Behavioural absorption of Black Swans: simulation with an artificial neural network

Manuscript of article, 6900 words

Author: Krzysztof Wasniewski, PhD

ORCID: https://orcid.org/0000-0003-0076-4804

Affiliation: The Andrzej – Frycz Modrzewski Krakow University, Department of Management, Krakow, Poland

Address: Gustawa Herlinga-Grudzińskiego 1, 30-705 Kraków, Poland

Email: kwasniewski@afm.edu.pl or krzysztof.wasniewski@gmail.com

Phone: +48 601 48 90 51

Funding information: research presented in the manuscript has been funded from a grant of the Ministry Of Science and Higher Education of The Republic of Poland (PL: Ministerstwo Nauki i Szkolnictwa Wyższego Rzeczpospolitej Polskiej), grant #: WZiKS/DS/7/2018-KON.

Abstract

This article attempts to formalize the Black Swan theory as a phenomenon of collective behavioural change. A mathematical model of collectively intelligent social structure, which absorbs random external disturbances, has been built, with a component borrowed from quantum physics, i.e. that of transitory, impossible states, represented by negative probabilities. The model served as basis for building an artificial neural network, to simulate the behaviour of a collectively intelligent social structure optimizing a real sequence of observations in selected variables of Penn Tables 9.1. The simulation led to defining three different paths of collective learning: cyclical adjustment of structural proportions, long-term optimization of size, and long-term destabilization in markets. Capital markets seem to be the most likely to develop adverse long-term volatility in response to Black Swan events, as compared to other socio-economic variables.

JEL: E01, E17, J01, J11

Keywords: collective intelligence, artificial intelligence, Black Swans,

Introduction

The present socio-economic context, with biologically induced uncertainty of the pandemic, once again encourages to study socio-economic resilience to events, which we tend to label as catastrophic and unpredictable. That environment seems to be suiting perfectly a reconsideration of the Black Swan theory (Taleb 2007; Taleb & Blyth 2011). The shock of today will subside, sooner or later, and the big question is: how are we, as civilization, going to absorb and own that experience? How are we collectively learning from it? Across the economic system, spots of new patterns can be observed, when businesses look as if they were bracing for events bound to occur, yet unknown in their exact nature and magnitude. The significant build-up of cash and cash-equivalent balances in the balance sheets of Big Tech companies is a good example of that just-in-case collective behaviour (e.g. https://www.microsoft.com/en-us/investor ; https://abc.xyz/investor/). The concept of Black Swans found its way even into meta-studies of scientific discovery and there, Black Swans are defined as breakthrough events which radically change the direction of research (Zeng et al. 2017).

Even very technical applications of the Black Swan theory, e.g. in finance, almost naturally lead to asking fundamental questions, such as the very definition of Black Swan events, and the exact way that markets absorb such sudden shock-generating events. Interestingly, quantitative methods such as quantile autoregression unit-root test, allow bringing some answers in that respect (Lin & Tsai 2019). In other words, Black Swans can be extracted from quantitative data by studying autoregressive patterns. From the formal-mathematical point of view, there is strong evidence that accurate capturing of Black-Swan-type events requires a shift in methodologies, away from smoothing and towards explicit inclusion of random disturbances (Prestwich 2019). Methodologically, Black-Swan events can be studied as conditional risks with unusual magnitude of damage (Rhee & Wu 2020), which, in turn, leads to studying conditional probabilities rather than tail probabilities. When translated into the broader context of global markets, Black Swan events can be reliably represented as a-cyclical and yet recurrent events of uncertain magnitude (Wang et al. 2019).

It is to keep in mind that Black Swan events largely depend on the observer, and on the pertinence of information that observer has. That pertinence, in turn, depends both on the access to information strictly speaking (i.e. on the observability of pertinent events), and on the capacity to incorporate new information into the existing culture (Hajikazemi et al. 2016). Recurrence of Black Swan events impacts organizations in two different, and almost opposite ways. Black Swans deplete our resources in the effort of response, and yet they are likely to make us more resilient in the future, as we learn to provide for contingencies (Flage & Aven

2015; Aven 2015; Hajikazemi et al. 2016 op. cit.). Interestingly, that point of view coincides with much more classical takes by, respectively, Talcott Parsons, and Herbert Simon. Collectively, we do more than just absorbing shocks. We are after something. Societies have orientations, expressed as values. Those orientations remain in loop with reality, which we keep conceptualizing, and keep producing behaviour based on that conceptualization (Parsons 1951). We build cultural frames of reference, and, with time, those frames absorb our experience of events such as Black Swans Rational economic decisions are supposed to happen in and be driven by a stable and clearly structured system of preferences, where human behaviour depends on the access to information, and on computational capacities. There is the scope of choices we are aware of and which we consciously ponder, and there a different scope of choices which really exist, partly out of our cognition (Simon 1955)

There is evidence that Black Swans are not entirely unpredictable. In financial markets, it is possible to delineate behavioural patterns that lead to building up uncertainties, which, in turn, tend to explode as apparent Black Swans (Chen & Huang 2018). With that in mind, recent developments in cloud computing and collection of data might allow speeding up our capacity to identify the occurrence of Black Swans in the so-far sequences of events, which, in turn, might facilitate prediction and resilience (Batrouni et al. 2018). Furthermore, Black Swan events play out by interaction with human response to exogenous stressors. The behavioural repertoire of that response is finite and otherwise quite easy to express as recurrent patterns, and that allows narrowing down decisional uncertainties in the presence of Black Swans (Ale et al. 2020).

The above reasoning leads to an interesting insight. Black Swan events are observable only to the extent that we have the conceptual apparatus to recognize and acknowledge them, thus to the extent of having some sort of dormant behavioural patterns, which match those events, for one, and which, nevertheless, remain dormant. That dormancy can be explained with a now-classical framework of game theory in the context of evolutionary biology (Hammerstein & Selten 1994). Behavioural patterns can be studied as strategies oriented on payoffs, and, among all the strategies present in the social space, some are clearly prevalent over others. Those others, the 'niche' strategies, match the empirical context with a rarity which pushes them into the fringe of probability. Yet, when the push comes to shove, thus when Black Swans flap their wings, those fringe behaviours suddenly become functional, and this is how we recognize and acknowledge Black Swans. The game-theoretic approach by Selten and Hammerstein offer another insight. The theory of social change and technological change involves two different interpretation of economic equilibrium as a general concept. In one perspective, economic equilibrium is the only sensible place to be, whence concepts such as 'Golden Rules' or 'Golden Paths'. Yet, another perspective is possible, where social change is a chain of equilibriums and disequilibriums, with the latter being the necessary spin and push of change. Adaptive social change can be interpreted as a recurrent rivalry between different strategies – involving both a choice of preferred outcomes and a repertoire of means to achieve them – with a 'niche' strategy corresponding to actions commendable in the presence of Black Swan events. This article investigates the unfolding of collective reaction to Black Swan events with a strong connection to that last viewpoint. Question: how exactly does it happen?

The model

Human societies can be represented as collectively intelligent phenomenological structures where the recurrent incidence of mutually coherent behavioural patterns yields observable socio-economic outcomes. The 'phenomenological' adjective means that, whilst the society is composed of individuals, it is structured by and into their recurrent, patterned behaviour. Mutual coherence of behavioural patterns means that social coordination happens through behavioural coupling between individuals, tacitly or explicitly.

Formally, a human society can be studied as a complex structure $\{SR, P_{SR}, LC_{SR}, O\}$, where SR, P_{SR}, LC_{SR} and O are component sets which represent, respectively: m social roles $SR = \{sr_1, sr_2, \ldots, sr_m\}$ available to individuals and occurring with corresponding probabilities $P_{SR} = \{p(sr_1), p(sr_2), \ldots, p(sr_m)\}$, which, in turn, happen with mutual lateral coherences (i.e. manifestations of lateral coupling) $LC_{SR} = \{lc(sr_1), lc(sr_2), \ldots, lc(sr_m)\}$, and are correlated with the achievement of k social outcomes $O = \{o_1, o_2, \ldots, o_k\}$.

The structure $\{SR, P_{SR}, LC_{SR}, O\}$ is collectively intelligent to the extent that learns, i.e. modifies its component subsets SR, P_{SR}, LC_{SR}, and O, by experimenting with many alternative versions of itself. Experimentation is sequential, i.e. each such alternative version of the structure $\{SR, P_{SR}, LC_{SR}, O\}$ occurs in a different moment t in time. Therefore, the process of collectively intelligent learning happens as a chain of states $\{SR(t), P_{SR}(t), LC_{SR}(t), O(t)\}$.

Consistently with the Interface Theory of Perception (Hoffman et al. 2015 op. cit., Fields et al. 2018 op. cit.), as well as the theory of Black Swans (Taleb 2007 op. cit.; Taleb & Blyth 2011 op. cit.), it is assumed that collective learning with respect to exogenous stressors can happen only if the structure absorbs those stressors into endogenous mechanisms of adaptation. Unacknowledged stressors, which remain purely exogenous, without being coupled with exogenous adaptation in the structure, are irrelevant.

Absorption of external stressors means their transformation into endogenous constraints, which are expressed, in the first place, as regards collective outcomes. Any given instance $\{SR(t), P_{SR}(t), LC_{SR}(t), O(t)\}$ of the structure $\{SR, P_{SR}, LC_{SR}, O\}$ pitches its real local outcomes $O(t)$ against their expected local state $E[O(t)]$. It is to stress that expected outcomes are essentially local, i.e. they are instrumental to absorbing external stressors. There are no grounds to assume something like a general state of expectation $E(O)$ in the structure $\{SR, P_{SR}, LC_{SR}, O\}$.

With the above-stated assumptions, the sequence of states from instance *{SR(t0), P_{SR}(t0), LC_{SR}(t0),O(t0)}* to *{SR(t), P_{SR}(t), LC_{SR}(t),O(t)}* can be studied as a Markov chain of states, which transform into each other through a σ-algebra. The current state *{SR(t), P_{SR}(t), LC_{SR}(t),O(t)}* and its expected outcomes *E[O(t)]* contain all the information from past learning, and therefore the local error in adaptation, i.e. $e(t) = \{E[O(t)] – O(t)\}*dO(t)$ - where *dO(t)* stands for the local derivative (local first moment) of *O(t)* - conveys all the information from past learning. Error *e(t)* in adaptation is factorised into a residual difference and a first moment, as it is assumed that any current state instance *{SR(t), P_{SR}(t), LC_{SR}(t), O(t)}* is essentially dynamic, i.e. on the move towards a subsequent state *{SR(t+1), P_{SR}(t+1), LC_{SR}(t+1), O(t+1)}* in the Markov chain. The use of Markov chains in this model is largely consistent with the author's understanding of the Interface Theory of Perception (Hoffman et al. 2015, Fields et al. 2018).

Theoretically, it is possible to assume $e(t) = 0$, with either $E[O(t)] = O(t)$ or $dO(t) = 0$, yet such a perfect standstill is rather a very special state than normal step in collective learning. A society experiencing $e(t) = 0$ does not learn anything new: all constraints are met, and no external stressors impose new constraints. The expected state of the structure *{SR, P_{SR}, LC_{SR}, O}* is $E[O(t)] \neq O(t)$ and $dO(t) \neq 0$, and therefore $\{SR(t), P_{SR}(t), LC_{SR}(t),O(t)\} = \{SR(t-1) + e(t-1), P_{SR}(t-1) + e(t-1), LC_{SR}(t-1) + e(t-1), O(t-1) + e(t-1)\}$ and $E[O(t)] = E[O(t-1)] + e(t-1)$. Collective learning is essentially incremental, and not revolutionary. Each consecutive state *{SR(t), P_{SR}(t), LC_{SR}(t),O(t)}* is a one-mutation neighbour of the immediately preceding state *{SR(t-1), P_{SR}(t-1), LC_{SR}(t-1),O(t-1)}* rather than its structural modification. Hence, we are talking about arithmetical addition rather than multiplication or division.

Constraints produced by the structure *{SR, P_{SR}, LC_{SR}, O}* in response to external stressors take two forms: recurrent and incidental. The former impact individual decisions to endorse a given social role, i.e. those decisions take into account the past state of the structure

$\{SR, P_{SR}, LC_{SR}, O\}$ and randomly distributed, current exogenous information $X(t)$. That random exogenous parcel of information affects all the people susceptible to endorse the given social role sr_i which, in turn, means arithmetical multiplication rather than addition, i.e. $P_{SR}(t) = X(t)*[P_{SR}(t-1) + e(t-1)]$.

Incidental exogenous stressors, thus events in the type of Black Swans, consist in short-term, violently disturbing events, likely to put some social roles extinct or, conversely, trigger into existence new social roles. Extinction of a social role means that its probability becomes null: $P(sr_i) = 0$. The birth of a new social role, on the other hand, means that some pre-existing skillsets gain social recognition from the distant social environment of people possessing them, and therefore turn into professions, crafts, business models etc.

Mathematically, it means that the set *SR* of social roles entails two subsets: active and dormant. Active social roles display $p(sr_i;t) > 0$, and, under the impact of a local, Black-Swan type event, they can turn $p(sr_i;t) = 0$. Dormant social roles are at $p(sr_i;t) = 0$ for now, and can turn into display $p(sr_i;t) > 0$ in the presence of a Black Swan.

In the presence of active recurrent stress upon the structure $\{SR, P_{SR}, LC_{SR}, O\}$, thus if we assume $X(t) > 0$, I can present a succinct mathematical example of Black-Swan-type exogenous disturbance, with just two social roles, sr_1 and sr_2. Before the disturbance, sr_1 is active and sr_2 is dormant. In other words, $P(sr_1; t -1)*X(t-1) > 0$ whilst $P(sr_2; t -1)*X(t-1) = 0$. With the component of learning by incremental error in a Markov chain of states, it means $[P(sr_1; t - 2) + e(t-2)]*X(t-1) > 0$ and $[P(sr_2; t -1) + e(t-2)]*X(t-1) = 0$, which logically equates to $P(sr_1; t - 2) > - e(t-2)$ and $P(sr_2; t -1) = - e(t – 2)$. After the disturbance, the situation changes dialectically, namely $P(sr_1; t -1)*X(t-1) = 0$ and $P(sr_2; t -1)*X(t-1) > 0$, implying that $P(sr_1; t - 2) = - e(t-2)$ and $P(sr_2; t -1) > - e(t – 2)$.

The above development leads to the issue of negative probabilities. With the assumptions stated above, in the local state $\{SR(t), P_{SR}(t), LC_{SR}(t), O(t)\}$ it is possible that some $p(sr_i) < 0$, i.e. $E[O(t-1)] > O(t-1)$, with $dO(t-1) > 0$. It is an otherwise frequent situation when the actual outcomes are below expectations, and yet display a positive gradient of change. The case of $p(sr_i) < 0$ is, technically, an impossible state. Can collective intelligence of a human society go into those impossible states? Quantum physics supply a possible interpretation in that respect. If the probability of an event is conditional on another probability, and this is precisely the case with the here-presented model, negative probability corresponds to an essentially intermediary state, i.e. a state impossible to hold or impossible to be verified directly (Feynman 1987; Curtright & Zachos 2001). That formal interpretation in quantum physics somehow mirrors the distinction between economic equilibrium, and the lack thereof, in the theory of economic cycles (Schumpeter 1939). In that perspective, states marked by $p(sr_i) < 0$ are the necessary spin and push in the long-term learning of the collectively intelligent social structure.

Economic sciences supply interesting stylized facts with respect to equilibriums. There are some fundamental proportions in societies, such as the average time worked per person per year, or the average consumption of energy per person per year, which we collectively adjust ourselves around without even noticing much of it, and yet there are visible trends of change in those proportions. Adjustment of size, would it be demographics or the gross real output, is a bit harder, in the sense that it entails occasional bumps and requires collective effort (e.g. proper economic policies). Adjustment in markets is probably the hardest and the bumpiest, whence the common observation that prices and their gradient of change are the beating pulse of the economic system and can become volatile under new external stressors. Hence, a working hypothesis is being formulated for the empirical application of the above-presented model: '**Under the sudden, Black-Swan-type impact of external stressors, human societies**

can develop one of the three possible paths of collectively intelligent learning: a) structural, cyclical adjustment without visible social change b) adjustment of size with visible social change and c) long-term destabilization'.

The dataset and the method of analysis

Penn Tables 9.1 (Feenstra et al. 2015) have been chosen as source of empirical data. It is assumed that each variable in that database is the outcome of a collectively intelligent social structure $\{SR, P_{SR}, LC_{SR}, O\}$, and, accordingly, each country-year observation in that variable is the expected local outcome $E[O(t)]$ coupled with a local instance $\{SR(t), P_{SR}(t), LC_{SR}(t), O(t)\}$. That coupling, together with the entire model developed in the preceding section, is translated into an artificial neural network designed by the author of this article. As the neural network requires a database with no empty cells inside, the initial contents of the Penn Tables 9.1 have been reduced to $N = 3006$ fully filled country-year observations. The basic cycle of learning applied to the neural network is precisely N = 3006 experimental instances.

The architecture of the network is based on two streams: the mainstream of adaptive walk in rugged landscape, with random recurrent disturbance, on the one hand, and incidental deep disturbance akin to Black-Swan-type events. For the sake of presentational clarity, the former is further designated as **the stream of learning**, whilst the latter is **the stream of disturbance**. Since the stream of disturbance is relatively simpler in its architecture, it is being presented in the first place, and, in the next presentational step, it is being incorporated into the stream of learning as regards the $\{SR, P_{SR}, LC_{SR}, O\}$ structure.

The stream of disturbance starts with the input neuron (input layer), which generates quasi-random numbers $X(t)$ between 0 and 1. This neuron corresponds to the strictly speaking occurrence of random disturbance. The next two neurons (layers) transform the raw happening into stimuli for the intelligent structure $\{SR, P_{SR}, LC_{SR}, O\}$. The first phase of transformation consists in giving to that raw event as many dimensions as there are input variables in the network, thus as many as there are social roles sr_i in the component set *SR* of the structure. Since that number, in the model, is '*m*', the second neuron of disturbance takes *X(t)* to the power *m*. The idea behind this logical step is that each social role in the structure $\{SR, P_{SR}, LC_{SR}, O\}$ absorbs just a local impact of the Black Swan. The third neuron transforms that multi-dimensional event into neural perception, through the function of hyperbolic tangent, or $exp\{2*[X(t)^m]\} – 1/ exp\{2*[X(t)^m]\} + 1$. For clarity, the value $exp\{2*[X(t)^m]\} – 1/ exp\{2*[X(t)^m]\} + 1$ is further called '**observable disturbance**' or *DO(t)*. Finally, the fourth neuron of disturbance compares *DO(t)* to the error *e(t-1)* generated in the preceding experimental round. If $DO(t) > e(t-1)$, the stream of disturbance feeds the disturbance strictly spoken to the stream of learning (the logic of that particular feed is explained further in this section). In the opposite case, thus when $DO(t) \le e(t-1)$, the stream of disturbance remains passive. With *X(t)* being quasi-randomly generated by the software supporting the network, the procedure 'Generate *X(t)* => Take *X(t)* to power m => take hyperbolic tangent of *X(t)* power m' generates incidental, essentially non-cyclical feed of disturbance. The reader can compare this approach to that proposed by Prestwich (2019 op. cit.).

In the stream of learning, the network contains $m = 40$ social roles in the component set *SR* of the structure, and therefore **the input layer of learning** consists in $m = 40$ probabilities $P_{SR} = \{p(sr_1), p(sr_2), ..., p(sr_{40})\}$. The $m = 40$ social roles SR are divided into 20 active social roles, with initial probabilities $p(sr_i,t0) > 0$ randomly generated in a standardized normal distribution, and 20 dormant social roles with initial probabilities $p(sr_i,t0) = 0$. Subsequent probabilities t1 ÷ t3005 are being fed forward from the output layer of the learning stream. That forward feed is conditional: at this point, the network starts generating conditional probabilities. That entails the interpretation developed earlier in the model: negative conditional probabilities are real, yet they designate essentially transitory states. If *DO(t) > e(t-*

1), then $p(sr_i,t) = 0$ for active social roles and $p(sr_i,t) = X(t)$ for dormant social roles, where $X(t)$ is once again a randomly generated value. If, on the other hand, $DO(t) \leq e(t-1)$, then $p(sr_i,t) = p(sri,t-1)+e(t-1)$ for the active social roles, and $p(sr_i,t) = 0$ in the dormant ones. This is the first impact of disturbance on learning: it breaks the replication of active social roles, breaks incremental learning, and triggers into activity dormant social roles.

Given further arrangement of subsequent layers, it can be assumed that each social role sr_i in the input layer is a separate neuron, as it separately projects into the **2nd layer**, which consists in another set of 40 neurons, each computing the lateral coherence $LC(sr_i)$ with $m - 1$ other social roles in the set *SR*, and that coherence is defined as the average Euclidean distance between $p(sr_i)$ and the probabilities $p(sr_j)$ of other social roles, as in equation (1) below. The use of Euclidean distance as measure of coherence inside a social structure is based on the swarm theory, where learning in a collective manifests as sequential shift between different strengths of behavioural coupling between individuals and groups (Stradner et al. 2013[1]). The 2nd layer of learning is not directly exposed to disturbance.

$$LC(sr_i) = \frac{\sum_{j=1}^{m-1} \sqrt[2]{[p(sr_i)-p(sr_j)]^2}}{m-1} \qquad (1)$$

Neurons from both the 1st and the 2nd layer project into the third one, which generates quasi-randomly weighed stimulus for neural activation. **That 3rd layer of learning** is also made of $m = 40$ neurons, which absorb the recurrent feed from layers 1 and 2, whilst generating another recurrent random factor $X(t)$, and absorbing the disturbance $DO(t)$. One remark is due, once again for clarity: random numbers $X(t)$ are generated both in the stream of disturbance and in the stream of learning, but they are independently random. In the third layer of learning, if $DO(t) > e(t-1)$, then stimulus $TI(sr_i,t) = X(t)*p(sr_i,t)$. Should $DO(t) \leq e(t-1)$, then $TI(sr_i,t) =$

[1] Stradner, J., Thenius, R., Zahadat, P., Hamann, H., Crailsheim, K., & Schmickl, T. (2013). Algorithmic requirements for swarm intelligence in differently coupled collective systems. Chaos, Solitons & Fractals, 50, 100-114

$X(t)*p(sr_i,t)*lc(sr_i,t-1)$. Given the logical structure of learning layers 1 and 2, the third layer reinforces the impact of the disturbance. When $DO(t) > e(t-1)$, the ad-hoc activated dormant social roles are fed into the neural activation function, next in line, with no link to lateral coherence, whilst the so-far active social roles remain zeroed. On the other hand, if disturbance remains at $DO(t) \leq e(t-1)$, dormant social roles remain dormant, and active social roles are fed into neural activation with correction for their lateral coherence.

The 3rd layer of learning projects into the **4th layer of learning**, which aggregates stimuli. It consists of one neuron which produces $h(t) = \sum_{(i=1 \to i=m)} TI(sr_i,t)$. The immediately following **5th neuron of learning** activates aggregate stimuli, using the through the function of hyperbolic tangent, thus it produces $tanh(t) = exp[2*h(t)] - 1/ exp[2*h(t)] + 1$. Finally, the 6th layer of learning produces the output of the network, in the form of error $e(t)$, through the equation $e(t) = \{E[O(t)] - tanh(t)\}*[1 - tanh(t)^2]$, and feeds that output to the next experimental round $t+1$.

<u>Simplified explanation of the neural network</u>

The stream of disturbance: $X(t) \Rightarrow X(t)^m \Rightarrow DO(t) = exp\{2*[X(t)^m]\} - 1/ exp\{2*[X(t)^m]\} + 1 \Rightarrow [DO(t) - e(t-1)] \Rightarrow$ **feed to the stream of learning**

The stream of learning : $[p(sr_i)] \Rightarrow [lc(sr_i)] \Rightarrow$ [if $DO(t) > e(t-1)$, then $TI(sr_i,t) = X(t)*p(sr_i,t)$] and [if $DO(t) \leq e(t-1)$, then $TI(sr_i,t) = X(t)*p(sr_i,t)*lc(sr_i,t-1)$] $\Rightarrow$ [$h(t) = \sum TI(sr_i,t)] \Rightarrow \{tanh(t) = exp[2*h(t)] - 1/ exp[2*h(t)] + 1\} \Rightarrow e(t) = \{E[O(t)] - tanh(t)\}*[1 - tanh(t)^2] \Rightarrow$ **feed forward to the next experimental round of learning**.

Empirical data from the $N = 3006$ fully filled country-year observations in Penn Tables 9.1 is fed into that neural network as $E[O(t)]$. This research is supposed to privilege the explanation of technological change, whilst staying consistent with the hypothesis phrased out in conclusion of the theoretical model. Therefore, among the variables of Penn Tables 9.1., the following ones have been selected as *E[O(t)]*. Acronyms in parentheses correspond to those used in Penn Tables 9.1:

a) Variables informative about structural proportions: average hours worked per person per year (*AVH*), average depreciation rate of the capital stock (*delta*), human capital index based on years of schooling and returns to education (*hc*), real internal rate of return (*IRR*), share of labour compensation in GDP at current national prices (*labsh*), Total Factor Productivity level at current PPPs (USA=1) (*CTFP*);

b) Variables informative about the absolute size of social structure: population (*POP*), output-side real GDP at chained PPPs (in mil. 2011US$) (*RGDPO*), capital stock at constant 2011 national prices (in mil. 2011US$) (*rnna*);

c) Variables informative about markets: price level of household consumption, price level of USA GDPo in 2011=1 (*PL_C*); price level of capital formation, price level of USA GDPo in 2011=1, (PL_I); price level of exports, price level of USA GDPo in 2011=1 (*PL_X*);

d) Control variable: randomly weighted compound of all variables, computed as $\sum[x_i*X(t)]$;

With each of the above-named variables as its *E[O(t)]*, the network did N = 3006 experimental rounds of learning, going through the corresponding country-year observations. Thus, 13 alternative networks have been created. Their learning process has been studied as regards three main aspects. Firstly, the gradient of error *e(t),* i.e. the capacity to reduce error, to magnify it, or to make it oscillate monotonously has been studied. In the lines of the model developed in earlier in this article, the gradient of error represents the way which the structure

$\{SR, P_{SR}, LC_{SR}, O\}$ absorbs external disturbances. The average lateral coherence $lc(sr_i,t)$ across all the social roles in simulation was the second essential observable. Its oscillation and drift show the degree to which the structure $\{SR, P_{SR}, LC_{SR}, O\}$ releases or tightens its internal coherence in the presence of random external stressor. Thirdly, and finally, the average expected probabilities $p(sr_i,t)$ are calculated, for each social role separately, across all the 3006 experimental rounds. This is informative about the expected state of the structure $\{SR, P_{SR}, LC_{SR}, O\}$ in the given empirical setting, i.e. with a given real socio-economic variable as its expected output.

Results

In the Appendix, at the end of this article, the reader can find links to 13 Excel files, freely downloadable, where each file corresponds to the neural network described in the preceding section working with a different output variable $E[O(t)]$. The author believes that the wealth of observations possible to make when directly studying the way those networks behave goes beyond the necessarily succinct summary of results given here below. Figures 1 ÷ 6, further in the Appendix, exemplify the distribution of error $e(t)$ over the entire cycle of N = 3006 experimental rounds of learning. Those error curves are curated as for their recurrent shape over repeated experiments with each network. In the author's opinion, that shape is the fundamental representation of the learning process observable in the structure $\{SR, P_{SR}, LC_{SR}, O\}$. Table 1, below, presents the expected average probabilities $p(sr_i)$ computed for each of the 40 social roles sr_i in the structure tested. Probabilities are presented separately for two groups of sr_i, i.e. the initially active ones as opposed to those initially dormant. Just as for graphs of error, these probabilities have been curated through multiple iterations of the experiment run with each respective output variable $E[O(t)]$.

Table 1

Output variable E[O(t)]	**Expected average *p(sr_i)* in the initially ACTIVE social roles**	**Expected average *p(sr_i)* in the initially DORMANT social roles**
AVH	0,35	0,04 ÷ 0,05
LABSH	0,38	0,04 ÷ 0,05
IRR	0,19	0,03 ÷ 0,04
CTFP	0,24	0,04
DELTA	0,31	0,03 ÷ 0,04
POP	0,07	0,05 ÷ 0,06
RGDPO	0,05	0,05 ÷ 0,06
RNNA	0,05	0,05 ÷ 0,06
HC	0,38	0,04 ÷ 0,05
PL_I	0,07	0,05
PL_C	0,25	0,03
PL_X	0,24	0,03 ÷ 0,04
Weighted compound index of all the variables	0,56	0,12 ÷ 0,13
Lateral average of initial probabilities in experimental round t0	**0,33 ÷ 0,35**	**0,00**

Empirical results brought by the here-discussed research can be roughly divided in two viewpoints. The expected state of the structure $\{SR, P_{SR}, LC_{SR}, O\}$, under different constraints of exogenously imposed expected outcomes *E[O(t)]*, is one perspective. The way that expected state comes into being, thus the strictly spoken behaviour of that structure, makes another type of insight.

The expected states of the structure $\{SR, P_{SR}, LC_{SR}, O\}$, as measured with average probabilities brought forth in Table 1, do differ according to the actual outcome being optimized. That disparity is much more pronounced as regards the initially active social roles, as compared to the dormant ones. Interestingly, disparity is almost inexistent inside those two categories, and, even more interestingly, any kind of differentiation happens only between the initially dormant social roles, whilst those initially active land as uniformly distributed. That seems counterintuitive. The initial state of active social roles is a quasi-random, Gaussian distribution, whilst dormant social roles start from a uniform distribution of null values. The neural network used to process those initial states somehow inverted their structures. The initially randomly distributed probabilities in active social roles end up as uniform, whilst the initially uniformly zeroed probabilities of dormant roles finish as slightly differentiated.

Thus, the network has the property of levelling out disparate states of the world, and conversely, adding some disparity to the uniform ones. On the other hand, as uniformized as $p(sr_i)$ become across the initially active social roles, they are truly disparate according to the empirical variable from Penn Tables 9.1., taken as the expected outcome *E[O(t)]*, thus the empirical base for assessing error *e(t)*. Three categories of experiments delineate themselves. Networks pegged on, respectively, the headcount of population, real output, real capital stock, and price level in capital formation, make the first category: they produce noticeably low expected $p(sr_i)$ in the initially active social roles, making them practically level with those initially dormant. These specific socio-economic outcomes have the property of spreading the occurrence of all 40 social roles almost evenly across the structure.

On the other end of the spectrum, the network pegged on a randomly weighted compound index made of all the variables in Penn Tables 9.1. produces the highest expected probabilities, across the board, especially in the initially active social roles. It is the only experiment which yields a two-digit expected probability in the occurrence of initially dormant social roles. Between those two poles, networks experimenting with AVH, LABSH, IRR, DELTA, HC, PL_C and PL_X produce two-digit probabilities in the initially active social roles, accompanied by one-digit probabilities in the initially dormant ones. Interestingly, networks optimizing AVH, LABSH, and HC, thus variables pertinent to the labour market, produce expected probabilities of the initially active social roles practically equal to their initial, lateral average, thus around 0,35. These three networks seem to be essentially conserving the occurrence of active social roles, and add to them a fringe occurrence of the initially dormant roles.

Expected probabilities produced by the network bring the first generalization: the network used in empirical research, and therefore the theoretical model which this network is based on, produce different expected states of the structure *{SR, P_{SR}, LC_{SR}, O}*. The way of achieving those expected states is the next point to discuss. The network which makes the structure *{SR, P_{SR}, LC_{SR}, O}* optimize **Total Factor Productivity** (CTFP) is probably the most intriguing among the 13 studied as regards the process of learning (see Figure 2 in the Appendix). Whilst it produces somehow median probabilities as the expected state of the structure, the way it arrives to that outcome is puzzling, even in the light of assumptions laid out in the theoretical model. Error amplifies all the long of the learning process and that is unusual for a simple network like this one. One is almost tempted to call that process 'unlearning'. The average lateral coherence $lc(sr_i)$ between social roles swings cyclically, and steadily increases across the 3006 rounds of learning. The cyclicality of $lc(sr_i)$ is, by the way, a common denominator across all the networks studied. It looks as if that specific intelligent structure was sequentially tightening its internal coherence and releasing it. The cycle is arhythmical, and yet visible. Both the graph of error *e(t)*, and the graph of average coherence $lc(sr_i)$ look, in this CTFP-pegged network, like Ito processes with a clear upwards drift.

One can observe a somehow weaker version of the that strange process of 'unlearning' when the network is optimizing **price level in capital formation** (PL_I). Error *e(t)* tends to widen its swing over the 3006 experimental rounds, yet there is no visible drift in it. Interestingly, other market-based variables - price level in household consumption (PL_C) and price level in exports (PL_X) – produce a different process of learning, with the graph of error much more similar to that produced by the network, when it is pegged on variables computed as structural proportions: AVH, LABSH, IRR, DELTA, and HC. These experiments yield a cyclical error, which neither widens nor narrows down its amplitude and follows something like a steady swing with no visible drift.

Experiments which force the network to optimize size-type variables, namely POP, RGDPO, and RNNA, are the only ones to produce a process possible to label as true learning, at least from the standpoint of artificial intelligence. The network pegged on population seems to be the most salient in this respect: error *e(t)* visibly narrows down, producing something akin equilibrium between the randomly disturbed structure $\{SR, P_{SR}, LC_{SR}, O\}$ and the expected output, thus the number of humans being around.

Final discussion

Results produced by experiments conducted with the neural network allow exploring and verifying the working hypothesis, formulated as conclusion of the theoretical model. Three different patterns emerge, indeed, corresponding to the three hypothetical paths of collectively intelligent learning: cyclical adjustment devoid of clear trajectory, long-term adjustment in size-related variables, with visible learning, and, finally, the path of destabilization under the repeated impact of random disturbances.

How exactly should the probabilities $p(sr_i)$ produced by the network be interpreted as social phenomena? Probabilities are likelihoods that a randomly chosen individual endorses a social role. Relatively high probabilities mean that individuals are likely to endorse many social roles at once. This is a society of quick learning across many, socially relevant skillsets. On the other hand, low probabilities correspond to a compartmentalized state of society, more likely to follow the model of one person endorsing just one social role.

The cyclically oscillating lateral coherence $lc(sr_i)$, observed across all the experiments, run with different variables as expected outcomes $E[O(t)]$, connects to the previously cited swarm theory. It seems that absorption of Black Swan events, in the presence of any desired outcome, requires a cycle of loosening and tightening in behavioural coupling between social roles. Once again, the experiment pegged on Total Factor Productivity (CTFP) is original in that respect, as, amidst the usual cyclical swing of *lc(sr*$_i$*)*, it shows an upwards drift in that metric, thus a downwards drift in the internal coherence of the structure *{SR, P_{SR}, LC_{SR}, O}*.

What is so exceptional about Total Factor Productivity CTFP, as measured in Penn Tables 9.1., to make it destabilize the neural network is such a peculiar way? Why does this specific experiment behave as if the metric of productivity was particularly sensitive to random disturbances induced by the network, thus as if it was a highly volatile market price? Tentatively, it can be assumed that Total Factor Productivity is greatly impacted by the price of capital goods, perhaps more than by anything else. That explanation would coincide with the observably similar characteristics of the experiment pegged on the price level of capital formation AKA PL_I. Still, that leaves us with two more puzzles. Why do networks pegged on variables pertinent to the labour market (AVH, LABSH, HC) behave so differently? Perhaps, in the presence of random disturbance, the labour market is more accommodative and sort of 'sucks in' the randomly happening Black Swans, whilst the market of capital goods is much more prone to going haywire? Furthermore, why do experiments pegged on other price levels, namely household consumption (PL_C), and exports (PL_X) behave as if these variables were structural proportions rather than market prices? Perhaps they are.

The theoretical model which served to run the empirical research is largely based on the theory of evolutionary games by Hammerstein & Selten (1994 op. cit.). Empirical research in itself allows formulating interesting hypotheses for future research in that respect. The occurrence of random disturbance in the type of Black Swan events seems to have the property of dragging dormant patterns of behaviour, thus niche strategies, out of the shadow, and that whatever kind of outcome is the social structure pursuing. Black Swan events could be seen as beneficial, as they force to value and acknowledge fringe skillsets.

Finally, empirical results can be translated into the actual social reality we are living at the moment of writing this article. Black Swans coming from biology (pandemic), from climate change or from political instability are the most likely to destabilize capital markets. As strange as it may sound, our lifestyles, as well as the size of our economy, are much less likely to go south under the impact of those events. In these respects, our social system seems to be much more absorptive and accommodative in relation to environmental stressors.

References

1) Ale, B. J., Hartford, D. N., & Slater, D. H. (2020). Dragons, black swans and decisions. Environmental research, 183, 109127. https://doi.org/10.1016/j.envres.2020.109127
2) Aven, T. (2015). Implications of black swans to the foundations and practice of risk assessment and management. Reliability Engineering & System Safety, 134, 83-91. https://doi.org/10.1016/j.ress.2014.10.004
3) Batrouni, M., Bertaux, A., & Nicolle, C. (2018). Scenario analysis, from BigData to black swan. Computer Science Review, 28, 131-139. https://doi.org/10.1016/j.cosrev.2018.02.001
4) Chen, D. H., & Huang, H. L. (2018). Panic, slash, or crash—Do black swans flap in stock markets?. Physica A: Statistical Mechanics and its Applications, 492, 1642-1663. https://doi.org/10.1016/j.physa.2017.11.087
5) Curtright, T., & Zachos, C. (2001). Negative probability and uncertainty relations. Modern Physics Letters A, 16(37), 2381-2385.
6) Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" American Economic Review, 105(10), 3150-3182, available for download at www.ggdc.net/pwt
7) Feynman, R. P. (1987). Negative probability. Quantum implications: essays in honour of David Bohm, 235-248,
8) Fields, C., Hoffman, D. D., Prakash, C., & Singh, M. (2018). Conscious agent networks: Formal analysis and application to cognition. Cognitive Systems Research, 47, 186-213. https://doi.org/10.1016/j.cogsys.2017.10.003
9) Flage, R., & Aven, T. (2015). Emerging risk–Conceptual definition and a relation to black swan type of events. Reliability Engineering & System Safety, 144, 61-67. https://doi.org/10.1016/j.ress.2015.07.008
10) Hajikazemi, S., Ekambaram, A., Andersen, B., & Zidane, Y. J. (2016). The Black Swan–Knowing the unknown in projects. Procedia-Social and Behavioral Sciences, 226, 184-192. https://doi.org/10.1016/j.sbspro.2016.06.178
11) Hammerstein, P., & Selten, R. (1994). Game theory and evolutionary biology. Handbook of game theory with economic applications, 2, 929-993.
12) Hoffman, D. D., Singh, M., & Prakash, C. (2015). The interface theory of perception. Psychonomic bulletin & review, 22(6), 1480-1506.
13) Lin, W. Y., & Tsai, I. C. (2019). Black swan events in China's stock markets: Intraday price behaviors on days of volatility. International Review of Economics & Finance, 59, 395-411. https://doi.org/10.1016/j.iref.2018.10.005
14) Prestwich, S. D. (2019). Tuning Forecasting Algorithms for Black Swans. IFAC-PapersOnLine, 52(13), 1496-1501. https://doi.org/10.1016/j.ifacol.2019.11.411
15) Rhee, S. G., & Wu, F. H. (2020). Conditional extreme risk, black swan hedging, and asset prices. Journal of Empirical Finance, 58, 412-435. https://doi.org/10.1016/j.jempfin.2020.07.002
16) Schumpeter, Joseph Alois. Business cycles. Vol. 1. New York: McGraw-Hill, 1939.
17) Simon, H. A. (1955). A behavioral model of rational choice. The quarterly journal of economics, 69(1), 99-118.
18) Stradner, J., Thenius, R., Zahadat, P., Hamann, H., Crailsheim, K., & Schmickl, T. (2013). Algorithmic requirements for swarm intelligence in differently coupled collective systems. Chaos, Solitons & Fractals, 50, 100-114

19) Talcott Parsons (1951;1991). The Social System. first published 1951, edition 1991, Routledge, ISBN 0-203-99295-4
20) Taleb, N. N. (2007). The black swan: The impact of the highly improbable (Vol. 2). Random house.
21) Taleb, N. N., & Blyth, M. (2011). The black swan of Cairo: How suppressing volatility makes the world less predictable and more dangerous. Foreign Affairs, 33-39.
22) Wang, Y., Cao, X., Sui, X., & Zhao, W. (2019). How do black swan events go global?- Evidence from US reserves effects on TOCOM gold futures prices. Finance Research Letters, 31. https://doi.org/10.1016/j.sbspro.2016.06.178
23) Zeng, C. J., Qi, E. P., Li, S. S., Stanley, H. E., & Fred, Y. Y. (2017). Statistical characteristics of breakthrough discoveries in science using the metaphor of black and white swans. Physica A: Statistical Mechanics and its Applications, 487, 40-46. https://doi.org/10.1016/j.physa.2017.05.041

Appendix

Here below you can find the list of links to Excel files with neural networks pegged on, as *E[O(t)]*, the individual variables used in the empirical research. The name of each file contains the name of the corresponding *E[O(t)]* variable:

i. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-AVH_Gaussian.xlsx
ii. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-CTFP_Gaussian.xlsx
iii. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-Delta_Gaussian.xlsx
iv. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-HC_Gaussian.xlsx
v. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-IRR_Gaussian.xlsx
vi. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-LABSH_Gaussian.xlsx
vii. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-PL_C_Gaussian.xlsx
viii. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-PL_I_Gaussian.xlsx
ix. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-PL_X_Gaussian.xlsx
x. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-POP_Gaussian.xlsx
xi. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-RGDPO_Gaussian.xlsx
xii. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-RNNA_Gaussian.xlsx
xiii. https://discoversocialsciences.com/wp-content/uploads/2020/10/Perceptron-Black-Swans-pegged-on-WEIGHED-COMPOUND_Gaussian.xlsx

Figure 1

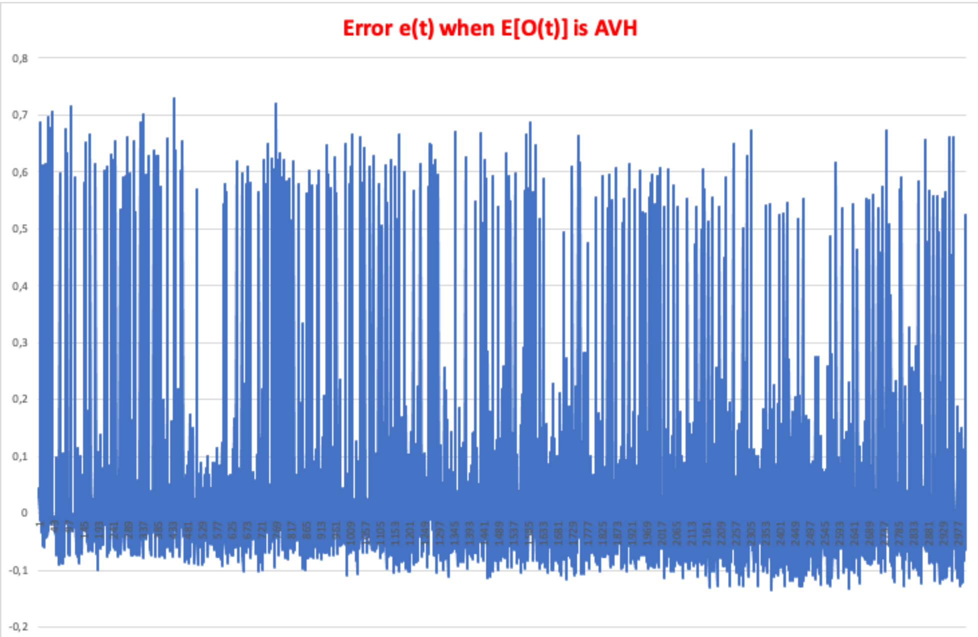

Figure 2

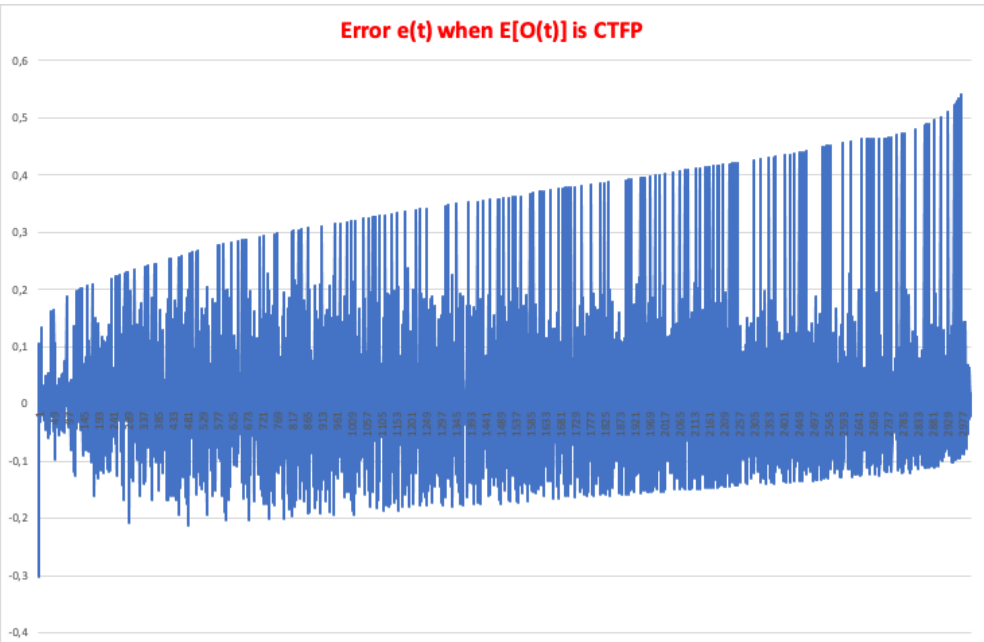

Figure 3

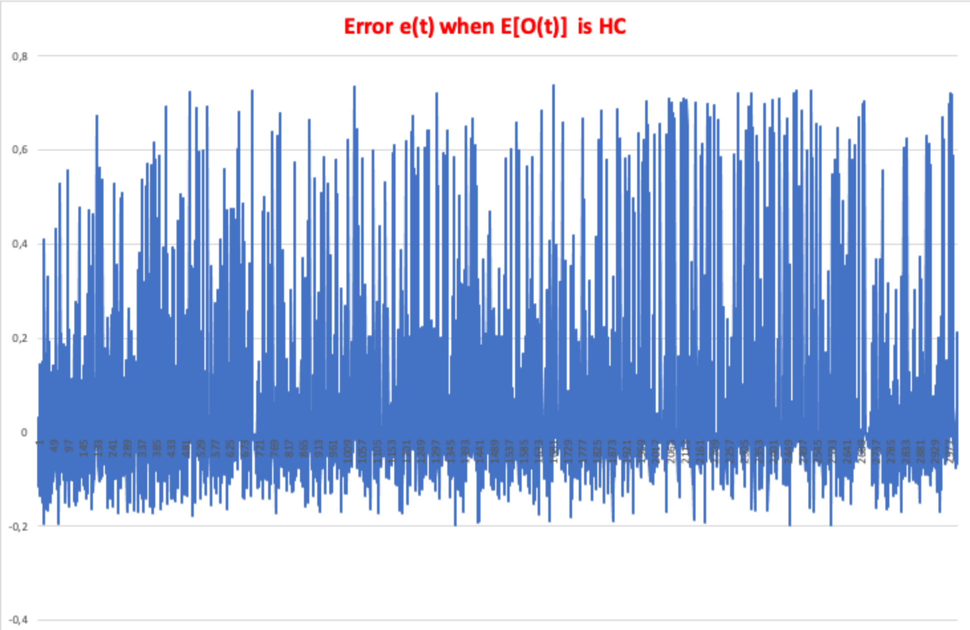

Figure 4

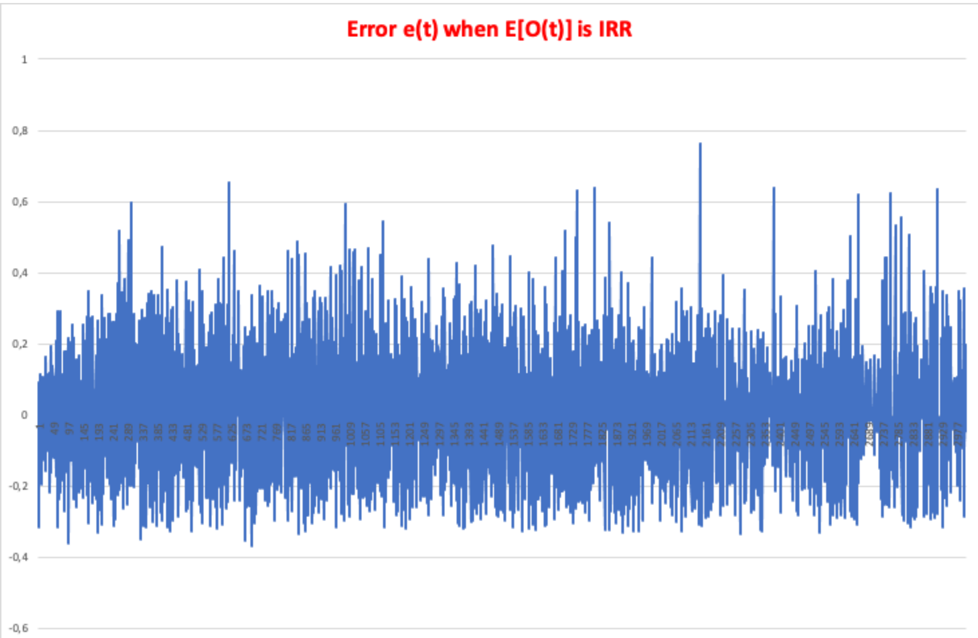

Figure 5

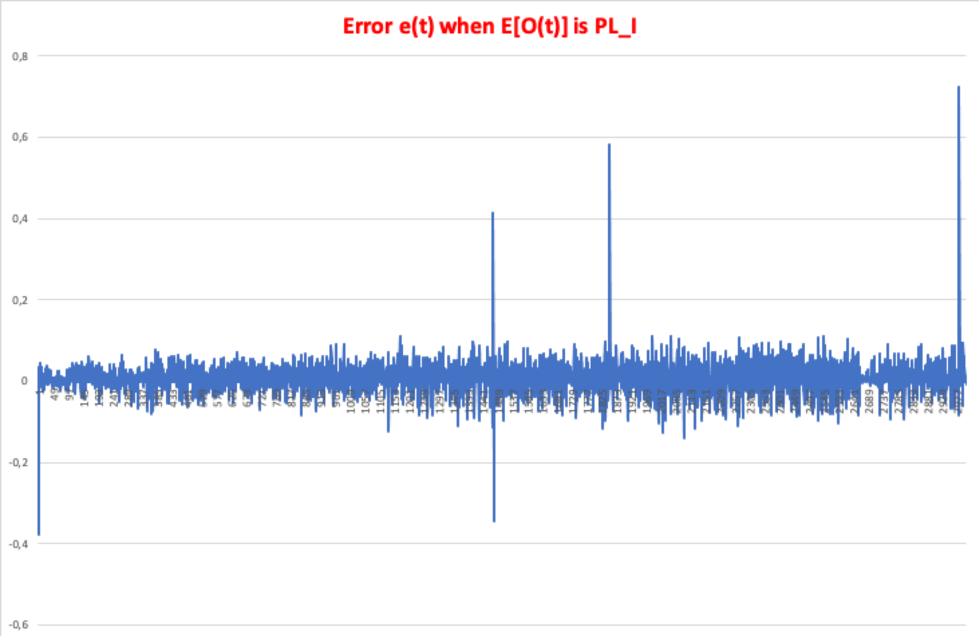

Figure 6