create **interesting dynamic strategies**.

Bottom-Up (BU): Any **deterministic** quantitative approach which assumes that **strategies interacting** create the **dynamics of the market** (*new angle*).

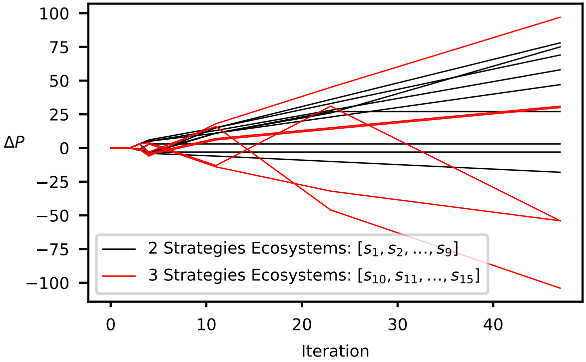


Can we use this new angle to **"solve"** the market?

The Scientific Method for "solving" the market

Question: What do we mean by "solve"?

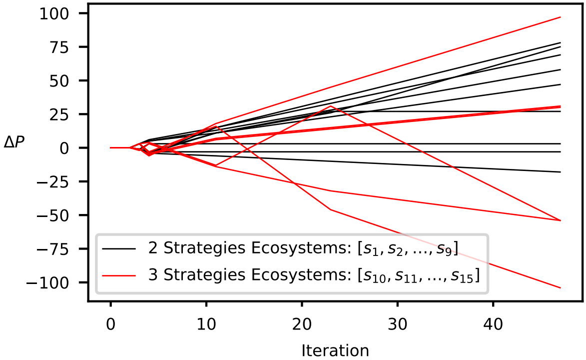
The Scientific Method for "solving" the market



Question: What do we mean by "solve"?

Answer: We want to look at market price processes and be able to tell what systematic strategies were involved, what their P&L has been and this at all times (including in the future): **ecosystem details**.

The Scientific Method for "solving" the market

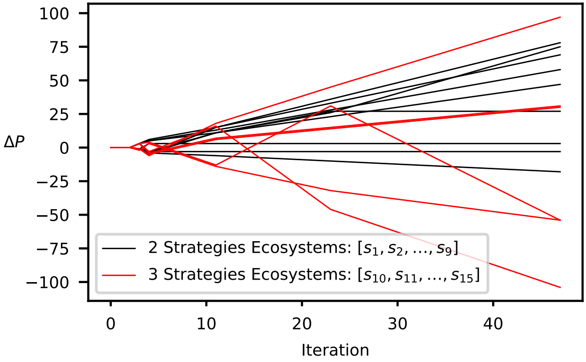


Question: What do we mean by "solve"?

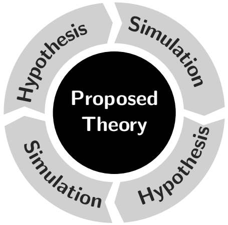
Answer: We want to look at market price processes and be able to tell what systematic strategies were involved, what their P&L has been and this at all times (including in the future): **ecosystem details**.

Caveat: Is the idea mad, ambitious or both?

The Scientific Method for "solving" the market



Scientific Method: A good theory can be simulated but simulations can also help bring intuition on what the theory might be [34].



Question: What do we mean by "solve"?

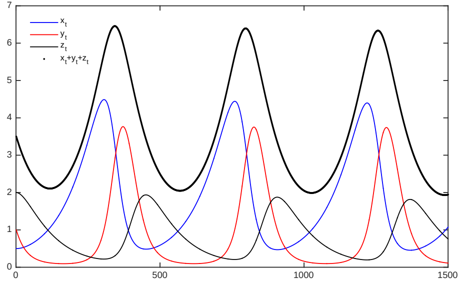
Answer: We want to look at market price processes and be able to tell what systematic strategies were involved, what their P&L has been and this at all times (including in the future): **ecosystem details**.

Caveat: Is the idea mad, ambitious or both?

Predator/Prey models

biological ecosystems **predator/prey** (PP) [35, 8] models (a, b, c, d, e, f and g are rate of growth or predation). The relationship between $x(t)$, $y(t)$ and $z(t)$ is **deterministic**:

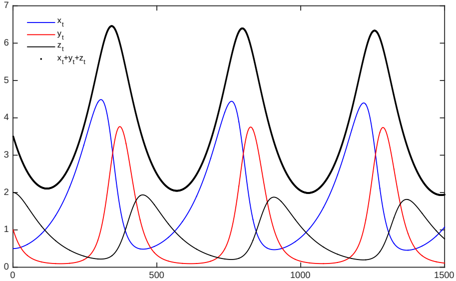
$$\begin{cases} \frac{dx(t)}{dt} & = ax(t) - bx(t)y(t) \\ \frac{dy(t)}{dt} & = -cy(t) + dx(t)y(t) - ey(t)z(t) \\ \frac{dz(t)}{dt} & = -fz(t) + gy(t)z(t) \end{cases}$$



Predator/Prey models

We can make the **hypotheses that the economical cycles** or oscillation in prices are due to the same type of disruptions that can occur in biological ecosystems **predator/prey** (PP) [35, 8] models (a, b, c, d, e, f and g are rate of growth or predation). The relationship between $x(t)$, $y(t)$ and $z(t)$ is **deterministic**:

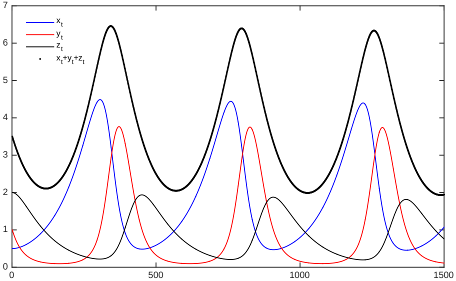
$$\begin{cases} \frac{dx(t)}{dt} & = ax(t) - bx(t)y(t) \\ \frac{dy(t)}{dt} & = -cy(t) + dx(t)y(t) - ey(t)z(t) \\ \frac{dz(t)}{dt} & = -fz(t) + gy(t)z(t) \end{cases}$$



Predator/Prey models

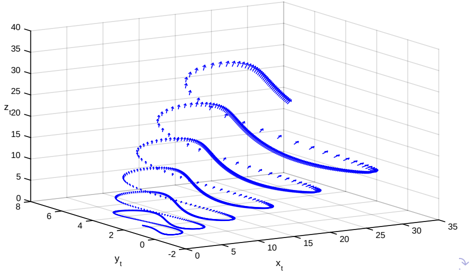
We can make the **hypotheses that the economical cycles** or oscillation in prices are due to the same type of disruptions that can occur in biological ecosystems **predator/prey** (PP) [35, 8] models (a, b, c, d, e, f and g are rate of growth or predation). The relationship between $x(t)$, $y(t)$ and $z(t)$ is **deterministic**:

$$\begin{cases} \frac{dx(t)}{dt} &= ax(t) - bx(t)y(t) \\ \frac{dy(t)}{dt} &= -cy(t) + dx(t)y(t) - ey(t)z(t) \\ \frac{dz(t)}{dt} &= -fz(t) + gy(t)z(t) \end{cases}$$



Answering if an **ecosystem** (or by extension financial market) composed of 3 strategies is **stable** would come to studying the Jacobian matrix J [8]. If if all **eigenvalues** of $J(x, y, z)$ have negative real parts then our system is asymptotically **stable**. Though simplistic, the model can easily be expanded to more **complex ecological niches**.

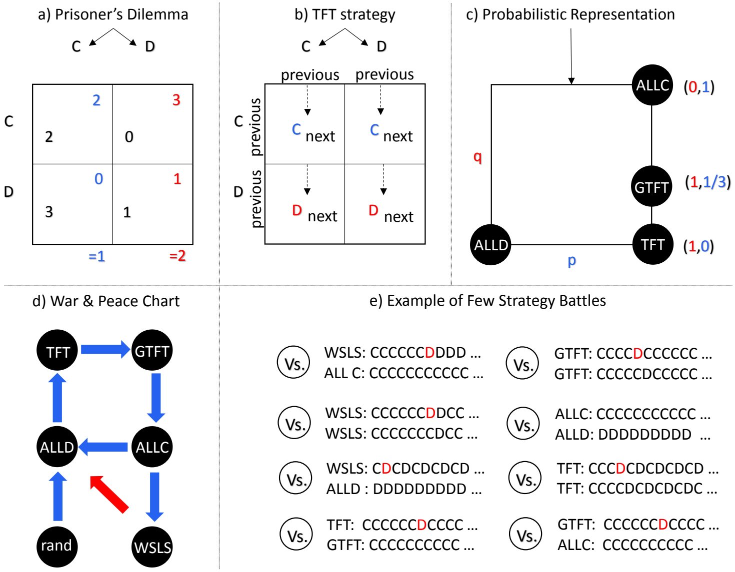
$$J(x, y, z) = \begin{bmatrix} a - by & -xb & 0 \\ yd & -c + dx - ez & -ye \\ 0 & -zg & -f + gy \end{bmatrix}$$



Evolutionary Dynamics

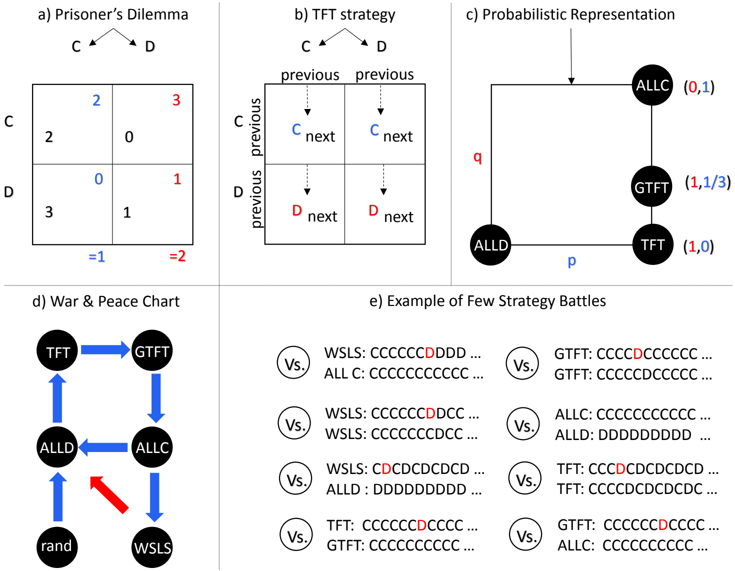
Axelrod's [3, 4] **computer tournament** and **Nowak**'s work on Evolutionary Dynamics (ED) [30] on **invasion** need to influence 21st century Quantitative Finance especially the BU approach.

Evolutionary Dynamics



Axelrod's [3, 4] **computer tournament** and **Nowak**'s work on Evolutionary Dynamics (ED) [30] on **invasion** need to influence 21st century Quantitative Finance especially the BU approach.

Evolutionary Dynamics



Axelrod's [3, 4] **computer tournament** and **Nowak**'s work on Evolutionary Dynamics (ED) [30] on **invasion** need to influence 21st century Quantitative Finance especially the BU approach.

The methodology in ED is interesting because the **strategies** are both **systematic & interacting** with each other (like it is the case in algo trading).

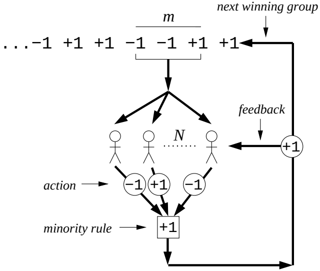
A first application in Economics: Minority Game

In the **Minority Game** [7], developped by Challet, Marsili and Zhang, players need to choose between two options $(+1, -1)$. Those who have selected the option chosen by the minority “win”.

A first application in Economics: Minority Game

In the **Minority Game** [7], developped by Challet, Marsili and Zhang, players need to choose between two options $(+1, -1)$. Those who have selected the option chosen by the minority "win".

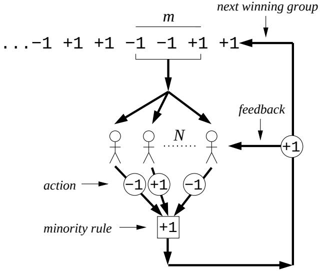
input	output
−1 −1 −1	−1
−1 −1 +1	−1
−1 +1 −1	+1
−1 +1 +1	−1
+1 −1 −1	−1
+1 −1 +1	+1
+1 +1 −1	−1
+1 +1 +1	+1

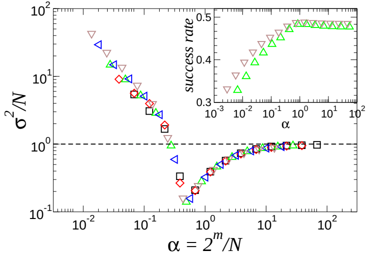


A first application in Economics: Minority Game

In the **Minority Game** [7], developped by Challet, Marsili and Zhang, players need to choose between two options $(+1, -1)$. Those who have selected the option chosen by the minority "win".

input	output
−1 −1 −1	−1
−1 −1 +1	−1
−1 +1 −1	+1
−1 +1 +1	−1
+1 −1 −1	−1
+1 −1 +1	+1
+1 +1 −1	−1
+1 +1 +1	+1



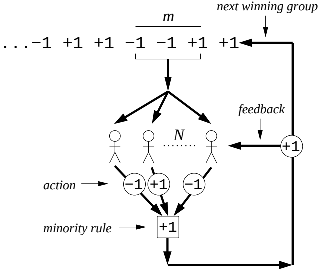


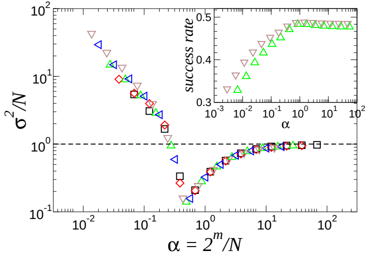
Physical Laws can emerge from these simple rules. We can observe that σ^2/N is only a function of $\alpha = 2m/N$ which considering the complexity of the interactions between the set of agents can be quite remarkable.

A first application in Economics: Minority Game

In the **Minority Game** [7], developped by Challet, Marsili and Zhang, players need to choose between two options $(+1, -1)$. Those who have selected the option chosen by the minority "win".

input	output
−1 −1 −1	−1
−1 −1 +1	−1
−1 +1 −1	+1
−1 +1 +1	−1
+1 −1 −1	−1
+1 −1 +1	+1
+1 +1 −1	−1
+1 +1 +1	+1





Criticism: Is this realistic for Economics? Maybe, but not for algorithmic trading (eg: TF strategy in a TF concentrated ecosystem)?

Physical Laws can emerge from these simple rules. We can observe that σ^2/N is only a function of $\alpha = 2m/N$ which considering the complexity of the interactions between the set of agents can be quite remarkable.

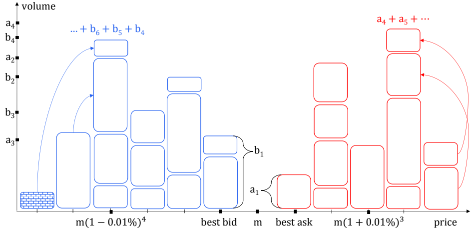
Bringing these Ideas into Electronic Trading

- As ML "translation" of ED and PP models, we have **Generative Adversarial Networks** (GANs) [14], introduced in 2014, usually **involve** a system of **two neural networks** competing in a zero-sum game settings. This process continues as long as needed since the lack of **data is no longer a problem**.

Bringing these Ideas into Electronic Trading

- As ML "translation" of ED and PP models, we have **Generative Adversarial Networks** (GANs) [14], introduced in 2014, usually **involve** a system of **two neural networks** competing in a zero-sum game settings. This process continues as long as needed since the lack of **data is no longer a problem**.

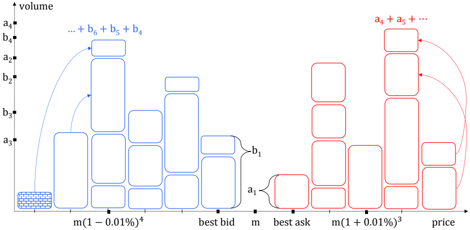
- The **communication tool** is the **order book** and the game is **not zero-sum game** in our research.

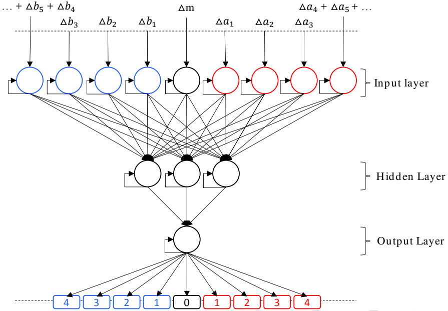


Bringing these Ideas into Electronic Trading

- As ML "translation" of ED and PP models, we have **Generative Adversarial Networks** (GANs) [14], introduced in 2014, usually **involve** a system of **two neural networks** competing in a zero-sum game settings. This process continues as long as needed since the lack of **data is no longer a problem**.

- The **communication tool** is the **order book** and the game is **not zero-sum game** in our research. The strategies are in **High Frequency Financial Funnel** (HFFF) format [23].

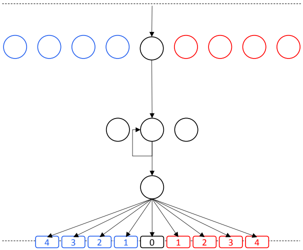




High Frequency Financial Funnel & Classic Strategies

HFFF can model financial strategies:

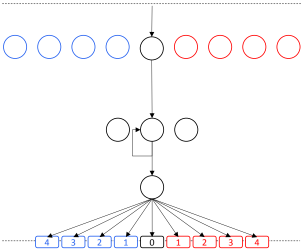
High Frequency Financial Funnel & Classic Strategies

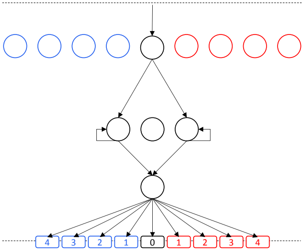


HFFF can model financial strategies:

1. Trend Following (TF)

High Frequency Financial Funnel & Classic Strategies

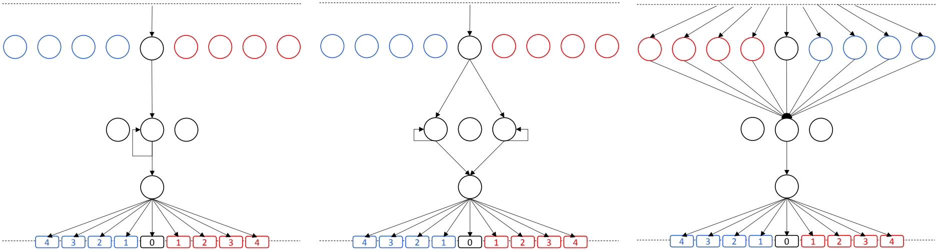




HFFF can model financial strategies:

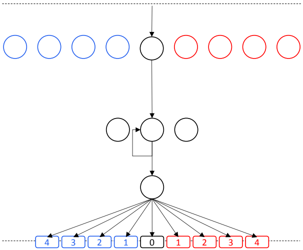
1. Trend Following (TF)
2. MACD

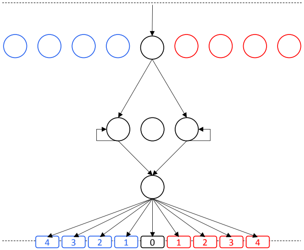
High Frequency Financial Funnel & Classic Strategies

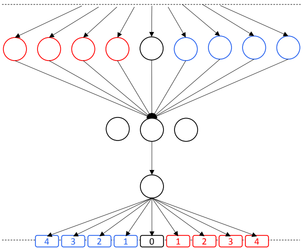


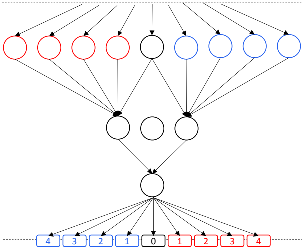
HFFF can model financial strategies:

1. Trend Following (TF)
2. MACD
3. MLR





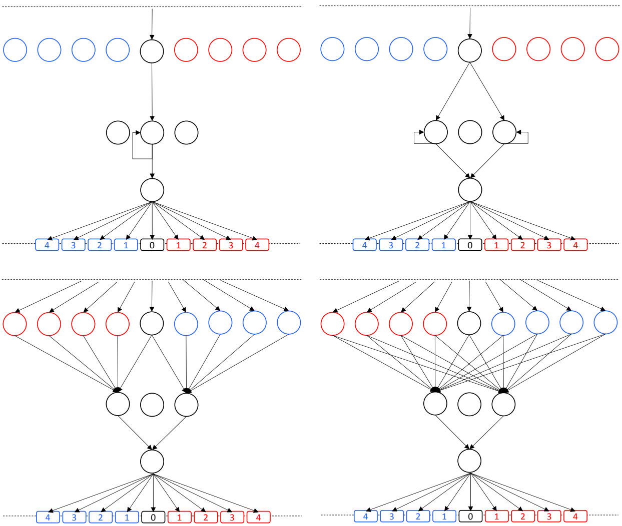




HFFF can model financial strategies:

1. Trend Following (TF)
2. MACD
3. MLR
4. XOR

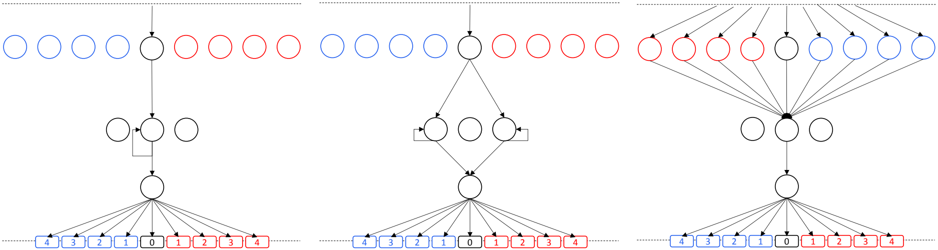
High Frequency Financial Funnel & Classic Strategies

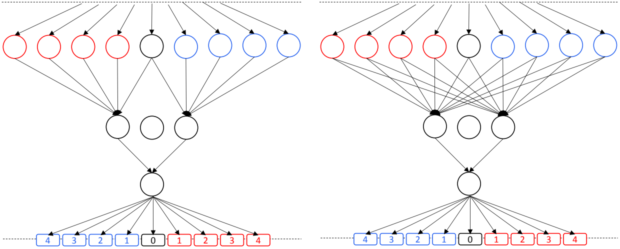


HFFF can model financial strategies:

1. Trend Following (TF)
2. MACD
3. MLR
4. XOR
5. Lasso Regressions etc ...

High Frequency Financial Funnel & Classic Strategies



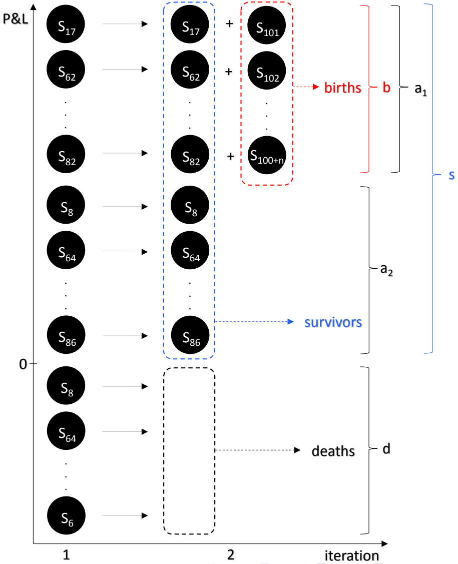


HFFF can model financial strategies:

1. Trend Following (TF)
2. MACD
3. MLR
4. XOR
5. Lasso Regressions etc ...

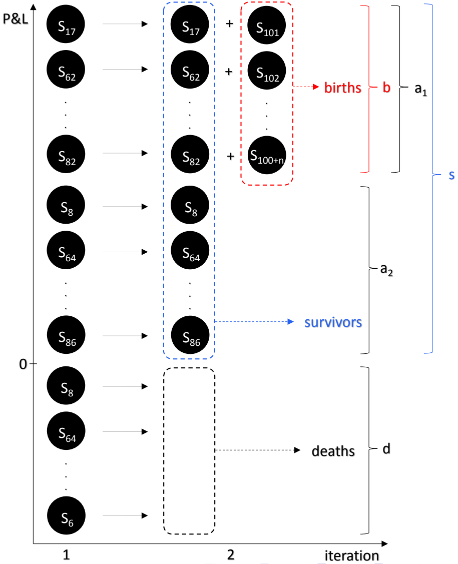
Architecture Complexity and **strategy sophistication** explains the **incentive for Deep Learning** (DL).
Paradoxically we witness **potential for regularization** as the network becomes more complex.

Genetic Algorithm & NN Complexity



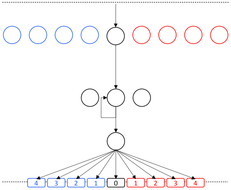
Genetic Algorithm & NN Complexity

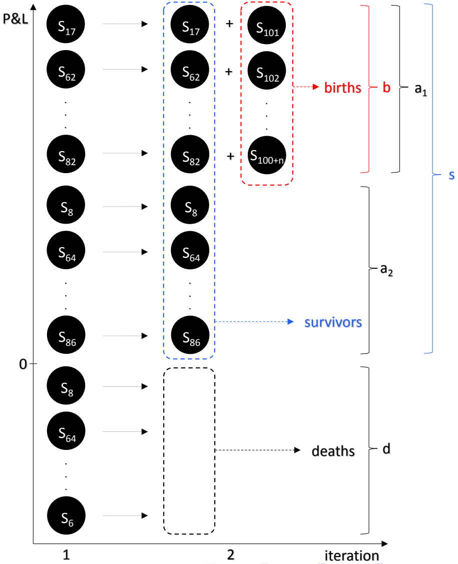
NN **depth & breadth** should contribute in **increasing the learning potential**.



Genetic Algorithm & NN Complexity

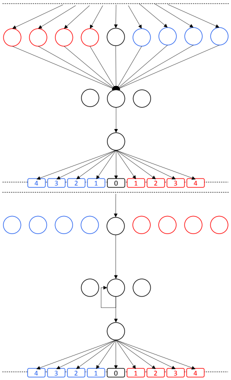
NN **depth & breadth** should contribute in **increasing the learning potential**. The TF (bottom) is less complex

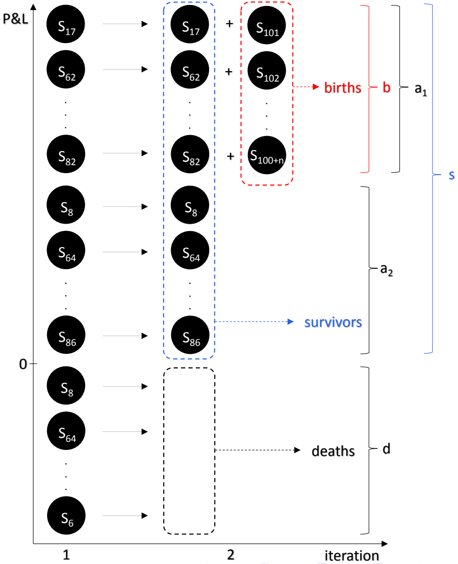




Genetic Algorithm & NN Complexity

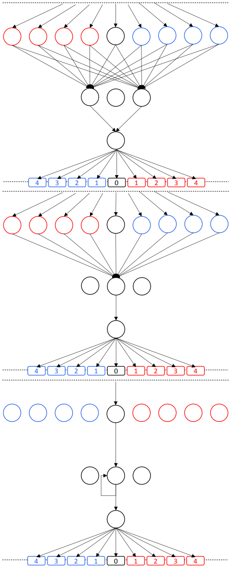
NN **depth & breadth** should contribute in **increasing the learning potential**. The TF (bottom) is less complex than the MLR (middle)

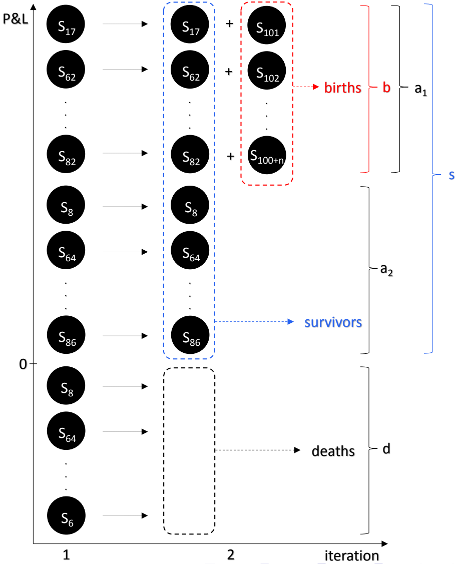




Genetic Algorithm & NN Complexity

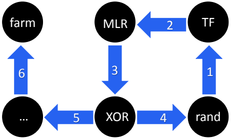
NN **depth & breadth** should contribute in **increasing the learning potential**. The TF (bottom) is less complex than the MLR (middle) which is less complex than the XOR (top) strategy.

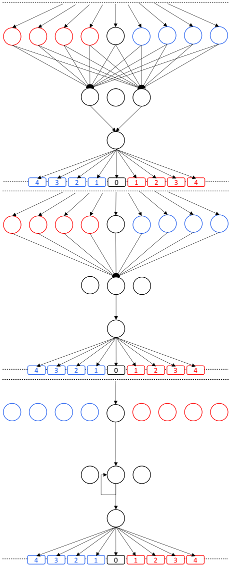


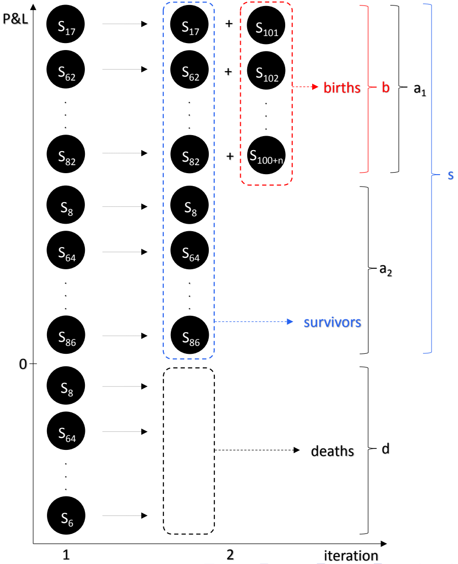


NN **depth & breadth** should contribute in **increasing the learning potential**. The TF (bottom) is less complex than the MLR (middle) which is less complex than the XOR (top) strategy.

Question: Can we make an **analogy to the predator prey ecosystem**? Do we get similar behaviour as the Lotka-Volterra equations [35, 19]?







The Path of Interaction

Answer: not quite but there are **few interesting links** (exponential grow of the smaller prey/self fulfilling strategies such as TF) but they are **many issues** (classification, timescale etc...).

The Path of Interaction

Answer: not quite but there are **few interesting links** (exponential grow of the smaller prey/self fulfilling strategies such as TF) but they are **many issues** (classification, timescale etc...).

The genetic algorithm presented in the previous slide creates complications (classification issues). This pushed us to study the bottom-up approach using concepts taken from evolutionary dynamics and created the the concept of **Path of Interaction**: table of 7 rows **documenting the interaction's details** (eg: table above).

The Path of Interaction

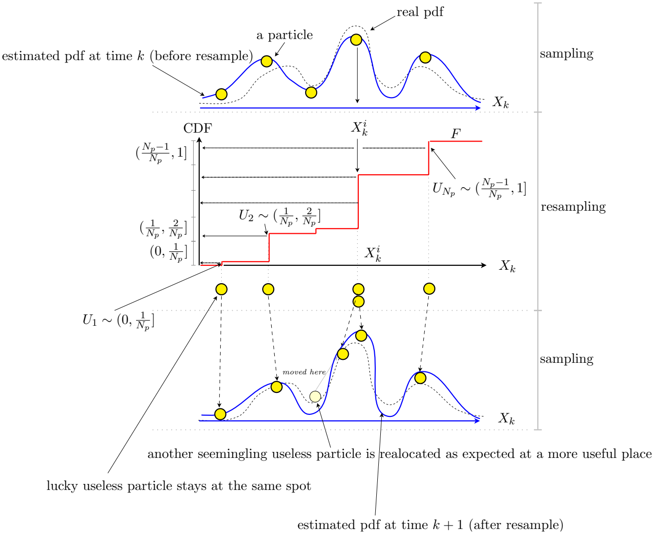
Answer: not quite but there are **few interesting links** (exponential grow of the smaller prey/self fulfilling strategies such as TF) but they are **many issues** (classification, timescale etc...).

Strategy	seed ⇈	TF1	TF2	TF1	TF2
Iteration	0		1		2
Signal	N/A	$+1$	$+1$	$+1$	$+1$
OB	${}_{0}\mathrm{P}^{1,1,1,1}$	${}_{0,0}\mathrm{P}^{1,1,1}$	${}_{0,0,0}\mathrm{P}^{1,1}$	${}_{0,0,0,0}\mathrm{P}^{1}$	${}_{0,0,0,0,0}\mathrm{P}$
Last Price	100	101	102	103	104
ΔOI	$+1$	-1	-2	-3	-4
ΔPrice	$+1$	$+1$	$+1$	$+1$	$+1$
P&L	$[0,0]$		$[1,0]$		$[2,1]$

The genetic algorithm presented in the previous slide creates complications (classification issues). This pushed us to study the bottom-up approach using concepts taken from evolutionary dynamics and created the the concept of **Path of Interaction**: table of 7 rows **documenting the interaction's details** (eg: table above).

Sequential Monte Carlo

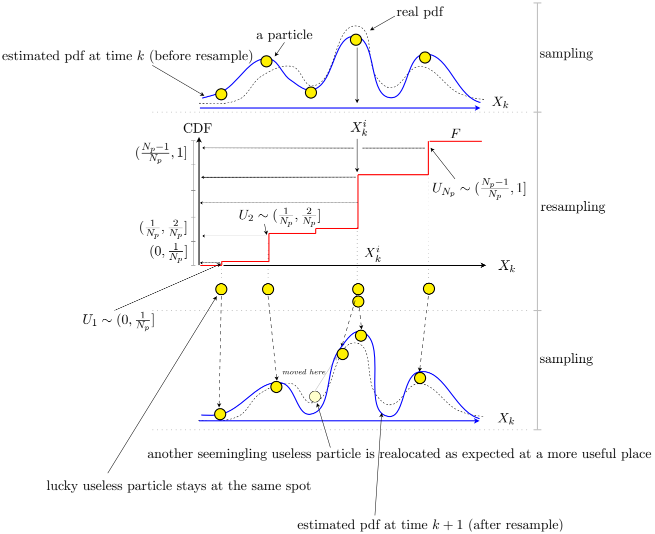
Sequential Monte Carlo (SMC) methods [10, 11, 18], also know as Particle Filter have emerged as a fashionable tool to **track scenarios** in the last 15 years [31, 32]. They are the sequential analogue of Markov Chain Monte Carlo (MCMC) methods and similar to importance sampling methods.



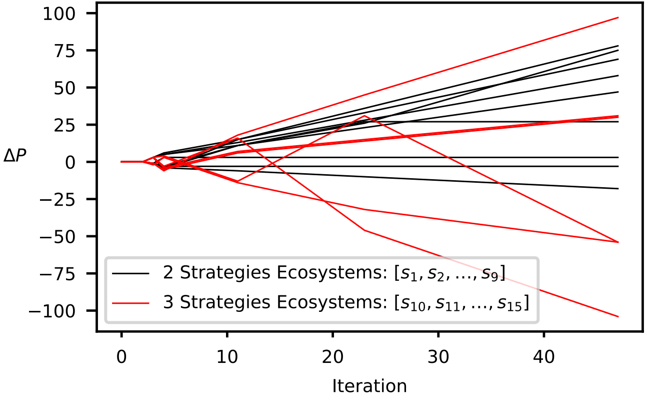
Sequential Monte Carlo

Sequential Monte Carlo (SMC) methods [10, 11, 18], also know as Particle Filter have emerged as a fashionable tool to **track scenarios** in the last 15 years [31, 32]. They are the sequential analogue of Markov Chain Monte Carlo (MCMC) methods and similar to importance sampling methods.

The aims of the PF is to **estimate the sequence of hidden parameters** (eg: the frequencies of certain types of strategies), **based on an indirect observations** (eg: the fluctuations of the market).

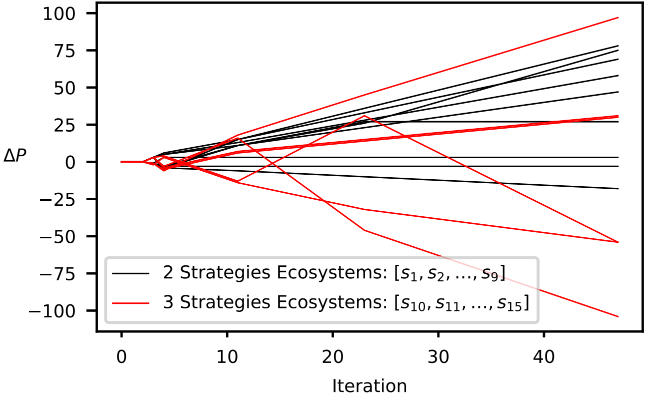


The random Brownian paths become deterministic



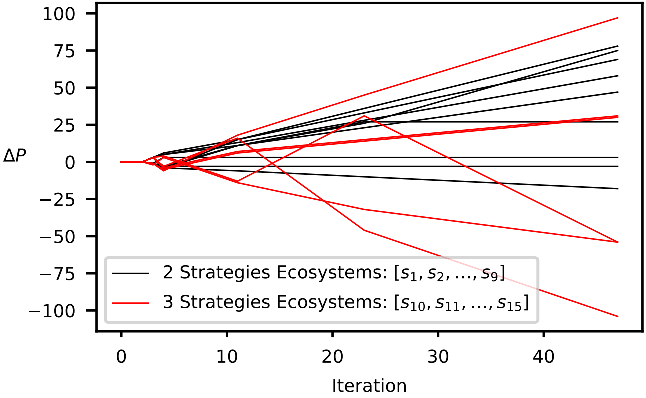
The random Brownian paths become deterministic

Each **Brownian Path Lookalike** is no longer stochastic but **deterministic**.



The random Brownian paths become deterministic

Each **Brownian Path Lookalike** is no longer stochastic but **deterministic**.



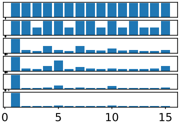
Each **deterministic path** corresponds to a **sequence of interaction** of several **strategies** for which the sequence and the **P&Ls** can be traced through our SMC methods by looking at the market **price only**.

PF assigns a probability for each ecosystem scenario

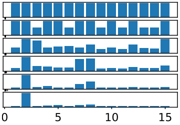
We have recorded **15 different scenarios** (ecosystem history) for the sake of this presentation, all of which are clearly detected after the 11th iteration [27].

PF assigns a probability for each ecosystem scenario

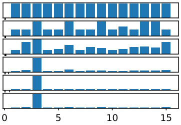
We have recorded **15 different scenarios** (ecosystem history) for the sake of this presentation, all of which are clearly detected after the 11th iteration [27].



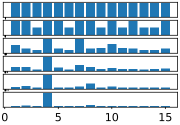
(a) Scenario s_1



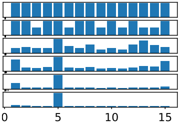
(b) Scenario s_2



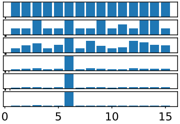
(c) Scenario s_3



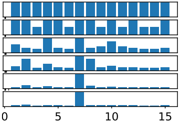
(d) Scenario s_4



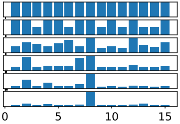
(e) Scenario s_5



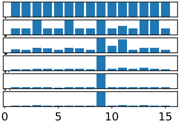
(f) Scenario s_6



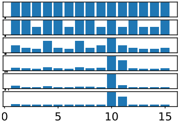
(g) Scenario s_7



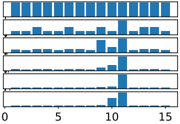
(h) Scenario s_8



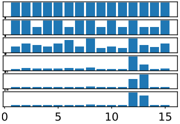
(i) Scenario s_9



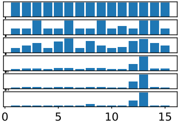
(j) Scenario s_{10}



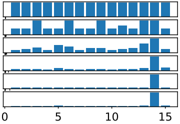
(k) Scenario s_{11}



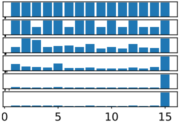
(l) Scenario s_{12}



(m) Scenario s_{13}



(n) Scenario s_{14}



(o) Scenario s_{15}

Model Assuming Data

Definition (Correlation Model)

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t\geq 0)}, \mathbb{P})$ the Correlation model is given by the set of two SDE's:

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t),$$

$$dY(t) = \mu Y(t)dt + \sigma Y(t)d\tilde{W}(t), \tag{1}$$

$$d\langle W, \tilde{W}\rangle_t = \rho dt, \tag{2}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ are drift and diffusion coefficients of asset price, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$.

Model Assuming Data

Definition (Correlation Model)

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t\geq 0)}, \mathbb{P})$ the Correlation model is given by the set of two SDE's:

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t),$$
$$dY(t) = \mu Y(t)dt + \sigma Y(t)d\tilde{W}(t), \qquad (1)$$
$$d\langle W, \tilde{W}\rangle_t = \rho dt, \qquad (2)$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ are drift and diffusion coefficients of asset price, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$.

Remark: few **assumptions** need to be fulfilled. Variance must remain constant (**homoscedasticity**) and the **returns independent**. If not instantaneous measured correlation is misleading with respect to long term correlation:

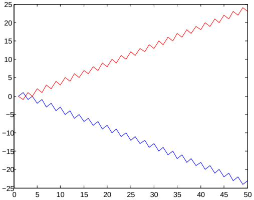
Model Assuming Data

Definition (Correlation Model)

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t\geq 0)}, \mathbb{P})$ the Correlation model is given by the set of two SDE's:

$$\begin{aligned} dX(t) &= \mu X(t)dt + \sigma X(t)dW(t), \\ dY(t) &= \mu Y(t)dt + \sigma Y(t)d\tilde{W}(t), \quad (1) \\ d\langle W, \tilde{W}\rangle_t &= \rho dt, \quad (2) \end{aligned}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ are drift and diffusion coefficients of asset price, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$.



Remark: few **assumptions** need to be fulfilled. Variance must remain constant (**homoscedasticity**) and the **returns independent**. If not instantaneous measured correlation is misleading with respect to long term correlation:

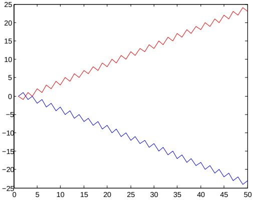
- Top Right Figure: $\hat{\rho} = -1$,

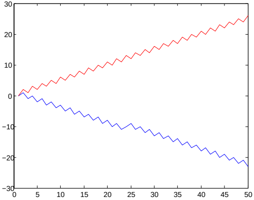
Definition (Correlation Model)

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t\geq 0)}, \mathbb{P})$ the Correlation model is given by the set of two SDE's:

$$\begin{aligned} dX(t) &= \mu X(t)dt + \sigma X(t)dW(t), \\ dY(t) &= \mu Y(t)dt + \sigma Y(t)d\tilde{W}(t), \qquad (1) \\ d\langle W, \tilde{W}\rangle_t &= \rho dt, \qquad (2) \end{aligned}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ are drift and diffusion coefficients of asset price, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$.





Remark: few **assumptions** need to be fulfilled. Variance must remain constant (**homoscedasticity**) and the **returns independent**. If not instantaneous measured correlation is misleading with respect to long term correlation:

- Top Right Figure: $\hat{\rho} = -1$,
- Middle Right Figure $\hat{\rho} = 0$,

Model Assuming Data

Definition (Correlation Model)

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t \geq 0)}, \mathbb{P})$ the Correlation model is given by the set of two SDE's:

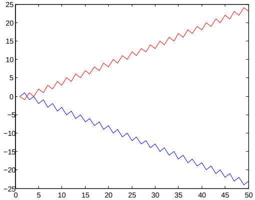
$$
\begin{aligned}
dX(t) &= \mu X(t)dt + \sigma X(t)dW(t), \\
dY(t) &= \mu Y(t)dt + \sigma Y(t)d\tilde{W}(t), &(1)\\
d\langle W, \tilde{W}\rangle_t &= \rho dt, &(2)
\end{aligned}
$$

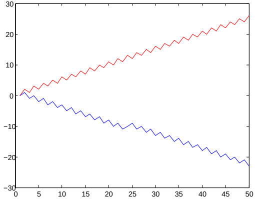
where $\mu \in \mathbb{R}$, $\sigma > 0$ are drift and diffusion coefficients of asset price, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$.

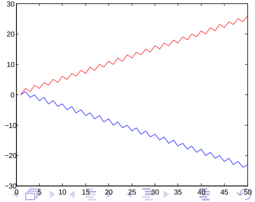
Remark: few **assumptions** need to be fulfilled. Variance must remain constant (**homoscedasticity**) and the **returns independent**. If not instantaneous measured correlation is misleading with respect to long term correlation:

- Top Right Figure: $\hat{\rho} = -1$,
- Middle Right Figure $\hat{\rho} = 0$,
- Bottom Right Figure $\hat{\rho} = +1$.

.







Data Disagrees with Assumptions

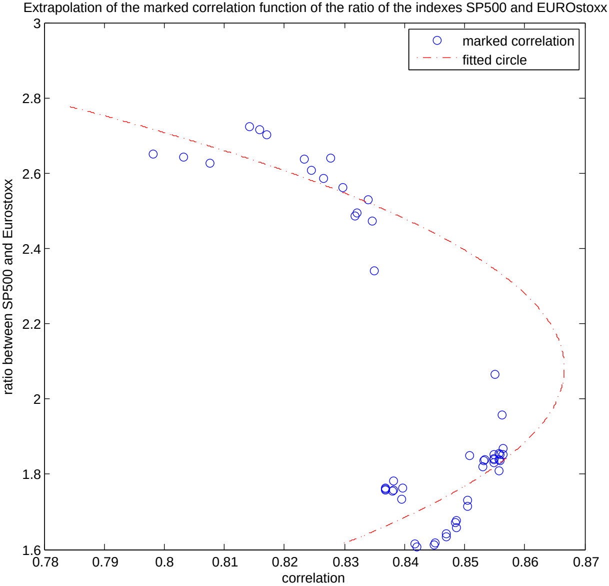
Data violates assumptions of the classic model (2):

- **directly** (this is not new: the returns are not iid and the variance is not constant),

Data Disagrees with Assumptions

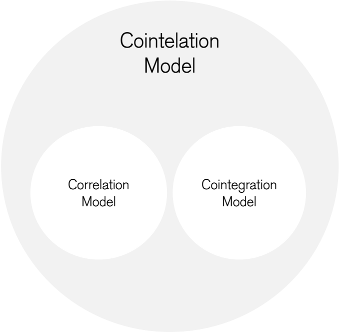
Data violates assumptions of the classic model (2):

- **directly** (this is not new: the returns are not iid and the variance is not constant),
- or **indirectly** (Figure on left hand side). Implied correlation is "marked", the same way implied volatility is. The ratio of two closely related underliers exhibits mean reversion in the minds of the traders risk managing these products.



Data Reassuming the Model

Intuitive Definition: Cointelation is a portmanteau neologism in finance, designed to signify a hybrid method between between **cointe**gration and corre**lation** models (Data-Driven adjustment to classic financial math models: data "reassuming the model").



Data Reassuming the Model

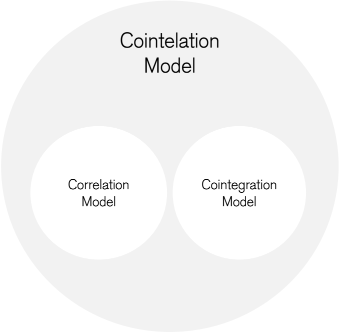
Definition (Cointelation Model)

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t\geq 0)}, \mathbb{P})$ the Cointelation model is given by the SDE's:

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t),$$
$$dY(t) = \kappa\left(X(t) - Y(t)\right)dt + \eta Y(t)d\tilde{W}(t),$$
$$d\langle W, \tilde{W}\rangle_t = \rho dt,$$

with $\mu \in \mathbb{R}$, $\sigma > 0$, $\eta > 0$ are drift and diffusion coefficients of asset price X, κ - the rate of mean reversion, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$. The process $(X(t))_{t\geq 0}$ is called the leading process, $(Y(t))_{t\geq 0}$ the lagging process.

Intuitive Definition: Cointelation is a portmanteau neologism in finance, designed to signify a hybrid method between between **cointe**gration and corre**lation** models (Data-Driven adjustment to classic financial math models: data "reassuming the model").



Data Reassuming the Model

Definition (Cointelation Model)

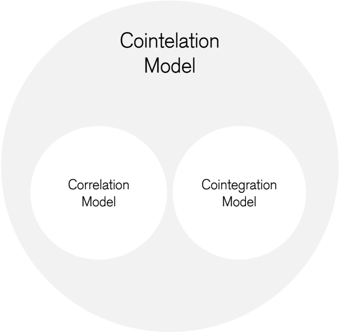
On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{(t \geq 0)}, \mathbb{P})$ the Cointelation model is given by the SDE's:

$$\begin{aligned} dX(t) &= \mu X(t)dt + \sigma X(t)dW(t), \\ dY(t) &= \kappa \left(X(t) - Y(t)\right) dt + \eta Y(t) d\tilde{W}(t), \\ d\langle W, \tilde{W} \rangle_t &= \rho dt, \end{aligned}$$

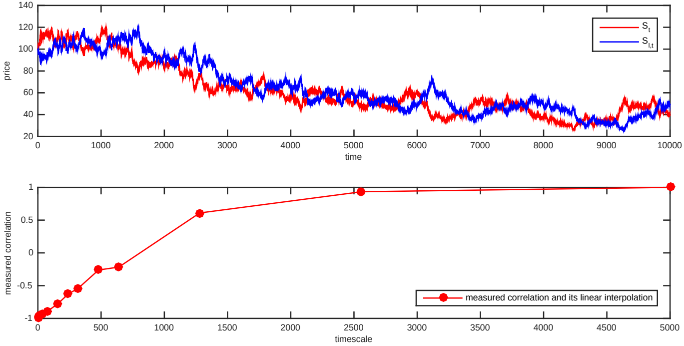
with $\mu \in \mathbb{R}$, $\sigma > 0$, $\eta > 0$ are drift and diffusion coefficients of asset price X, κ - the rate of mean reversion, $\tilde{W}(t)$ and $W(t)$ are two correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$. The process $(X(t))_{t \geq 0}$ is called the leading process, $(Y(t))_{t \geq 0}$ the lagging process.

Remark: We note that setting $\kappa = 0$ yields the classic correlation model. Conversely, setting $\rho = 0$ yields the cointegration model.

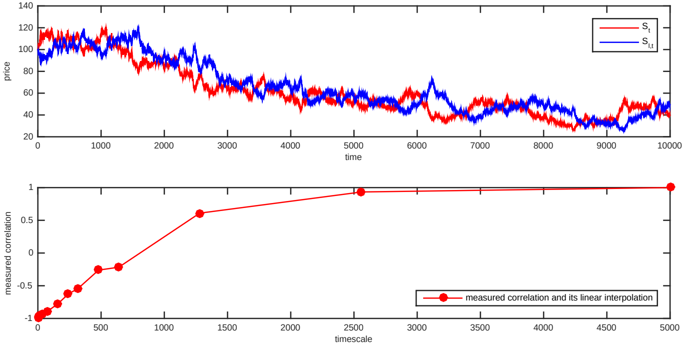
Intuitive Definition: Cointelation is a portmanteau neologism in finance, designed to signify a hybrid method between between **cointe**gration and corre**lation** models (Data-Driven adjustment to classic financial math models: data "reassuming the model").



Interesting Property: the Inferred Correlation



Interesting Property: the Inferred Correlation



Remark: Cointelation model **can hit the whole correlation spectrum** depending on timescale.

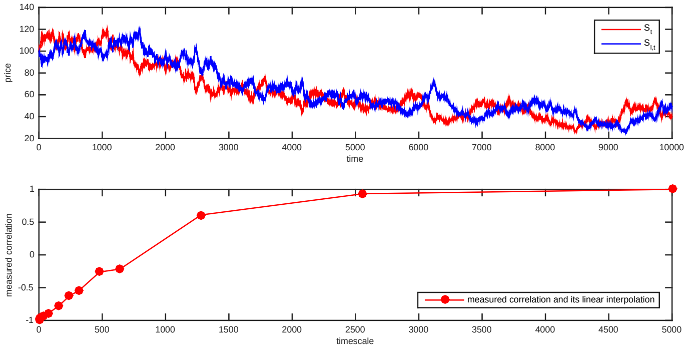
Interesting Property: the Inferred Correlation

Definition (Inferred Correlation)

Considering, the dynamics of equation (3), the **Inferred Correlation** formula is given by (3).

$$\rho^*_\tau \approx \rho + (1-\rho)\left[1-\exp\left(-\kappa\lambda(\tau-1)\right)\right] \quad (3)$$

where $\rho^*_\tau = \mathbb{E}[\sup_{0<t\leq\tau} \rho_t]$, $\tau \in \mathbb{Z}^*$, $\theta \in [0,1]$ and λ constant.



Remark: Cointelation model **can hit the whole correlation spectrum** depending on timescale.

Application: Socially Responsible Finance

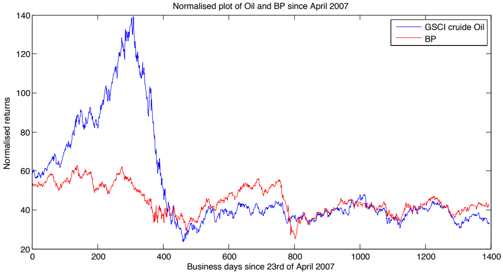
Socially Responsible & Consumer Finance is a wave of quant finance that gains momentum after each crisis but quickly runs unfortunately out of fuel:

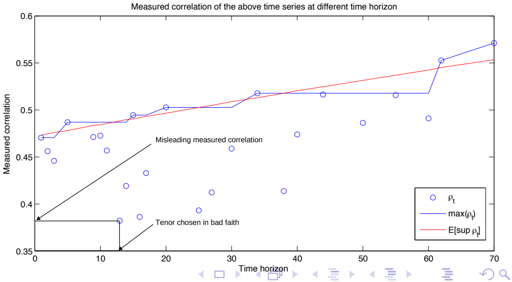
- **Situation**: Equity/Commodities salesman trying to convince clients who have a portfolio in the Commodities/Equities asset class to diversify by buying salesman's products.
- **Objective**: show small correlation to suggest diversification benefits

Application: Socially Responsible Finance

Socially Responsible & Consumer Finance is a wave of quant finance that gains momentum after each crisis but quickly runs unfortunately out of fuel:

- **Situation**: Equity/Commodities salesman trying to convince clients who have a portfolio in the Commodities/Equities asset class to diversify by buying salesman's products.
- **Objective**: show small correlation to suggest diversification benefits
- **Algorithm**: Find the smallest measured correlation as a function of timescale.

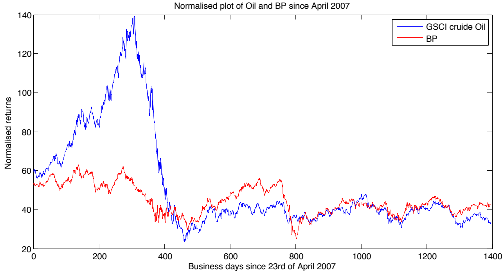


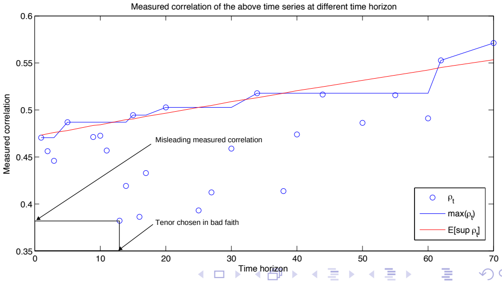


Application: Socially Responsible Finance

Socially Responsible & Consumer Finance is a wave of quant finance that gains momentum after each crisis but quickly runs unfortunately out of fuel:

- **Situation**: Equity/Commodities salesman trying to convince clients who have a portfolio in the Commodities/Equities asset class to diversify by buying salesman's products.
- **Objective**: show small correlation to suggest diversification benefits
- **Algorithm**: Find the smallest measured correlation as a function of timescale.
- **Penalty**: Violation of the Misleading Statement & Action, and Client Best Interest Rules.

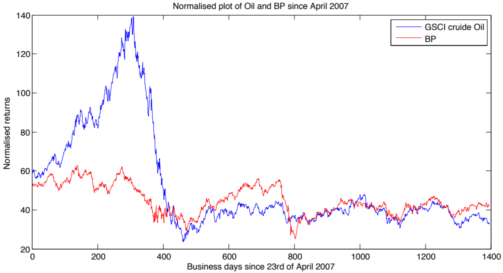


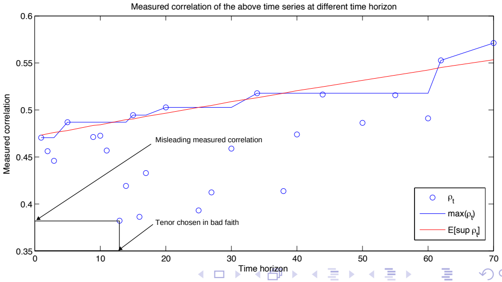


Application: Socially Responsible Finance

Socially Responsible & Consumer Finance is a wave of quant finance that gains momentum after each crisis but quickly runs unfortunately out of fuel:

- **Situation**: Equity/Commodities salesman trying to convince clients who have a portfolio in the Commodities/Equities asset class to diversify by buying salesman's products.
- **Objective**: show small correlation to suggest diversification benefits
- **Algorithm**: Find the smallest measured correlation as a function of timescale.
- **Penalty**: Violation of the Misleading Statement & Action, and Client Best Interest Rules.
- **Solution**: Inferred Correlation.





Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

- ... we can notice that we can decompose the Cointelation model in two familiar models:

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

- ... we can notice that we can decompose the Cointelation model in two familiar models:
    - Markowitz **Mean-Variance Analysis** (ρ-centered sub strategy):

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

- ... we can notice that we can decompose the Cointelation model in two familiar models:
  - Markowitz **Mean-Variance Analysis** (ρ-centered sub strategy):
  - Ornstein–Uhlenbeck "like" **Stochastic Control** (κ-centered sub strategy):

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

- ... we can notice that we can decompose the Cointelation model in two familiar models:
  - Markowitz **Mean-Variance Analysis** (ρ-centered sub strategy):
  - Ornstein–Uhlenbeck "like" **Stochastic Control** (κ-centered sub strategy):
- The optimization process involves a dynamics switching strategy between both approaches

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

- ... we can notice that we can decompose the Cointelation model in two familiar models:
  - Markowitz **Mean-Variance Analysis** (ρ-centered sub strategy): **Easy**
  - Ornstein–Uhlenbeck "like" **Stochastic Control** (κ-centered sub strategy): **Difficult**
- The optimization process involves a dynamics switching strategy between both approaches

$$w_t^* = \max_{h(t)} \left[2\tau E(r_p) - \sigma^2(r_p)\right] 1_{\left\{\left|\kappa\left(\frac{X_t - Y_t}{Y_t}\right)\right| \le \mu\right\}} + \sup_{\pi(t)\in\mathcal{A}(0,v_0)} \mathbb{E}\left[\frac{1}{\gamma}(V^\pi(T))^\gamma\right] 1_{\left\{\left|\kappa\left(\frac{X_t - Y_t}{Y_t}\right)\right| > \mu\right\}}$$

Portfolio Optimization

We consider the portfolio of:

- **two stocks** X_t and Y_t with price dynamics following the cointelation model (3).
- we would like to leverage as much as possible on methods already in place ...

- ... we can notice that we can decompose the Cointelation model in two familiar models:
  - Markowitz **Mean-Variance Analysis** (ρ-centered sub strategy): **Easy**
  - Ornstein–Uhlenbeck "like" **Stochastic Control** (κ-centered sub strategy): **Difficult**
- The optimization process involves a dynamics switching strategy between both approaches

$$w_t^* = \max_{h(t)} \left[2\tau E(r_p) - \sigma^2(r_p)\right] 1_{\left\{\left|\kappa\left(\frac{X_t - Y_t}{Y_t}\right)\right| \leq \mu\right\}} + \sup_{\pi(t) \in \mathcal{A}(0, v_0)} \mathbb{E}\left[\frac{1}{\gamma}(V^\pi(T))^\gamma\right] 1_{\left\{\left|\kappa\left(\frac{X_t - Y_t}{Y_t}\right)\right| > \mu\right\}}$$

The **red part** does not have a closed for solution: **the pure Classic Financial Mathematics does not have a solution**.

Portfolio Optimization: ML/FM Hybrid Method

The way to handle the **red part** part is to consider the function $G(t, v, s)$ such that $G \in C^{1,2}(Q)$, the Hamilton-Jacobi-Bellman (**HJB**) equation corresponding to stochastic control problem is

$$\frac{\partial G}{\partial t}(t, v, s) + \sup_{\pi} \mathcal{L}^{\pi} G(t, v, s) = 0, \tag{4}$$

Portfolio Optimization: ML/FM Hybrid Method

The way to handle the **red part** part is to consider the function $G(t,v,s)$ such that $G \in C^{1,2}(Q)$, the Hamilton-Jacobi-Bellman (**HJB**) equation corresponding to stochastic control problem is

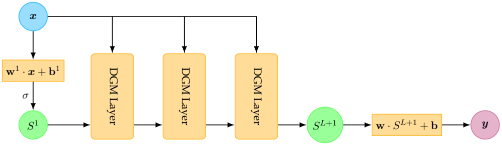
$$\frac{\partial G}{\partial t}(t,v,s) + \sup_{\pi} \mathcal{L}^{\pi} G(t,v,s) = 0, \qquad (4)$$

Through a series of educated guess (1st replacing sup with π^* then using the separation ansatz), we turn an equation of **4** to one of **2 unknowns**

The way to handle the **red part** part is to consider the function $G(t,v,s)$ such that $G \in C^{1,2}(Q)$, the Hamilton-Jacobi-Bellman (**HJB**) equation corresponding to stochastic control problem is

$$\frac{\partial G}{\partial t}(t,v,s) + \sup_{\pi} \mathcal{L}^{\pi} G(t,v,s) = 0, \qquad (4)$$

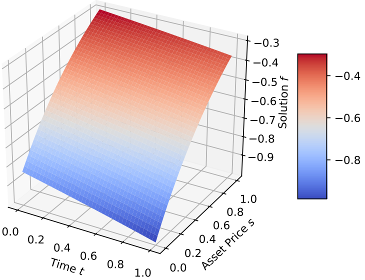
Through a series of educated guess (1st replacing sup with π^* then using the separation ansatz), we turn an equation of **4** to one of **2 unknowns** which we can approximate using **Deep Learning** [1, 33]: it works for the **Merton non-linear PDE**.

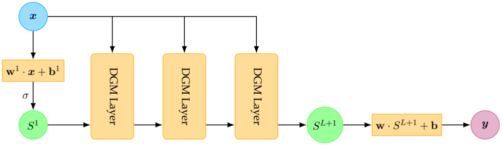


The way to handle the **red part** part is to consider the function $G(t, v, s)$ such that $G \in C^{1,2}(Q)$, the Hamilton-Jacobi-Bellman (**HJB**) equation corresponding to stochastic control problem is

$$\frac{\partial G}{\partial t}(t, v, s) + \sup_{\pi} \mathcal{L}^{\pi} G(t, v, s) = 0, \tag{4}$$

Through a series of educated guess (1st replacing sup with π^* then using the separation ansatz), we turn an equation of **4** to one of **2 unknowns** which we can approximate using **Deep Learning** [1, 33]: it works for the **Merton non-linear PDE**.

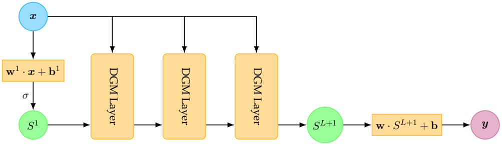


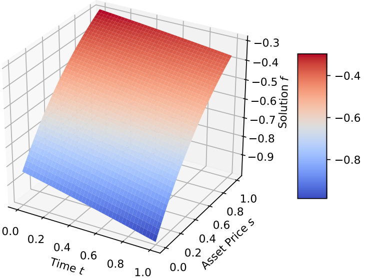


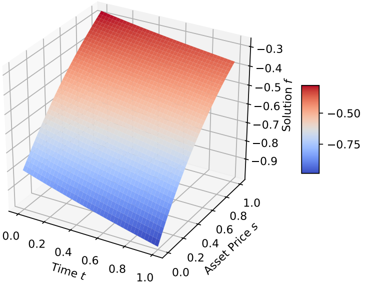
The way to handle the **red part** part is to consider the function $G(t,v,s)$ such that $G \in C^{1,2}(Q)$, the Hamilton-Jacobi-Bellman (**HJB**) equation corresponding to stochastic control problem is

$$\frac{\partial G}{\partial t}(t,v,s) + \sup_{\pi} \mathcal{L}^{\pi} G(t,v,s) = 0, \qquad (4)$$

Through a series of educated guess (1st replacing sup with π^* then using the separation ansatz), we turn an equation of **4** to one of **2 unknowns** which we can approximate using **Deep Learning** [1, 33]: it works for the **Merton non-linear PDE**.





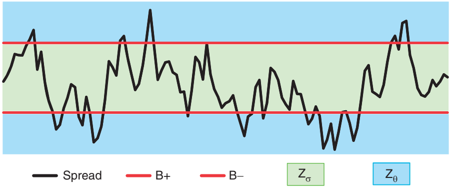


Portfolio Optimization: Pure ML Approximation

We can use clustering in order to study the P&L of n (eg: 4) strategies through time.

Portfolio Optimization: Pure ML Approximation

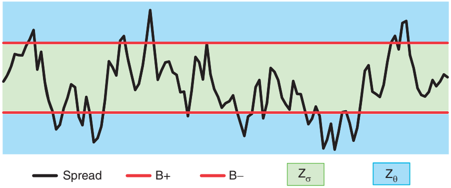
We can use clustering in order to study the P&L of *n* (eg: 4) strategies through time.



Portfolio Optimization: Pure ML Approximation

We can use clustering in order to study the P&L of n (eg: 4) strategies through time.

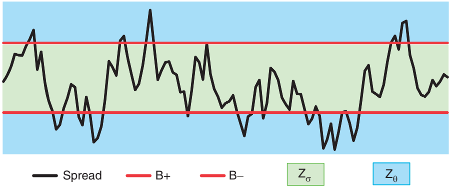
- Strategy S^{++} in which we are long both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{++}_{[a_i,b_i],t}$.



Portfolio Optimization: Pure ML Approximation

We can use clustering in order to study the P&L of n (eg: 4) strategies through time.

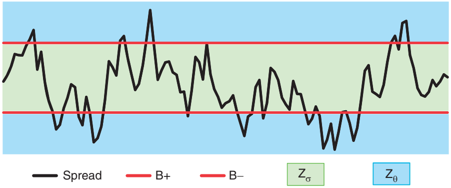
- Strategy S^{++} in which we are long both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{++}_{[a_i,b_i],t}$.
- Strategy S^{+-} in which we are long X and short Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{+-}_{[a_i,b_i],t}$.



Portfolio Optimization: Pure ML Approximation

We can use clustering in order to study the P&L of n (eg: 4) strategies through time.

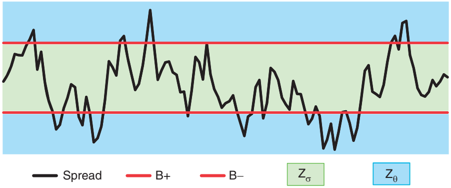
- Strategy S^{++} in which we are long both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{++}_{[a_i,b_i],t}$.
- Strategy S^{+-} in which we are long X and short Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{+-}_{[a_i,b_i],t}$.
- Strategy S^{-+} in which we are short X and long Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{-+}_{[a_i,b_i],t}$.



Portfolio Optimization: Pure ML Approximation

We can use clustering in order to study the P&L of n (eg: 4) strategies through time.

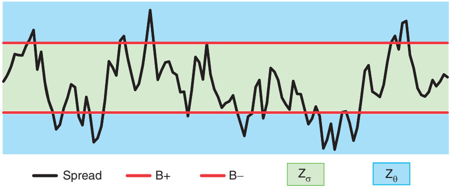
- Strategy S^{++} in which we are long both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{++}_{[a_i,b_i],t}$.
- Strategy S^{+-} in which we are long X and short Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{+-}_{[a_i,b_i],t}$.
- Strategy S^{-+} in which we are short X and long Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{-+}_{[a_i,b_i],t}$.
- Strategy S^{--} in which we are short both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{--}_{[a_i,b_i],t}$.

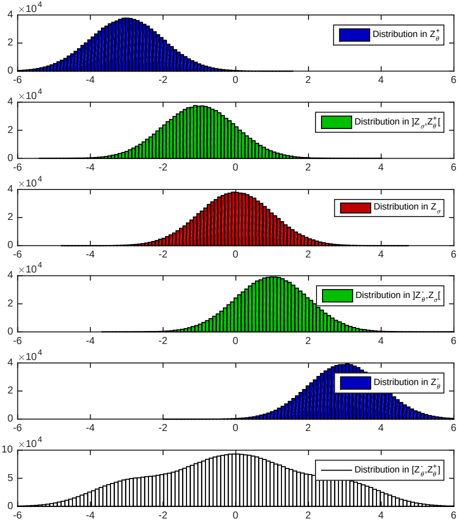


Portfolio Optimization: Pure ML Approximation

We can use clustering in order to study the P&L of n (eg: 4) strategies through time.

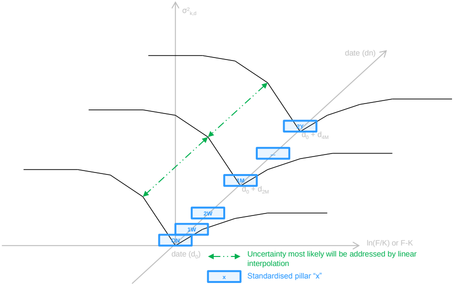
- Strategy S^{++} in which we are long both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{++}_{[a_i,b_i],t}$.
- Strategy S^{+-} in which we are long X and short Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{+-}_{[a_i,b_i],t}$.
- Strategy S^{-+} in which we are short X and long Y at time t within $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{-+}_{[a_i,b_i],t}$.
- Strategy S^{--} in which we are short both X and Y at time t within bands $[a_i, b_i]$, $i \in \mathbb{N}$, and with P&L $V^{--}_{[a_i,b_i],t}$.





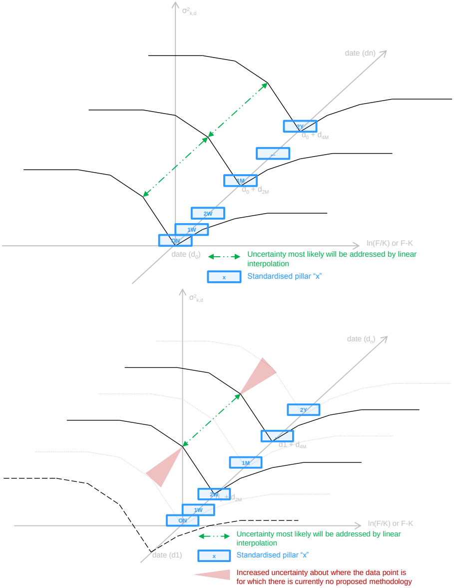
Remark: We can have as many bands (strategies) as we have weight proportions.

Anomaly Detection & Volatility Surface de-Arbitraging



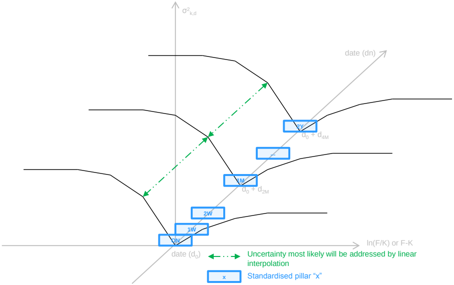
Normalizing the data coming from the markets (left figure)

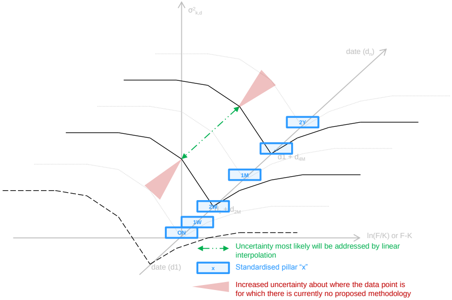
Anomaly Detection & Volatility Surface de-Arbitraging



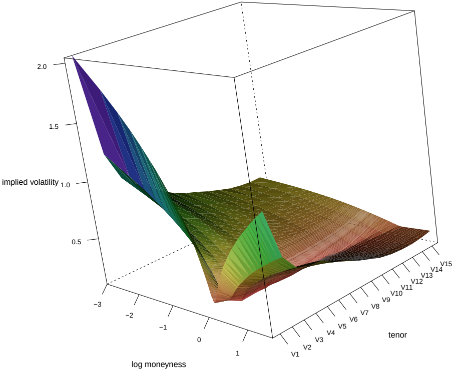
Normalizing the data coming from the markets (left figure) in **rolling contract form**

Anomaly Detection & Volatility Surface de-Arbitraging





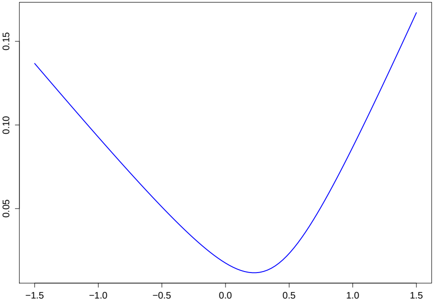
Normalizing the data coming from the markets (left figure) in **rolling contract form** into a coherent **fixed pillars implied volatility surface** (bottom right figure) presents challenges.



Anomaly Detection & Volatility Surface de-Arbitraging

The classic arbitrage conditions (butterfly & calendar spread) have been replaced with more elegant models (e.g. $\forall K, \forall T, |T\partial_K \sigma^2(K,T)| \leq 4$) that were nevertheless not sufficient and the arrival of Big Data in the wings exposed these limitations in the Financial Mathematics models.

The classic arbitrage conditions (butterfly & calendar spread) have been replaced with more elegant models (e.g. $\forall K, \forall T, |T\partial_K \sigma^2(K,T)| \leq 4$) that were nevertheless not sufficient and the arrival of Big Data in the wings exposed these limitations in the Financial Mathematics models.



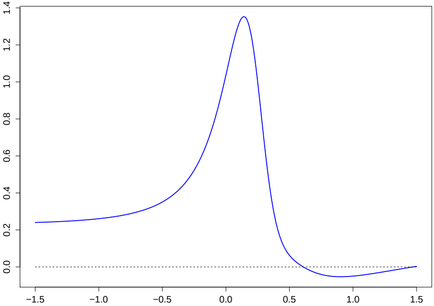
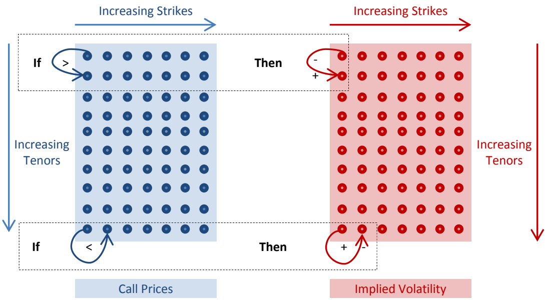


Figure: Vogt's total variance example verifying $b(1+|\rho|) \leq \frac{4}{T}$ (left figure: with the x axis being the log-moneyness and the y axis being the implied variance) and the corresponding $\partial^2_{K,K} BS\left(\sigma^2(K,T)\right)$ approximating the (supposed) always positive pdf (right figure: with the x axis being the log-moneyness and the y axis being the non normalized pdf).

Anomaly Detection & Volatility Surface de-Arbitraging

The work of this chapter revolves around **anomaly detection** in the context of the implied volatility surface trying to use the **classic methods**: butterfly arbitrage (or call spread) in equation (5b) and calendar spread arbitrage in equation (5c) and the **more advanced methods** such as in equation (5e) (a **necessary but not sufficient** condition).

Anomaly Detection & Volatility Surface de-Arbitraging



The work of this chapter revolves around **anomaly detection** in the context of the implied volatility surface trying to use the **classic methods**: butterfly arbitrage (or call spread) in equation (5b) and calendar spread arbitrage in equation (5c) and the **more advanced methods** such as in equation (5e) (a **necessary but not sufficient** condition).

$$\text{solve: } \hat{\sigma}_t(\tau, d) = \underbrace{\arg\min}_{\tilde{\sigma}_t(\tau,d)} \sum_{\tau} \sum_{d} \left[C\Big(\sigma_{i,t}(\tau, d)\Big) - C\Big(\tilde{\sigma}_t(\tau, d)\Big)\right]^2 \quad (5a)$$

subject to: $\forall \tau$ and $\forall K$

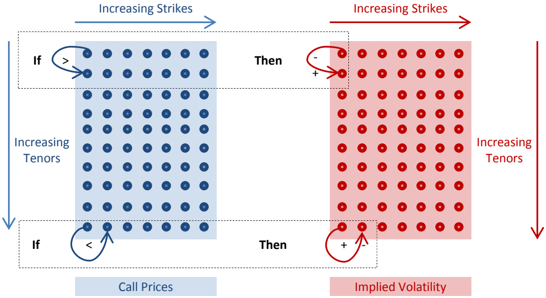
$$C\Big(K - \Delta, \sigma_0(K - \Delta, \tau)\Big) - C\Big(K, \sigma_0(K, \tau)\Big) \geq 0 \quad (5b)$$

$$C\Big(K, \tau + \Delta, \sigma_0(K, \tau + \Delta)\Big) - C\Big(Ke^{-r\Delta}, \sigma_0(Ke^{-r\Delta}, \tau)\Big) \geq 0 \quad (5c)$$

but not: (5d)

$$\forall K, \forall T, |T \partial_K \sigma^2(K, T)| \leq 4 \quad (5e)$$

Anomaly Detection & Volatility Surface de-Arbitraging



The work of this chapter revolves around **anomaly detection** in the context of the implied volatility surface trying to use the **classic methods**: butterfly arbitrage (or call spread) in equation (5b) and calendar spread arbitrage in equation (5c) and the **more advanced methods** such as in equation (5e) (a **necessary but not sufficient** condition). We also discuss some of the **idiosyncratic asset class related differences** that require modifications in the optimization process. Finally we discuss few **trading idea by simplifying the IVP** model.

solve: $\hat{\sigma}_t(\tau, d) = \underbrace{\arg\min}_{\tilde{\sigma}_t(\tau,d)} \sum_\tau \sum_d \left[C\Big(\sigma_{i,t}(\tau, d)\Big) - C\Big(\tilde{\sigma}_t(\tau, d)\Big)\right]^2$ (5a)

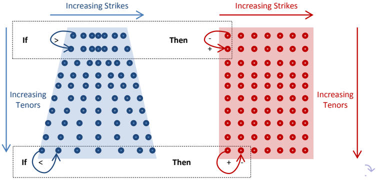
subject to: $\forall \tau$ and $\forall K$

$C\Big(K - \Delta, \sigma_0(K - \Delta, \tau)\Big) - C\Big(K, \sigma_0(K, \tau)\Big) \geq 0$ (5b)

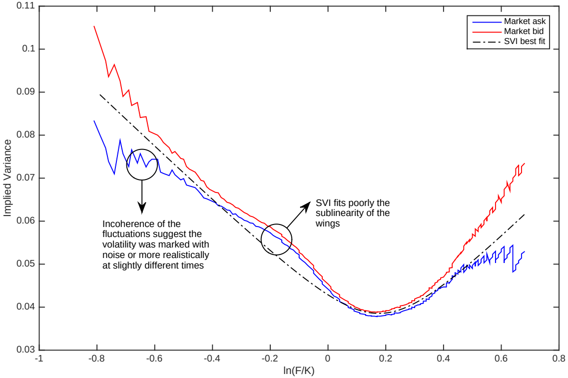
$C\Big(K, \tau + \Delta, \sigma_0(K, \tau + \Delta)\Big) - C\Big(Ke^{-r\Delta}, \sigma_0(Ke^{-r\Delta}, \tau)\Big) \geq 0$ (5c)

but not: (5d)

$\forall K, \forall T, |T \partial_K \sigma^2(K, T)| \leq 4$ (5e)



Big Data Changing the Vanilla Options Landscape

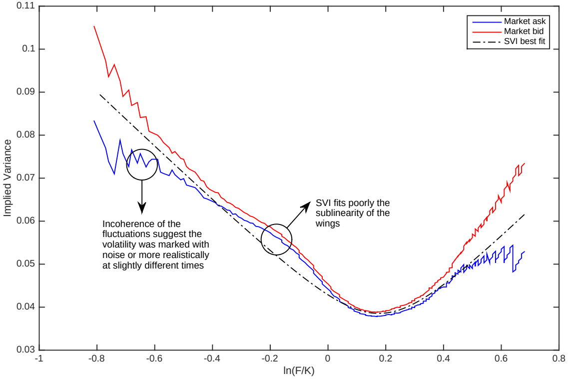


- exposed the limitations of both the **wings** (e.g. **SVI**) and of the **liquidity** in the options market,

We saw in the **Options market**:

- a step back in **complexity** (from exotics to vanilla) but with more **coherent** pricing
- a step forward in **liquidity** modelling

Big Data Changing the Vanilla Options Landscape

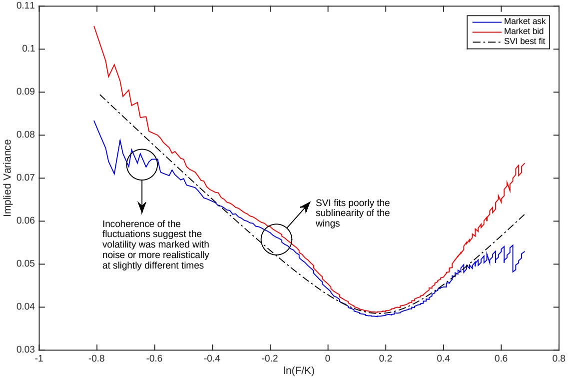


- exposed the limitations of both the **wings** (e.g. **SVI**) and of the **liquidity** in the options market,
- the need for **proxying** when the data is scarse.

We saw in the **Options market**:

- a step back in **complexity** (from exotics to vanilla) but with more **coherent** pricing
- a step forward in **liquidity** modelling
- the need (in CCPs), to map **risk factors** to economical concepts.

Big Data Changing the Vanilla Options Landscape



$$dS_t = \sqrt{v_t} S_t dW_t^1, \quad S_0 \in \mathbb{R}_+^* \tag{6a}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2, \quad v_0 \in \mathbb{R}_+^* \tag{6b}$$

The **rise of big data**:

- exposed the limitations of both the **wings** (e.g. **SVI**) and of the **liquidity** in the options market,
- the need for **proxying** when the data is scarce.
- the **Heston** (Stochastic Volatility) model and the local volatility model and the need for **harmonizing** these two concepts.

We saw in the **Options market**:

- a step back in **complexity** (from exotics to vanilla) but with more **coherent** pricing
- a step forward in **liquidity** modelling
- the need (in CCPs), to map **risk factors** to economical concepts.

Big Data Changing the Vanilla Options Landscape

Big Data exposed the limitations of the **SVI** in the wings and the **subprime crisis** of 2007 exposed the need to incorporate **liquidity** directly in the options model.

Big Data Changing the Vanilla Options Landscape

We define the Implied Volatility surface Parametrization (IVP) split with its mid in equation (7) with the **downside transform** in equation (7b) **enhancing the SVI**,

$$\sigma^2_{IVP,o,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{o,\tau} - m_\tau \right) + \sqrt{\left(z_{o,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] \quad (7a)$$

$$z_{o,\tau} = \frac{k}{\beta_{o,\tau}^{1+4|k-m|}} \quad (7b)$$

Big Data exposed the limitations of the **SVI** in the wings and the **subprime crisis** of 2007 exposed the need to incorporate **liquidity** directly in the options model. The **IVP** addresses these two points

Big Data Changing the Vanilla Options Landscape

We define the Implied Volatility surface Parametrization (IVP) split with its mid in equation (7) with the **downside transform** in equation (7b) **enhancing the SVI**,

$$\sigma^2_{IVP,o,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{o,\tau} - m_\tau \right) + \sqrt{\left(z_{o,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] \quad (7a)$$

$$z_{o,\tau} = \frac{k}{\beta_{o,\tau}^{1+4|k-m|}} \quad (7b)$$

and its liquidity parameters in equation (8). Parameters ψ represents the **Wings Curvature**, α represents the **ATM Spread**. The latter two parameters can be defined in terms of functions to **accommodate the position size** in which case the market dept is controlled with the η parameters.

$$\sigma^2_{IVP,+,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{+,\tau} - m_\tau \right) + \sqrt{\left(z_{+,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] + \alpha_\tau(p) \quad (8a)$$

$$z_{+,\tau} = z_{o,\tau}[1 + \psi_\tau(p)] \quad (8b)$$

$$\sigma^2_{IVP,-,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{-,\tau} - m_\tau \right) + \sqrt{\left(z_{-,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] - \alpha_\tau(p) \quad (8c)$$

$$z_{-,\tau} = z_{o,\tau}[1 - \psi_\tau(p)] \quad (8d)$$

$$\alpha_\tau(p) = \alpha_{0,\tau} + (a_\tau - \alpha_{0,\tau})(1 - e^{-\eta_{\alpha_\tau} p}) \quad (8e)$$

$$\psi_\tau(p) = \psi_{0,\tau} + (1 - \psi_{0,\tau})(1 - e^{-\eta_{\psi_\tau} p}) \quad (8f)$$

Big Data exposed the limitations of the **SVI** in the wings and the **subprime crisis** of 2007 exposed the need to incorporate **liquidity** directly in the options model. The **IVP** addresses these two points

Big Data Changing the Vanilla Options Landscape

We define the Implied Volatility surface Parametrization (IVP) split with its mid in equation (7) with the **downside transform** in equation (7b) **enhancing the SVI**,

$$\sigma^2_{IVP,o,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{o,\tau} - m_\tau \right) + \sqrt{\left(z_{o,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] \tag{7a}$$

$$z_{o,\tau} = \frac{k}{\beta_{o,\tau}^{1+4|k-m|}} \tag{7b}$$

and its liquidity parameters in equation (8). Parameters ψ represents the **Wings Curvature**, α represents the **ATM Spread**. The latter two parameters can be defined in terms of functions to **accommodate the position size** in which case the market dept is controlled with the η parameters.

$$\sigma^2_{IVP,+,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{+,\tau} - m_\tau \right) + \sqrt{\left(z_{+,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] + \alpha_\tau(p) \tag{8a}$$

$$z_{+,\tau} = z_{o,\tau}[1 + \psi_\tau(p)] \tag{8b}$$

$$\sigma^2_{IVP,-,\tau}(k) = a_\tau + b_\tau \left[\rho_\tau \left(z_{-,\tau} - m_\tau \right) + \sqrt{\left(z_{-,\tau} - m_\tau \right)^2 + \sigma^2_\tau} \right] - \alpha_\tau(p) \tag{8c}$$

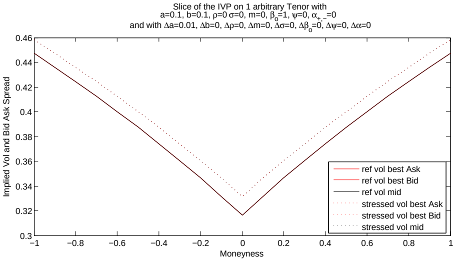
$$z_{-,\tau} = z_{o,\tau}[1 - \psi_\tau(p)] \tag{8d}$$

$$\alpha_\tau(p) = \alpha_{0,\tau} + (a_\tau - \alpha_{0,\tau})(1 - e^{-\eta_{\alpha_\tau} p}) \tag{8e}$$

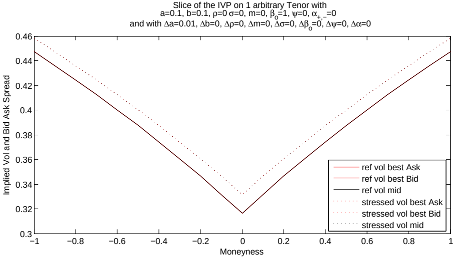
$$\psi_\tau(p) = \psi_{0,\tau} + (1 - \psi_{0,\tau})(1 - e^{-\eta_{\psi_\tau} p}) \tag{8f}$$

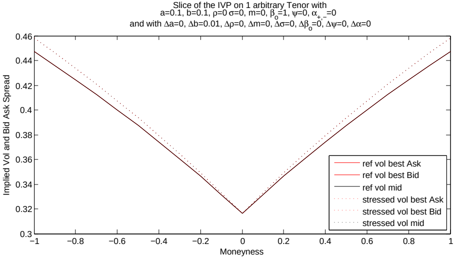
Big Data exposed the limitations of the **SVI** in the wings and the **subprime crisis** of 2007 exposed the need to incorporate **liquidity** directly in the options model. The **IVP** addresses these two points as well as allow for additional benefits:

- **Proxying**
- **Backtesting** weaponry on complex strategies.

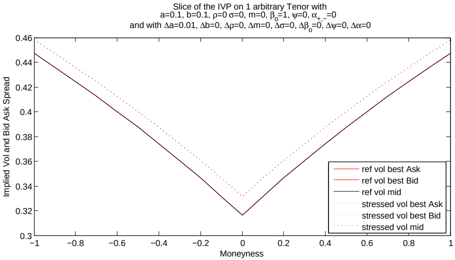

Slice of the IVP on 1 arbitrary Tenor with
a=0.1, b=0.1, ρ=0 σ=0, m=0, β_o=1, ψ=0, α_{+,−}=0
and with Δa=0.01, Δb=0, Δρ=0, Δm=0, Δσ=0, Δβ_o=0, Δψ=0, Δα=0
Implied Vol and Bid Ask Spread
0.46
0.44
0.42
0.4
0.38
0.36
0.34
0.32
0.3
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
Moneyness
ref vol best Ask
ref vol best Bid
ref vol mid
stressed vol best Ask
stressed vol best Bid
stressed vol mid

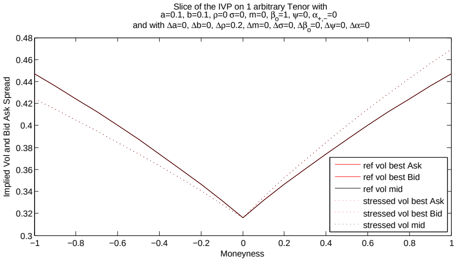
Big Data, Proxying & Handling Dimensionality for Options

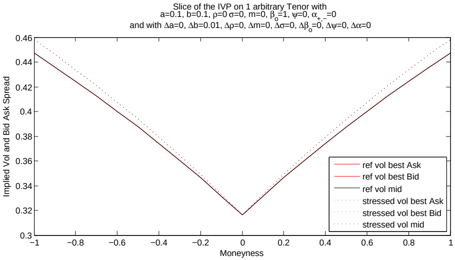


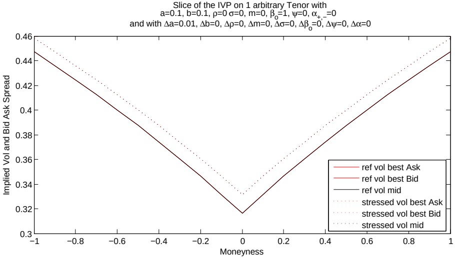


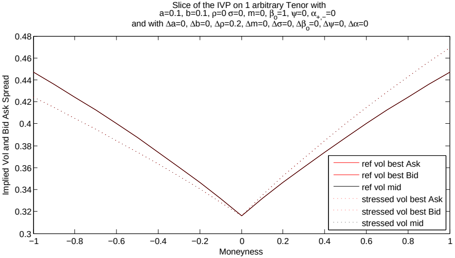
Big Data, Proxying & Handling Dimensionality for Options

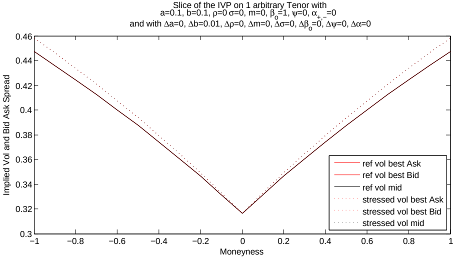


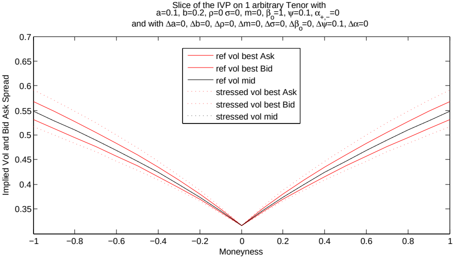





Slice of the IVP on 1 arbitrary Tenor with
a=0.1, b=0.1, ρ=0 σ=0, m=0, β_o=1, ψ=0, α_{+,-}=0
and with Δa=0.01, Δb=0, Δρ=0, Δm=0, Δσ=0, Δβ_o=0, Δψ=0, Δα=0
Implied Vol and Bid Ask Spread
Moneyness
ref vol best Ask
ref vol best Bid
ref vol mid
stressed vol best Ask
stressed vol best Bid
stressed vol mid


Slice of the IVP on 1 arbitrary Tenor with
a=0.1, b=0.1, ρ=0 σ=0, m=0, β_o=1, ψ=0, α_{+,-}=0
and with Δa=0, Δb=0, Δρ=0.2, Δm=0, Δσ=0, Δβ_o=0, Δψ=0, Δα=0
Implied Vol and Bid Ask Spread
Moneyness
ref vol best Ask
ref vol best Bid
ref vol mid
stressed vol best Ask
stressed vol best Bid
stressed vol mid

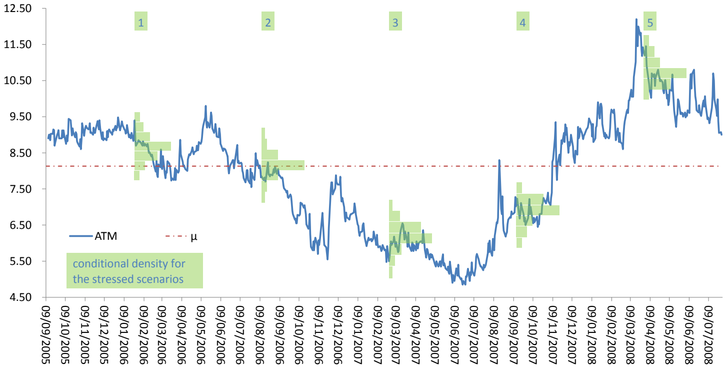

Slice of the IVP on 1 arbitrary Tenor with
a=0.1, b=0.1, ρ=0 σ=0, m=0, β_o=1, ψ=0, α_{+,-}=0
and with Δa=0, Δb=0.01, Δρ=0, Δm=0, Δσ=0, Δβ_o=0, Δψ=0, Δα=0
Implied Vol and Bid Ask Spread
Moneyness
ref vol best Ask
ref vol best Bid
ref vol mid
stressed vol best Ask
stressed vol best Bid
stressed vol mid


Slice of the IVP on 1 arbitrary Tenor with
a=0.1, b=0.2, ρ=0 σ=0, m=0, β_o=1, ψ=0.1, α_{+,-}=0
and with Δa=0, Δb=0, Δρ=0, Δm=0, Δσ=0, Δβ_o=0, Δψ=0.1, Δα=0
Implied Vol and Bid Ask Spread
Moneyness
ref vol best Ask
ref vol best Bid
ref vol mid
stressed vol best Ask
stressed vol best Bid
stressed vol mid

Clustering for Distribution & Regime Change Forecasting

We show how **clustering can help enhance MF** and therefore the two fields can be apposed instead of opposed in the context of modelling risk factors (RF) which behave elements of mean reversion (Spread, Options RF). More specifically we look at how **we can free oneself with the assumptions of SDEs** to construct a general clustering methodology. This can allow us to construct concepts like the **Anticipative VaR** (a leading regime change) as opposed to Responsive VaR (a lagging regime change).

Clustering for Distribution & Regime Change Forecasting



In the left figure, we apply clustering in order to classify **dynamic zones** in which the **returns act differently**. For example when the underlier is significantly above its mean, the forecasted distribution is normally distributed with however a negative mean (vice versa when the underlier is below its historical mean). The distribution is symmetric when at the long term mean.

We show how **clustering can help enhance MF** and therefore the two fields can be apposed instead of opposed in the context of modelling risk factors (RF) which behave elements of mean reversion (Spread, Options RF). More specifically we look at how **we can free oneself with the assumptions of SDEs** to construct a general clustering methodology. This can allow us to construct concepts like the **Anticipative VaR** (a leading regime change) as opposed to Responsive VaR (a lagging regime change).

Clustering for Distribution & Regime Change Forecasting

In order to **Reconcile** discordant instructions of our regulators to create a risk measure which is **responsive** but **stable** at the same time we propose the **Responsible VaR**, a risk measure responsive on the upside but stable on the downside. We give a couple of examples (figures on the left) of complex portfolio (straddle) backtests in which we modify the λ to control the stability on the downside.

Clustering for Distribution & Regime Change Forecasting

In order to **Reconcile** discordant instructions of our regulators to create a risk measure which is **responsive** but **stable** at the same time we propose the **Responsible VaR**, a risk measure responsive on the upside but stable on the downside. We give a couple of examples (figures on the left) of complex portfolio (straddle) backtests in which we modify the λ to control the stability on the downside.

$$\alpha = \int_{-\infty}^{\nu_t^+} p_t(x)dx \tag{9a}$$

$$\tilde{\nu}_0^+ = \nu_0^+ \tag{9b}$$

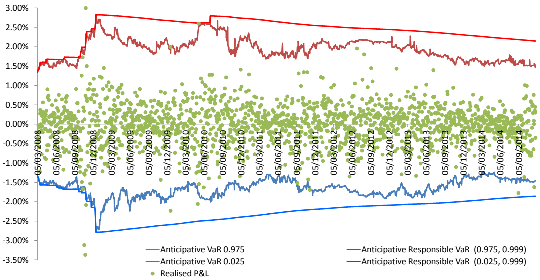
$$\tilde{\nu}_t^+ = \max\left(\nu_t^+, \lambda\tilde{\nu}_{t-1}^+ + (1-\lambda)\nu_t^+\right) \tag{9c}$$

$$1-\alpha = \int_{\nu_t^-}^{+\infty} p_t(x)dx \tag{9d}$$

$$\tilde{\nu}_0^- = \nu_0^- \tag{9e}$$

$$\tilde{\nu}_t^- = \min\left(\nu_t^-, \lambda\tilde{\nu}_{t-1} + (1-\lambda)\nu_t^-\right) \tag{9f}$$

Clustering for Distribution & Regime Change Forecasting



In order to **Reconcile** discordant instructions of our regulators to create a risk measure which is **responsive** but **stable** at the same time we propose the **Responsible VaR**, a risk measure responsive on the upside but stable on the downside. We give a couple of examples (figures on the left) of complex portfolio (straddle) backtests in which we modify the λ to control the stability on the downside.

$$\alpha = \int_{-\infty}^{\nu_t^+} p_t(x)dx \quad (9a)$$

$$\tilde{\nu}_0^+ = \nu_0^+ \quad (9b)$$

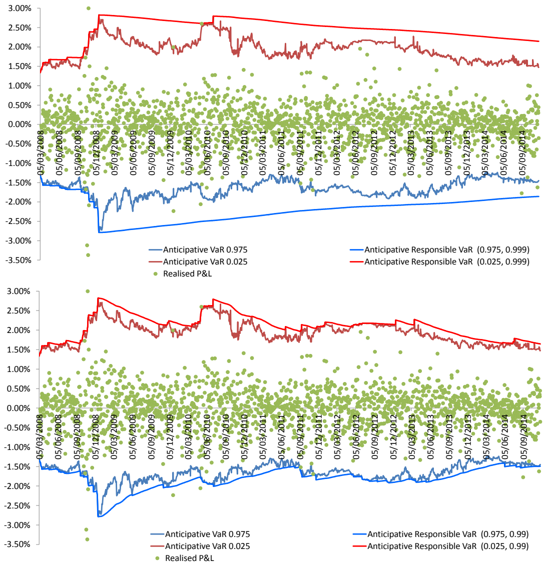
$$\tilde{\nu}_t^+ = \max\left(\nu_t^+, \lambda\tilde{\nu}_{t-1}^+ + (1-\lambda)\nu_t^+\right) \quad (9c)$$

$$1 - \alpha = \int_{\nu_t^-}^{+\infty} p_t(x)dx \quad (9d)$$

$$\tilde{\nu}_0^- = \nu_0^- \quad (9e)$$

$$\tilde{\nu}_t^- = \min\left(\nu_t^-, \lambda\tilde{\nu}_{t-1} + (1-\lambda)\nu_t^-\right) \quad (9f)$$

Clustering for Distribution & Regime Change Forecasting



In order to **Reconcile** discordant instructions of our regulators to create a risk measure which is **responsive** but **stable** at the same time we propose the **Responsible VaR**, a risk measure responsive on the upside but stable on the downside. We give a couple of examples (figures on the left) of complex portfolio (straddle) backtests in which we modify the λ to control the stability on the downside.

$$\alpha = \int_{-\infty}^{\nu_t^+} p_t(x)dx \tag{9a}$$

$$\tilde{\nu}_0^+ = \nu_0^+ \tag{9b}$$

$$\tilde{\nu}_t^+ = \max\left(\nu_t^+, \lambda\tilde{\nu}_{t-1}^+ + (1-\lambda)\nu_t^+\right) \tag{9c}$$

$$1-\alpha = \int_{\nu_t^-}^{+\infty} p_t(x)dx \tag{9d}$$

$$\tilde{\nu}_0^- = \nu_0^- \tag{9e}$$

$$\tilde{\nu}_t^- = \min\left(\nu_t^-, \lambda\tilde{\nu}_{t-1} + (1-\lambda)\nu_t^-\right) \tag{9f}$$

Summary & Results: Part I (Opposition)

- We pointed to a puzzling observation: **Flash Crash**. More specifically as we have seen that during few the crash, fascinating **patterned oscillations** occurred in the commodities market.

Summary & Results: Part I (Opposition)

- We pointed to a puzzling observation: **Flash Crash**. More specifically as we have seen that during few the crash, fascinating **patterned oscillations** occurred in the commodities market. These oscillations **cannot be explained by the Top-Down** assumption in Quantitative Finance (e.g. the Brownian motion). We have proposed to study these oscillations with the **Bottom-Up** approach instead.

Summary & Results: Part I (Opposition)

- We pointed to a puzzling observation: **Flash Crash**. More specifically as we have seen that during few the crash, fascinating **patterned oscillations** occurred in the commodities market. These oscillations **cannot be explained by the Top-Down** assumption in Quantitative Finance (e.g. the Brownian motion). We have proposed to study these oscillations with the **Bottom-Up** approach instead.
- We expressed, **classic Financial Strategies in HFFF format** and shown the incentive for going from **from a simple perceptron, to shallow and finally deep learning**.

Summary & Results: Part I (Opposition)

- We pointed to a puzzling observation: **Flash Crash**. More specifically as we have seen that during few the crash, fascinating **patterned oscillations** occurred in the commodities market. These oscillations **cannot be explained by the Top-Down** assumption in Quantitative Finance (e.g. the Brownian motion). We have proposed to study these oscillations with the **Bottom-Up** approach instead.
- We expressed, **classic Financial Strategies in HFFF format** and shown the incentive for going from **from a simple perceptron, to shallow and finally deep learning**.
- We also established **connections** (though not complete) to fields that are traditionally associated **to mathematical biology**, namely **predator-prey models** and **evolutionary dynamics**.

Summary & Results: Part I (Opposition)

- We pointed to a puzzling observation: **Flash Crash**. More specifically as we have seen that during few the crash, fascinating **patterned oscillations** occurred in the commodities market. These oscillations **cannot be explained by the Top-Down** assumption in Quantitative Finance (e.g. the Brownian motion). We have proposed to study these oscillations with the **Bottom-Up** approach instead.
- We expressed, **classic Financial Strategies in HFFF format** and shown the incentive for going from **from a simple perceptron, to shallow and finally deep learning**.
- We also established **connections** (though not complete) to fields that are traditionally associated **to mathematical biology**, namely **predator-prey models** and **evolutionary dynamics**.
- This was done in order to express the bottom-up approach at the infinitesimal level. More specifically we developed the concept of **Path of Interaction** in an **HFTE Game**.

Summary & Results: Part I (Opposition)

- We pointed to a puzzling observation: **Flash Crash**. More specifically as we have seen that during few the crash, fascinating **patterned oscillations** occurred in the commodities market. These oscillations **cannot be explained by the Top-Down** assumption in Quantitative Finance (e.g. the Brownian motion). We have proposed to study these oscillations with the **Bottom-Up** approach instead.
- We expressed, **classic Financial Strategies in HFFF format** and shown the incentive for going from **from a simple perceptron, to shallow and finally deep learning**.
- We also established **connections** (though not complete) to fields that are traditionally associated **to mathematical biology**, namely **predator-prey models** and **evolutionary dynamics**.
- This was done in order to express the bottom-up approach at the infinitesimal level. More specifically we developed the concept of **Path of Interaction** in an **HFTE Game**.
- Finally we looked at tracking methods using **MTT**.

Future Research (Part I: Opposition)

- **Classification Simplification**: the direct simulation approach of an HFTE [23] creates situations in which two very different architectures yield the exact same function.

Future Research (Part I: Opposition)

- **Classification Simplification**: the direct simulation approach of an HFTE [23] creates situations in which two very different architectures yield the exact same function.
- **Complex Food Webs**: We need to bring the complexity of our state space to the level of a complex food web. Additional strategies must be incorporated and more HFTE games must be included in our database of scenarios and we need to incorporate **Birth & Death Processes**.

Future Research (Part I: Opposition)

- **Classification Simplification**: the direct simulation approach of an HFTE [23] creates situations in which two very different architectures yield the exact same function.
- **Complex Food Webs**: We need to bring the complexity of our state space to the level of a complex food web. Additional strategies must be incorporated and more HFTE games must be included in our database of scenarios and we need to incorporate **Birth & Death Processes**.
- **Order-Book Dynamics**: Many of the markets are driven by different rules for the OB.

Future Research (Part I: Opposition)

- **Classification Simplification**: the direct simulation approach of an HFTE [23] creates situations in which two very different architectures yield the exact same function.
- **Complex Food Webs**: We need to bring the complexity of our state space to the level of a complex food web. Additional strategies must be incorporated and more HFTE games must be included in our database of scenarios and we need to incorporate **Birth & Death Processes**.
- **Order-Book Dynamics**: Many of the markets are driven by different rules for the OB.
- **Increased HFFF complexity does not equate to Invasion**: A clear picture did not necessarily emerge from the first simulations.

Future Research (Part I: Opposition)

- **Classification Simplification**: the direct simulation approach of an HFTE [23] creates situations in which two very different architectures yield the exact same function.
- **Complex Food Webs**: We need to bring the complexity of our state space to the level of a complex food web. Additional strategies must be incorporated and more HFTE games must be included in our database of scenarios and we need to incorporate **Birth & Death Processes**.
- **Order-Book Dynamics**: Many of the markets are driven by different rules for the OB.
- **Increased HFFF complexity does not equate to Invasion**: A clear picture did not necessarily emerge from the first simulations.
- **Diversity as it relates to Stability**: In biology diversity in an ecosystem leads to its instability [35, 8] but what about Finance?.

Future Research (Part I: Opposition)

- **Classification Simplification**: the direct simulation approach of an HFTE [23] creates situations in which two very different architectures yield the exact same function.
- **Complex Food Webs**: We need to bring the complexity of our state space to the level of a complex food web. Additional strategies must be incorporated and more HFTE games must be included in our database of scenarios and we need to incorporate **Birth & Death Processes**.
- **Order-Book Dynamics**: Many of the markets are driven by different rules for the OB.
- **Increased HFFF complexity does not equate to Invasion**: A clear picture did not necessarily emerge from the first simulations.
- **Diversity as it relates to Stability**: In biology diversity in an ecosystem leads to its instability [35, 8] but what about Finance?.
- **Is TF the most "moral" strategy**? TF is similar to the TFT (it replicates but adapts).

Summary & Results: Part II (Apposition)

- **Descriptive approach** [28] of the market with the **Cointelation** diffusion model helps us understand **misleading risk measures** and introduced **inferred correlation** [21] as a **conservative alternative**.

Summary & Results: Part II (Apposition)

- **Descriptive approach** [28] of the market with the **Cointelation** diffusion model helps us understand **misleading risk measures** and introduced **inferred correlation** [21] as a **conservative alternative**.
- We also showed how **clustering** can help us in the context of **portfolio optimization** and **risk** management and more specifically the concept of **Anticipative Responsible VaR** [25]

Summary & Results: Part II (Apposition)

- **Descriptive approach** [28] of the market with the **Cointelation** diffusion model helps us understand **misleading risk measures** and introduced **inferred correlation** [21] as a **conservative alternative**.
- We also showed how **clustering** can help us in the context of **portfolio optimization** and **risk** management and more specifically the concept of **Anticipative Responsible VaR** [25]
- We have also shown how **hybrid methodology** between classic FM and ML can **outperform their individual sums**. More specifically, we try to solve our nonlinear partial differential equation with a **deep learning** [33] to **solve an SDE problem**.
- Pointed to **errors** in the FM literature when it comes to Implied Volatility **arbitrage** modelling and introduced **de-arbing** method,

Summary & Results: Part II (Apposition)

- **Descriptive approach** [28] of the market with the **Cointelation** diffusion model helps us understand **misleading risk measures** and introduced **inferred correlation** [21] as a **conservative alternative**.
- We also showed how **clustering** can help us in the context of **portfolio optimization** and **risk** management and more specifically the concept of **Anticipative Responsible VaR** [25]
- We have also shown how **hybrid methodology** between classic FM and ML can **outperform their individual sums**. More specifically, we try to solve our nonlinear partial differential equation with a **deep learning** [33] to **solve an SDE problem**.
- Pointed to **errors** in the FM literature when it comes to Implied Volatility **arbitrage** modelling and introduced **de-arbing** method,
- Enhanced the SVI [24, 22] with the **IVP** model [22, 26] designed to adjust exposed data driven limitation of the latter (**wings and liquidity**).

Future Research (Part II: Apposition)

- There are few extensions or improvements that can be performed on the optimization process for cointelated pairs research. We can first ask ourselves the question of the **n-Cointelated** case.

Future Research (Part II: Apposition)

- There are few extensions or improvements that can be performed on the optimization process for cointelated pairs research. We can first ask ourselves the question of the **n-Cointelated** case.
- **Testing of the IVP proxying methodology with quality data** (eg: does the $\rho_{S\&P}$ parameter compare well with the ρ_{SS}?).

Future Research (Part II: Apposition)

- There are few extensions or improvements that can be performed on the optimization process for cointelated pairs research. We can first ask ourselves the question of the **n-Cointelated** case.
- **Testing of the IVP proxying methodology with quality data** (eg: does the $\rho_{S\&P}$ parameter compare well with the ρ_{SS}?).
- **Application** of the above two (Cryptocurrency Option's market proxy; Bitcoin vs Altcoins) to the world of **Cryptocurrency**.

Future Research (Part II: Apposition)

- There are few extensions or improvements that can be performed on the optimization process for cointelated pairs research. We can first ask ourselves the question of the **n-Cointelated** case.
- **Testing of the IVP proxying methodology with quality data** (eg: does the $\rho_{S\&P}$ parameter compare well with the ρ_{SS}?).
- **Application** of the above two (Cryptocurrency Option's market proxy; Bitcoin vs Altcoins) to the world of **Cryptocurrency**.
- Current **De-Arbitraging** methodology is **not robust in between pillars** (eg: Interpolation).

Future Research (Part II: Apposition)

- There are few extensions or improvements that can be performed on the optimization process for cointelated pairs research. We can first ask ourselves the question of the **n-Cointelated** case.
- **Testing of the IVP proxying methodology with quality data** (eg: does the $\rho_{S\&P}$ parameter compare well with the ρ_{SS}?).
- **Application** of the above two (Cryptocurrency Option's market proxy; Bitcoin vs Altcoins) to the world of **Cryptocurrency**.
- Current **De-Arbitraging** methodology is **not robust in between pillars** (eg: Interpolation).
- **Particle Filter for Implied Volatility MTT**: very complex co-movements of 3 or more parameters are not taken into account. Though, these do not matter for vanilla options, they may matter for more complex exotics.

Future Research (Part II: Apposition)

- There are few extensions or improvements that can be performed on the optimization process for cointelated pairs research. We can first ask ourselves the question of the **n-Cointelated** case.
- **Testing of the IVP proxying methodology with quality data** (eg: does the $\rho_{S\&P}$ parameter compare well with the ρ_{SS}?).
- **Application** of the above two (Cryptocurrency Option's market proxy; Bitcoin vs Altcoins) to the world of **Cryptocurrency**.
- Current **De-Arbitraging** methodology is **not robust in between pillars** (eg: Interpolation).
- **Particle Filter for Implied Volatility MTT**: very complex co-movements of 3 or more parameters are not taken into account. Though, these do not matter for vanilla options, they may matter for more complex exotics.
- **Additional Liquidity issues** for Implied Volatility (the ATM Bid Ask is asset class sensitive) which makes the model less universal.

Future Research (Part II: Apposition)

- **Harmonizing Stochastic & Local Volatility**: We have seen that both the Heston and SVI models are popular in the industry and converge asymptotically to each other [13]: see Equation (10). **Are their limitations linked**?

$$dS_t = \sqrt{v_t} S_t dW_t^1, \qquad S_0 \in \mathbb{R}_+^* \tag{10a}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma v_t^{\frac{1}{2}} dW_t^2, \qquad v_0 \in \mathbb{R}_+^* \tag{10b}$$

$$d\left\langle W^1, W^2 \right\rangle_t = \rho dt, \tag{10c}$$

$$v(k,t) \to a + b[\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}] \tag{10d}$$

Future Research (Part II: Apposition)

- **Harmonizing Stochastic & Local Volatility**: We have seen that both the Heston and SVI models are popular in the industry and converge asymptotically to each other [13]: see Equation (10). **Are their limitations linked**?

$$dS_t = \sqrt{v_t} S_t dW_t^1, \qquad S_0 \in \mathbb{R}_+^* \tag{10a}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma v_t^{\frac{1}{2}} dW_t^2, \qquad v_0 \in \mathbb{R}_+^* \tag{10b}$$

$$d\left\langle W^1, W^2 \right\rangle_t = \rho dt, \tag{10c}$$

$$v(k,t) \to a + b[\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}] \tag{10d}$$

- **IVP and Assumed Correlation** of Equation (11) the answer?

$$dS_t = \sqrt{v_t} S_t dW_t^1, \qquad S_0 \in \mathbb{R}_+^* \tag{11a}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma \sqrt{v_t} dW_t^2, \qquad v_0 \in \mathbb{R}_+^* \tag{11b}$$

$$d\left\langle W^1, W^2 \right\rangle_t = \rho(t, S_t)dt, \tag{11c}$$

$$\rho(t, S_t) = \rho_+(t) + [\rho_-(t) - \rho_+(t)]\,[1 - \exp(-\beta(t)|S_t - K|)] \tag{11d}$$

$$v(k,t) \to a + b[\rho(z-m) + \sqrt{(z-m)^2 + \sigma^2}] \tag{11e}$$

Closing Statement with Q&A

– Russell L. Ackoff

[1]a quote that is generally attributed to George Box

[2]to mean, less useful for practitioners.

Closing Statement with Q&A

"The more efficient you are at doing the wrong thing, the wronger you become.

– Russell L. Ackoff

[1] a quote that is generally attributed to George Box

[2] to mean, less useful for practitioners.

Closing Statement with Q&A

"The more efficient you are at doing the wrong thing, the wronger you become. If you do the right thing wrong and correct it, you get better."

– Russell L. Ackoff

[1] a quote that is generally attributed to George Box

[2] to mean, less useful for practitioners.

Closing Statement with Q&A

"The more efficient you are at doing the wrong thing, the wronger you become. If you do the right thing wrong and correct it, you get better."

– Russell L. Ackoff

To that extend we know that **all models are wrong, but some are useful**[1] and in that spirit we have arguably done the **right thing wronger**[2] in the **1st** part of the thesis but the **wrong thing righter** in the **2nd** part.

[1] a quote that is generally attributed to George Box

[2] to mean, less useful for practitioners.

Al-Aradi, A., Correia, A., Naiff, D., and Jardim, G.
Solving nonlinear and high-dimensional partial differential equations via deep learning.

Al-Jarrah, O., Yoo, P., Muhaidat, S., Karagiannidis, G., and Taha, K.
Efficient machine learning for big data: A review.
87–93.

Axelrod, R.
The evolution of cooperation.
Basic Books (1984).

Axelrod, R.
The complexity of cooperation: Agent-based models of competition and collaboration.
Princeton University Press (1997).

Bouchaud, J.-P.
Economics needs a scientific revolution.
Nature 455 (2008), 1181.

Buchanan, M.
Meltdown modelling: Could agent-based computer models prevent another financial crisis?
Nature 460 (2009), 680–682.

Challet, D., Marsili, M., and Zhang, Y.-C.
Minority Games: Interacting Agents In Financial Markets.
Oxford Finance Series, 2014.

Chauvet, E., Paullet, J. E., Previte, J. P., and Walls, Z.
A Lotka-Volterra three-species food chain.
Mathematics Magazine 75 (2002), 243–255.

Cukier, K., and Mayer-Schoenberger, V.
The rise of big data: How it's changing the way we think about the world.
Foreign Affairs 92, 3 (2013), 28–40.

Doucet, A., de Freitas, N., and Gordon, N.

Sequential monte carlo in practice.
New York: Springer-Verlag (2001).

Doucet, A., Godsill, S., and Andrieu, C.
On sequential monte carlo sampling methods for bayesian filtering.
Statistics and Computing 10, 9 (2000), 197–208.

Farmer, J. D., and Foley, D.
The economy needs agent-based modelling.
Nature 460 (2009), 685–686.

Gatheral, J., and Jacquier, A.
Convergence of Heston to SVI.
Quantitative Finance 11 (2011), 1129–1132.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.
Generative adversarial networks.

Johnson, T. C.
Finance and mathematics: Where is the ethical malaise?
The Mathematical Intelligencer 37, 4 (Dec 2015), 8–11.

Lee, J.-G., and Minseo, K.
Geospatial big data: Challenges and opportunities.

Leonard, C.
Apres la crise, l'enseignement de la finance repense.

Liu, J., and Chen, R.
Sequential monte carlo methods for dynamic systems.
Journal of the American Statistical Association.

Lotka, A. J.
Elements of physical biology.
Williams & Wilkins Co. (1925).

MAHDAVI-DAMGHANI, B.
UTOPE-ia.
Wilmott Magazine 60 (2012), 28–37.

MAHDAVI-DAMGHANI, B.
The non-misleading value of inferred correlation: An introduction to the cointelation model.
Wilmott 67 (2013), 50—-61.

MAHDAVI-DAMGHANI, B.
Introducing the implied volatility surface parametrisation (IVP):application to the fx market.

MAHDAVI-DAMGHANI, B.
Introducing the HFTE model: a multi-species predator prey ecosystem for high frequency quantitative financial strategies.
Wilmott Magazine 89 (2017), 52–69.

MAHDAVI-DAMGHANI, B., AND KOS, A.
De-arbitraging with a weak smile.
Wilmott Magazine (2013).

MAHDAVI-DAMGHANI, B., AND ROBERTS, S.
A proposed risk modeling shift from the approach of stochastic differential equation towards machine learning clustering: Illustration with the concepts of anticipative & responsible VaR.
University of Oxford - Oxford-Man Institute of Quantitative Finance (2017).

MAHDAVI-DAMGHANI, B., AND ROBERTS, S.
Machine learning techniques for deciphering implied volatility surface data in a hostile environment: Scenario based particle filter, risk factor decomposition & arbitrage constraint sampling.
University of Oxford - Oxford-Man Institute of Quantitative Finance (2018).

MAHDAVI-DAMGHANI, B., AND ROBERTS, S.
A bottom-up approach to the financial markets agent-based quantitative algorithmic strategies: Ecosystem, dynamics & detection.

MAHDAVI-DAMGHANI, B., WELCH, D., O'MALLEY, C., AND KNIGHTS, S.

The misleading value of measured correlation.
Wilmott 62 (2012), 64—-73.

Nanex.
Strange days june 8'th, 2011 - natgas algo, 2011.

Nowak, M.
Evolutionary dynamics : exploring the equations of life.
2006.

Sidenbladh, H.
Multi-target particle filtering for the probability hypothesis density.
Proceedings 6th International Conference on Information Fusion (2003), 800–ñ806.

Sidenbladh, H., and Wirkander, S.
Particle filtering for finite random sets.
IEEE Transactions on Aerospace and Electronic Systems (2003).

Sirignano, J., and Spiliopoulos, K.
Dgm: A deep learning algorithm for solving partial differential equations.
Journal of Computational Physics 375 (2018), 1339 – 1364.

Villani, C.
Cedric villani: la simulation numérique est un enjeu majeur pour la societe, 2012.

Volterra, V.
Variazioni e fluttuazioni del numero d'individui in specie animali conviventi.
Mem. Acad. Lincei Roma (1926), 31–113.