# Deciphering Price Formation in the High Frequency Domain:

Systems & Evolutionary Dynamics as keys for construction of the High Frequency Trading Ecosystem.

## (working paper)

Babak Mahdavi-Damghani*    and Stephen Roberts[3]

Oxford-Man Institute of Quantitative Finance

*Abstract*—In this paper we propose a new approach to studying electronic trading & systemic risk by re-introducing the High Frequency Trading Ecosystem (HFTE) model [65]. We specify an approach in which agents interact through a topological structure designed to address the complexity demands of most common high frequency strategies but designed randomly at inception. This strategy ecosystem is then studied through a simplified genetic algorithm. The results open up intriguing social and regulatory implications which we propose to study through tracking methodologies and results from theoretical biology. The proposed simplifications aim to bring rigour compared to our first attempt at studying the problem.

*Keywords:* High Frequency Trading Ecosystem (HFTE), High Frequency Financial Funnel (HFFF), Multi-Target Tracking (MTT), Stability of Financial Systems, Markov Chain Monte Carlo (MCMC), Reversible-Jump, RJ-MCMC, Data Analysis and Patterns in Data, Electronic Trading, Systemic Risk, High Frequency Trading, Game Theory, Machine Learning, Predator Prey Models.

## I. INTRODUCTION

### A. Historical Context

After the subprime crisis of 2008 and the resulting social uproar, governments strongly pushed the regulators to develop more efficient risk monitoring systems[1]. The new candidate sector under question was that of algorithmic systematic trading which led to the flash crash of May 6, 2010, in which the Dow Jones Industrial Average lost almost 10% of its value in matter of minutes. However, the current state of the art risk models are the ones inspired by the last subprime crisis and are essentially models of networks in which each node can be impacted by the connected nodes through contagion [37] and is better suited to lower frequency models. Indeed, on 06/08/2011 a seemingly relatively unnoticed event occurred on the natural gas commodities market. We say "relatively unnoticed" simply because the monetary impact was limited and finance is unfortunately an industry in which warning signs are usually dismissed until it is too late. We can see from Figure 1 that clearly something non-random is occurring. This feeling is exacerbated by the strong intuition that only interacting agents falling into some sort of quagmire could yield such series of increasing oscillations followed by
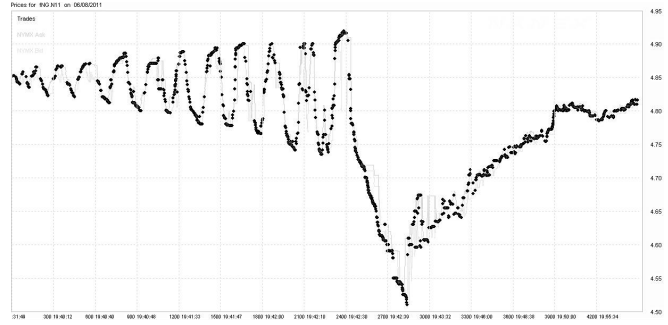


Fig. 1. Natural Gas flash crash of 06/08/2011 [76]

a mini crash. Indeed, commodities has historically been seen as a physical market, this in turn meaning that the prices are driven by supply and demand of commodities which can be consumed, stored and/or produced. This particular point is a unique feature compared to the other markets (Equities, FX, or Rate). Also this Figure 1 suggests that the common, though perhaps a bit lazy view, that crashes occur through totally unpredictable [105] events may not be true for algorithmic trading.

### B. Scientific method & parallel to Conway's Game of Life

In this paper we take an approach, similar in methodology, to Conway's Game of Life [27], a four-rule cellular automaton. We apply Conway's methodology to the world of High Frequency Trading (HFT) while adjusting some of the idiosyncratic parts of the exercise. As a reminder, Conway's Game of Life assumes that complexity in an ecosystem[2] arises from simple rules. For instance, these rules can lead to different families of automatons (when iterations are increased and the seed is random) such as:

- Stable forms[3]. Intuitively the reader may guess that the concept of financial stability may be related through a similar methodology.
- Oscillating forms: for example the "Blinker", the "Toad", the "Beacon[4]", the "Pulsar" and the "Pentadecathlon[5]". The concept of financial cycles, or HF

---

[2] "Ecosystem" and "Market" are used interchangeable in this paper.

[3] eg: the "Block", the "Beehive", the "Loaf", the "Boat" etc...

[4] all three, examples of two period iteration

[5] three and fifteen period iterations respectively.

oscillations such as of Figure 1 may be induced through a similar methodology.
- Moving forms[6] with different sizes and speeds[7].

The parallel to the world of quantitative financial strategies would be the following few points:
- interacting agents lead to market price fluctuations and, more specifically, their interaction determines the stability or instability of the market depending on what the market participating strategies involved as well as the evolving order-book.
- the market will follow the rules of a zero-player game[8] with, however, random seeds.
- agents (eg: strategies) will follow a rules for their births and deaths.

## C. Market & Orderbooks

*1) Caveat:* this paper assumes a simplification of the market: that is one product into one single possible market with few market participants who are unable to cheat the system through technology. In reality there exists a plethora of products in many markets in multiple geographical locations and the Securities & Exchange Commission (SEC) and the Financial Conduct Authority (FCA) expose new stories of cheats on daily basis. This approach may seem overly simplistic, but, we will see that this simplistic rule abiding approach may open up a new perceptive towards how people see and may want to take actions on the market.
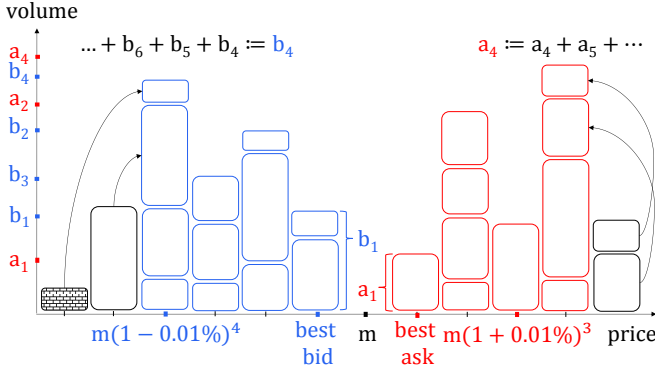


Fig. 2. Order-book visual representation

*2) Description:* traditional order books consists of a list of orders that a trading venue (such as an exchanges) uses to record the market participants' interests in a particular financial product. Typically a rule-based algorithm records these interests, taking into account the price & the volume proposed (on either side of the bid ask) as well as the time that interest was recorded.

**Definition** We define $a_1^t$ and $b_1^t$ the best ask & bid total volumes at time $t$. By extension $a_i^t$, $b_i^t$ with $i \in \{1, 2, 3, 4\}$

correspond to total volume at the relevant depth in the order book with the special case where $i = 4$ which represents the total volume at the 4th level and beyond. We will call $m_t$ the mid price of the best bid/ask at time $t$.

Figure 2 represent an order book which the previous definition aims at describing.

*3) Variable Definition:*

**Definition** We will label by $\{y_i\}_{i=0}^{n-1}$ the price process of interest, $i \in [0, n]$ its discretized 500ms snapshots with $i = 0$ being the most recent snapshot and $i = n$ its most distant snapshot. Moreover we will assume here that 500ms is enough time for the trading system to take the data, reformat it, analyze it as well allow the relevant strategy to take actions. Similarly we will define $\{x_{j,1}, x_{j,2}, \ldots x_{j,p}\}_{j=i+1}^{n}$ the relevant, $p$ leading indicators to the price dynamic of interest.

**Remark** We will assume that the leading indicators[9] for the price process can only be taken from the order book which is a reasonable assumption in the higher frequencies. Some accepted leading indicators are listed below:
- The price of the underlier itself
- The accumulated volume at different market depths of the order book (4 of the bid side and 4 on the ask side for a total of 9 leading indicators with the price process: see Figure 2).

## D. Problem Formulation & Agenda

*1) Problem Formulation:* the connection between machine learning and high frequency trading (HFT) has long been implicitly established via the numerous systematic trading positions available in most job searching tools (eFinancialCareers, LinkedIn, etc). It is, however, unclear which of the numerous machine learning techniques is most relevant to what high frequency traders wish to accomplish. The field of machine learning itself is rich; genetic algorithms, algorithmic game theory, state space models, Kalman filters, sequential Monte Carlo methods, support vector machines, neural networks or even a simple multi-linear regression are some of the key words mentioned. However, what most of these methodologies have in common is that they assume a pattern inherent to the market itself as opposed to taking the market as a consequence of the strategies composing this market[10].

**Remark** An interesting analogy can be made with respect to how the gene centered view of evolution (as opposed to the individual centered view of evolution) completely reshuffled our understanding of natural selection and gave the opportunity to see altruism at a different light. By analogy, we are trying to communicate the idea that the market centered view of financial systems is the wrong way to understand the fluctuation of the market and that the

---

[6]eg: the "Glider" and the "LWSS (Lightweight spaceship)" etc...

[7]idiosyncratic properties from the Game of Life, which parallel to our problem is not necessarily transferable.

[8]Meaning that its evolution is determined by its initial state, requiring no further input.

[9]A leading indicator is a measurable financial/economic factor that changes before the variable which is the object of the forecast (eg: a price) starts to follow a particular pattern.

[10]Top-Down vs Bottom-Up approach

strategy centered view of the financial system provides the opportunity to look at the market differently.

*2) Agenda:* We will first go through a literature review of mathematical methods for tracking is section II, a shorter and perhaps less broad literature review of relevant theoretical biology in section III. We will then summarize the relevant points associated to the High Frequency Trading Ecosystem (HFTE) model recently introduced [65] in section IV. Finally in section V-B we will discuss the current and future anticipated research associated to the problem of interest.

## II. REVIEW OF INFERENCE AND DYNAMICAL MODELS

### A. Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) algorithms [72] sample from a probability distribution based on a Markov chain that has a desired equilibrium distribution, the quality of the sample improving at each additional iteration. We will see next few version of the MCMC algorithm.

*1) Metropolis-Hastings algorithm:* The Metropolis-Hastings algorithm is a MCMC method that aims at obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult [73] and initially advertised of high dimensions. We will see in the next few algorithm examples that the methodology is now classified as useful for low dimensional problems. At each iteration $x_t$, the proposal next point $x'$ is sampled through a proposed distribution $g(x'|x_t)$. We then calculate:

- with $a_1 = \frac{P(x')}{P(x_t)}$ is the the probability ratio between the proposed sample and the previous sample,
- and $a_2 = \frac{g(x_t|x')}{g(x'|x_t)}$, the ratio of the proposal density in both directions[11].

and set $a = \max(a_1 a_2, 1)$, we then accept $x_{t+1} = x'$ if $r \sim U[0,1] \geqslant a$ which essentially means that if $a = 1$, accept is always true otherwise you accept with a probability $a_1 a_2$. The algorithm works best if the proposal distribution is similar to the real distribution. Note that the seed is slowly forgotten as the number of iterations increases.

*2) Gibbs sampling:* Perhaps one of the simplest MCMC algorithms, the Gibbs Sampling (GS) algorithm was introduced in by Geman & Geman [29] with the application of image processing. Later it was discussed in the context of missing data problems [106]. The benefice of the Gibbs algorithm for Bayesian analysis was demonstrated in Tanner and Wong [106]. To define the Gibbs sampling algorithm, let the set of full conditional distributions be: $\pi(\psi_1|\psi_2, \ldots, \psi_p)$, $\ldots$, $\pi(\psi_d|\psi_1, \psi_2, \ldots, \psi_{d-1}, \psi_{d+1}, \ldots, \psi_p)$, $\ldots$, $\pi(\psi_p|\psi_1, \ldots, \psi_{p-1})$. One cycle of the GS, described in algorithm 1, is completed by sampling $\{\psi_k\}_{k=1}^p$ from the mentioned distributions, in sequence and refreshing the conditioning variables. When $d$ is set to 2 we obtain the two block Gibbs sampler described by Tanner & Wong [106]. If we take general conditions, the chain generated by the GS converges to the target density as the number of iterations

---

**Algorithm 1** GIBBS-SAMPLING($\boldsymbol{\psi}_1^{(0)}, \ldots, \boldsymbol{\psi}_p^{(0)}$)

---

**Require:** Specify an initial value $\boldsymbol{\psi}^{(0)} = \left(\boldsymbol{\psi}_1^{(0)}, \ldots, \boldsymbol{\psi}_p^{(0)}\right)$
**Ensure:** $\{\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}, \ldots, \boldsymbol{\psi}^{(M)}\}$

1: **for** $j = 1, 2, \ldots, M$ **do**
2:      Generate $\boldsymbol{\psi}_1^{(j+1)}$ from $\pi\left(\boldsymbol{\psi}_1|\boldsymbol{\psi}_2^{(j)}, \boldsymbol{\psi}_3^{(j)}, \ldots, \boldsymbol{\psi}_p^{(j)}\right)$
3:      Generate $\boldsymbol{\psi}_2^{(j+1)}$ from $\pi\left(\boldsymbol{\psi}_2|\boldsymbol{\psi}_1^{(j+1)}, \boldsymbol{\psi}_3^{(j)}, \ldots, \boldsymbol{\psi}_p^{(j)}\right)$
4:      $\vdots$
5:      Generate $\boldsymbol{\psi}_d^{(j+1)}$ from $\pi(\boldsymbol{\psi}_d|\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{d-1}, \boldsymbol{\psi}_{d+1}, \ldots, \boldsymbol{\psi}_p)$.
6:      $\vdots$
7:      Generate $\boldsymbol{\psi}_p^{(j+1)}$ from $\pi\left(\boldsymbol{\psi}_p|\boldsymbol{\psi}_1^{(j+1)}, \ldots, \boldsymbol{\psi}_{p-1}^{(j+1)}\right)$
8:      Return the values $\{\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}, \ldots, \boldsymbol{\psi}^{(M)}\}$
9: **end for**

---

goes towards infinity. The main drawback with this method however is its relative computational heavy aspect because of the burn-in period.

*3) Hamiltonian Monte Carlo:* Hamiltonian Monte Carlo [21], sometimes also referred to[12] as hybrid Monte Carlo is an MCMC method for obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult. It serves to address the limitations of the Metropolis-Hastings algorithm by adding few more parameters that aim is to reduce the correlation between successive samples using a Hamiltonian evolution process and also by targeting states with a higher acceptance rate.

*4) Ordered Overrelaxation:* Overrelaxation is usually a term associated with a Gibbs Sampler but in the context of this subsection we discuss Ordered Overrelaxation. The methodology aims at addressing the slowness associated in performing a random walk with inappropriately selected step sizes. The latter problem was addressed by incorporating a momentum parameter which consist of sampling $n$ random variables (20 is considered a good [63] number for $n$), sorting them from biggest to smallest, looking where $x_t$ ranks, say at $p$'s position, amongst the $n$ variables and the picking $n-p$ for the subsequent sample $x_{t+1}$ [77]. This form of optimal "momentum" parameter design is a central pillar of research in MCMC.

*5) Slice sampling:* Slice sampling is one of the remarkably simple methodologies [77] of MCMC which can be considered as a mix of Gibbs sampling, Metropolis-Hastings and rejection sampling methods. It assumes that the target density $P^*(x)$ can be evaluated at any point $x$ but is more robust compared to the Metropolis-Hastings especially when if comes to step size. Like rejection sampling it draws samples from the volume under the curve. The idea of the algorithm is that it switches vertical and horizontal uniform

---

[11]equal to 1 is the proposal density is symmetric

[12]though more in the past.

sampling by starting horizontally, then vertically performing "slices" based on the current vertical position. MacKay made good contributions in its visual [63] representation.

*6) Multiple-try Metropolis:* One way to address the curse of dimensionality is the Multiple-try Metropolis which can be though of as a enhancement of the Metropolis-Hastings algorithm. The former allows multiple trials at each point instead of one by the latter. By increasing both the step size and the acceptance rate, the algorithm helps the convergence rate of the sampling trajectory [60]. The curse of dimensionality is another central area of research for MCMCs.

*7) Reversible-Jump:* Another variant of the Metropolis-Hastings, and perhaps most promising methodology when it comes to our application is the Reversible-jump MCMC (RJ-MCMC) developed by Green [35]. One key factor or RJ-MCMC is that it is designed to address changes of dimensionality issues. In our case, as we saw in section III of "Paper Format" document, we face a dual type issues around change of dimensionality. The first being the frequency of each strategy in an ecosystem and the second element being the HFFF[13] which branching structure and size changes as a function of the strategy[14]. More formally. Let us define $n_m \in N_m = \{1, 2, \ldots, I\}$, as our model indicator and $M = \bigcup_{n_m=1}^{I} \mathbb{R}^{d_m}$ the parameter space whose number of dimensions $d_m$ is function of model $n_m$ (with our model indicators not needing to be finite). The stationary distribution is the joint posterior distribution of $(M, N_m)$ that takes the values $(m, n_m)$. The proposal $m'$ can be constructed with a mapping $g_{1mm'}$ of $m$ and $u$, where $u$ is drawn from a random component $U$ with density $q$ on $\mathbb{R}^{d_{mm'}}$. The move to state $(m', n'_m)$ can thus be formulated as $(m', n'_m) = (g_{1mm'}(m, u), n'_m)$. The function $g_{mm'} := (m, u) \mapsto (m', u')$, with $(m', u') = (g_{1mm'}(m, u), g_{2mm'}(m, u))$ must be one to one and differentiable, and have a non-zero support: $\text{supp}(g_{mm'}) \neq \varnothing$, in order to enforce the existence of the inverse function $g_{mm'}^{-1} = g_{m'm}$, that is differentiable. Consequently $(m, u)$ and $(m', u')$ must have the same dimension, which is enforced if the dimension criterion $d_m + d_{mm'} = d_{m'} + d_{m'm}$ is verified ($d_{mm'}$ is the dimension of $u$). This criterion is commonly referred to as dimension matching. Note that if $\mathbb{R}^{d_m} \subset \mathbb{R}^{d_{m'}}$ then the dimensional matching condition can be reduced to $d_m + d_{mm'} = d_{m'}$, with $(m, u) = g_{m'm}(m)$. The acceptance probability is given by $a(m, m') = \min\left(1, \frac{p_{m'm}p_{m'}f_{m'}(m')}{p_{mm'}q_{mm'}(m,u)p_m f_m(m)}\left|\det\left(\frac{\partial g_{mm'}(m,u)}{\partial(m,u)}\right)\right|\right)$, where $p_m f_m$, the posterior probability is given by $c^{-1}p(y|m, n_m)p(m|n_m)p(n_m)$ with $c$ being the normalising constant. Many problems in data analysis require the unsupervised partitioning. Roberts, Holmes and Denison [89] re-considered the issue of data partitioning from an information-theoretic viewpoint and shown that minimisation of partition entropy may be used to evaluate the most probable set of data generators which can be employed

---

[13]See figure 6 of "Paper Format" document for more information.
[14]See in section III and Figures 6, 9, 10 and 11.

using a RJ-MCMC.

*B. Dynamical Linear Methods*

Multi-Target Tracking (MTT) which deals with state space estimation of moving targets has applications in different fields [6], [57], [103], the most intuitive ones being perhaps of radar and sonar function.

*1) Kalman Filter:* The Kalman Filter (KF) is a mathematical tool which purpose is to make the best estimation in a Mean Square Error (MSE) sense of some dynamical process, $(x_k)$, perturbed by some noise and influenced by a controlled process. For the sake of our project we will assume that the controlled process is null but will still incorporate it in the general state in order to fully understand the model. The estimation is done via observations which are functions of these dynamics $(y_k)$. Roweis and Ghahramani made a quality review [91] of the topic. The dynamics of the KF is usually referred in the literature as $x_k$ and given by equation (1).

$$x_k = F_k x_{k-1} + B_k u_k + w_k \qquad (1)$$

with $F_k$ is the state transition model which is applied to the previous state $x_{k-1}$; $B_k$ is the control-input model which is applied to the vector $u_k$ (often taken as the null vector); $w_k$ is the process noise which is assumed to be drawn from a zero mean multivariate normal distribution with covariance $Q_k$ and $w_k \sim N(0, Q_k)$. At time $k$ an observation of $x_k$, $y_k$ is made according to equation (2).

$$y_k = H_k x_k + v_k \qquad (2)$$

where $H_k$ is the observation model which maps the true state space into the observed space. $v_k$ is the observation noise which is assumed to be zero mean Gaussian white noise with $v_k \sim N(0, R_k)$. We also assume that the noise vectors $(\{w_1, \ldots, w_k\}, \{v_1 \ldots v_k\})$ at each step are all assumed to be mutually independent ($cov(v_k, w_k) = 0$ for all $k$). The KF being a recursive estimator, we only need the estimated state from the previous time step and the current measurement to compute the estimate for the current state. $\hat{x_k}$ will represent the estimation of our state $x_k$ at time up to $k$. The state of our filter is represented by two variables: $\hat{x}_{k|k}$, the estimate of the state at time $k$ given observations up to and including time $k$; $P_{k|k}$, the error covariance matrix (a measure of the estimated accuracy of the state estimate). The KF has two distinct phases: Predict and Update. The predict phase uses the state estimate from the previous timestep to produce an estimate of the state at the current timestep. In the update phase, measurement information at the current timestep is used to refine this prediction to arrive at a new, more accurate state estimate, again for the current timestep. The formula for the updated estimate covariance above is only valid for the optimal Kalman gain. Usage of other gain values require a more complex formula. Below we present a partial proof of the KF algorithm [47], [48].
*Proof:* The second line of the algorithm is derived the following way: $\hat{x}_{k|k-1} = \mathrm{E}[x_k] = \mathrm{E}[F_k x_{k-1} + B_k u_k + w_k] = F_k \hat{x}_{k-1|k-1} + B_{k-1}u_{k-1}$. The third line of the algorithm is derived the following way: $P_{k|k-1} =$

**Algorithm 2** KALMAN-FILTER($w$)

**Require:** array of weights $w_1^N$
**Ensure:** array of weights $w_1^M$ resampled

1: //**Predicted state**:
2: $\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1}$
3: $P_{k|k-1} = F_k P_{k-1|k-1} F_k^{\mathrm{T}} + Q_{k-1}$
4: //**Update state**:
5: //Innovation (or residual)
6: $\tilde{y}_k = y_k - H_k \hat{x}_{k|k-1}$
7: //Covariance
8: $S_k = H_k P_{k|k-1} H_k^T + R_k$
9: //Optimal Kalman gain
10: $K_k = P_{k|k-1} H_k^{\mathrm{T}} S_k^{-1}$
11: //Updated state estimate
12: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k$
13: //Updated estimate covariance
14: $P_{k|k} = (I - K_k H_k) P_{k|k-1}$

---

$$\mathrm{E}\left[x_k x_k\right] = \mathrm{E}\left[F_k \underbrace{x_{k-1|k-1} x_{k-1|k-1}^{\mathrm{T}}}_{\mathrm{E}\left[x_{k-1|k-1} x_{k-1|k-1}^{\mathrm{T}}\right] = P_{k|k-1}} F_k^{\mathrm{T}}\right]$$

$$+ 2\underbrace{\mathrm{E}\left[F_k x_{k-1|k-1} B_k u_k\right]}_{0} + 2\underbrace{\mathrm{E}\left[F_k x_{k-1|k-1} w_k\right]}_{0} +$$

$$2\underbrace{\mathrm{E}\left[B_k u_k w_k\right]}_{0} + \underbrace{\mathrm{E}\left[w_k w_k^{\mathrm{T}}\right]}_{Q_k} = F_k P_{k-1|k-1} F_k^{\mathrm{T}} + Q_{k-1}.$$ The

8th line is derived the following way: $S_k = \mathrm{E}\left[y_k y_k\right] =$

$$\mathrm{E}\left[H_k \underbrace{x_k x_k^{\mathrm{T}}}_{\mathrm{E}\left[x_{k-1|k-1} x_{k-1|k-1}^{\mathrm{T}}\right] = P_{k|k-1}} H_k^{\mathrm{T}}\right] + 2\underbrace{\mathrm{E}\left[H_k x_k v_k\right]}_{0} +$$

$$\underbrace{\mathrm{E}\left[v_k v_k\right]}_{R_k} = H_k P_{k|k-1} H_k^{\mathrm{T}} + R_k.$$

As for the Kalman Gain, we first rearrange some of the equations in a more useful form. First, with the error covariance $P_{k|k}$ as above $P_{k|k} = \mathrm{cov}(x_k - \hat{x}_{k|k})$ and substitute in the definition of $\hat{x}_{k|k}$ $P_{k|k} = \mathrm{cov}(x_k - (\hat{x}_{k|k-1} + K_k \tilde{y}_k))$ and substitute $\tilde{y}_k$. $P_{k|k} = \mathrm{cov}(x_k - (\hat{x}_{k|k-1} + K_k(y_k - H_k \hat{x}_{k|k-1})))$. $P_{k|k} = \mathrm{cov}(x_k - (\hat{x}_{k|k-1} + K_k(H_k x_k + v_k - H_k \hat{x}_{k|k-1})))$, now by collecting the error vectors we get $P_{k|k} = \mathrm{cov}((I - K_k H_k)(x_k - \hat{x}_{k|k-1}) - K_k v_k)$. Given that the measurement error $v_k$ is uncorrelated with the other terms, we have $P_{k|k} = \mathrm{cov}((I - K_k H_k)(x_k - \hat{x}_{k|k-1})) + \mathrm{cov}(K_k v_k)$, now by the properties of vector covariance this becomes $P_{k|k} = (I - K_k H_k)\mathrm{cov}(x_k - \hat{x}_{k|k-1})(I - K_k H_k)^{\mathrm{T}} + K_k \mathrm{cov}(v_k) K_k^{\mathrm{T}}$ which, using our invariance on $P_{k|k-1}$ and the definition of $R_k$ becomes $P_{k|k} = (I - K_k H_k) P_{k|k-1}(I - K_k H_k)^{\mathrm{T}} + K_k R_k K_k^{\mathrm{T}}$. This rearrangement is known in the literature as the Joseph form of the covariance equation, which is true independently of $K_k$. Now if $K_k$ is the optimal Kalman gain, we can simplify further. The Kalman filter is a minimum MSE estimator.

The error is $x_k - \hat{x}_{k|k}$. We would like to minimize the expected value of the square of the magnitude of this vector, $\mathrm{E}\left[|x_k - \hat{x}_{k|k}|^2\right]$. This idea is equivalent to minimizing the trace of the posterior estimate covariance matrix $P_{k|k}$. By expanding out the terms in the equation above and rearranging, we get: $P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^{\mathrm{T}} K_k^{\mathrm{T}} + K_k(H_k P_{k|k-1} H_k^{\mathrm{T}} + R_k) K_k^{\mathrm{T}} = P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^{\mathrm{T}} K_k^{\mathrm{T}} + K_k S_k K_k^{\mathrm{T}}$. The trace is minimized when the matrix derivative is zero: $\frac{\partial \, \mathrm{tr}(P_{k|k})}{\partial \, K_k} = -2(H_k P_{k|k-1})^{\mathrm{T}} + 2 K_k S_k = 0$. Solving this for $K_k$ yields the Kalman gain: $K_k S_k = (H_k P_{k|k-1})^{\mathrm{T}} = P_{k|k-1} H_k^{\mathrm{T}}$ $K_k = P_{k|k-1} H_k^{\mathrm{T}} S_k^{-1}$. This optimal Kalman gain, is the one that yields the best estimates when used. The formula used to calculate the posterior error covariance can be simplified when the Kalman gain equals the optimal value derived above. Multiplying both sides of our Kalman gain formula on the right by $S_k K_k^T$, it follows that $K_k S_k K_k^T = P_{k|k-1} H_k^T K_k^T$. Referring back to our expanded formula for the posterior error covariance, $P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^T K_k^T + K_k S_k K_k^T$ we find that the last two terms cancel out, giving $P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} = (I - K_k H_k) P_{k|k-1}$. This formula is low latency and thus usually used. One should keep in mind that it is only correct for the optimal gain though. ∎

*2) Extended Kalman Filter:* The EKF is essentially an approximation of the KF for non-severely-non-linear models which linearises about the current mean and covariance, so that the state transition and observation models need not be linear functions of the state but may instead be differentiable functions. The dynamics and measurements of this equation is presented in (3).

$$\begin{cases} x_k & = f(x_{k-1}, u_k) + w_k \\ y_k & = h(x_k) + v_k \end{cases} \tag{3}$$

The algorithm is very similar to the one described in

---

**Algorithm 3** EXTENDED-KALMAN-FILTER($w$)

**Require:** array of weights $w_1^N$
**Ensure:** array of weights $w_1^M$ resampled

1: //**Predicted state**:
2: $\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}, u_k)$
3: $P_{k|k-1} = F_k P_{k-1|k-1} F_k^{\mathrm{T}} + Q_{k-1}$
4: //**Update state**:
5: //Innovation (or residual)
6: $\tilde{y}_k = y_k - h(\hat{x}_{k|k-1})$
7: //Covariance
8: $S_k = H_k P_{k|k-1} H_k^{\mathrm{T}} + R_k$
9: //Optimal Kalman gain
10: $K_k = P_{k|k-1} H_k^{\mathrm{T}} S_k^{-1}$
11: //Updated state estimate
12: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k$
13: //Updated estimate covariance
14: $P_{k|k} = (I - K_k H_k) P_{k|k-1}$

---

algorithm 2 but with couple of modifications highlighted

below algorithm 3[15].

*Proof:* The proof for algorithm 3 is very similar to the proof of algorithm 2 with couple of exceptions. First, $F_k$ and $H_k$ are approximations of first order of $F_k$ and $H_k$. Second, we get a truncation error which can be bounded and satisfies the inequality known as Cauchy's estimate: $|R_n(x)| \leqslant M_n \frac{r^{n+1}}{(n+1)!}$, here $(a - r, a + r)$ is the interval where the variable $x$ is assumed to take its values and $M_n$ positive real constant such that $|f^{(n+1)}(x)| \leqslant M_n$ for all $x \in (a-r, a+r)$. $M_n$ gets bigger as the curvature or non-linearity gets more severe. When this error increases it is possible to improve our approximation at the cost of complexity by increasing by one degree our Taylor approximation, i.e: $F_k = \frac{\partial f}{\partial x}\big|_{f(\hat{x}_{k-1|k-1}, u_k)} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}\big|_{f(\hat{x}_{k-1|k-1}, u_k)^2}$ and $H_k = \frac{\partial h}{\partial x}\big|_{f(\hat{x}_{k|k-1})} + \frac{1}{2} \frac{\partial h}{\partial x}\big|_{f(\hat{x}_{k|k-1})^2}$. ∎

**Remark** Though the EKF tries to address some of the limitations of the KF by relaxing some of the linearity constraints it still needs to assume that the underlying function dynamics are both known and derivable. This particular point is not at all desirable in many applications.

*C. Dynamical Non-linear methods*

*1) Sequential Monte Carlo methods:* Sequential Monte Carlo methods (SMC) [18], [59] known alternatively as Particle Filters (PF) [31], [52] or also seldom CONDENSATION [45], are statistical model estimation techniques based on simulation. They are the sequential (or 'on-line') analogue of Markov Chain Monte Carlo (MCMC) methods and similar to importance sampling methods. If they are elegantly designed they can be much faster than MCMC. Because of their non linear quality they are often an alternative to the Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF). They however have the advantage of being able to approach the Bayesian optimal estimate with sufficient samples. They are technically more accurate than the EKF or UKF. The aims of the PF is to estimate the sequence of hidden parameters, $x_k$ for $k = 1, 2, 3, \ldots$, based on the observations $y_k$. The estimates of $x_k$ are done via the posterior distribution $p(x_k|y_1, y_2, \ldots, y_k)$. PF do not care about the full posterior $p(x_1, x_2, \ldots, x_k|y_1, y_2, \ldots, y_k)$ like it is the case for the MCMC or importance sampling (IS) approach. Let's assume $x_k$ and the observations $y_k$ can be modeled in the following way:

- $x_k|x_{k-1} \sim p_{x_k|x_{k-1}}(x|x_{k-1})$ and with given initial distribution $p(x_1)$.
- $y_k|x_k \sim p_{y|x}(y|x_k)$.
- equations (4) and (5) gives an example of such system.

$$x_k = f(x_{k-1}) + w_k \qquad (4)$$
$$y_k = h(x_k) + v_k \qquad (5)$$

It is also assumed that $cov(w_k, v_k) = 0$ or $w_k$ and $v_k$ mutually independent and iid with known probability density

---
[15]Note that here $F_k = \frac{\partial f}{\partial x}\big|_{\hat{x}_{k-1|k-1}, u_k}$ and $H_k = \frac{\partial h}{\partial x}\big|_{\hat{x}_{k|k-1}}$.

functions. $f(\cdot)$ and $h(\cdot)$ are also assumed known functions. Equations (4) and (5) are our state space equations. If we define $f(\cdot)$ and $h(\cdot)$ as linear functions, with $w_k$ and $v_k$ both Gaussian, the KF is the best tool to find the exact sought distribution. If $f(\cdot)$ and $h(\cdot)$ are non linear then the Kalman filter (KF) is an approximation. PF are also approximations, but convergence can be improved with additional particles. PF methods generate a set of samples that approximate the filtering distribution $p(x_k|y_1, \ldots, y_k)$. If $N_P$ in the number of samples, expectations under the probability measure are approximated by equation (6).

$$\int f(x_k)p(x_k|y_1, \ldots, y_k)dx_k \approx \frac{1}{N_P} \sum_{L=1}^{N_P} f(x_k^{(L)}) \qquad (6)$$

Sampling Importance Resampling (SIR) is the most com-

---

**Algorithm 4** RESAMPLE($w$)

**Require:** array of weights $w_1^N$
**Ensure:** array of weights $w_1^M$ resampled

1: $u^0 \sim \mathcal{U}[0, 1/M]$
2: **for** $m = 1$ to $N$ **do**
3: $\quad i^{(m)} \leftarrow \left\lfloor (w_n^{(m)} - u^{(m-1)}m) \right\rfloor + 1$
4: $\quad u^{(m)} = u^{(m)} + \frac{i^{(m)}}{M} - w_n^{(m)}$
5: **end for**

---

monly used PF algorithm, which approximates the probability measure $p(x_k|y_1, \ldots, y_k)$ via a weighted set of $N_P$ particles

$$\left( w_k^{(L)}, x_k^{(L)} \right) \; : \; L = 1, \ldots, N_P \qquad (7)$$

The importance weights $w_k^{(L)}$ are approximations to the relative posterior probability measure of the particles such that $\sum_{L=1}^{P} w_k^{(L)} = 1$. SIR is a essentially a recursive version of importance sampling. Like in IS, the expectation of a function $f(\cdot)$ can be approximated like described in equation (8).

$$\int f(x_k)p(x_k|y_1, \ldots, y_k)dx_k \approx \sum_{L=1}^{N_P} w^{(L)} f(x_k^{(L)}) \qquad (8)$$

The algorithm performance is dependent on the choice of the proposal distribution $\pi(x_k|x_{1:k-1}, y_{1:k})$ with the optimal proposal distribution being $\pi(x_k|x_{0:k-1}, y_{0:k})$ in equation (9).

$$\pi(x_k|x_{1:k-1}, y_{1:k}) = p(x_k|x_{k-1}, y_k) \qquad (9)$$

Because it is easier to draw samples and update the weight calculations the transition prior is often used as importance function.

$$\pi(x_k|x_{1:k-1}, y_{1:k}) = p(x_k|x_{k-1})$$

The technique of using transition prior as importance function is commonly known as Bootstrap Filter and Condensation Algorithm. Figure 3 gives an illustration of the algorithm just described. Note that on line 5 of algorithm 5, $\hat{w}_k^{(L)}$, simplifies to $w_{k-1}^{(L)} p(y_k|x_k^{(L)})$, when $\pi(x_k^{(L)}|x_{1:k-1}^{(L)}, y_{1:k}) = p(x_k^{(L)}|x_{k-1}^{(L)})$.

6

**Algorithm 5** SMC($w$)

---

**Require:** array of weights $w_p^N$, $\pi(x_k|x_{1:k-1}^{(L)}, y_{1:k})$
**Ensure:** array of weights $w_p^N$ resampled

---

1: **for** $L = 1$ to $N_P$ **do**
2: $\quad x_k^{(L)} \sim \pi(x_k|x_{1:k-1}^{(L)}, y_{1:k})$
3: **end for**
4: **for** $L = 1$ to $N_P$ **do**
5: $\quad \hat{w}_k^{(L)} = w_{k-1}^{(L)} \frac{p(y_k|x_k^{(L)})p(x_k^{(L)}|x_{k-1}^{(L)})}{\pi(x_k^{(L)}|x_{1:k-1}^{(L)}, y_{1:k})}$
6: **end for**
7: **for** $L = 1$ to $N_P$ **do**
8: $\quad w_k^{(L)} = \frac{\hat{w}_k^{(L)}}{\sum_{J=1}^{P} \hat{w}_k^{(J)}}$
9: **end for**
10: $\hat{N}_{eff} = \frac{1}{\sum_{L=1}^{P}\left(w_k^{(L)}\right)^2}$
11: **if** $\hat{N}_{eff} < N_{thr}$ **then**
12: $\quad$ resample: draw $N_P$ particles from the current particle set with probabilities proportional to their weights. Replace the current particle set with this new one.
13: $\quad$ **for** $L = 1$ to $N_P$ **do**
14: $\quad\quad w_k^{(L)} = 1/N_P$.
15: $\quad$ **end for**
16: **end if**

---

*2) Resampling Methods:* Resampling methods are usually used to avoid the problem of weight degeneracy in our algorithm. Avoiding situations where our trained probability measure tends towards the Dirac distribution must be avoided because it really does not give much information on all the possibilities of our state. There exists many different resampling methods, Rejection Sampling, Sampling-Importance Resampling, Multinomial Resampling, Residual Resampling, Stratified Sampling, and the performance of our algorithm can be affected by the choice of our resampling method. The stratified resampling proposed by Kitagawa [53] is optimal in terms of variance. Figure 3 gives an illustration of the Stratified Sampling and the corresponding algorithm is described in algorithm 4. We see at the top of the figure 3 the discrepancy between the estimated pdf at time $t$ with the real pdf, the corresponding CDF of our estimated PDF, random numbers from $[0,1]$ are drawn, depending on the importance of these particles they are moved to more useful places.

*3) Importance Sampling :* Importance sampling (IS) was first introduced in [69] and was further discussed in several books including in [38]. The objective of importance sampling is to sample the distribution in the region of importance in order to achieve computational efficiency via lowering the variance. The idea of importance sampling is to choose a proposal distribution $q(x)$ in place of the true, harder to sample probability distribution $p(x)$. The main constraint is related to the support of $q(x)$ which is assumed to cover that

of $p(x)$. In equation (10) we write the integration problem in the more appropriate form.

$$\int f(x)p(x)dx = \int f(x)\frac{p(x)}{q(x)}q(x)dx \qquad (10)$$

In IS the number, $N_p$, usually describes the number of independent samples drawn from $q(x)$ to obtain a weighted sum to approximate $\hat{f}$ in equation (11).

$$\hat{f} = \frac{1}{N_p}\sum_{i=1}^{N_p} W\left(x^{(i)}\right) f\left(x^{(i)}\right) \qquad (11)$$

where $W(x^{(i)})$ is the Radon-Nikodym derivative of $p(x)$ with respect to $q(x)$ or called in engineering the importance weights (equation (12)).

$$W\left(x^{(i)}\right) = \frac{p\left(x^{(i)}\right)}{q\left(x^{(i)}\right)} \qquad (12)$$

If the normalizing factor for $p(x)$ is not known, the importance weights can only be evaluated up to a normalizing constant, as equation (13).

$$W\left(x^{(i)}\right) \propto p\left(x^{(i)}\right) q\left(x^{(i)}\right) \qquad (13)$$

To ensure that $\sum_{i=1}^{N_p} W(x^{(i)}) = 1$, we normalize the importance weights to obtain equation (14).

$$\hat{f} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p} W\left(x^{(i)}\right) f\left(x^{(i)}\right)}{\frac{1}{N_p}\sum_{i=1}^{N_p} W\left(x^{(i)}\right)} = \frac{1}{N_p}\sum_{i=1}^{N_p} \tilde{W}\left(x^{(i)}\right) f\left(x^{(i)}\right) \qquad (14)$$

where $\tilde{W}\left(x^{(i)}\right) = \frac{W\left(x^{(i)}\right)}{\sum_{i=1}^{N_p} W\left(x^{(i)}\right)}$ are called the normalized importance weights. The variance of importance sampler estimate [12] in equation (14) is given:

$$
\begin{aligned}
Var_q[\hat{f}] &= \frac{1}{N_p} Var_q[f(x)W(x)] \\
&= \frac{1}{N_p} Var_q[f(x)p(x)/q(x)] \\
&= \frac{1}{N_p} \int [\frac{f(x)p(x)}{q(x)} - \mathbb{E}_p[f(x)]]^2 q(x)dx \\
&= \frac{1}{N_p} \int [(\frac{(f(x)p(x))^2}{q(x)}) \\
&\quad -2p(x)f(x)\mathbb{E}_p[f(x)]]dx \\
&\quad + \frac{(\mathbb{E}_p[f(x)])^2}{N_p} \\
&= \frac{1}{N_p} \int [(\frac{(f(x)p(x))^2}{q(x)})]dx - \frac{(\mathbb{E}_p[f(x)])^2}{N_p}
\end{aligned}
$$

The variance can be reduced when an appropriate $q(x)$ is chosen to either match the shape of $p(x)$ so as to approximate the true variance; or to match the shape of $|f(x)|p(x)$ so as to further reduce the true variance.
*Proof:* $\frac{\partial Var_q[\hat{f}]}{\partial q(x)} = -\frac{1}{N_p}\int[(\frac{(f(x)p(x))^2}{q(x)^2})]dx = -\frac{1}{N_p}\int[(\frac{(f(x)p(x))^2}{q(x)q(x)})]dx$. $q(x)$ having the constraint of being a probability measure that is $\int_{-\infty}^{+\infty} p(x)dx = 1$, we find that $q(x)$ must match the shape of $p(x)$ or of $|f(x)|p(x)$. $\blacksquare$
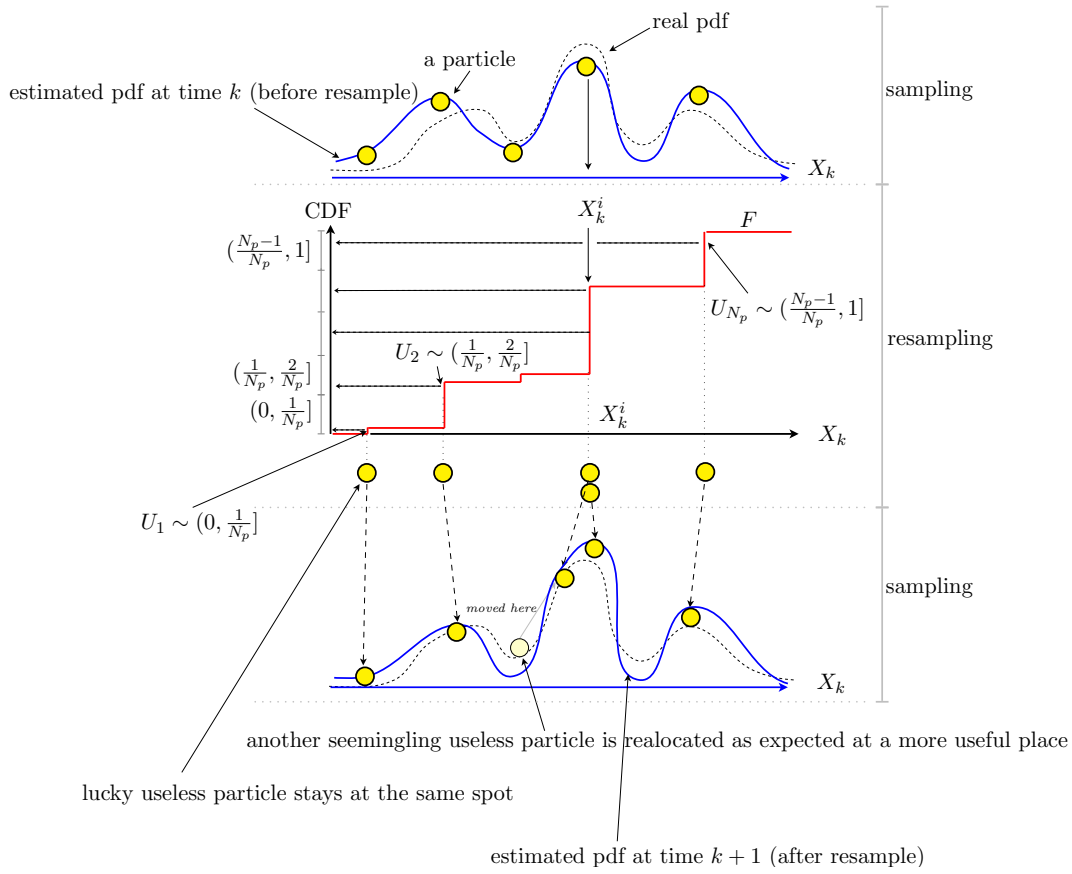
Fig. 3. Stratified Sampling illustration

## D. Scenario Tracking Algorithm

*1) Context:* Recently, SMC methods [17], [18], [58], especially when it comes to the data association issue, have been developed. Particle Filters (PF) [31], [52], have recently become a popular framework for Multi Target Tracking (MTT), because able to perform well even when the data models are nonlinear and non-Gaussian, as opposed to linear methods used by the classical methods like the KF/EKF [39]. Given the observations and the previous target state information SMC can employ sequential importance sampling recursively and update the posterior distribution of our target state. The Probability Hypothesis Density (PHD) filter [97], [99], [67], which combines the Finite Set Statistics (FISST), an extension of Bayesian analysis to incorporate comparisons between different dimensional state-spaces, and the SMC methods, was also proposed for joint target detection and estimation [79]. The M-best feasible solutions is also a new useful finding in SMC [79], [54], [8], [55], [9]. Articles [98], [100] were proposed to cope with both the multitarget detection and tracking scenario but according to Ng, Li, Godsill, and Vermaak [78] they are not robust if the environment becomes more noisy and hostile, such as having a higher clutter density and a low probability of target detection. To cope with these problems a hybrid approach and it extensions [78] were implemented. The aim of these methods is to stochastically estimate the number of targets and therefore the multitarget state. The soft-gating approach described in [80] is an attempt to address the complex measurement-to-target association problem. To solve this issue of detection in the presence of spurious objects a new SMC algorithm is presented in [56]. That method provided a solution to deal with both time-varying number of targets, and measurement-to-target association issues.

*2) Time-varying number of targets & measurement-to-target association:* Currently, tracking for multiple targets has a couple of major challenges that are yet to be answered efficiently. The first of these two main challenges is the modelling of the time-varying number of targets in an environment high in clutter density and low in detection probability (hostile environment). To some extend the PHD filter [68], [98], [100], based on the FISST , has proved ability in dealing with this problem with unfortunately a significant degradation of its performance when the environment is hostile [78]. The second main challenge is the measurement-to-target association problem. Because there is an ambiguity between whether the observation consists of measurements originating from a true targets or a clutter point, it becomes obviously essential to identify which one is which. The typical and popular approach to solve this issue is the Joint Probabilistic Data Association (JPDA) [6], [25]. Its major drawback though is that its tracks tend to coalesce when targets are closely spaced [24] or intertwined.

This problem has been, however, partially studied. Indeed the sensitivity of the track coalescence may be reduced if we use a hypothesis pruning strategy [10], [40]. Unfortunately the track swap problems still remain. Also performance of the EKF [39] is known to be limited by the linearity of the data model on the contrary to SMC based tracking algorithms developed by [44], [34], [33], [41]. This issue of data association can also be sampled via Gibbs sampling [41]. Also because target detection and initialization were not covered by this framework algorithms developed in [107] were suggested in order to improve detection and tracking performance. The algorithm suggested in [107] combines a deterministic clustering algorithm for the target detection issue. This clustering algorithm enabled to detect the number of targets by continuously monitoring the changes in the regions of interest where the moving targets are most likely located. Another approach in [92] combines the track-before-detect (TBD) and the SMC methods to perform joint target detection and estimation, where the observation noise is Rayleigh distributed but, according to [92], this algorithm is currently applicable only to single target scenario. Solutions to the data association problem arising in unlabelled measurements in a hostile environment and the curse of dimensionality arising because of the increased size of the state-space associated with multiple targets were given in [107]. In [107], a couple of extensions to the standard known particle filtering methodology for MTT was presented. The first extension was referred to as the Sequential Sampling Particle Filter (SSPF), sampled each target sequentially by using a factorisation of the importance weights. The second extension was referred by the Independent Partition Particle Filter (IPPF), makes the hypothesis that the associations are independent. Real world MTT problems are usually made more difficult because of couple of main issues. First realistic models have usually a very non-linear and non-Gaussian target dynamics and measurement processes therefore no closed-form expression can be derived for the tracking recursions. The most famous closed form recursion leads to the KF [2] and arises when both the dynamic and the likelihood model are chosen to be linear and Gaussian. The second issue with real world problem is due to the poor sensors targets measurements labeling which leads to a combinatorial data association problem that is challenging in a hostile environment. The complexity of the data association problem may be enhanced by the increase in probability of clutter measurements in lieu of a target in areas rich in multi-path effects. We have seen that the KF is limited in modeling non linearity because of its linear properties but it is still an interesting tool as an approximation mean like it has been done with the EKF [2] which capitalizes on linearity around the current state in non linear models. Logically the performance of the EKF decreases as the non-linearity increases. The Unscented Kalman Filter (UKF) [46] was created to answer this problem. The method maintains the second order statistics of the target distribution by recursively propagating a set of carefully selected sigma points. The advantage of this method is that it does not require linearisation

as well as usually yields more robust estimates. Models with non-Gaussian state and/or observation noise were initially studied and partially solved by the Gaussian Sum Filter (GSF) [1] . That method approximates the non-Gaussian target distribution with a mixture of Gaussians but suffers when linear approximations are required similarly to the EKF. Also, over time we experience a combinatorial growth in the number of mixture components which ultimately leads to eliminate branches to keep control of an exponential explosion as iterations go forward. Another option that does not require any linear approximations like it is the case with the EKF or the GSF was proposed [51]. In this case the non-Gaussian state is approximated numerically with a fixed grid, using Bayes' rule, the prediction step is integrated numerically. Unfortunately because the computational cost of the integration explodes with the dimension of the state-space the method becomes useless for dimensions larger than four [107]. For non-linear and non-Gaussian models, generally speaking SMC's have become popular user friendly numerical techniques that approximate Bayesian recursions for MTT. Its popularity is mainly due to flexibility, relative simplicity as well as efficiency. The method models the posterior distribution with a set of particles with an associated weights more or less big relative to the particle importance and are propagated and adjusted throughout iterations. The very big advantage with SMC method is that the computational complexity does not become exorbitant with an increase in the dimension of the state-space [51]. It has been defined in [107] that there exists numerous strategies available to solve the data association problem but they could be categorized as either single frame assignment methods, or multi-frame assignment methods. The multi-frame assignment problem can be solved using Lagrangian relaxation [87]. Another algorithm the Multiple Hypotheses Tracker (MHT) [88] tries to keep track of all the possible association hypotheses over time which makes it awkward as the number of associations hypotheses grows exponentially with each iteration.

*3) The problem of pruning:* The Nearest Neighbor Standard Filter (NNSF) [6] links each target with the closest measurement in the target space. This simplistic method has the flaws that one may assume it has, that is the method suppresses many feasible hypotheses. The Joint Probabilistic Data Association Filter (JPDAF) [6], [25] is more interesting in this respect as it does not do as much pruning or pruning only infeasible hypotheses. The parallel filtering algorithm goes through the remaining hypotheses and adjusts the corresponding posterior distribution. Its principal deficiency is that the final estimate looses information because, to maintain tractability, the corresponding estimate is distorted to a single Gaussian. This problem however has been identified and strategies have been suggested to address this shortcoming. For example [83], [93] proposed strategies to instead reduce the number of mixture components in the original mixture to a tractable level. This algorithm unfortunately only partially solved the problem as many feasible hypotheses may still be pruned away. The Probabilistic Multiple Hypotheses Tracker (PMHT) [28], [104] takes as hypothesis that the association

variables to be independent and avoids the problems of reducing our state space. This leads to an incomplete data problem that, however may be solved using the Expectation Maximisation (EM) algorithm [16]. Unfortunately the PMHT is not suitable for sequential applications because considered a batch strategy. Moreover [109] has shown that the JPDA filter outperforms the PMHT and we have seen earlier the shortcomings of the JPDAF. Recently strategies have been proposed to combine the JPDAF with particle techniques to address the general non-linear and non-Gaussian models [96], [95], [26], [49] issue of approximation of linearity failing when the dynamic of measurement functions are severely non-linear. The feasibility of multi-target tracking with SMC has first been described in [3], [32] but the simulations dealt only with a single target. In the article [42] the distribution and the hypotheses of the association is computed using a Gibbs sampler, [29] at each iterations. This method, similar to the one described in [15], uses MCMC [30] to compute the associations between image points within the framework of stereo reconstruction. Because they are iterative in nature and take an unknown number of iterations to converge. These MCMC strategies though, are not always suitable for online applications. Doucet [33] presents a method where the associations are sampled from a well chosen importance distribution. Although intuitivly appealing it is, however, reserved to Jump Markov Linear Systems (JMLS) [19]. The follow up of this strategy, based on the UKF and the Auxiliary Particle Filter (APF) [85], so that applicable to Jump Markov Systems (JMS) is presented in [20]. Similar in [44], particles of the association hypotheses are generated via an optimal proposal distribution. SMC have also been applied to the problem of MTT based on raw measurements [11], [94]. We have seen that the MTT algorithms suffers from exponential explosion that is as the number of targets increases, the size of our state spaces increases exponentially. Because pruning is not always efficient it may commonly occur that particles contain a mixture of good estimates for some target states, and bad estimates for other target states. This problem has been first acknowledged in [82], and where a selection strategy is addressed to solve this problem. In [107] a number of particle filter based strategies for MTT and data association for general non-linear and non-Gaussian models is presented. The first, is referred to as the Monte Carlo Joint Probabilistic Data Association Filter (MC-JPDAF) and presented by the authors as a generalization of the strategy proposed in [96], [95] to multiple observers and arbitrary proposal distributions. Two extensions to the standard particle filtering methodology for MTT is developed. The first strategy is presented by the authors as an exact methodology that samples the individual targets sequentially by utilizing a factorization of the importance weights, called the Sequential Sampling Particle Filter (SSPF). The second strategy presented in [107] assumes the associations to be independent over the individual target, similar to the approximation made in the PMHT, and implies that measurements can be assigned to more than one target. This assumption claims that it effectively removes all dependencies between

the individual targets, leading to an efficient component-wise sampling strategy to construct new particles. This approach was named Independent Partition Particle Filter (IPPF). Their main benefit is that as opposed to the JPDAF, neither approach requires a gating procedure like in [44].

### III. RELEVANT THEORETICAL BIOLOGY REVIEW

Our first few simulations [65] are unfortunately non conclusive because some of the computing and statistical optimality elements were dismissed in order to keep intuition high.

#### A. Theoretical Biology & Predator/Prey models

More specifically few elements of our simulations resembled well know biological models. To bring context it was discussed in the 1960s [36] that complexity in an ecosystem insures its stability or to keep the same jargon "communities not being sufficiently complex to damp out oscillations" [23], [43] have a higher likelihood of vanishing. It also is widely accepted, in the context of ecosystem simulation, that complexity should always arise from simplicity [70], [14]. The diversity-stability debate in the context of ecosystem modeling has been ongoing since the 1950s [71] with no consensus being ever reached. It was initially believed [71], [62], [22] nature was infinitely complex and therefor more diverse ecosystem should insure more stability. This assertion was however ultimately challenged through rigorous mathematical specification [70], [110], [84] in the 1970s and 1980s by using Lotka-Volterra's Predator/Prey model initially published in the 1920's [108], [61] with similar "non-intuitive" results. More recently the work has been extended to small ecosystems of three-species food chain [13]. The intuitive 3 species example we have chosen to discuss is the one containing Sharks (chosen to be the $z$ parameter), Tuna (chosen to be the $y$ parameter) and Small Fishes (chosen to be the $x$ parameter), the idea being that tunas eat small fishes which in turn are eaten by sharks. Without loss of generality sharks are assumed to die of natural causes and their decomposing bodies go on to feed the small fishes. The set of differential equations has been summarized in equation (15).

$$\begin{cases} \frac{dx}{dt} & = ax - bxy \\ \frac{dy}{dt} & = -cy + dxy - eyz \\ \frac{dz}{dt} & = -fz + gyz \end{cases} \quad (15)$$

where $a$ is the natural growth rate of species $x$ in the absence of predator, $d$ the one of $y$ in the absence of $z$. We also have $b$ representing the negative predation effect of $y$ on $a$ and $e$ the one of $z$ on $y$. We also have $g$ which mirrors the efficiency of reproduction of $z$ in the presence of prey $y$. Note that we assume that $x$ never dies of natural causes (if it's too old then it can't run fast enough to outrun predator $y$) but this is not the case for $z$ since it is an alpha predator and therefore needs some natural death rate which is symbolized by $f$. This relatively simple system of three equations has been studied extensively [71] for stability. For example figure 5 represents a particular instance in which the system is unstable. Indeed,

we can notice that the oscillations between the 3 species increases to the point, here not shown, where the amplitudes are so big that $z$ goes instinct and at which point $x$ and $y$ start oscillating, with however a constant amplitude. We refer the motivated reader back to the original papers [71] for the other cases and interesting idiosyncratic properties. One interesting point to notice is that in cases of "relative best stability", in which $a = b = c = d = e = f = g = 1\%$ from figure 4, we have oscillation which are stable through time with the highest peek from the ultimate prey $(x)$ coming first with the highest peek and the the one of the ultimate predator $(z)$ coming last but with the smallest amplitude. This suggest that sophisticated working trading strategies[16] need enough prey like strategies[17] in the same ecosystem to get them to be profitable. One other interesting observation is that the total ecosystem population as depicted in the thick black line from the same figure suggest that it itself oscillates which may not necessarily be intuitive. Indeed one could have speculated that the loss of a species directly benefits the other and that therefore the total population should stay constant. This interesting observation suggest that the oscillations of a financial market may likewise be subject of similar dynamics: a financial ecosystem may go through periods in which it thrives followed by period in which it declines, the economy itself is cyclical with, some may argue oscillations which are more and more important like one depicted by the unstable ecosystem from figure 5. The stunning similarities of the
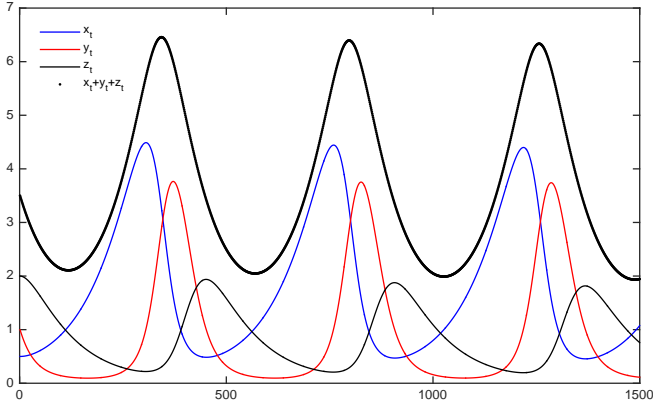
this paper is the one of systemic risk. Given that this paper proposes that the fluctuations of the markets are linked to the frequency of the strategies composing the ecosystem of the market, we propose a model which would take advantage of this assumptions to build original high level regulations. The exercise would consist of monitoring these strategies interactions and flag the market when necessary. Suppose now that we label strategies of figure 13, 14 and 18 by respectively $x$, $y$ and $z$ and we use equation (15). If we can somehow infer what the frequency of $x$, $y$ and $z$ are in the ecosystem, then we can study whether or not the ecosystem is stable [13]. Returning to the actual mathematical study of the stability of the financial market, determining a market composed of 3 strategies is stable requires studying the Jacobian matrix $J$ from equation (16).

$$
J(x, y, z) = \begin{bmatrix} a - by & -xb & 0 \\ yd & -c + dx - ez & -ye \\ 0 & -zg & -f + gy \end{bmatrix} \quad (16)
$$

By examining the eigenvalues of $J(x, y, z)$ we can indirectly gain information around the equilibrium of our financial system at the regulatory level[18]. More specifically if all eigenvalues of $J(x, y, z)$ have negative real parts then our system is asymptotically stable. Figure 5 gives an illustration of a situation in which one of the eigenvalues is negative. Many questions could be raised here: how can the regulators
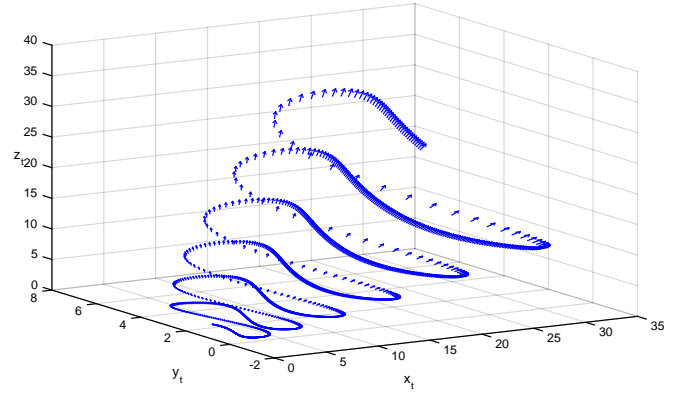


Fig. 4. The Lotka-Volterra three-species food chain equation 15 with $x_1 = 0.5$, $y_1 = 1$, $z_1 = 2$ and $a = b = c = d = e = f = g = 1\%$



Fig. 5. The Lotka-Volterra three-species food Chain equation 15 with $x_1 = 0.5$, $y_1 = 1$, $z_1 = 2$, $a = b = c = d = e = f = 1\%$ and $g = 1.6\%$

competitive resource driven biological ecosystem along with some compelling similarities in some of its cyclical behavior makes the Lotka-Volterra n-species food chain equation an interesting candidate when it comes to studying the stability of the financial market in the context of the HFTE because of its systematic rule based approach and zero sum game like roots. However, these are hypothesis that we need to check more rigorously.

*1) Regulatory Implications:* The second and last immediate application we will take a look at in the context of

gain information on the parameters composing systems of equation (15)? Also the market has surely more than 3 types of strategies, how many exactly? Are these strategies easily classifiable in terms of prey, predator and super predator or can you find more subtle instances? It is very likely that trading desks especially in the high frequency domain refuse to provide their sets of strategies for the regulators to study the Jacobian matrix in order to take the relevant actions[19].

---

[16]perhaps from top algorithmic desks in top tier investment banks?
[17]perhaps the retail clients of the world?

[18]we assume for the sake of this example that we only have 3 strategies
[19]instruct the trading desks to increase or decrease their notionals so as to enforce a manual intervention for the sake of the market's stability

## B. Optimal Control Theory

The Hamilton-Jacobi-Bellman (HJB) partial differential equation [7] was developed in 1954 and is widely considered as a central theme of optimal control theory. Its solutions is the value function giving the minimum cost for a given dynamical system and its associated cost function. Solved locally, the HJB is a necessary condition, but when over the entire of state space, it is referred to as necessary and sufficient for an optimum. Its method can be generalized to stochastic systems. Its discrete version is referred to as the Bellman equation and its continuous version, the Hamilton-Jacobi equation.

*1) Optimal Control Formalization:* Formally we consider the problem in deterministic optimal control over the time period $[0, T]$:

$$V(x(0), 0) = \min_u \left\{ \int_0^T C[x(t), u(t)] \, dt + D[x(T)] \right\} \quad (17)$$

where $C[]$ is the scalar cost rate function, $D[]$ is the utility at the final state, $x(t)$ the system state vector with $x(0)$ usually given, and finally $u(t)$ where $0tT$ is called the control vector we aim at finding. The system of equation is also subject to $\dot{x}(t) = F[x(t), u(t)]$ where $F[]$ is a deterministic vector describing the evolution of the state vector over time.

*2) Partial Differential Equation Specification:* The HJB partial differential equation is given by:

$$\dot{V}(x, t) + \min_u \{\nabla V(x, t) \cdot F(x, u) + C(x, u)\} = 0 \quad (18)$$

subject to the terminal condition $V(x, T) = D(x)$. $V(x, t)$, commonly known as the Bellman value function (our unknown scalar) represents the cost incurred from starting in $x$ at time $t$ and controlling the system optimally until $T$.

*3) Equation derivation:* $V(x(t), t)$ is the optimal cost-to-go function, then by Bellman's principle of optimality from time $t$ to $t + dt$, we have $V(x(t), t) = \min_u \left\{ V(x(t + dt), t + dt) + \int_t^{t+dt} C(x(s), u(s)) \, ds \right\}$. The Taylor expansion of the first term is $V(x(t + dt), t + dt) = V(x(t), t) + \dot{V}(x(t), t) \, dt + \nabla V(x(t), t) \cdot \dot{x}(t) \, dt + o(dt)$ where $(o)(dt)$ denotes the higher order terms of the Taylor expansion. Canceling $V(x(t), t)$ on both sides and dividing by $dt$, and taking the limit as $dt$ approaches zero, we obtain the HJB equation. Its resolutions is done backwards in time which can be extended to its stochastic version. In this latter case we have $\min_u \mathbb{E} \left\{ \int_0^T C(t, X_t, u_t) \, dt + D(X_T) \right\}$, with this time $(X_t)_{t \in [0,T]}$ being stochastic and needing optimization and $(u_t)_{t \in [0,T]}$ the control process. By first using Bellman and then expanding $V(X_t, t)$ with Ito's rule, one finds the stochastic HJB equation $\min_u \{\mathcal{A}V(x, t) + C(t, x, u)\} = 0$ where $\mathcal{A}$ represents the stochastic differentiation operator, and subject to the terminal condition $V(x, T) = D(x)^{20}$.

---

[20] the randomness has disappeared.

## C. Game Theoretical Approach

Another area of investigation is the one of Game Theory. Broadly speaking the prisoner's dilemma (PD) can be formalized into a matrix[21] of 2 by 2 with CC, CD, DC and DD with respective payoffs (2,2), (0,3), (3,0) and (1,1). The reason why this game theory concept is within the family of dilemmas is because although the prisoners clearly should cooperate here, given that they do not know what the other is going to do, by expectation (with equal probability for a C and a D) any user should deceit given that the expectation of the payoff for a deceit is 2 as opposed to a 1 for a cooperation.

*1) Axelrod's computer tournament:* however this dilemma presented in the previous subsection proved to shuffle the rules of payoff strategy optimality when the game became iterative, Robert Axelrod main contribution to the field. Indeed Axelrod [4], [5] designed a computer tournament which aim was to take a look at what strategy would prevail in an iterative format. In that occasion he invited few Mathematicians, Computer Scientists, Economists and Political Scientists to code a strategy they believed could win such tournament with the constraints of a PD rules in which it is not known when the tournament will stop[22]. Many strategies were thrown into this ecosystem in form of a tournament ranging from being simplistic like "Always Deceit" (AD) strategy[23] to many other more complicated strategies which generic representation can be looked at in Figure 6b). Surprisingly the Tit For Tat (TFT) strategy came at the top of this tournament. The TFT is considered in the literature to be a nice strategy, meaning that it is never the first to deceit (its first move is by design to be a C), but it is also a strategy that is able to retaliate in situation in which it was previously deceited. Finally, it is a strategy that is able to forgive meaning that if it sees that the adversary has decided to cooperate after a deceit, then he switches back to a C.

*2) Evolutionary Dynamics:* Martin Nowak [81] recently enhanced some of Axelrod's work by introducing new strategies and further developing the concepts of invasion/dominance[24] within a competitive strategic ecosystem. For instance as we can see from Figure 6d) that some strategies invade others but these latter strategies can be in turn invaded by other ones which in turn can be invaded by the very first strategy mentioned and induce cycles[25]. Indeed an ecosystem composed of a set of unbiased random strategies (that would randomly C or D) would invite the invasion of an ALLD (always defect) kind. In term the

---

[21] Figure 6a)

[22] eg: it is by expectation best to deceit if one plays the PD only once. By iteration he should always deceit on the last move, but knowing this, the adversary should also deceit. Using this logic each player should deceit on the next to the last move and the same logic kicks in and very quickly one is led to arrive to the conclusion that he/she should deceit from the very first move.

[23] or its mirror: the AC "Always Cooperate" (AC) strategy

[24] by extension when applied to finance some strategies may dominate and invade others.

[25] economical cycles for example when applied to our primary problem

frequency of ALLD would take the ecosystem which would invite the TFT strategy which would benefit from the mutual cooperation when within the same proximity etc ... Figure (6) exposes how some of these strategies may interact with each other. The following additional information may help in refreshing what some of these acronyms mean:

- TFT (*Tit of Tat*) developed in the previous section
- GTFT (*Generous Tit of Tat*) which makes it slightly less grudge prone compared to the TFT as it only deceits for 2 successive D's from the opponent.
- WSLS (*Win-Stay, Lose-Shift*) that outperforms tit-for-tat in the Prisoner's Dilemma game [81], [101]
- ALLD (*Always Deceits*) which is self explanatory
- ALLC (*Always Cooperates*) which is also self explanatory
- rand (*Random Strategy*) which outputs a C or a D with equal probability.

The main takeaway from this parallel was to expose how the rise and fall of strategies can easily be engineered through simple systematic rules based on an ecosystem and how complexity can be induced from simple rules.

## IV. HIGH FREQUENCY TRADING ECOSYSTEM

Recently, the concept of ecosystems of strategies [50] was introduced. Though the idea had great potential, the paper assumes a set of static strategies which does offer to some extent an interesting snapshot of the market but does not offer:

- a history for this snapshot,
- an inspiring future for the field,
- a topology for these strategies (in the form of a DNA),
- a sense of how to study the stability of the ecosystem,
- insight about how this should impact the regulatory horizon,
- a connection to other fields[26] with concepts and properties that could be used to increase our mathematical weaponry.

**Definition** We call **HFTE** the **H**igh **F**requency **T**rading **E**cosystem model which attempt is to answer the points raised.

### A. Network & learning potential

Two important milestones in Machine Learning are worth remembering, as they shed light on why the core building blocks of our HFTE model is a certain way. First, Warren McCulloch and Walter Pitts [86] introduced their threshold logic model in 1943 which is agreed to have guided the research in network topology as it relates to artificial intelligence for more or less a decade. Second, Rosenblatt [90], formally introduced the perceptron concept in 1962 though some early stage work had started in the 1950s. The idea of the perceptron was one in which the inputs $x_1$ and $x_2$

[26]eg: Game Theory, Mathematical Biology, Signal Processing

could act as separators[27] and therefore a direct equivalence could be made to the multi-linear regression which we will elaborate on more in details is section IV-B.2. One observed limitation of the perceptron as described by Rosenblatt, in 1969, was that a simple yet critical well known functions such as the XOR function could not be modeled [74]. This resulted in a loss of interest in the field until it was shown that a Feedforward Artificial Neural Network (ANN) with two or more layers could in fact model these functions. Added, to this we have the well known overfitting [102] problems when it comes to supervised learning which has been there since inception though regular progress is being made in that domain without real breakthrough.

### B. The Funnel

The Funnel, introduced by Martin Nowak [81], represents the simplest possible network to model (therefore which minimizes overfitting) the key functions of our application. The area of evolutionary graph theory is quite rich, and graphs provide interesting properties. We can formalize the learning process from all of our strategies using the topology of Figure 7 by providing a set $\mathcal{T}$, as described by equation (19) of weights corresponding to all the possible weights of this particular figure.

$$\mathcal{T} \triangleq \left\{ \begin{array}{ll} \cup_{j\in[1,9]} w^i_{\bar{s},j} & \cup_{j\in[1,9]} w^i_{s,j}, \\ \cup_{j\in[1,9],i\in[1,3]} w^{h_1}_{\bar{s},i,j} & \cup_{j\in[1,9],i\in[1,3]} w^{h_1}_{s,i,j}, \\ \cup_{j\in[1,3]} w^{h_2}_{\bar{s},j} & \cup_{j\in[1,3]} w^{h_2}_{s,j}, \\ w^o_{\bar{s},j\in[1,9]} & w^o_{s,j\in[1,9]} \end{array} \right\}$$
(19)

with $w^i$, $w^h$ and $w^o$, respectively the weights associated to the input, hidden and output layers.

**Remark** Note that in the context of this paper we have chosen to work with Martin Nowak's [81] funnel, as Figure 7. This topological structure offers the advantage of making some interesting bridges between the worlds of:

- information theory since it also resembles the classic structure of a Neural Network and can therefore easily accommodate the mapping of classic and less classic financial strategies,
- evolutionary dynamics since Moran-like Processes [75] can easily be formalized,
- biology since it is a potent amplifier of selection [81][28].

We will conclude this subsection by providing a definition of the High Frequency Financial Funnel.

**Definition** We define the **H**igh **F**requency **F**inancial **F**unnel (HFFF) to be a network structure with 9 inputs, 3 hidden layers and 1 output layer. Each node connects to the next

[27]the exact research was one in which the methodology acted as a $1, 0$ through a logistic activation function $f(x) = \frac{1}{1+e^{-x}}$ as opposed to a linear one. However that small distinction is not significant enough in the context to delve too much into it but deserved a clarification in the footnotes.

[28]Indeed, as we will see in section IV-B.1 its simplest structure (the EWMA) serves as pillar to the section IV-B.2 (MLR) which itself does the same for the XOR strategy. So we have this incremental complexity in the network that corresponds to an incremental complexity in information processed.
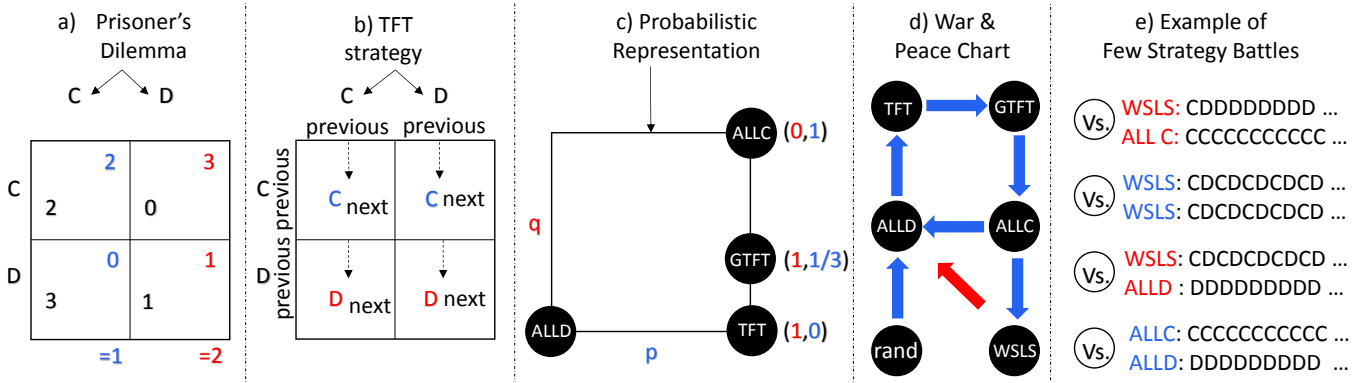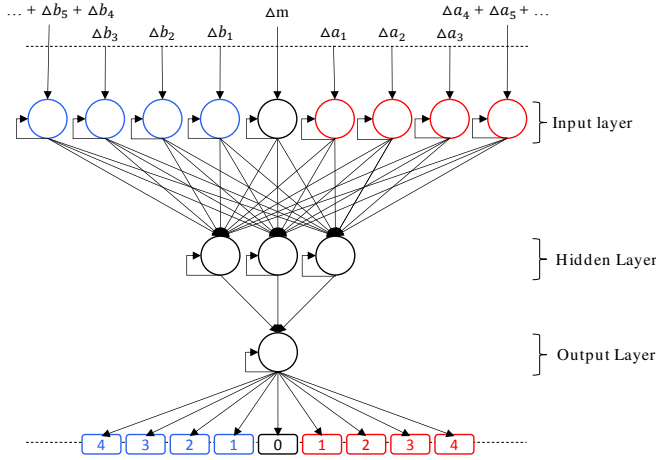
Fig. 6. Some classic Game theory representations [81].



Fig. 7. The High Frequency Financial Funnel

layer and to itself. Each self connection will be labelled by $w_s$ and the others by $w_{\bar{s}}$. We will admit that $w_{\bar{s}} \sim \mathcal{U}[-1,1]$ and that $w_s \sim \mathcal{U}[0,1]$ hence:

$$w_x \sim \mathcal{U}[-1_{x=\bar{s}}, 1] \qquad (20)$$

*1) The Trend Following Topology:* a very common trading strategy is trend following (TF). The idea of TF is that if the price has been going a certain way (eg: up or down) in the recent past, then it is more likely to follow the same trend in the immediate future.

**Definition** The mathematical formulation of TF can be diverse but in the context of this paper we use an exponentially weighted moving average (EWMA), formally described by equation (21),

$$\hat{x}_t = (1-\lambda)x_t + \lambda\hat{x}_{t-1}, \ \lambda \in [0,1] \qquad (21)$$

in which $\lambda$ represents the smoothness parameter with $\lambda \in [0,1]$.

**Remark** The lower the magnitude of $\lambda$, the more the next value will be conditional to the previous value. Conversely, the higher $\lambda$, the more the future value will be function to the long term trend. The idea being that through a simple

filtering process, the noise is extracted from the signal which then returns a clean time series $\hat{x}_t$.

**Proposition** The HFFF can model trend following strategies.

*Proof:* Simply set $\cup_{j\in[1,4]}w_{\bar{s},j}^i = 0$, $\cup_{j\in[1,4]}w_{s,j}^i = 0$, $\cup_{j\in[6,9]}w_{\bar{s},j}^i = 0$, $\cup_{j\in[6,9]}w_{s,j}^i = 0$ $\cup_{j\in[1,4],i\in[1,3]}w_{\bar{s},i,j}^{h_1} = 0$, $\cup_{j\in[1,4],i\in[1,3]}w_{s,i,j}^{h_1}$ $\cup_{j\in[6,9],i\in[1,3]}w_{\bar{s},i,j}^{h_1} = 0$, $\cup_{j\in[6,9],i\in[1,3]}w_{s,i,j}^{h_1}$, $w_{\bar{s},3}^h = 0$, $w_{s,1}^h = 0$ and $w_{s,3}^h = 0$. ∎

**Remark** The proof is visually illustrated by Figure 13 (the weight equal to 0 have not been represented).

*2) Multi Linear Regression Topology:* the Multi Linear Regression (MLR) is another well known strategy traders have been using for a time in the industry.

**Definition** Given a data set $\{y_i, x_{i-1,1}, \ldots, x_{i-1,9}\}_{i=1}^n$, where $n$ is the sample size, $\{\beta_i\}_{i=1}^9$, the weight of the explanatory variables and $y_i$ the output, then our MLR is formalized by

$$
\begin{aligned}
y_i &= \beta_1 x_{i-1,1} + \cdots + \beta_9 x_{i-1,9} + \varepsilon_i \qquad (22) \\
&= \mathbf{x}_{i-1}^{\mathrm{T}}\beta + \varepsilon_i, \ i = 1, \ldots, n
\end{aligned}
$$

**Proposition** The HFFF can model multi linear regression like strategies.

*Proof:* Simply set $\cup_{j\in[1,4]}w_{\bar{s},j}^i = 0$, $\cup_{j\in[1,4]}w_{s,j}^i = 0$, $\cup_{j\in[6,9]}w_{\bar{s},j}^i = 0$, $\cup_{j\in[6,9]}w_{s,j}^i = 0$, $w_{\bar{s},1}^h = 0$, $w_{\bar{s},3}^h = 0$, $w_{s,1}^h = 0$, $w_{s,3}^h = 0$. ∎

**Remark** We will make 4 remarks:

- MLR is illustrated in Figure 14 (weights equal to 0 have not been represented).
- Logistic or weighted MLR can be modelled through the topology of Figure 14 by simply changing respectively the activation function (from linear to logistic) and the weights.

*3) XOR Topology:* How is the XOR function relevant to HFT? Let's look at the following known HF rational.

**Definition** If we define the Open Interest (OI) as being the total volume left on the order book then it is known that when:

- the price and the OI are rising then the market is bullish,
- the price is rising but the OI is falling then the market is bearish,
- the price is falling but the OI is rising then the market is bearish,
- the price and OI are both falling then the market is bullish.

**Remark** These 4 market situations can be summarized by table IV-B.3.

| Open Interest | Price | Combined Symbol | Signal |
|---|---|---|---|
| Rising | Rising | $\uparrow\uparrow$ | Buy |
| Rising | Falling | $\uparrow\downarrow$ | Sell |
| Falling | Rising | $\downarrow\uparrow$ | Sell |
| Falling | Falling | $\downarrow\downarrow$ | Buy |

TABLE I

THE RELATIONSHIP BETWEEN OPEN INTEREST (OI), PRICE (I) & SIGNAL FOR XOR STRATEGY [65]

**Proposition** The HFFF can model XOR like strategies.

*Proof:* Simply set $\cup_{j\in[1,4]}w_{\bar{s},j}^i = 0$, $\cup_{j\in[1,4]}w_{s,j}^i = 0$, $\cup_{j\in[6,9]}w_{\bar{s},j}^i = 0$, $\cup_{j\in[6,9]}w_{s,j}^i = 0$, $w_{\bar{s},1}^i = 0$, $w_{\bar{s},3}^i = 0$, $w_{s,1}^h = 0$, $w_{s,3}^h = 0$. ∎

**Remark** We will make the following 2 observations:

- The preceding proof is visually illustrated by Figure 15 (the weights equal to 0 have not been represented).
- The XOR topology can be designed in various ways.

*4) Execution strategy:* to make the problem more realistic, one needs to formalize an execution strategy which would apply to all strategies, but still be rule based and function of its topology. In this paper we will take the simple approach in which all strategies have that same execution strategy.The idea of this algorithm will be that:

- the execution strategy will be subject to a certainty-like function,
- certainty will be decided by the historical returns from the relevant topology split into intervals,
- since the decision needs to be made and that data comes regularly a rolling percentile function should be used.

In this context our algorithm returns a value between 0 and 9, the 9 explanatory variables of our HFFF and corresponding to all of the admissible actions in our order book. The tested input is compared against the current output as it compares to the historical outputs and returns the corresponding percentile which then goes on populating the order-book. Given that no history exists in the first iteration and that the first few iterations are not significant, we will randomize the first $R_n$ iterations.

*C. Genetic Algorithm as a means to study the market through time*

We will take a look at a couple of methods to study the market through time. We first take, in this subsection, an approach with the objective to gain intuition in order to strategize with respect to future research and then a second method which is mathematically more optimal in lieu of the RJ-MCMC mentioned in the literature review from the "Report Format Document".In this "intuitive" section we specify the genetic algorithm which we have used to study our problem with intuition in mind as opposed to optimality.

*1) Looping & Fitting Function:* Throughout this subsection we will refer to Micro and Macro increments.

**Definition** We will define two types of iterations:

- the first type being **Micro** corresponding to an infinitesimal increment in our environment, namely an increment in which a strategy $S$ analyses and in turn changes the order book by placing a order itself.
- the second type being **Macro**, corresponding to a generational increment in our environment, namely *a certain equal number* of Micro increments per strategy leading to a calculation of a Profit and Loss (P&L) and a survival process[29] based on this P&L.

We will label as $N_k$ the number of total live strategies, $N_k^e$ the number of t**r**end following like strategies, $N_k^m$ the number of **m**ulti-linear regretion like strategies, $N_k^r$ the number of xo**r** like strategies and $N_k^o$ the number of *other unclassified* strategies[30]. The relationship between these entities can be summarized by equation (23).

$$N_k = N_k^e + N_k^m + N_k^r + N_k^o \qquad (23)$$

A strategy $\mathcal{S}$ will consist of a topology $\mathcal{T}$, a rolling P&L $\mathcal{P}$ and a common orderbook $\mathcal{O}$ as shown by equation (24).

$$\mathcal{S} \triangleq \{\mathcal{P}, \mathcal{T}, \mathcal{O}\}. \qquad (24)$$

*2) Survival & birth processes:* the survival, death & birth processes are a set of processes which impact the number of live strategies $N_k$ at any time $k$ the following way. If we call $\mathcal{S}_{N_k} = S_{(1)}, S_{(2)}, \ldots, S_{(n)}, S_{(n+p)}, \ldots, S_{(N_k)}$, the strategies ranked with respect to their P&L from highest to lowest, we will admit the following definitions:

**Definition** The Survivor set[31] is the set of strategies with a positive P&L. Namely if $\mathcal{S}_a = S_{(1)}, S_{(2)}, \ldots, S_{(s)}$ with $S_{(s)} \geqslant 0$ and $S_{(s+1)} < 0$. We will subdivide this set by distinguishing:

- secondary survivors set with cardinality $a_2 = \lfloor \frac{s}{2} \rfloor$, survive without reproducing
- primary survivors set with cardinality $a_1 = s - a_2$, survive and have one offspring with a "slightly different DNA" in form of a conditional resampling of their topology.

---

[29]explained next
[30]This label will be the same in section III-A.
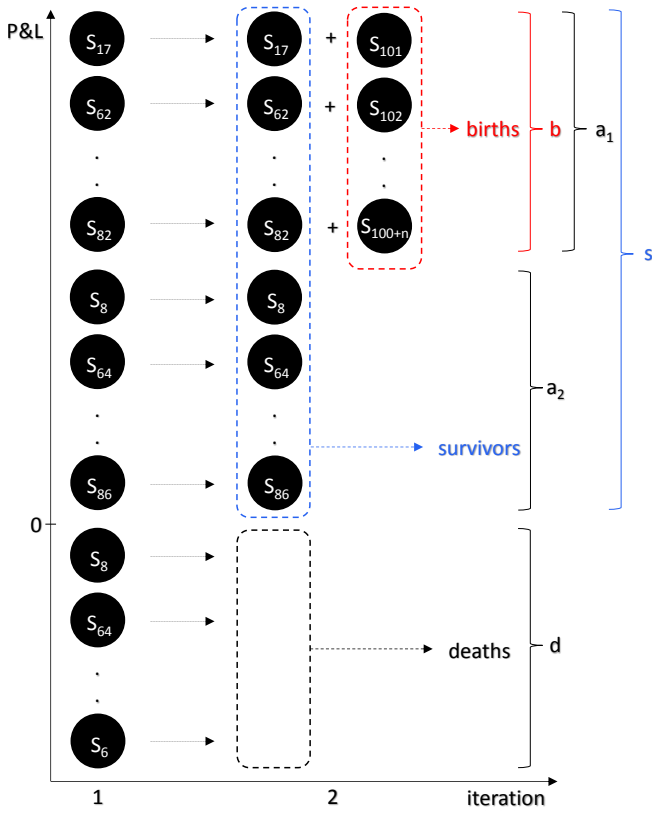[31]or alternatively alive process

Fig. 8.   Illustration for the Death and Birth processes in our GA

**Definition** We will call the Birth process, the set of rules conducting the selection of top strategies and their reproduction with mutations. The protocol starts by selecting the ranked first half of survived strategies. Namely, if $a_1 = b = \left\lfloor \frac{s}{2} \right\rfloor$ the strategies $\mathcal{S}_1 \ldots \mathcal{S}_{a_1}$ will both survive and reproduce and create a set of equal size but with a slightly different topology and with cardinality $b = a_1$.

**Definition** We define the Death process, the set of protocols aiming at eliminating part of the strategies from our ecosystem, more specifically, the set of strategies with a negative P&L. Namely if $\mathcal{S}_d = S_{(s+1)}, S_{(s+2)}, \ldots, S_{(N_k)}$ will disappear from the market at Macro iteration $k + 1$.

**Remark** We can easily see that $s = a_1 + a_2$, $a_1 \geqslant a_2$, $a_1 = b$. Figure 8 illustrates these definitions.

*3) Inheritance with Mutations :* the intuition about the mutation process is that each birth is a function of a successful strategy (the positive P&L of parents $\mathcal{S}_1 \ldots \mathcal{S}_{a_1}$) and should resemble the single parent[32] which produced it. We have seen in section IV that the DNA of our strategies is essentially their topology $\mathcal{T}$ (which is itself a combination of weights). We will therefore concentrate on performing the re-sampling on the weights of the offspring. The reason why this distribution is interesting[33] is that:

  • is defined in a closed interval [0,1] and can therefore

---

[32]no crossover in this model
[33]though, again not optimal

be rescaled easily through a change of variable to [-1,1], an interval which is a basic way of formalizing a normalized importance of each node in the topology decision making of Figure 7.

  • on the contrary to the uniform distribution, it is more flexible and offers a broad range of interesting shapes allowing the possibility to code a conditional resampling model and therefore make clever proximity changes around the symbolic levels: $-1$, $0$ and $1$. This way we can prevent too large deviations and rather select small incremental changes and intuitively follow the principles of selection. We can see that the $\mathcal{B}eta(x, 1, 7)$ or $\mathcal{B}eta(1 - x, 1, 7)$ both concentrate a great deal of the distribution towards 0 and 1 respectively. Likewise $\mathcal{B}eta(x, 3, 7)$ and $\mathcal{B}eta(x, 5, 7)$ provide a more Gaussian like distribution towards in between zones which is what we want.

*D. Preliminary results*

*1) Classification Issue:* One of the very first issues we came across was the problem of strategy classification. Indeed besides the obvious pitfalls associated to resampling on continuous distribution we also have several instances where different networks give the very same results. Few examples of this problem are given by Figures 16, 17, and 18.

*2) Simulation issues:* Besides the classification issues just mentioned which, to an extend was already a show stopper, as a mean of rigorous proof, we faced a great deal of programming issues mainly associated to the complexity of the computing object oriented exercise.

*3) Few encouraging notes:* Few of the simulations seemed to indicated a positive correlation between TF and MLR strategies on strictly increasing or strictly decreasing markets. When the market trend happened to be less clear the correlation between their growth rate seemed less significant and perhaps even negative. It was difficult to make a proper quantification of these observations due to the low speed of each simulation and also because the observation conditional to market tendencies came a posteriori of the simulations. Not enough simulations were performed to really be able to assert the mentioned relationship definitively. Similar results were found as for the relationship between XOR and MLR strategies, though with even less significance. Few simulations were actually such that, the results remind us to the Lotka-Volterra 3-predator-prey model with however a great deal of noise and unclassified strategies. These limitations forced us to review and simplify the problem in such a way as to improve the chance of computation, to decipher the relationship between these strategies at the ecosystem level and try to prove our results with more rigor.

*4) hypothesis:* Our seminal paper exposed some of the relationship between traditional strategies in the financial markets and speculated through incomplete simulations the theory of strategy invasion for which couple of example have be summarized in figure 9 which acronyms are reminded in table IV-D.4.   Finally our first paper [65] ended with the
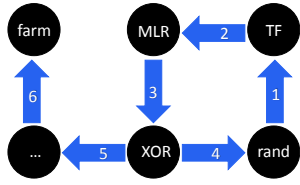
Fig. 9. Illustration for a possible strategy invasion we described [65]



Fig. 10. MCMC for our 4 states rand, TF, MLR & XOR organized in terms incremental architectural complexity

| Name | Meaning | Summary |
|------|---------|---------|
| rand | random | buy or sale randomly |
| TF | Trend Following | Follow trend with small lag |
| MLR | Multi-Linear-Regression | buy/sell according to the betas |
| XOR | XOR | follow table IV-B.3 |
| farm | farmer | strategy with self awareness |
| OIP | Open Interest Price | 4 joint cases ($\Uparrow$, $\uparrow\downarrow$, $\downarrow\uparrow$, $\Downarrow$) |

TABLE II

ACRONYMS & THEIR MEANING FOR THE MAIN STRATS DESCRIBED [65]

proposed a "Diversity & the Financial Markets" conjecture below.

**Conjecture** Diversity in financial strategies in the market leads to its instability.

**Remark** Note this conjecture can be studied indirectly or at least intuitively using some of the finding in mathematical biology. More specifically the one associated with diversity in ecosystem and stability[34].

## V. CONCLUSION

### A. Summary

We have started this paper by pointing to a puzzling observation from the newly born high frequency commodities market which because of its extreme youth and therefore immaturity makes it a great case study for a high frequency market at inception and therefore for our purpose. More specifically as we have seen with Figure 1, on 06/08/2011, fascinating patterned oscillations occurred in the commodities market. We have proposed in this paper that these oscillations are due to the interactions of the different strategies participating in the market and contributing to the fluctuations of the market. We have done a literature review of mathematical methods for tracking in section II and of predator/prey and evolutionary dynamics results in section III. We propose to use these two fields as leverage keys to help us decipher the dynamics of the HFTE model for which we have summarized the main results in section IV. Finally we have discussed some of the current open problems and ideas we have to developing this paper further in section V-B.

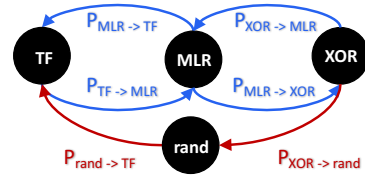[34]Though no consensus is reached there either.

### B. Current & Future Research

Our first few simulations opened our eyes up to issues associated to optimality and need for more scientific rigor.

*1) Classification Simplification:* As mentioned before the direct simulation approach is too challenging and the results perhaps too convoluted to filter out the essence of the paper. For this reason we propose to study the problem using fixed topologies, of Figures 13 14 and 15, but with possibility to transition into each other and also a random state, with a "jump" as opposed to small incremental changes in the network topology of section IV-C. This will be done via a transition probability as illustrated by figure 10. The rational of this choice is that the incremental change in topology between states also corresponds to an increased subtlety in the information processed.

**Remark** It has been speculated that the need for a bigger brain in men is partly due to the need for human to elaborate deceitful strategies with their rivals and cooperative strategies with their allies. It is therefore not entirely ridiculous to associate increased neural network branching (to be roughly understood as increase in cranial size) with increased strategy complexity. However, increased intelligence does not necessarily equate to survival as we can see in the shark population, considered like an apex predator but with a relatively small brain that has not evolved for millennia.

*2) Invasion flow charts:* One way to control our simulation dimensionality issue is to try to express our strategies in win/lose matrix leading to an invasion, Nowak's [81] proposed methodology for versions of strategies battling in the context of the Tit for Tat like summarized in Figure 6. In order to do this rigorously. Let's first go through few definitions.

**Definition** We will call an **dynamic mini order-book** $o$, the sequence of static snapshots of the order-book $^{a_2,a_1}\mathtt{M}_{b_1,b_2}$ of asked and bid volumes $a_i/b_i$ where $i$ corresponds to the depth of the order book and $\mathtt{M}$ its mid price.

**Definition** We will call an **environment** of size $i$ a set of strategy, $S = s^a, s^b, \ldots, s^i$ of topology spanning the one from figure 7 composed of an order-book, $^{a_2,a_1}\mathtt{M}_{b_1,b_2}$.

**Definition** We will call an **HFTE Game** the sequence of environments composed of 2 strategies, $S = s^a, s^b, \ldots, s^i$ of topology spanning the one from figure 10 with a dynamic mini order-book and P&L.

The tables from appendix III give few examples of HFTE Games.

**Definition** We will call a strategy, $s$ **invasive** with respect to an environment, $e$ when the P&L of $s$ increases in an HFTE Game.

**Proposition** The MLR strategy is invasive with respect to TF.

*Proof:* The way to prove this assertion is to test it through the 4 possible order-book seeds. The iterations from each steps have been included in appendixes II. ∎

**Definition** We will call a strategy, $s$ **self-fulfilling** when it is invasive with respect to itself.

**Proposition** The TF strategy is self-fulfilling.

*Proof:* The way to prove this assertion is to test it through the 4 possible order-book seeds. The iterations from each steps have been included in appendixes I. ∎

**Remark** The intuition we had [65] around the TF acting like a prey increasing exponentially in the absence of predator is arguably confirmed.

**Definition** We will call a strategy, $s^a$ **weakly dominant** with respect to a strategy $s^b$ in an environment when, with the 4 seeds, the average P&L of $s^a$ is bigger to the one of $s^b$.

**Conjecture** The XOR strategy is weakly dominant compared to strategy MLR.

**Remark** The dominance relationship between the following strategies is not yet fully understood:
- MLR with XOR when MLR takes actions first,
- XOR with MLR when XOR takes actions first,
- TF with XOR when TF takes actions first,
- XOR with TF when XOR takes actions first,

One of the objectives of future research is to rigorously formalize these 4 open problems like we have done with tables of appendixes I, II or III.

*3) New Simulation:* In this new approach we try to substitute our GA by a MCMC and assume that each successful strategy makes 2 copies in each macro iteration[35], one of itself and one following an incremental complexity described by figure 10. For example a random strategy which is successful becomes a TF. A TF which is successful becomes a MLR and so on. In order to close the system we will assume like it was the case in the seminal paper that the ecosystem also contains random strategies which have the property to:
- invite the invasion of TF strategies but
- XOR strategies, assumed the most advanced turn into so much complexion that they can be considered back to a random state.

These assumptions are summarized by the Markov Chain of figure 11. Note that this is consistent with the increased

architectural complexity spanning from the HFFF for each of these strategies. For example like we have seen in figure 13 and 14, the MLR uses the TF architecture with additional OI information. Similarly figures 14 and 14 show that the XOR strategy uses an additional layer on top of the MLR architecture. We try here to understand whether increased architectural complexity translates into P&L success especially in a dynamic environment. Like we have seen in section II-A, another variant of the Metropolis-Hastings, and perhaps most promising methodology when it comes to our application is the Reversible-jump MCMC (RJ-MCMC) developed by Green [35] and Roberts [89]. One key factor of RJ-MCMC is that it is designed to address changes of dimensionality issues. In our case, we face a dual type issues around change of dimensionality. The first being the frequency of each strategy in an ecosystem and the second element being the HFFF which branching structure and size changes as a function of the strategy. In order to illustrate this point, the reader may want to take a look at Figures 7, 13, 14 and 15. The details of the simulations haven't yet been formalized. An idea we currently have is to reduce the types of ecosystem to a collection of at least 13 spaces illustrated by figure 11.
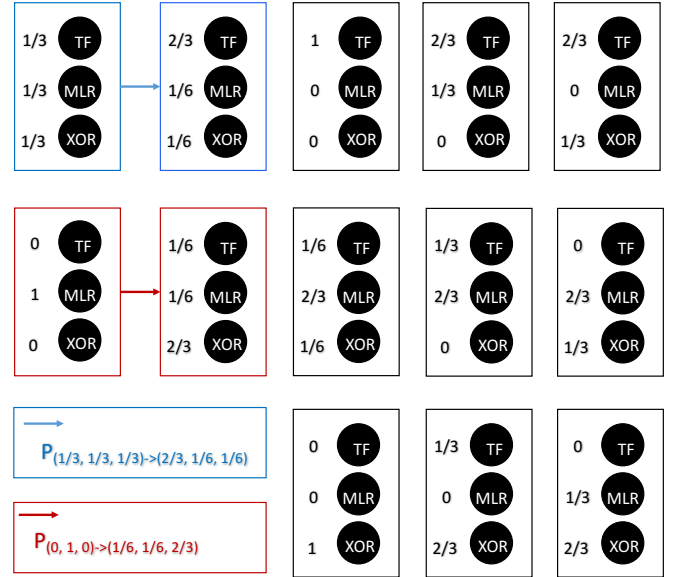


Fig. 11. Incomplete Space of boxed ecosystems scenarios defined by a triplet of strategies with an associated proportion in an ecosystem with few examples of transition probabilities symbolized by an arrow in anticipation of a future Reversible jump MCMC simulation.

*4) The regulatory aspect:* as we have seen in section III-A.1 the regulatory implications from this research naturally invites us to explore a research project in which we would try to *guess*[36] the frequency of each types of strategies using the LotkaVolterra multi species models as likelihood functions. We propose to use a particle filter on scenarios to achieve this point. We will discuss this particular point in a subsequent paper.

---

[35]Please see our original simulation [65] if you are unfamiliar with the jargon.

[36]given that we cannot ask the market participants to provide us with their strategies
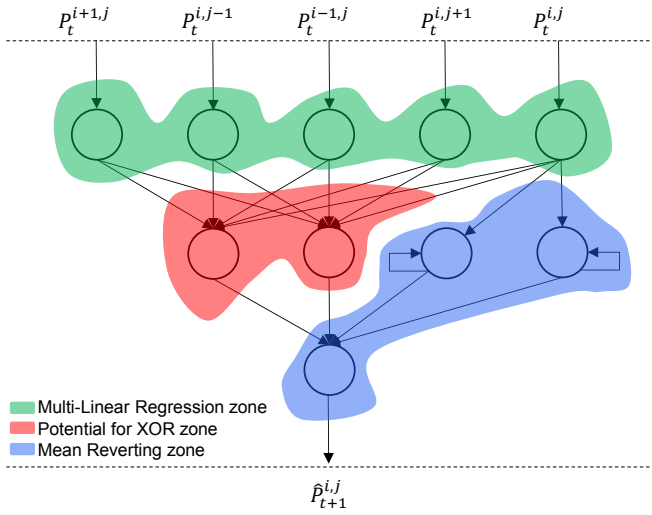
Fig. 12. Illustration of an ANN aiming at predicting the price of an option based on its adjacent points.

*5) The options market:* we have recently introduced a new parametrisation of the implied volatility surface [64], [66] and have established that de-arbing is a convoluted mathematical optimization.

For the sake of making the notation a bit more intuitive, we use the notation of equation (25).

$$P_t^{i-1,j} := C_t(Ke^{-r2\Delta}, T - \Delta) \tag{25a}$$

$$P_t^{i,j-1} := C_t(K - \Delta)e^{-r\Delta}, T) \tag{25b}$$

$$P_t^{i,j} := C_t(Ke^{-r\Delta}, T) \tag{25c}$$

$$P_t^{i,j+1} := C_t(K + \Delta)e^{-r\Delta}, T) \tag{25d}$$

$$P_t^{i+1,j} := C_t(K, T + \Delta) \tag{25e}$$

where $C_t(.)$ representing the call price under the relevant asset class assumption[37]. We aim at studying the Bayesian probability problem of equation (26).

$$p\Big(\mathbf{P}^{i,j} - l(\mathbf{P}^{i,j})\Big| l(\mathbf{F})\Big) \tag{26}$$

where, $\mathbf{F} = \{\mathbf{P}^{i,j}, \mathbf{P}^{i-1,j}, \mathbf{P}^{i,j-1}, \mathbf{P}^{i,j+1}, \mathbf{P}^{i+1,j}\}$ in the discrete space, $\mathbf{P} = P_{\tau \in [1,t]}$ and $l$ represents the lag inducing function such that $l(P_{t+1}) = (P_t)$. The implied volatility is a very different product than spot because it has a tendency to mean revert, it is very much subject to what the adjacent points are doing and reacts in a lower frequency than spot. Our aim will be to study the HFTE in light of these observations. However it is interesting to already notice that these observations could be addressed by a modification of the HFFF (Figure 12). Following the rational from section IV we need to create an learning architecture that would incorporate the following observations:

- Presumably the price point $P_t^{i,j}$ can be best approximated by the 4 adjacent points, a simple MLR[38] can be used to model this idea (green part of Figure 12)

- The second point to notice is that each point of the implied volatility is a mean reverting stochastic process and this can be modeled in terms of network architecture by a spread of EWMA[39] (blue part of Figure 12)
- At least one hidden layer to address some of the economical drivers leading to a need for an architecture that could learn XOR like functions like we saw could sometimes be necessary in algorithmic trading from Table IV-B.3 (red part of Figure 12).

**Remark** Note that the XOR like functions may not be as necessary as the dynamics of spot since vol may be driven by economical factors that are different especially if we study the problem at different timescales. This suggests that the red part of Figure 12 may, at the end, be the identity function. For the sake of keeping that door open though, we have left it in our network topology.

We will also see how the parameters introduced in the newly published IVP model [64] can contribute in fine tuning the learning process as well as it execution strategy which requires insight around liquidity. Following the rational from section IV we need to create a learning architecture that would incorporate:

- Presumably the price point $P_t^{i,j}$ can be best approximated by the 4 adjacent points, a simple MLR[40] can be used to model this idea (green part of Figure 12)
- The second point to notice is that each point of the implied volatility is a mean reverting stochastic process and this can be modeled in terms of network architecture by a spread of EWMA[41] (blue part of Figure 12)
- At least one hidden layer to address some of the economical drivers leading to a need for an architecture that could learn XOR like functions like we saw could sometimes be necessary in algorithmic trading from Table IV-B.3 (red part of Figure 12).

## REFERENCES

[1] D.L. Alspach and H.W. Sorenson. Nonlinear bayesian estimation using gaussian sum approximation. *IEEE Transactions on Automatic Control*, 17(4):439–448, 1972.
[2] and Anderson, B.D.O. Optimal filtering. *Englewood Cliffs: Prentice-Hall*, 1979.
[3] D. Avitzour, W. Burgard, and D. Fox. Stochastic simulation bayesian approach to multitarget tracking. *IEE Proceedings on Radar and Sonar Navigation*, 142(2):41–44, 1995.
[4] Robert Axelrod. The evolution of cooperation. *Basic Books*, 1984.
[5] Robert Axelrod. The complexity of cooperation: Agent-based models of competition and collaboration. *Princeton University Press*, 1997.
[6] Y. Bar-Shalom and T. E. Fortmann. Tracking and data association. *Academic Press*, 1988.
[7] Richard Bellman. Dynamic programming and a new formalism in the calculus of variations. *Proceedings of the National Academy of Sciences of the United States of America*, 40(4):231–235, 1954.
[8] D.P. Bertsekas. Auction algorithms for network flow problems: A tutorial introduction. *Computational Optimization and Applications*, 1:7–66, 1992.
[9] S. Blackman and R. Popoli. Design and analysis of modern tracking systems. *Norwood, MA: Artech House*, 1999.

---

[37]eg: Log-Normal diffusion in Equities, Normal diffusion for rates and Garman Kohlhagen for FX.
[38]$y_i = \beta_1 x_{i-1,1} + \cdots + \beta_9 x_{i-1,9} + \varepsilon_i$ from equation (22)

[39]$\hat{x}_t = (1 - \lambda)x_t + \lambda\hat{x}_{t-1}$, $\lambda \in [0,1]$ from equation (21)
[40]$y_i = \beta_1 x_{i-1,1} + \cdots + \beta_9 x_{i-1,9} + \varepsilon_i$ from equation (22)
[41]$\hat{x}_t = (1 - \lambda)x_t + \lambda\hat{x}_{t-1}$, $\lambda \in [0,1]$ from equation (21)

[10] H.A.P. Blom and E.A. Bloem. Probabilistic data association avoiding track coalescence. *IEEE Tr. Automatic Control*, 45:247–259, 2000.

[11] Y. Boers, J.N. Driessen, F. Verschure, W. P. M. H. Heemels, and A. Juloski. A multi target track before detect application. *in Proceedings of the IEEE Workshop on Multi-Object Tracking*, 2003.

[12] J.A. Bucklew. Large deviation techniques in decision, simulations, and estimation. *Wiley*, 1990.

[13] Erica Chauvet, Joseph E. Paullet, Joseph P. Previte, and Zac Walls. A Lotka-Volterra three-species food chain. *Mathematics Magazine*, 75:243–255, 2002.

[14] John Horton Conway. *On Numbers and Games*. 1976.

[15] F. Dellaert, S.M. Seitz, C. Thorpe, and S. Thrun. Em, mcmc, and chain flipping for structure from motion with unknown correspondence. *Machine Learning*, 50(1-2):45–71, 2003.

[16] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.

[17] A. Doucet, N. de Freitas, and N. Gordon. Sequential monte carlo in practice. *New York: Springer-Verlag*, 2001.

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

[19] A. Doucet, N. Gordon, and V. Krishnamurthy. Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on Signal Processing*, 49(3):613–624, 2001.

[20] A. Doucet, B. Vo, C. Andrieu, and M. Davy. Particle filtering for multi-target tracking and sensor management. *in Proceedings of the 5th International Conference on Information Fusion*, 2002.

[21] Simon Duane, A.D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216–222, 1987.

[22] C. S. Elton. Ecology of invasions by animals and plants. *Chapman & Hall, London*, pages 228–233.

[23] C. S. Elton. The ecology of invasions by animals and plants. *Methuen*, 1958.

[24] R.J. Fitzgerald. Track biases and coalescence with probabilistic data association. *IEEE Tr. Aerospace*, 21:822–825, 1985.

[25] T. Fortmann, Y. Bar-Shalom, and Scheffe M. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE Journal of Oceanic Engineering*, 8:173–184, 1983.

[26] O. Frank, J. Nieto, J. Guivant, and S. Scheding. Multiple target tracking using sequential monte carlo methods and statistical data association. *in Proceedings of the IEEE / RSJ International Conference on Intelligent Robots and Systems*, 2003.

[27] Martin Gardner. Mathematical games. *Scientific American*, 1970.

[28] H. Gauvrit, J-P. Le Cadre, and C. Jauffret. A formulation of multitarget tracking as an incomplete data problem. *IEEE Transactions on Aerospace and Electronic Systems*, 33(4):1242–1257, 1997.

[29] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, 1984.

[30] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. Markov chain monte carlo in practice. *Chapman and Hall*, 1996.

[31] N. Gordon, D. Salmond, and A. Smith. Novel approach to non-linear/non-gaussian bayesian state estimation. *IEE Proceedings-F*, 140(2):107–113, 1993.

[32] N. J. Gordon. A hybrid bootstrap filter for target tracking in clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 33(1):353–358, 1997.

[33] N.J. Gordon and A. Doucet. Sequential monte carlo for maneuvring target tracking in clutter. *Proceedings SPIE*, pages 493–500, 1999.

[34] N.J. Gordon, D.J. Salmond, and D. Fisher. Bayesian target tracking after group pattern distortion. *O. E. Drummond, editor, Signal and Data Processing of Small Targets, SPIE 3163*, pages 238–248, 1997.

[35] Peter J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[36] Andrew G. Haldane and Robert M. May. Systemic risk in banking ecosystems. *Nature*, 469:351–355, 2011.

[37] Rama Cont Hamed Amini and Andrea Minca. Resilience to contagion in financial networks. *Mathematical Finance*, 26:329–365, 2016.

[38] J.E. Handschin and D.Q. Mayne. Monte carlo techniques to estimate conditional expectation in multi-state non-linear filtering. *Int. J. Contr.*, 9(5):547–559, 1969.

[39] S. Haykin. Adaptive filter theory. *Englewood Cliffs, NJ:Prentice Hall*, 4, 2000.

[40] A.P. Henk, H.A.P. Blom, and E.A. Bloem. Joint imm and coupled pda to track closely spaced targets and to avoid track coalescence. *Proceedings of the Seventh International Conference on Information Fusion*, 1:130–137, 2004.

[41] C. Hue, J-P. Le Cadre, and P. Perez. Sequential monte carlo methods for multiple target tracking and data fusion. *IEEE Tr. Signal Processing*, Feb(50):309–325, 2002.

[42] C. Hue, J-P. Le Cadre, and P. Perez. Tracking multiple objects with particle filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 3(38):791–812, 2002.

[43] G. E. Hutchinson. Homage to santa rosalia, or why are there so many kinds of animals? *Nature*, 93:145–159, 1959.

[44] N. Ikoma and S.J. Godsill. Extended object tracking with unknown association, missing observations, and clutter using particle filters. *Proceedings of the 2003 IEEE Workshop on Statistical Signal Processing*, September:485–488, 2003.

[45] M. Isard and A. Blake. Condensation conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[46] S.J. Julier and J. K. Uhlmann. A new extension of the kalman filter to nonlinear systems. *Proceedings of AeroSense: The 11th International Symposium on Aerospace / Defence Sensing, Simulation and Controls, vol. Multi Sensor Fusion, Tracking and Resource Management II*, 1997.

[47] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82:35–45, 1960.

[48] R. E. Kalman and Bucy R. S. New results in linear filtering and prediction theory. *Transactions of the ASME - Journal of Basic Engineering*, 83:95–107, 1961.

[49] R. Karlsson and F. Gustafsson. Monte carlo data association for multiple target tracking. *in Proceedings of the IEE Seminar Target Tracking: Algorithms and Applications*, page 13/113/5, 2001.

[50] S. Mankad; G. Michailidis; A. Kirilenko. Discovering the ecosystem of an electronic financial market with a dynamic machine-learning method. *Algorithmic Finance*, 2:2:151–165, 2013.

[51] G. Kitagawa. Non-gaussian state-space modeling of nonstationary time series (with discussion). *Journal of the American Statistical Association*, 82(400):1032–1063, 1987.

[52] G. Kitagawa. Monte carlo filter and smoother for non-gaussian non-linear state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996.

[53] Genshiro Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996.

[54] J. Larsen and I. Pedersen. Experiments with the auction algorithm for the shortest path problem. *Nordic Journal of Computing*, 6(4):403–421, 1998.

[55] R.D. Leone and D. Pretolani. Auction algorithms for shortest hyperpath problems. *Society for Industrial and Applied Mathematics*, 11(1):149–159, 2001.

[56] Jack Li, William Ng, Simon Godsill, and Jaco Vermaak. Online multitarget detection and tracking using sequential monte carlo methods. *Department of Engineering University of Cambridge*.

[57] X.R. Li, Y. Bar-Shalom, and T. Kirubarajan. Estimation, tracking and navigation: Theory, algorithms and software. *New York: John Wiley and Sons*, June, 2001.

[58] J. Liu and R. Chen. Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association*.

[59] J.S. Liu and R. Chen. Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93:1032–1044, 1998.

[60] Jun S. Liu, Faming Liang, and Wing Hung Wong. The multiple-try method and local optimization in metropolis sampling. *Journal of the American Statistical Association*, 95(449):121–134, 2000.

[61] Alfred J. Lotka. Elements of physical biology. *Williams & Wilkins Co.*, 1925.

[62] R. H. MacArthur. Fluctuations of animal populations and a measure of community stability. *Ecology*, 36:533536, 1955.

[63] David MacKay. Information theory, inference, and learning algorithms. *Cambridge University Press*, 2003.

[64] Babak Mahdavi-Damghani. Introducing the implied volatility surface parametrisation (IVP):application to the fx market. 2015.

[65] Babak Mahdavi-Damghani. Introducing the HFTE model: a multi-species predator prey ecosystem for high frequency quantitative financial strategies. *Wilmott*, 89:52–69, 2017.

[66] Babak Mahdavi-Damghani and Andrew Kos. De-arbitraging with a weak smile. *Wilmott Magazine*, 2013.

[67] R. P. S. Mahler. Multitarget bayes filtering via firstorder multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4), 2003.

[68] R.P.S. Mahler. Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4), 2003.

[69] A. Marshall. The use of multi-stage sampling schemes in monte carlo computations. *in Symposium on Monte Carlo Methods, M. Meyer Ed. New York: Wiley*, pages 123–140, 1956.

[70] Robert M. May. Simple mathematical models with very complicated dynamics. *Nature*, pages 459–467, 1976.

[71] Kevin Shear McCann. The diversitystability debate. *Nature*, pages 228–233.

[72] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, and Augusta H. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087, 1953.

[73] David D. L. Minh and Do Le (Paul) Minh. Understanding the hastings algorithm. *Communications in Statistics - Simulation and Computation*, 44(2):332–349, 2015.

[74] Marvin Minsky and Seymour Papert. *Perceptrons: An Introduction to Computational Geometry*. 2nd edition with corrections, first edition 1969 edition, 1972.

[75] Patrick Alfred Pierce Moran. The statistical processes of evolutionary theory. *Oxford: Clarendon Press*, 1962.

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

[77] Radford M. Neal. Suppressing random walks in markov chain monte carlo using ordered overrelaxation. *University of Toronto, Department of Statistics*, 1995.

[78] W. Ng, J. Li, S. Godsill, and J. Vermaak. A hybrid approach for online joint detection and tracking for multiple targets. *University of Cambridge, Department of Engineering*, 2005.

[79] William Ng, Jack Li, Simon Godsill, and Jaco Vermaak. A hybrid approach for online joint detection and tracking for multiple targets. *IEEEAC paper,Department of Engineering, University of Cambridge*, 1, 2004.

[80] William Ng, Jack Li, Simon Godsill, and Jaco Vermaak. Multiple target tracking using a new soft-gating approach and sequential monte carlo methods. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 4:1049–1052, 2005.

[81] Martin Nowak. *Evolutionary dynamics : exploring the equations of life*. 2006.

[82] M. Orton and W. Fitzgerald. A bayesian approach to tracking multiple targets using sensor arrays and particle filters. *IEEE Transactions on Signal Processing*, 50(2):216–223, 2002.

[83] L.Y. Pao. Multisensor multitarget mixture reduction algorithms for target tracking. *AIAA Journal of Guidance, Control and Dynamics*, 17, 1994.

[84] S. L. Pimm and J. H. Lawton. On feeding on more than one trophic level. *Nature*, 275:542–544, 1978.

[85] M. K. Pitt and N. Shephard. Filtering via simulation: Auxiliary particle filter. *Journal of the American Statistical Association*, 94:590–599, 1999.

[86] Warren McCulloch; Walter Pitts. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.

[87] R.L. Popp, T. Kirubarajan, and K.R. Pattipati. Survey of assignment techniques for multitarget tracking. *in Multitarget/Multisensor Tracking: Applications and Advances III, Y. Bar-Shalom and W. D. Blair, Eds. Artech House,*, 2000.

[88] D. Read. An algorithm for tracking multiple targets. *IEEE Transactions on Automation and Control*, 24(6):84–90, 1979.

[89] S. J. Roberts, C. Holmes, and D. Denison. Minimum-entropy data partitioning using reversible jump markov chain monte carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):909–914, Aug 2001.

[90] Frank Rosenblatt. *Principles of neurodynamics: perceptrons and the theory of brain mechanisms*. 1962.

[91] S. Roweis and Z. Ghahramani. A unifying review of linear gaussian models. *Neural Comput*, 11(2):305–345, 1999.

[92] M.G. Rutten, N.J. Gordon, and S. Maskell. Particle-based track-before-detect in rayleigh noise. *Proceedings of SPIE Conference on Signal Processing of Small Targets*, 2004.

[93] D.J. Salmond. Mixture reduction algorithms for target tracking in clutter. *in Signal and Data Processing of Small Targets, SPIE 1305, O. E. Drummond, Ed.*, pages 434–445, 1990.

[94] D.J. Salmond and H. Birch. A particle filter for track-before-detect. *in Proceedings of the American Control Conference*, pages 3755–3760, 2001.

[95] D. Schulz, W. Burgard, and D. Fox. People tracking with mobile robots using sample-based joint probabilistic data association filters. *in Proceedings of the IEEE International Conference on Robotics and Automation*, 22(2), 2003.

[96] D. Schulz, W. Burgard, D. Fox, and A.B. Cremers. Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. *in Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1665–1670, 2001.

[97] H. Sidenbladh. Multi-target particle filtering for the probability hypothesis density. *Proceedings 6th International Conference on Information Fusion*, pages 800–806, 2003.

[98] H. Sidenbladh. Multi-target particle filtering for the probability hypothesis density. *Proceedings 6th International Conference on Information Fusion*, pages 800–806, 2003.

[99] H. Sidenbladh and S.L. Wirkander. Particle filtering for finite random sets. *IEEE Transactions on Aerospace and Electronic Systems*, 2003.

[100] H. Sidenbladh and S.L. Wirkander. Particle filtering for finite random sets. *IEEE Transactions on Aerospace and Electronic Systems*, 2003.

[101] Marton Nowak; Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364:56–58, 1993.

[102] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, and Ruslan Salakhutdinov Ilya Sutskever and. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.

[103] L.D. Stone, C.A. Barlow, and T.L. Corwin. Bayesian multiple target tracking.

[104] R.L. Streit and T. E. Luginbuhl. Maximum likelihood method for probabilistic multi-hypothesis tracking. *in Signal and Data Processing of Small Targets, SPIE 2235, O. E. Drummond, Ed.*, 1994.

[105] Nassim Nicholas Taleb. *The Black Swan*. 2007.

[106] M.A. Tanner and W.H. Wong. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82:528–549, 1987.

[107] J. Vermaak, S. Godsill, and P. Perez. Monte carlo filtering for multi-target tracking and data association. *Signal Processing Engineering Depratmenet, University of Cambridge*, 2004.

[108] Vito Volterra. Variazioni e fluttuazioni del numero dindividui in specie animali conviventi. *Mem. Acad. Lincei Roma*, pages 31–113, 1926.

[109] P. Willett, Y. Ruan, and R. Streit. Pmht: Problems and some solutions. *IEEE Transactions on Aerospace and Electronic Systems*, 38(3):738–754, 2002.

[110] P. Yodzis. The stability of real ecosystems. *Nature*, 289:674–676, 1981.

| Strategy Acting | TF1 | TF2 | TF1 | TF2 |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,+1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{0,0}M_{1,1}$ | $^{1,0,0}M_{1,1}$ | $^{0,0,0}M_{1,1}$ |
| $\Delta$OI | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| $\Delta$Price (bps) | +1 | +1 | 0 | +1 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {4,2} |

TABLE III

INTERACTION PATH FOR 2 TF STRATEGIES INTERACTING WITH SEED ⇈

| Strategy Acting | TF1 | TF2 | TF1 | TF2 |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,+1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{0,0}M_{1,1}$ | $^{1,0,0}M_{1,1}$ | $^{0,0,0}M_{1,1}$ |
| $\Delta$OI | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| $\Delta$Price (bps) | +1 | +1 | 0 | +1 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {4,2} |

TABLE IV

INTERACTION PATH FOR 2 TF STRATEGIES INTERACTING WITH SEED ⇅

| Strategy Acting | TF1 | TF2 | TF1 | TF2 |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,+1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{0,0}M_{1,1}$ | $^{1,0,0}M_{1,1}$ | $^{0,0,0}M_{1,1}$ |
| $\Delta$OI | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| $\Delta$Price (bps) | +1 | +1 | 0 | +1 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {4,2} |

TABLE V

INTERACTION PATH FOR 2 TF STRATEGIES INTERACTING WITH SEED ⇵

| Strategy Acting | TF1 | TF2 | TF1 | TF2 |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,+1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{0,0}M_{1,1}$ | $^{1,0,0}M_{1,1}$ | $^{0,0,0}M_{1,1}$ |
| $\Delta$OI | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| $\Delta$Price (bps) | +1 | +1 | 0 | +1 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {4,2} |

TABLE VI

INTERACTION PATH FOR 2 TF STRATEGIES INTERACTING WITH SEED ⇊

| Strategy Acting | TF | MLR | TF | MLR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,+1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{0,0}M_{1,1}$ | $^{1,0,0}M_{1,1}$ | $^{1,0,0}M_{0,1}$ |
| $\Delta$OI | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{+1}{4}$ |
| $\Delta$Price (bps) | +1 | +1 | 0 | +4 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {-4,1} |

TABLE VII

INTERACTION PATH FOR A TF MEETING AN MLR WITH SEED ⇈

| Strategy Acting | TF | MLR | TF | MLR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,+1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{0,0}M_{1,1}$ | $^{1,0,0}M_{1,1}$ | $^{1,0,0}M_{0,1}$ |
| $\Delta$OI | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{+1}{4}$ |
| $\Delta$Price (bps) | +1 | +1 | 0 | +4 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {-4,1} |

TABLE VIII

INTERACTION PATH FOR A TF MEETING AN MLR WITH SEED ⇅

| Strategy Acting | TF | MLR | TF | MLR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {-1,0} | {-1,-1} | {-2,-1} | {-2,+1} |
| Order-book | $^{1,1}M_{0,1}$ | $^{1,1}M_{0,0}$ | $^{1,1}M_{0,0,1}$ | $^{1,0}M_{0,0,1}$ |
| $\Delta$OI | $\frac{+1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ | $\frac{-1}{4}$ |
| $\Delta$Price (bps) | -1 | -1 | 0 | -4 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {-4,1} |

TABLE IX

INTERACTION PATH FOR A TF MEETING AN MLR WITH SEED ⇵

| Strategy Acting | TF | MLR | TF | MLR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {-1,0} | {-1,-1} | {-2,-1} | {-2,+1} |
| Order-book | $^{1,1}M_{0,1}$ | $^{1,1}M_{0,0}$ | $^{1,1}M_{0,0,1}$ | $^{1,0}M_{0,0,1}$ |
| $\Delta$OI | $\frac{+1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ | $\frac{-1}{4}$ |
| $\Delta$Price (bps) | -1 | -1 | 0 | -4 |
| P&L{TF,MLR} | {+1,0} | {2,1} | {2,1} | {-4,1} |

TABLE X

INTERACTION PATH FOR A TF MEETING AN MLR WITH SEED ⇊

| Strategy Acting | TF | XOR | TF | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,-1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{1,0}M_{0,1}$ | $^{1,0}M_{0,0}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| ΔPrice (bps) | +1 | -2 | 0 | +4 |
| P&L{TF,XOR} | {+1,0} | {-2,+1} | {-1,+2} | {-2,+2} |

TABLE XI

INTERACTION PATH FOR A TF MEETING AN XOR WITH SEED ⇈

| Strategy Acting | TF | XOR | TF | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,-1} | {+2,+1} | {+2,-1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{1,0}M_{0,1}$ | $^{1,0}M_{0,0}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| ΔPrice (bps) | +1 | -2 | 0 | +4 |
| P&L{TF,XOR} | {+1,0} | {-2,+1} | {-1,+2} | {-2,+2} |

TABLE XII

INTERACTION PATH FOR A TF MEETING AN XOR WITH SEED ↑↓

| Strategy Acting | TF | XOR | TF | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {-1,0} | {-1,+1} | {-2,-1} | {-2,+1} |
| Order-book | $^{1,1}M_{0,1}$ | $^{1,0}M_{0,1}$ | $^{0,0}M_{0,1}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{+1}{4}$ | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ |
| ΔPrice (bps) | +1 | -2 | 0 | +4 |
| P&L{TF,XOR} | {+1,0} | {-2,+1} | {-1,+2} | {-2,+2} |

TABLE XIII

INTERACTION PATH FOR A TF MEETING AN XOR WITH SEED ↓↑

| Strategy Acting | TF | XOR | TF | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {-1,0} | {-1,+1} | {-2,-1} | {-2,+1} |
| Order-book | $^{1,1}M_{0,1}$ | $^{1,0}M_{0,1}$ | $^{0,0}M_{0,1}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{+1}{4}$ | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ |
| ΔPrice (bps) | +1 | -2 | 0 | +4 |
| P&L{TF,XOR} | {+1,0} | {-2,+1} | {-1,+2} | {-2,+2} |

TABLE XIV

INTERACTION PATH FOR A TF MEETING AN XOR WITH SEED ⇊

| Strategy Acting | MLR | XOR | MLR | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,-1} | {-1,+1} | {0,+1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{1,0}M_{0,1}$ | $^{1,0}M_{0,0}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| ΔPrice (bps) | +1 | -1 | -1 | +2 |
| P&L{TF,XOR} | {+1,0} | {0,+1} | {1,+2} | {0,+1} |

TABLE XV

INTERACTION PATH FOR A MLR MEETING AN XOR WITH SEED ⇈

| Strategy Acting | MLR | XOR | MLR | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {+1,0} | {+1,-1} | {-1,+1} | {0,+1} |
| Order-book | $^{1,0}M_{1,1}$ | $^{1,0}M_{0,1}$ | $^{1,0}M_{0,0}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{-1}{4}$ | $\frac{+1}{4}$ | $\frac{+1}{4}$ | $\frac{-1}{4}$ |
| ΔPrice (bps) | +1 | -1 | -1 | +2 |
| P&L{TF,XOR} | {+1,0} | {0,+1} | {1,+2} | {0,+1} |

TABLE XVI

INTERACTION PATH FOR A MLR MEETING AN XOR WITH SEED ↑↓

| Strategy Acting | MLR | XOR | MLR | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {-1,0} | {-1,+1} | {+1,-1} | {0,-1} |
| Order-book | $^{1,1}M_{0,1}$ | $^{1,0}M_{0,1}$ | $^{0,0}M_{0,1}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{+1}{4}$ | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ |
| ΔPrice (bps) | -1 | +1 | +1 | -2 |
| P&L{TF,XOR} | {+1,0} | {0,+1} | {1,+2} | {0,+1} |

TABLE XVII

INTERACTION PATH FOR A MLR MEETING AN XOR WITH SEED ↓↑

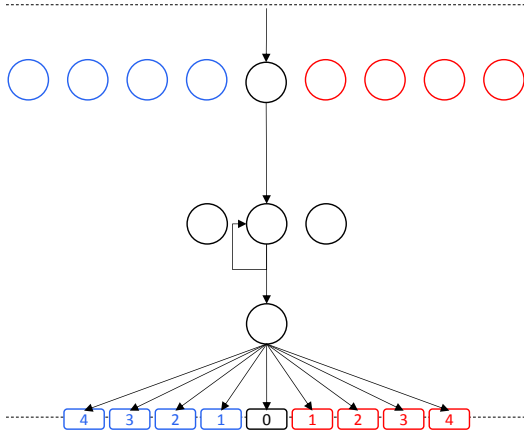| Strategy Acting | MLR | XOR | MLR | XOR |
|---|---|---|---|---|
| Iteration | 1 | 2 | 3 | 4 |
| Signal | {-1,0} | {-1,+1} | {+1,-1} | {0,-1} |
| Order-book | $^{1,1}M_{0,1}$ | $^{1,0}M_{0,1}$ | $^{0,0}M_{0,1}$ | $^{0,0}M_{0,0}$ |
| ΔOI | $\frac{+1}{4}$ | $\frac{-1}{4}$ | $\frac{-1}{4}$ | $\frac{+1}{4}$ |
| ΔPrice (bps) | -1 | +1 | +1 | -2 |
| P&L{TF,XOR} | {+1,0} | {0,+1} | {1,+2} | {0,+1} |

TABLE XVIII

INTERACTION PATH FOR A MLR MEETING AN XOR WITH SEED ⇊

Fig. 13.   The EWMA strategy translated in terms of network topology (the weight equal to 0 have not been represented)
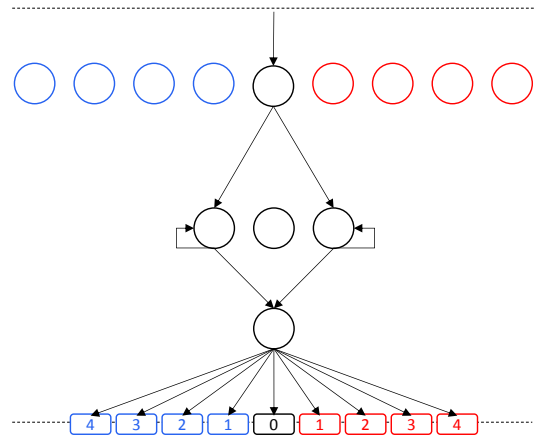


Fig. 16.   The difference of two EWMAs strategies translated in terms of network topology (the weights equal to 0 have not been represented)
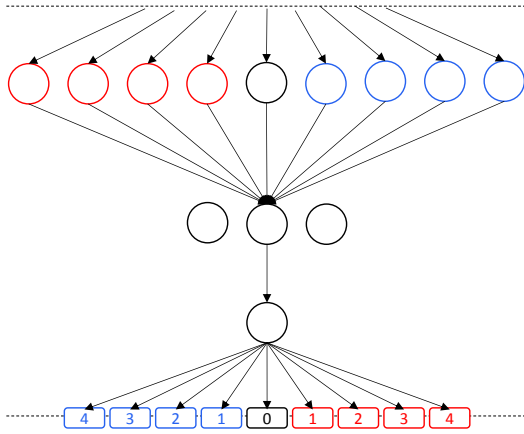


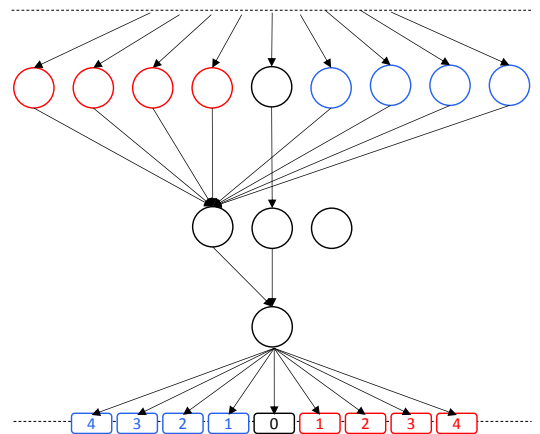Fig. 14.   The MLR strategy translated in terms of network topology



Fig. 17.   Another MLR strategy translated in terms of network topology
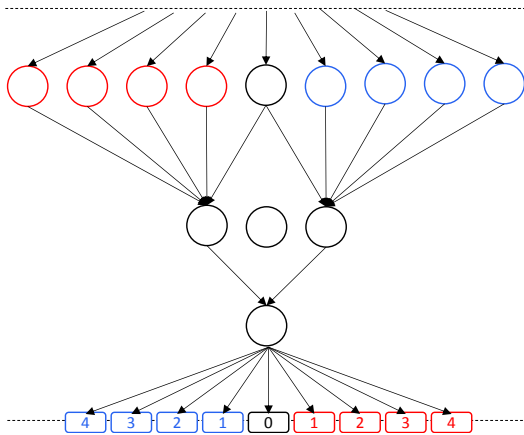


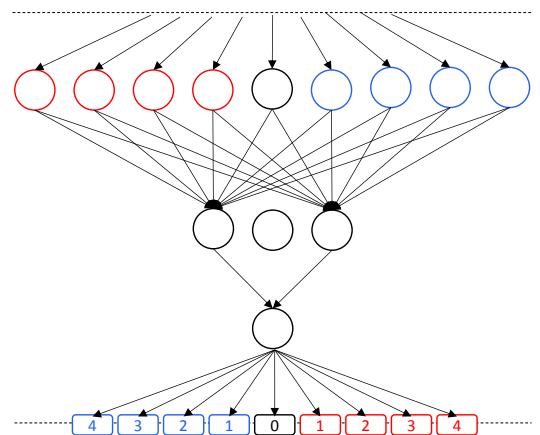Fig. 15.   Another XOR strategy translated in terms of network topology



Fig. 18.   The XOR strategy translated in terms of network topology