



Co-evolutionary Dynamics in a Simulation of Interacting Financial-Market Adaptive Automated Trading Systems

Dave Cliff¹

¹Department of Computer Science, University of Bristol, Bristol BS8 1UB, U.K.

Abstract

In present-day major financial markets around the world, adaptive automated trading systems are responsible for many more transactions than are human traders, and human traders have largely been replaced by trading machines that can process super-human quantities of data and react to market events at super-human speeds: this paper reports new results from a high-fidelity simulation of such highly automated financial markets, populated by minimal adaptive trading strategies, and presents results showing that competition between trading systems leads to unstable market dynamics and convergence on economically suboptimal outcomes. The simulations reported here used the public-domain open-source market simulator BSE, which offers an accurate model of a contemporary limit-order-book financial exchange, which was then extended by adding a collection of individual adaptive automated trading entities (“traders”), each of which uses a simple stochastic hill-climbing optimizer to adapt its zero-intelligence trading strategy over time, trying always to improve profitability. The traders interact and transact with each other on sub-second timescales, but this paper focuses on the simulated market’s dynamics over extremely long periods of times (hundreds of days, during which time many hundreds of millions of transactions can occur) and shows that while the system can remain in seemingly stable states for protracted periods, the overall long-term dynamics of the system can be unstable and economically inefficient. The extended BSE source-code written for these simulations is freely available on GitHub, for use by other researchers.

Keywords: Financial Markets; Automated Trading; Agent-Based Model; Zero Intelligence; Co-Evolutionary Dynamics.

1. Introduction

In attempting to understand and predict the fine-grained dynamics of financial markets, there is a long tradition of studying simulation models of such markets. Simulation studies nicely complement the two primary alternative lines of enquiry: analysis of real market data recorded at fine-grained temporal resolution, as is studied in the branch of finance known as *market microstructure*; and running carefully planned experiments where human subjects interact in artificial markets under controlled laboratory conditions, i.e. *experimental economics*. Simulation modelling of financial markets very often involves creating agent-based models (ABMs) that populate a market mechanism with some number of *trader-agents*: autonomous entities that have “agency” in the sense that they are em-

powered to buy and/or sell items within the particular market mechanism that is being simulated. This approach, known as *agent-based computational economics* (ACE), has a history stretching back for more than 30 years: for surveys see Farmer et al. (2005); Ladley (2012); Axtell and Farmer (2018). Over that multi-decade history, a small number of specific zero-intelligence (ZI) and/or minimal-intelligence trader-agent algorithms, i.e. precise mathematical and procedural specifications of simple but surprisingly effective trading strategies, have been frequently used for modelling various aspects of financial markets, and the convention that has emerged is to refer to each such strategy via a short sequence of letters, an acronym or abbreviation reminiscent of a stock-market ticker-symbol. Notable trading strategies in this literature include (in



chronological sequence): SNPR (AKA *Kaplan's Sniper*, as described in Rust et al. (1992)), ZIC (*Zero Intelligence Constrained*: Gode and Sunder (1993)); ZIP (*Zero Intelligence Plus*: Cliff (1997)); GD (*Gjerstad-Dickhaut*: Gjerstad and Dickhaut (1998)); MGD (*Modified GD*: Tesauro and Das (2001)); GDX (*GD eXtended*: Tesauro and Bredin (2002)); HBL (*Heuristic Based Learning*: Gjerstad (2003)) and AA (*Adaptive Aggressive*: Vytelingum et al. (2008)).

Of these, the seminal ZIC by Gode and Sunder (1993) is noteworthy for being both highly stochastic and extremely simple, and yet it gives surprisingly human-like market dynamics; GD and ZIP were the first two strategies to be demonstrated as superior to human traders, a fact first established in a landmark paper by IBM researchers Das et al. (2001) (see also: De Luca and Cliff (2011a,b); De Luca et al. (2011)), which is now commonly pointed to as initiating the rise of algorithmic trading in real financial markets; and until very recently AA was widely considered to be the best-performing strategy in the public domain. With the exception of SNPR and ZIC, all later strategies in this sequence are adaptive, using some kind of machine learning (ML) or artificial intelligence (AI) method to modify their responses over time, better-fitting their trading behavior to the specific market circumstances that they find themselves in, and details of these algorithms were often published in major AI/ML conferences and journals.

In a recent paper, Cliff (2021) I introduced a *parameterised-response zero-intelligence* (PRZI) trading algorithm, in which the “urgency” of trader i is controlled by a single continuous scalar strategy-value $s_i \in [-1.0, +1.0,] \in \mathbb{R}$: if $s_i = 0.0$, a PRZI trader behaves identically to ZIC, which generates quote-prices as draws from a uniform random distribution; but as s_i is moved closer to $+1.0$ its trading activity becomes more urgent (i.e., the trader alters the distribution that its quote-prices are drawn from, thereby biasing the prices it quotes in the market toward values more likely to be snapped up by a willing counterparty, but yielding less profit for the trader); and as s_i is moved closer to -1.0 the trading strategy becomes more “relaxed” (i.e., it biases its quote-prices toward make more more profit, which are hence less likely to be attractive to potential counterparties). At the extremes, when a PRZI trader i has $s_i = -1.0$ its trading strategy is equivalent to the maximally-relaxed SHVR strategy introduced in Cliff (2012); and when $s_i = +1.0$ the PRZI trader is acting as the maximally urgent GVWY strategy, also introduced in Cliff (2012) and explained further in Cliff (2018). PRZI was developed for use in a variety of contexts in simulation modelling of contemporary financial markets, one of which was to explore the co-evolutionary dynamics of markets in which all traders are simultaneously adapting their trading strategy, each trying to improve their own profitability, but each burdened by the uncertainty and complexity of trying to adapt to a market environment in which every other trader is also adapting, also changing its strategy in real time, all the time.

This paper presents the first results from such simulation studies, populating the public-domain BSE market simulator (an open-source ABM of a limit-order-book financial exchange: see Cliff (2012, 2018)) with adaptive PRZI traders, and studying the long-term dynamics of co-evolutionary markets, exploring whether they converge on stable equilibria, and/or whether they converge on economically efficient outcomes. The results presented here demonstrate that the dynamics can be unstable, with the system’s constant co-evolutionary “progress” eventually leading in cycles, back to states that it had previously evolved away from; results shown here also demonstrate that the system can evolve to economically stable but sub-optimal sets of conditions. Avenues for further research are discussed at the end of this paper. As supplementary background material, Appendix A gives a brief introduction to the *Recurrence Plot* visualisation technique used here; and Appendix B gives an overview of the BSE simulator’s system architecture.

2. Background

In the market-simulation literature surveyed here, typically each trader in the market is assigned a role, either *buyer* or *seller*, with the number of buyers being N_{Buy} , the number of sellers N_{Sell} , and the total number of traders in the market being $N_T = N_{Buy} + N_{Sell}$. Each buyer(seller) is periodically assigned an order to buy(sell) a fixed quantity (very often a single item) of the single arbitrary abstract commodity that is being traded on the market, along with a private (secret, known only to that trader) *limit price* denoted by λ , which for a buyer(seller) is the maximum(minimum) unit-price that they can buy(sell) the commodity at. Traders accrue *profit* (sometimes referred to as *utility* or *surplus*): an individual trader i ’s profit π_i on a transaction at price p is given by $\pi_i = |p - \lambda_i|$.

If we were to allow *only* a single PRZI trader i to adaptively vary its s_i value, trying to find the best setting of s_i relative to whatever distribution of $s_{j \neq i}$ values is present in the market (i.e., relative to the current mix of other strategies in the market) then we could say that i is *evolving* its value of s_i to try to find an optimum, the most profitable setting for its strategy parameter, given the unchanging set of fixed strategies that it is pitted against in the market. But when *every* PRZI trader in the market is simultaneously adapting its s -value, the system is *co-evolutionary* because what is an optimal setting of the s parameter for any one trader will likely depend on the s -values currently chosen by many or perhaps all of the other traders in the market. That is, the profitability of i is dependent not only on its own strategy value s_i but also on many or perhaps all other $s_{j \neq i}$ values in play at any particular time, and in principle all the strategy values will be altering all the time.

A primary motivation for studying such co-evolutionary markets with adaptive PRZI traders is the desire to move beyond prior studies of markets populated by adaptive automated traders in which the

“adaptation” merely involves selecting between one of typically only two or three fixed strategies (as in, e.g., Walsh et al. (2002); Vytelingum et al. (2008); Vach (2015)). The aim here is to create minimal model markets in which the space of possible ZI strategies is infinite, as a better approximation to the situation in real financial markets with high degrees of automated trading.

Prior researchers’ concentration on markets in which the traders can choose one of only two or three fixed strategies can be traced back to the sequence of publications that launched the trading strategies MGD, GDX, and AA (i.e., Tesouro and Das (2001); Tesouro and Bredin (2002); Vytelingum et al. (2008)), and the papers in which these strategies were shown to outperform human traders (i.e. Das et al. (2001); De Luca and Cliff (2011a,b)). All of these works relied on comparing the strategy of interest with a small number of other strategies in a series of carefully devised experiments: e.g., GDX was introduced in Tesouro and Bredin (2002), and was compared only to ZIP and GD.

In aiming for a fair and informative comparison, experimenters were immediately faced with issues in *design of experiments* (see e.g. Montgomery (2019)): how best to compare strategy S_1 with strategies S_2 and S_3 (and S_4 and S_5 and so on), given the finite time and compute-power available for simulation studies, and the need to control for the inherent noise in the simulated market systems.

Early comparative studies such as Das et al. (2001) limited themselves to running experiments that studied the performance of a selection of trading strategies in three fixed experiment designs: *homogeneous* (in which the market is populated entirely by traders of a single strategy-type); *one-in-many* (OIM: in which a homogeneous market was altered so that all the traders were of strategy type S_1 except one, which was of type S_2); and *balanced-group* (BG: in which there was a 50:50 split of S_1 and S_2 , balanced across buyers and sellers, with allocation of limit-prices set in such a way that for each trader of type S_1 with a limit price of λ_1 there would be a corresponding trader of type S_2 also assigned a limit price of λ_1). There were good reasons for this experiment design, and the results were informative, but they rested on only ever comparing two strategies S_1 and S_2 in markets with a total number of traders N_T where the ratio of $S_1:S_2$ was one of either $N_T:0$ (i.e., homogeneous); or $(N_T - 1):1$ (i.e., OIM); or $\frac{N_T}{2}:\frac{N_T}{2}$ (i.e., BG). This approach left open the question of whether the performance witnessed in one of these three special cases generalised to other possible ratios, other relative proportions of the two strategies in the market.

A method by which that open question could be resolved was developed by Walsh et al. (2002) who borrowed the technique of *replicator dynamics analysis* (RDA) from evolutionary game theory (see e.g. Maynard Smith (1982)). In a typical RDA, the population of traders is initiated with some particular ratio of the N_S strategies being compared, and the traders are allowed to interact in the market as per usual, but every now and again an individual trader will be selected via some stochastic process and will be allowed to

mutate its current strategy S_i to one of the other available strategies $S_{j \neq i}$ if that new strategy appears to be more profitable than S_i . In this way, given enough time, the market system can be started with any possible ratio of the N_S strategies, and in principle it can evolve from that starting point through other system state-vectors (i.e., other ratios of the N_S strategies) to any other possible ratio of those strategies. However in practice the nature of the evolutionary trajectories of the system, i.e. the paths traced by the time-series of state-vectors of the system, will be determined by the profitability of the various strategies that are in play: some points in the state-space (i.e., some particular ratios of N_S strategies) will be unprofitable *repellers*, with the evolutionary system evolving away from them; others will be profitable *attractors*, with the system converging towards them; and if the system converges to a stable attractor then it’s at an *equilibrium point*, or potentially on a repeating sequence of equilibrium points, i.e. a *limit cycle*. Walsh et al’s 2002 paper showed the results of RDA for market systems in which $N_S = 3$, comparing the trading strategies GD, SNPR, and ZIP, and visualised the evolutionary dynamics as plots of the two-dimensional *unit simplex*, an equilateral triangular plane with a three-variable barycentric coordinate frame.

Similar plots of the evolutionary dynamics on the 2D unit simplex were subsequently used by other authors when comparing trading strategies: see e.g. Vytelingum et al. (2008); Vach (2015), and those authors also limited themselves to studies in which the traders in the market could switch between one of only $N_S = 3$ different discrete strategies. And, in this strand of research, three-way comparisons seem to then have become the method of choice primarily because evolutionary trajectories through state-space, and the location and nature of any attractors and repellers on the space, is readily renderable as a 2D simplex when dealing with a $N_S = 3$ system, but rapidly gets very difficult, to the point of impracticability, as soon as $N_S > 3$. Higher-dimensional simplices are mathematically well-defined, but very difficult to visualise: the four-variable simplex is a 3D volume, a tetrahedron; and more generally the N_S -variable simplex is an $N_S - 1$ -dimensional volume – so if we wanted to study the evolutionary dynamics of a six-strategy system, we would need to find a way of usefully rendering projections of the 5-D simplex, or we need to find alternative methods of visualisation and analysis.

However, as first shown by Vach (2015) and later confirmed in more detailed studies by Snashall and Cliff (2019); Rollins and Cliff (2020) and Cliff and Rollins (2020), when the complete state-space of all possible ratios of discrete strategies is exhaustively explored, the dominance hierarchies indicated by the simple OIM/BG analyses are sometimes overturned. That is, if strategy S_1 outperformed strategy S_2 in both the OIM and the BG tests, that would usually be taken as evidence that S_1 generally outperformed S_2 , that S_1 was “dominant” in that sense; but actually if markets were set up with some ratio of $S_1:S_2$ other than the OIM or BG ratios, then in those markets

S_2 would dominate S_1 – that is, the direction of the dominance relationship between S_1 and S_2 can often depend on the ratio of $S_1:S_2$, their relative proportions of the overall population. Furthermore, while S_1 might dominate S_2 in two-strategy experiments (i.e., where $N_S = 2$), plausibly S_2 would dominate S_1 in experiments where values of $N_S > 2$: the indications are that as yet there is no single master-strategy that dominates all others in all situations; what strategy is best will depend on the specific circumstances.

By populating a model market entirely with adaptive PRZI traders we create a minimal test-bed for exploring issues of market efficiency and stability in situations where all traders are simultaneously co-evolving in an infinite continuous space of strategies. The state at time t of such a market with N_T traders in it can be characterised as an N_T -dimensional vector of s -values, denoted by $\vec{S}(t)$, identifying a single point in the N_T -dimensional hypercube that is the space of all possible system states, and that point will move over time as the traders each adapt their s values. We can attempt to identify attractors and repellers in this hypercube, but we will need new visualisation techniques: we'll need to leave simplices behind.

There are many ways in which a PRZI trader could be made to dynamically adapt its s -value in response to market conditions. Here, in the spirit of minimalism associated with studies of ZI traders, I use a crude and simple stochastic hill-climbing algorithm, of the sort that might be found as an introductory illustrative straw-man sketch in the opening chapter of a book on machine learning. To keep with the tradition of naming ZI/MI trading algorithms with short acronyms, I'll refer to this **PRZI Stochastic Hill-Climber** as PRSH (pronounced “pursh”). PRSH is defined in Section 3, and then some illustrative baseline results from experiments in which a *single* PRSH trader adapting in markets where all other traders are playing fixed strategies are presented in Section 4. After that, Section 5 shows results from experiments in which *all* traders are PRSH, and hence in which the market is maximally co-evolutionary. The Python source-code for PRSH has been released as free-to-use open-source, in BSE (see Cliff (2012)) to enable other researchers to replicate and extend the preliminary results shown here.

3. PRSH: a minimal PRZI Stochastic Hill-Climber

At any time t , a PRSH trader i has a set S_{i,t_m} that was created at time $t_m \leq t$ and that consists of $k \in \mathbb{Z}^+$ different PRZI strategy values s_{0,t_m} to s_{k-1,t_m} (i.e., $|S| = k > 1$). Although t is continuous in this model, alterations to $S_{i,t}$ happen only occasionally. After an initialisation step in which the k strategies are each assigned a value $s_{i,t_0} \in [-1, +1] \in \mathbb{R}$ via a *genesis* function $\mathcal{G}(\cdot)$, PRSH enters into an infinite loop: let t_m denote the time at which a new iteration of the loop is initiated; in each cycle of the loop a PRSH trader first *evaluates* each of its k strategies in turn, trading with each of them as the sole exclusive strategy for at least a minimum period of time Δ_t , such that all k have been eval-

uated by time $t_n \geq t_m + k\Delta_t$; after that, it *rank*s the strategies by some performance or *fitness* metric \mathcal{F} , and copies the top-ranked strategy (the *elite*) at time t_n into s_{0,t_n} ; it then creates $k - 1$ new ‘mutants’ of s_{0,t_n} , via a stochastic *mutation* function $\mathcal{M}(s_{0,t_n})$, and this set of new strategies $s_{j,t_n:1 \leq j \leq k-1}$ then replaces the old S_{i,t_m} , becoming S_{i,t_n} , at which point it loops back for the next iteration (and hence in that next iteration the value t_m is what was t_n in the prior iteration).

This definition leaves the experimenter free to decide certain key details when implementing PRSH:

- The choice of k and of Δ_t together determine the speed of adaptation: PRSH will generate a new $S_{i,t}$ at most once every $k\Delta_t$ seconds: i.e., $k\Delta_t$ is the minimum time-period between successive mutations, where each mutation is an *adaptive step* on the underlying *fitness landscape*. If you want a PRSH to make N_{steps} adaptive steps on the fitness landscape in the course of an experiment, that experiment needs to run for $> k\Delta_t N_{\text{steps}}$ seconds.
- Exactly how the set S_0 is created at initialisation is left open. Naturally $s_{i,0} = \mathcal{U}(-1, +1) \in \mathbb{R}; i \in \{0, \dots, k-1\}$ is the least constrained, but there may be circumstances where it is informative to use some other method, e.g. $s_{i,0} = c; \forall i$ for some constant c such as zero or ± 1 .
- The stochastic function $\mathcal{M} : [-1, +1] \in \mathbb{R} \mapsto [-1, +1] \in \mathbb{R}$ that creates new mutants of the elite s_{0,t_k} is similarly unspecified. Treating each mutation as the addition of a random draw from a distribution with zero mean and nonzero variance makes intuitive sense, and then either truncating or using ring-arithmetic to ensure that the function maps to $[-1, +1]$. In the experiments shown below, $\mathcal{M}(s_{0,t_k}) = s_{0,t_k} + \mathcal{N}(0, \sigma)$ with $\sigma = 0.01$. Plausibly a simulated-annealing approach could be introduced, steadily reducing σ as time progresses, but that is not explored here.
- For $k > 2$, questions immediately arise over what is the best way of generating the k mutants. For instance if $k = 3$ we could arrange a set of two different \mathcal{M} functions, one per mutant, such that $s_{1,t_k} < s_{0,t_k}$ and $s_{2,t_k} > s_{0,t_k}$ and hence PRSH is always sampling s -values at random magnitudes either side of the current elite strategy; and for $k = 5$ we could similarly arrange the mutants such that two are generated either side of the elite, one a small random distance away, and the other a much larger random distance away; such decisions are left as an implementation issue. In the work reported here we simply generate $k - 1$ mutants via \mathcal{M} with no additional constraints.
- Finally, each iteration of the loop requires deciding which of the k strategies is the current elite, via the fitness function \mathcal{F} , and there are many possible ways to do that. The method used here was to rank the k strategies at time t_k by the amount of profit generated per unit of time, denoted by PPS (profit per second), such that the elite s_{0,t_k} strategy has the highest PPS. To help avoid the hill-climber from becoming trapped on local maxima, if the difference between the PPS scores of

the two highest-ranked s -values in S is less than some threshold ϵ_s , then one of the two is chosen at random to be the elite for that iteration of the loop.

In essence, PRSH with k strategies is a very primitive k -armed bandit, and all of the extensive multi-armed bandit (MAB) literature (such as Gittins et al. (2011); Myles White (2012); Lattimore and Szepesvari (2020)) is potentially of relevance here, but ignored: again, the intention here is not to create the best adaptive-PRZI trader, instead it is merely to have a simple minimal adaptive-PRZI algorithm to act as a proof of concept and to enable an initial set of exploratory and illustrative experiments involving populations of adaptive-PRZI traders: PRSH does that job.

4. Evolution of Strategy in a Single PRSH Trader

Before studying co-evolving populations of PRSH traders, it is informative to explore situations in which there is only a single PRSH trader in the market, and all other traders are one or more of the three ZI strategies that are spanned by PRSH/PRZI, i.e. GVWY, SHVR, and ZIC. In such situations we can talk of how the PRSH trader's strategy evolves over time, but not of co-evolution because the rest of the traders in the market are non-adaptive. A single-PRSH-trader market is sufficiently simple that it eases the introduction of concepts that become significantly more complex in fully co-evolutionary markets.

First, we can visualise the fitness landscape for a single PRSH trader by setting up a market in which, purely for the sake of generating appropriate visualization data, we give the PRSH a large k , and initialize S_0 to a set of regularly-spaced $s_{i,0}$ values across the range $[-1, +1]$, and then plot the PPS fitness of each strategy in the first evaluation. Specifically, set: $S_0 = \{s_{i,0} : s_{i,0} = \frac{2i}{k-1} - 1; i \in \{0, \dots, k-1\}\}$ And let $\Delta_s = 2/(k-1)$, the step-size in our mapping of the fitness landscape. So for example with $k = 21$ we have $\Delta_s = 0.1$ and $S_0 = \{-1, -0.9, -0.8, \dots, +0.9, +1.0\}$.

For brevity, and without loss of generality, the discussion that follows in the rest of this section concentrates only on the case of a single PRSH seller in a market that is otherwise entirely populated by traders running nonadaptive strategies. The arguments that are made here for a single PRSH seller could just as easily be made for a single PRSH buyer, but to do both here would be overkill.

Figure 1 shows fitness landscapes plotted at $\Delta_s = 0.05$ for a single PRSH seller when all other traders in the market are either (from top to bottom) SHVR, ZIC, or GVWY: i.e., a progression from all other traders in the market being maximally relaxed (SHVR) through to maximally urgent (GVWY). In all experiments reported in this paper, all buyers had the same limit price λ_b and all sellers had the same limit price $\lambda_s > \lambda_b$, i.e. the supply and demand schedules were 'box' style, with perfect elasticity of supply and of demand. When generating the landscapes for SHVR and ZIC the number of buyers (N_{Buy}) and the number of sellers (N_{Sell}) were each 30, i.e. $N_T = 60$, but in the

landscape for GVWY results from $N_T = 60$ are overlaid with additional results from IID repetitions of the same experiment where $N_T = 30$ and where $N_T = 120$ (in each case $N_{Buy} = N_{Sell} = N_T/2$), to demonstrate that the overall shape of the fitness landscape varies very little with respect to the $N_T = 60$ case when the number of traders is halved or doubled. As can be seen from Figure 1, in the single-PRSH case the fittest (most profitable) strategies – i.e., the global maxima – are all at the high end of the range, at or close to $s = +1$, but in each landscape there is also a local maxima at/near $s = -1$.

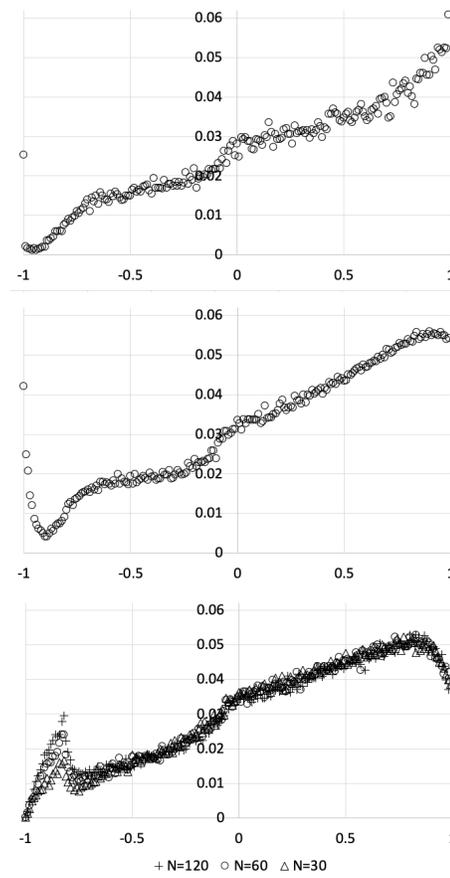


Figure 1. Fitness landscapes for a single PRSH seller in a market where all other traders are homogeneously playing the same fixed strategy: horizontal axis is PRSH strategy value s ; vertical axis is profit per second (PPS) recorded by the single PRSH trader using s as its strategy. Strategy evaluation time Δ_t is 7200s. Data points are plotted at strategy-steps of $\Delta_s = 0.01$. Upper graph is when all other traders are playing the fixed SHVR strategy; middle graph is when all other traders are ZIC; lower graph is when all other traders are GVWY. In the lower graph only, data is shown for IID repetitions of the experiment with the number of traders in the market (denoted by N_T) being set to 30 (data-points marked by open triangles), 60 (marked by open circles), and 120 (marked by plus-symbols).

The GVWY fitness landscape for a single PRSH seller shown at the bottom of Figure 1 clearly has a global maximum at $s \approx 0.8$. If the PRSH adaptation mechanism is operating as intended, when the single PRSH seller is ini-

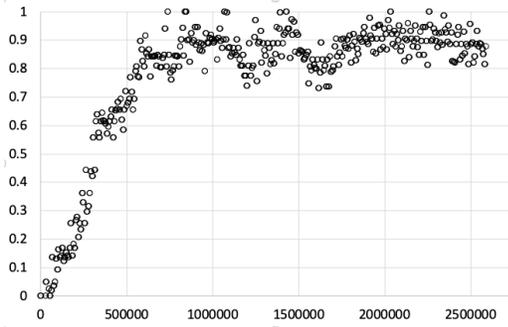


Figure 2. Hourly strategy value over 30 days of round-the-clock trading for a single PRSH seller in a market populated with 29 GVWY sellers and 30 GVWY buyers: horizontal axis is time in seconds; vertical axis is the PRSH trader’s strategy value s , which is initialized at the start of the experiment to $s = 0$, i.e. to the ZIC strategy. The s -value evolves steadily toward a range of values close to the global optimum strategy identified in the bottom fitness-landscape plot of Figure 1, and then stabilises to that range of values for the remainder of the experiment.

tialised with $s = 0$ and allowed to adapt for sufficiently long then its s value should converge to roughly 0.8, and then hold at that value. To demonstrate this, Figure 2 shows the PRSH trader’s s value, plotted once per hour, in a simulation of 30 continuous days of 24-hour trading: as can be seen, from its initial value of zero there is a steady rise in s over the first $\approx 750,000$ sec of trading (i.e., roughly the first 8.5 days), after which the system stabilises to s -values that noisily fluctuate around the 0.85 level. To smooth out some of the noise, define \hat{s} as the 12-hour simple moving average of the raw hourly s data: Figure 3 shows the \hat{s} line for the raw hourly data shown in Figure 2, along with \hat{s} lines from a further four IID repetitions of the same experiment. For the discussion that follows, let’s call trader i ’s \hat{s}_i value at the end of an experiment the *terminal strategy* for i in that experiment, and define the set \widehat{S}_T as the set of terminal strategies from a population of PRSH traders that have co-evolved in a particular market environment. For the current discussion of the merely evolutionary (i.e., not co-evolutionary) adaptation of single PRSH traders, we can fill \widehat{S}_T with the set of terminal strategy values arising from N_R IID repetitions of a particular experiment: in Figure 3, we have $N_R = 5$ and $\widehat{S}_T = \{0.86, 0.87, 0.88, 0.88, 0.93\}$. As N_R takes on larger values, it is natural to summarise values in the terminal strategy set \widehat{S}_T as a frequency histogram or kernel density estimate, and from there to note whether the distribution of values in the terminal strategy set is unimodal or multimodal, either by eyeballing the distribution or density estimate, or by applying a test of modality such as those proposed by Hartigan and Hartigan (1985) or Chasani and Likas (2022).

5. Strategy Co-Evolution in All-PRSH Markets

As a first illustration of the dynamics of a fully co-evolutionary ZI market system, Figure 4 shows the \hat{s}_i val-

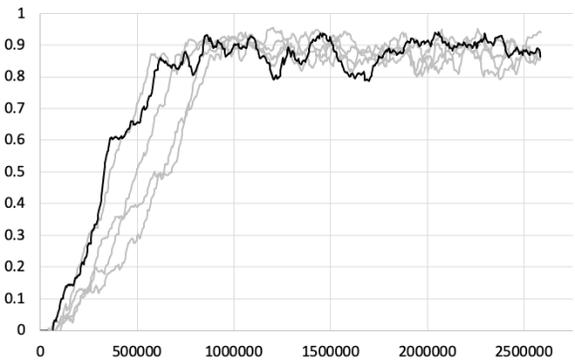


Figure 3. Smoothed PRSH strategy values from multiple 30-day experiments, each with a single PRSH seller in a market populated by 29 GVWY sellers and 30 GVWY buyers: horizontal axis is time in seconds; vertical axis is 12-hour moving-average strategy value (denoted by \hat{s}). Black line is the \hat{s} trace for the raw hourly s -data shown in Figure 2; the four grey lines are each the \hat{s} traces from four IID repetitions of the same experiment. After 100,000 seconds (roughly 11 days) of trading, all five \hat{s} traces have evolved to a steady state close to the global optimum strategy identified in the bottom fitness-landscape plot of Figure 1, and remain clustered around that value for the remainder of the experiment. The set of final \hat{s} values recorded at the end of each experiment is referred to as the *terminal strategy set*, denoted by \widehat{S}_T . Here, $\widehat{S}_T = \{0.86, 0.87, 0.88, 0.88, 0.93\}$: see text for further discussion.

ues over time for a 30-day experiment in which the market is populated by 30 PRSH sellers and 30 PRSH buyers, all of which are initialized to have $s_{i,0} = 0$: i.e. an experiment directly comparable to the results from the zero-initialized single-PRSH system explored in the previous section, except that here the fitness landscape for any one trader will depend on the distribution of strategy-values for all the other traders in the market, and in which the fitness landscape will be varying over time, in principle altering each time any one PRSH trader changes its strategy to a new value. Again, a \widehat{S}_T terminal strategy set can be assembled from the final \hat{s}_i values of the individual traders that co-evolved against each other in the single market experiment: the corresponding terminal strategy set distribution is again unimodal: in this experiment, all sellers converge on strategy-values in $[\approx +0.55, \approx +0.85]$; multiple IID repetitions of this market experiment generate much the same results.

Further investigation reveals that the unimodal distribution of terminal strategies in experiments like the one illustrated in Figure 4 is an artefact of the decision to initialize all traders with $s_{i,0} = 0$: if instead we set $s_{i,0} = \mathcal{U}(-1.0, +1.0)$ so that the initial set of strategy values in the population of traders is uniformly distributed over the entire range of possible strategies, we see qualitatively different results: for both the buyers and the sellers the distribution of terminal strategy values is then multimodal.

The development of multimodal terminal strategy distributions is not the only change resulting from switching the initial state from $s_{vi} = 0.0$ to $s_{vi} = \mathcal{U}(-1.0, +1.0)$. In Figure 4, over the 30 simulated days, the dynamics of the system’s co-evolution through strategy space are bipha-

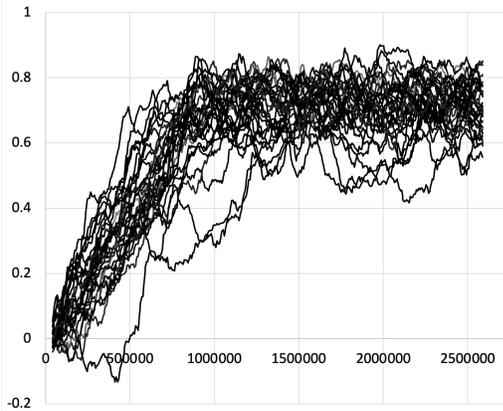


Figure 4. Smoothed ($\widehat{s}_{i,t}$) strategy values for each of 30 PRSH sellers in a market experiment lasting for 30 days of continuous trading, where all traders are initialized to have $s_{i,0} = 0$. Horizontal axis is time in seconds; vertical axis is the 12-hour moving average strategy $\widehat{s}_{i,t}$ of individual traders. The co-evolutionary dynamic is biphasic: in the initial “adaptive transient” phase over the ≈ 12 days (i.e., $\approx 1,000,000$ seconds) the system settles to a unimodal steady-state centered on $s_i \approx 0.7$; in the steady-state phase the strategy values of individual traders rise and fall but the overall distribution does not vary significantly.

sic: an initial phase of roughly 12 days in which all traders increased their s values from zero to ≈ 0.7 ; followed by a steady-state phase lasting for the remainder of the experiment where the population of s values wandered randomly around the 0.7 level. In contrast, when $s_{vi} = \mathcal{U}(-1.0, +1.0)$ the system shows no such long-term stability over the same time-period, as is illustrated in Figure 5 and explained in the caption to that figure: even after the system’s distribution of strategies has been relatively stable for a period of nine days, an equilibrium or stasis in which the traders have each executed roughly 150,000 transactions, chance co-evolutionary interactions can result in the stasis ending and the system entering a fresh period in which the strategies are in flux.

To illustrate the longer-term dynamics of this system, Fig. 6 shows buyer-strategy co-evolutionary time series similar to that illustrated in Fig. 5 from eight IID repetitions of an experiment that lasted 10 times longer, i.e. 300 simulated days. As is clear from the figure, although stable modes do occur in each experiment, individual trader’s strategy-values will sometimes transition from one mode to another, with no clear pattern or predictability to the timing and/or direction of these transitions.

Thus far, to save space, only the co-evolutionary trajectories of the strategies in the population of buyers have been shown. Naturally, each of the eight buyer-strategy time-series graphs shown in Figure 6 has a corresponding seller-strategy time-series graph, but in this specific set of experiments there was much less variation in the outcomes for the seller population: rather than showing all eight, Figure 7 shows one representative example; qualitatively, the other seven are all essentially identical to this.

The co-evolutionary dynamics of strategy values in these model markets is not the only factor of interest: an-

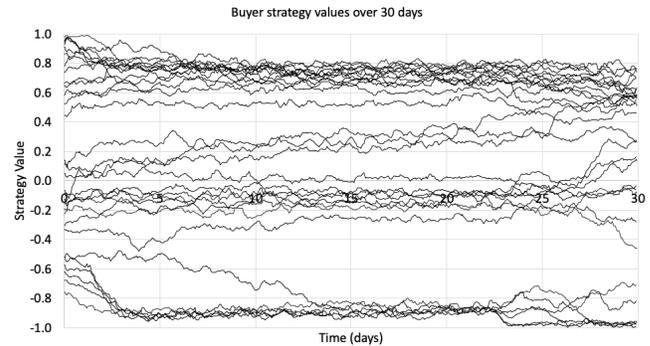


Figure 5. Smoothed ($\widehat{s}_{i,t}$) strategy values for each of 30 PRSH buyers in a market experiment lasting for 30 days of continuous trading, simulated at 60Hz time-resolution, where all traders are initialized to have $s_{i,0} = \mathcal{U}(-1.0, +1.0)$. Horizontal axis is time t , with a vertical gridline every 5 days; vertical axis is the 12-hour moving average strategy $\widehat{s}_{i,t}$ of individual traders, with horizontal gridlines at s intervals of 0.2: for $t \geq 0.5$ days (i.e., 12 hours) the trader’s average strategy value over the preceding 12 hours is plotted; for $t < 0.5$ days the trader’s average strategy since the start of the experiment is plotted. By roughly Day 13 the system has settled into a state that then persists as a temporary equilibrium or stasis until roughly Day 22: during the equilibrium phase the modes are at roughly $s = -0.9$ ($n = 6$), $s = -0.1$ ($n = 8$), $s = +0.3$ ($n = 3$), and $s = +0.7$ ($n = 13$). After that, the equilibrium “punctuates”, entering a new phase where first the mode at -0.9 loses its stability, then the mode at $+0.3$ seems to merge up into the mode that was at $+0.7$ but which now seems to be generally heading lower, and then the mode at -0.1 seems to dissipate in various directions. In the nine-days stasis/equilibrium, each trader would execute approximately 150,000 transactions. Clearly the dynamics have not reached a stable state after 30 days of trading, and longer simulations should be explored.

other equally significant concern is the efficiency of the markets populated by traders with co-evolving strategies: something that is illustrated in Figure 8 which shows, for each of the eight 300-day experiments illustrated in Figure 6, the total surplus/profit extracted by the traders. Data-lines show collective total profit extracted by the 30 buyers (denoted here as π_B), collective total profit extracted by the 30 sellers (denoted here as π_S), and total profit extracted by the entire set of 60 traders (denoted here as $\pi_T = \pi_B + \pi_S$). In each case, after the initial adaptive transient over the first 50 days or less, the buyers’ and seller’s profit levels stabilise to an approximately constant-sum relationship, where if π_B goes up then π_S goes down, and *vice versa*. The sum π_T that the two populations’ profit-levels add up to is notably unvarying within any one experiment, but the value that π_T settles on varies across experiments: for example, the experiments at upper-left and lower-left both have $\pi_T \approx 93 - 95$, whereas the upper-right and the left-hand experiment in the third row from the top both never see π_T go above 90. The underlying reason for this variation in total profit extracted is illuminated in Figure 9, which shows the inverse relationship between the number of traders with ‘relaxed’ strategy values ($s_i < 0$) and the total profit extracted: the more relaxed traders present in the market, the less profit extracted; despite their constant striving to improve profitability, traders with strategy values in the relaxed mode seem to be stuck on a local maximum in the fitness landscape.

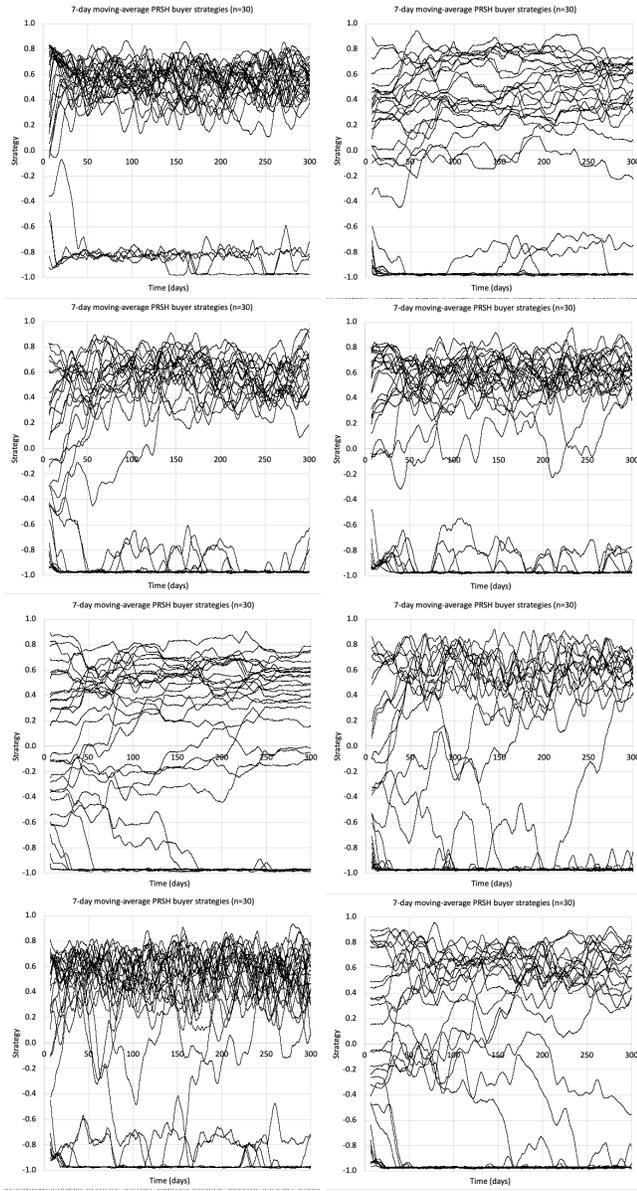


Figure 6. Results from eight IID experiments each otherwise the same as that illustrated in Figure 5 but instead continued for 300 days. Data lines show smoothed $(\bar{s}_{i,t})$ strategy values for each of 30 PRSH buyers in a market experiment over 300 days of continuous trading, simulated at 60Hz time-resolution, where all traders are initialized to have $s_{i,0} = \mathcal{U}(-1.0, +1.0)$. Horizontal axis is time t , with a vertical gridline every 50 days; vertical axis is the 7-day moving average strategy $\bar{s}_{i,t}$ of individual traders, with horizontal gridlines at s intervals of 0.2. The upper four graphs appear to show that, after an initial adaptive transient phase, the population of traders settles into a steady-state bimodal distribution; but the lower four graphs show that the system does not always quickly converge to such a steady-state distribution and that co-evolutionary interactions can result in major changes in the strategy distributions (e.g., a trader switching from one mode to another) even after 200 or more days of continuous trading, a period over which each trader would execute roughly 3,500,000 transactions. See text for further discussion.

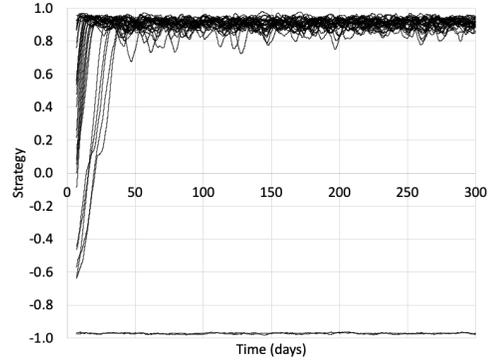


Figure 7. Time-series of co-evolving seller strategies from one of the eight experiments for which the buyer strategies were illustrated in Figure 6: qualitatively, all eight experiments have time series essentially the same as this one, so only the one is illustrated here. The vast majority of sellers rapidly shift their strategy-values to around +0.9, but in any one experiment a small number of sellers instead settle on strategy values close to -1.0. In all cases, these two modes are stable for the remainder of the duration of the experiment.

Although the time-series of co-evolving strategy values and histograms of strategy frequency distributions have served the purposes of this discussion thus far, there is a need for more sophisticated visualization and analysis techniques. Our very first studies of co-evolutionary dynamics with a preliminary $k = 2$ PRSH-like system, reported in Alexandrov et al. (2022) (which summarises results from Alexandrov (2021) and Figueroa (2021)) explored the prospects of producing phase portraits, graphical characterisations of the global dynamics of the system, for market sessions in which there are only two evolving traders, each adjusting their s -values with the intent of improving their profitability, while all other traders play fixed strategies: in such a two-PRSH market the phase-space of interest is two-dimensional, just the two evolving strategies, and hence very easy to plot as a 2D graphic. But for the all-PRSH $N_T = 60$ market sessions studied here, we need a useful way of plotting the trajectory of the dynamical system through its 60-dimensional real-valued phase-space: that is, the strategy vector $\vec{S}(t) \in [-1.0, +1.0]^{N_T} \in \mathbb{R}^{N_T}$.

Thankfully, in recent decades researchers in physics have developed a set of visualisation and analysis tools and techniques for such high-dimensional real-valued dynamical systems: the dynamics of such systems can be characterised visually, as a square array of pixels, via the creation of a *recurrence plot* (RP), which will often display macro-scale features that are obvious to the human eye; and then straightforward image-processing techniques can be used to generate quantitative statistics that summarise the nature of the RP and the features within it, an approach known as *Recurrence Quantification Analysis* (RQA). For readers unfamiliar with RPs and RQA, Appendix A presents a brief introduction.

Figure 10 shows an RP for a single $N_T = 60$ all-PRSH market session lasting for 7 days of continuous round-

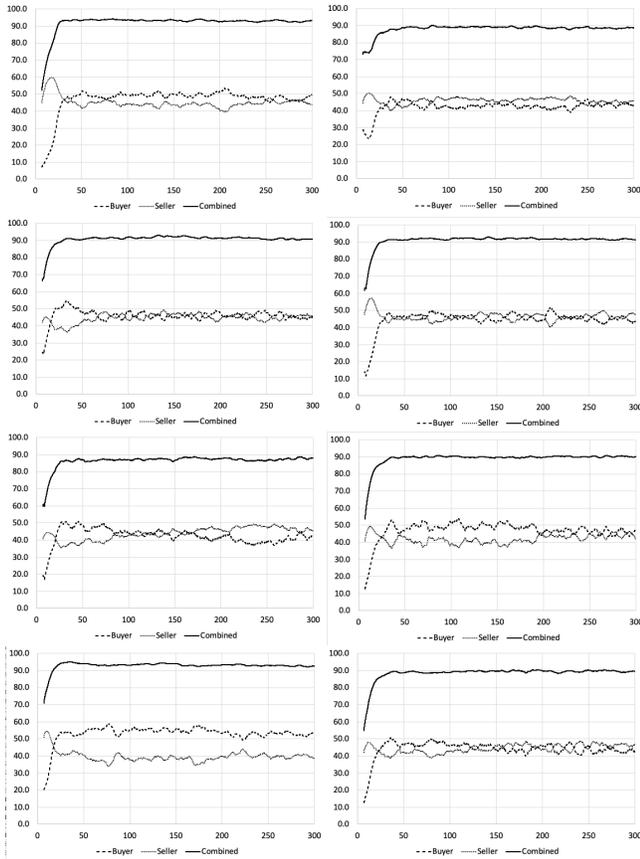


Figure 8. Total extraction of surplus/profit for the eight 300-day experiments illustrated in Figure 6: horizontal axis is time in days; vertical axis is total profit extracted by a group of traders. Data-lines show collective total profit extracted by the 30 buyers, collective total profit extracted by the 30 sellers, and total profit extracted by the entire set of 60 traders. In each case, after the initial adaptive transient over the first 50 days or less, the buyers' and seller's profit levels stabilise to an approximately constant-sum relationship, where if buyers' profits go up then sellers' profits go down, and vice versa. The sum that the two populations' profit-levels adds up to is notably unvarying within any one experiment, but varies across experiments: for example, the experiments at upper-left and lower-left both have the sum consistently around 93-95, whereas the upper-right and the left-hand experiment in the third row from the top both never see their sum go above 90. See text for further discussion.

the-clock trading, with the strategy-vector $\vec{S}(t)$ recorded hourly, resulting in a 168×168 -pixel plot (i.e., $7 \times 24 = 168$) where the time-difference between rows and columns is one hour. In all the RPs plotted here, $\vec{S}(t)$ is considered a recurrence of the state $\vec{S}(t - \Delta_t)$ when $|\vec{S}(t) - \vec{S}(t - \Delta_t)| < \epsilon$, using $\epsilon = \sqrt{60 \times 0.05^2} = 0.387$: the maximum distance possible in this system (i.e., the *diameter of the phase-space* in the terminology of the physics literature) is $\sqrt{60 \times 2^2} = 15.492$ (e.g., if $[\vec{S}(t)]_i = +1.0; \forall i$ and $[\vec{S}(t - \Delta_t)]_i = -1.0; \forall i$), so the value of ϵ used here is $\approx 2.5\%$ of the maximum distance.

As is clear from visual inspection of the RP in Fig. 10, there are almost always recurrences to the left and below the diagonal line of identity (LOI) and these recurrences are typically short-lasting, being roughly 10 pixels or less (i.e., 10 hours or less) in the first 50 hours of the session,

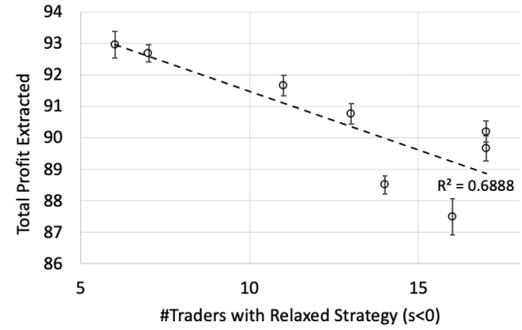


Figure 9. Inverse relationship between the percentage of traders in the market playing relaxed strategies (i.e., $s_i < 0$) and total profit extracted by all the traders in the market: horizontal axis is percentage of traders with $s_i < 0$; vertical axis is profit extracted over the final 50 days' trading (i.e., days 250-300) in the experiment. Markers show the arithmetic mean over that period, with error bars at \pm one standard deviation, for the eight experiments illustrated in Figure 6. The dashed line shows linear regression; $R^2 \approx 70\%$.

and then lengthening as the session continues, such that by the end of the session the recurrences are recorded as far-distant as roughly 48 hours previously. A commonly-used RQA summary statistic for this kind of observation is the *trapping time* (denoted by TT : see Appendix A for the definition): for the RP in Fig. 10, the overall $TT \approx 7.25$ hours: i.e., the system typically spends 7.25 hours within ϵ distance of any particular $\vec{S}(t)$, before co-evolution drives it away from that area of phase-space; and, given the large areas of unshaded area in the RP, we can see that once it co-evolves away from a particular state after a few hours, it never returns to that state (i.e., no further recurrences are recorded), indicating *acyclic* evolution – i.e., continuous “progress” of the co-evolutionary dynamic.

Figure 11 shows a set of six RPs, from six IID market sessions with all parameters set to the same values as used in the experiments illustrated in Figures 6 and 8, except these six experiments have each been left to run for 1,500 days. As before, $\vec{S}(t)$ data is recorded hourly, and the traders interact second-by-second simulated at 60Hz, trading around the clock, 24hrs/day; and hence these RPs in their full incarnation are 36000×36000 pixels (i.e., $1500 \times 24 = 36000$), which of necessity are then down-sampled for printed reproduction here. As is discussed in the caption to Figure 11, five of the six sessions show clear evidence of the co-evolutionary process being *cyclic*, in the sense that the system is continuously co-evolving, taking a very large sequence of adaptive steps in the 60-dimensional strategy-space, but eventually it returns to points in strategy space that it previously occupied at an earlier time in the session. And, surprisingly, the path-length of these cyclic transits can be extremely long: more than 1,000 days in one instance. And remember that each trading day in the session is simulated at 24hrs/day, at 60 frames per second resolution (i.e., the simulation timestep is 0.0167s), so the 1,000-day cycle occurred after 5.18Bn timesteps, during which more than a billion transactions

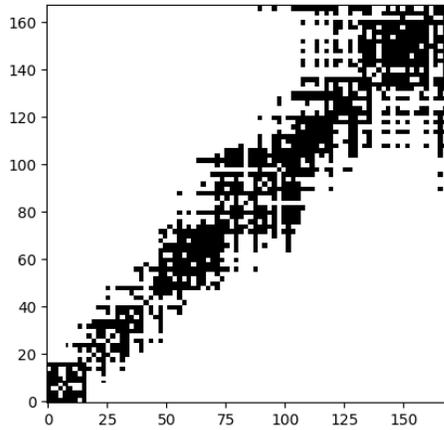


Figure 10. Example recurrence plot (RP) for a PRSH co-evolutionary market session: in this experiment (as with the experiments illustrated in Figures 6 and 8) there are 30 PRSH buyers and 30 PRSH sellers (i.e., $N_T = 60$) each co-evolving their individual strategy s -values, so the collective state of the system of co-evolving strategy values at time t is a strategy-vector $\vec{S}(t) \in [-1.0, +1.0]^{60} \in \mathbb{R}^{60}$. The traders interact continuously, simulated at 60Hz, trading around the clock 24 hours per day, but the \vec{S} strategy-vector is recorded only once every hour. This RP shows the first 7 days of the market session (i.e., $7 \times 24 = 168$ hours): numeric labels on the axes are hour-number. The state $\vec{S}(t)$ is considered a recurrence of the state $\vec{S}(t - \Delta_t)$ when $|\vec{S}(t) - \vec{S}(t - \Delta_t)| < \epsilon$, using $\epsilon = \sqrt{60 \times 0.05^2} = 0.387$ (here the diameter of the phase space, is $\sqrt{60 \times 2^2} = 15.492$, so the value of ϵ used here is $\approx 2.5\%$ of that diameter). See text for further discussion.

will probably have taken place. Simulations run for shorter durations would not have revealed these long-term cycles.

6. Discussion and Conclusion

The results presented here are the first from market simulations populated wholly by co-evolving parameterised-response zero-intelligence (PRZI) traders which adapt their strategies over time using stochastic hill-climbing (i.e., PRSH), and they demonstrate that such minimally simple models can exhibit surprisingly rich dynamics, over extremely long timescales, and also that the stable attractors in strategy-space are often neither at the extreme points of the range (i.e., $s_i = \pm 1.0$) nor at the mid-point ($s_i = 0.0$) but instead are at ‘hybrid’ points along the strategy-space, resulting in trading behaviors (quote-price distributions) with no precedents in the prior literature. There are a wide range of factors that could be explored in further work. For example: the particular form of adaptation used here, the simple stochastic hill-climber of PRSH, is likely to affect the co-evolutionary dynamics; i.e., it might be more likely to result in traders being stuck on local maxima in the fitness landscape, in comparison to other more sophisticated adaptation/optimisation techniques: in Cliff (2022), a follow-on to this paper, I discuss the use of *differential evolution* (see e.g., Storn and Price (1997); Bilal et al. (2020)) instead of stochastic hill-climbing, which better avoids local maxima. Also the nature of the supply and demand curves in the market can

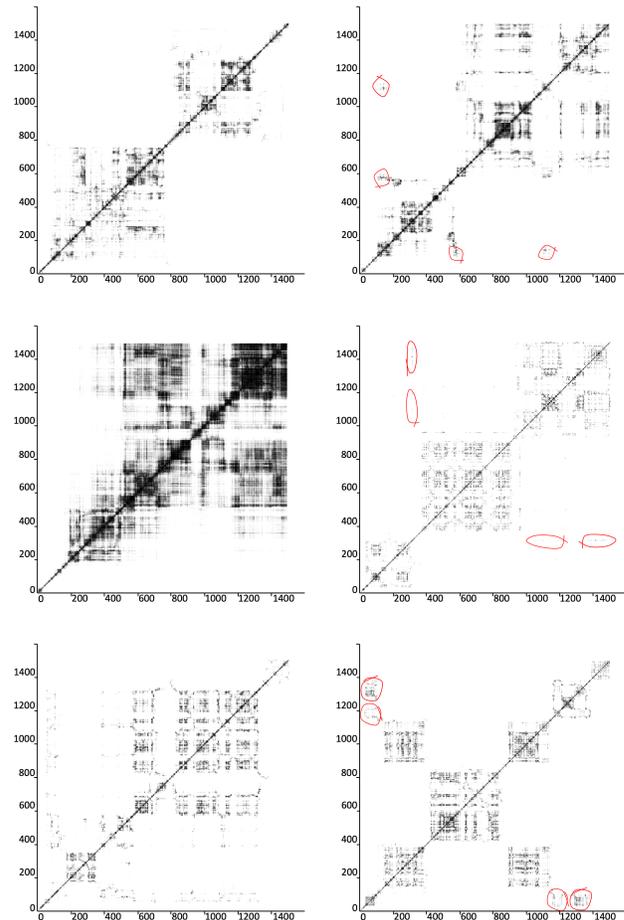


Figure 11. Recurrence Plots (RPs) for six IID market sessions, each running for 1,500 simulated days of continuous (24hr/day) trading, each simulated at 60Hz, and each involving multiple transactions per second, i.e. involving on the order of one billion transactions per 1,500-day session. Simulating each market session took approximately 280 hours of wall-clock continuous CPU time on a 16GB Apple Mac Mini (M1 Silicon, 2020), with data frames recorded once per simulated hour, yielding complete RPs that are $36,000 \times 36,000$ pixels. For each RP, the numeric labels on both axes shows the number of days elapsed. The RP at upper-left shows the population of traders drifting in one region of strategy space over days ≈ 100 to ≈ 200 , then another region over days ≈ 200 to ≈ 700 , before evolving into a new region that holds from days ≈ 900 to ≈ 1300 , and then continuing to evolve along a transient into previously unvisited areas of strategy space: this can reasonably be described as *acyclic* evolution. However in all five of the other sessions, there are clear recurrences, i.e. evidence of *cyclic* evolution: in the plot at mid-left, the region of strategy-space visited around days ≈ 300 to ≈ 500 is revisited in days ≈ 1200 to ≈ 1500 ; in the plot at lower-left, the region of strategy-space first visited over days ≈ 10 to ≈ 100 is revisited sporadically around roughly days 400–600, 700–900, and 1000–1300 as evidenced by the corresponding thin “trail of dust” in the RP; for the three plots in the right-hand column, regions of strategy-space first visited in the opening 100–300 days are returned to after many hundreds of days spent in other regions: the recurrences have been highlighted with ellipses. The lower-right plot is notable in that it shows a recurrence after a transit of more than 1,000 days of co-evolution.

be expected to affect the dynamics: in the experiments reported here, there was an obvious asymmetry in response, with the vast majority of the population of sellers rapidly co-evolving to be super-urgent (as shown in Figure 7) and the buyers then co-evolving toward multi-modal distributions of mainly relaxed strategies in response; with a different supply/demand schedule, this asymmetry could plausibly be reversed. Future papers will explore these and other issues.

A. Brief Introduction to Recurrence Plots

For the benefit of any readers unfamiliar with the recurrence plots (RPs) used in Figures 10 and 11, the diagrams in Figures 12 and 13 illustrate key aspects of this visualization technique for characterising high-dimensional dynamical systems: in their simplest incarnation, RPs are square arrays of cells or pixels, that are binary-shaded (e.g.: the pixels are either black or white), with a cell at column c and row r (denoted here by $C_{c,r}$) being shaded if the state of the system at the time associated with row r is a recurrence of a previously-observed system state that occurred at the time associated with column c ; otherwise unshaded.

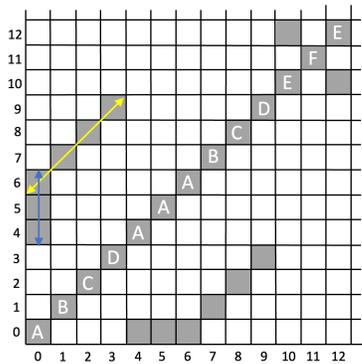


Figure 12. Illustrative synthetic recurrence plot (RP) for a dynamical system that starts at time $t = 0$ in state A and then over the next 12 timesteps transitions through the following sequence of states: B, C, D, A, A, A, B, C, D, E, F, E. Let $C_{c,r}$ denote the cell/pixel at column c and row r : the cell is shaded if the state of the system at time $t = r$ is a recurrence of the state of the system at time $t = c$, and is otherwise unshaded. For D -dimensional dynamical systems where the state of the system at time t is $\vec{s}(t) \in \mathbb{R}^D$, recurrence is usually defined to occur when the distance $|\vec{s}(r) - \vec{s}(c)| < \epsilon$ for some suitably small ϵ . By convention, the RP origin point is at lower left, and the diagonal line of cells $C_{c,r:c=r}$ is referred to as the *Line of Identity* (LOI); cells on the LOI are shaded because the distance from any state to itself is zero. The LOI divides the RP into two right-triangles with mirror-symmetric patterns of blank and shaded cells. The figure shows two key features in RPs: the *diagonal line* of four shaded cells (i.e., $C_{0,6}$, $C_{1,7}$, $C_{2,8}$ and $C_{3,9}$) starting at time $t = 6$ when the state sequence A-B-C-D recurs, having first occurred at times $t = 0$ to $t = 3$; and the *vertical line* of three shaded cells (i.e., $C_{0,4}$, $C_{0,5}$, and $C_{0,6}$) starting at time $t = 4$ where the state A recurs three times, having first occurred at time $t = 0$.

In systems where the state at any one time is one of a small number of discrete values, recurrence would usually be defined as strict equality of states. But in many dynamical systems of practical interest, the system state at time t is a D -dimensional real-valued vector $\vec{s}(t)$, and for creating an RP any subsequent state $\vec{s}(t + \Delta_t)$ that is within a D -dimensional solid hypersphere (i.e., a D -ball) centered on $\vec{s}(t)$ with radius ϵ is considered to be a recurrence of $\vec{s}(t)$. Naturally, the choice of ϵ is significant: if too large, each new state is registered as a recurrence of all previous states; if too small, it is possible that no recurrences are ever recorded. The RP origin point is normally displayed at lower left, and the diagonal line of cells $C_{c,r:c=r}$, referred to as the *Line of Identity* (LOI), is always shaded because the distance from any state to itself is zero.

ical systems of practical interest, the system state at time t is a D -dimensional real-valued vector $\vec{s}(t)$, and for creating an RP any subsequent state $\vec{s}(t + \Delta_t)$ that is within a D -dimensional solid hypersphere (i.e., a D -ball) centered on $\vec{s}(t)$ with radius ϵ is considered to be a recurrence of $\vec{s}(t)$. Naturally, the choice of ϵ is significant: if too large, each new state is registered as a recurrence of all previous states; if too small, it is possible that no recurrences are ever recorded. The RP origin point is normally displayed at lower left, and the diagonal line of cells $C_{c,r:c=r}$, referred to as the *Line of Identity* (LOI), is always shaded because the distance from any state to itself is zero.

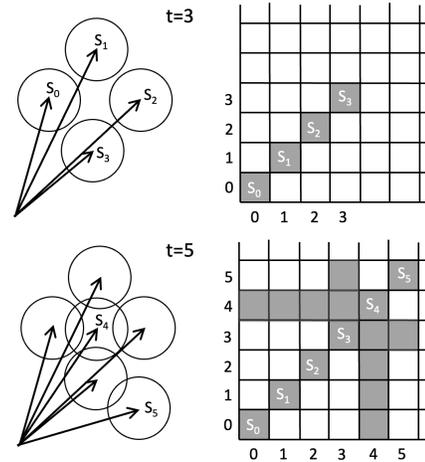


Figure 13. Illustrative synthetic recurrence plot (RP) for a D -dimensional dynamical system with state vector $\vec{s}(t) \in \mathbb{R}^D$ that starts at time $t = 0$ in state $\vec{s}(0) = S_0$ and then over the next three timesteps transitions through states S_1 to S_3 with no recurrences. The upper pair of figures, labelled $t = 3$, illustrates the set of non-recurring state-vectors on the left, and the corresponding RP on the right. Here the end-point of each state-vector is the centre of a D -ball (i.e., a solid D -dimensional hypersphere) of diameter ϵ , such that if any two balls intersect then the distance between the two vector end-points must be less than ϵ , which is thus counted as a recurrence. As there have been no recurrences by $t = 3$, the RP plot only shows shaded cells on the LOI. The lower pair of figures, labelled $t = 5$, illustrates the situation after the system has transitioned through state S_4 to state S_5 : the ball for S_4 intersected with the balls for each of states S_0 to S_3 , so the single state S_4 is recorded as a recurrence of each of the states S_0 to S_3 , giving rise to a horizontal line of recurrences on the RP at cells $C_{0,4} - C_{3,4}$; then S_5 intersects only with S_3 , shown on the RP as a single shaded cell at $C_{3,5}$.

Once an $N \times N$ RP is created, summary statistics can be calculated by doing simple image-processing such as computing the frequency distribution of lengths of vertical and diagonal lines in the RP, and then calculating summary statistics from those distributions: this approach is known as *Recurrence Quantification Analysis* (RQA). For example, the *trapping time* statistic (conventionally denoted by TT), given $P(v)$ the frequency distribution of vertical lines of length v in the RP, measures the RP's average length of vertical lines at least as long as v_{\min} (usually $v_{\min} = 2$):

$$TT = \left(\sum_{v=v_{\min}}^N vP(v) \right) / \left(\sum_{v=v_{\min}}^N P(v) \right)$$

So for example if an RP has a TT of 6, and the time delta between successive rows/columns on the RP is one hour, then the trapping time is six hours, indicating that on average the system remains within ϵ of any particular state for six hours.

For further details of RPs and RQA, see e.g. Eckmann et al. (1987); Marwan et al. (2007); Webber and Marwan (2015).

B. Bristol Stock Exchange: System Architecture

Figure 14 shows a schematic illustration of the overall architecture of the *Bristol Stock Exchange* (BSE) simulation of a contemporary fully electronic financial exchange, as was used in the experiments reported here.

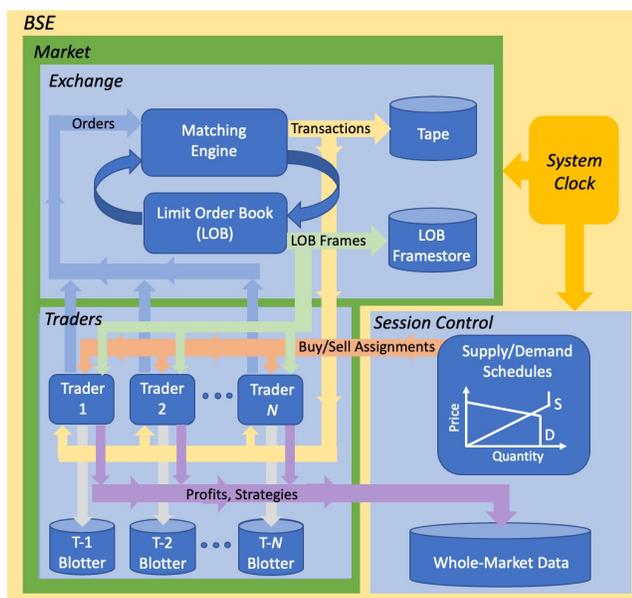


Figure 14. Schematic architecture of the BSE financial-market simulator: see text for further explanation.

BSE simulates a *market*, composed of an *exchange* and some number N of *traders* which each interact with the exchange. Any one simulation of a market session proceeds according to BSE’s *system clock*, which provides a unified time signal to all elements of the simulation. Separate from the simulation of the market is BSE’s *session control* logic, which determines the (potentially time-dependent) market’s *supply and demand schedule*: this is used to issue *assignments* to the traders, i.e. allocations of cash and limit-prices to buyers, and allocations of stock and limit-prices to sellers – this is the simulation’s correlate of real-world market *customer orders* coming from customers to sales-traders who are responsible for working each customer order. The session-control logic is also responsible for recording whole-market data, such as the profits and strategy-values of each trader in the market, as were vi-

sualised in the graphs and plots earlier in this paper.

The exchange receives orders from the traders: *bids* from buyers; *asks* from sellers. When each order arrives at the exchange, it is processed by the *matching engine*, attempting to find one or more matching bids for a newly-arrived ask, or one or more matching asks for a newly-arrived bid. It does this by comparing the new order to those earlier orders, as yet unfulfilled, that are “resting” at the exchange and which are summarised in aggregated and anonymized form on the exchange’s *limit order book* (LOB). If a new order can be matched with one or more existing orders on the LOB then the matching orders are removed from the LOB, and the new order plus its counterparty orders from the LOB are recorded as fulfilled, resulting in a transaction taking place. When a transaction occurs, its details are written to the exchange’s public record of transactions which is commonly referred to as the exchange’s *tape* – the tape records transactions and also other notable market events, such as cancellations of existing orders. When a transaction occurs, the exchange also notifies the traders concerned, adjusting their cash balances appropriately. The BSE exchange also can be configured to write the state of the LOB at any one instance (referred to as a *LOB frame*) to an external record, the *LOB framestore*, for subsequent analysis.

Each of the N traders in the market receives occasional fresh assignments from the session control, all have the same view of the LOB data published by the exchange, and when a trader is involved in a transaction it receives notification of the relevant details from the exchange’s matching engine. Each trader is able to send orders to the exchange, and to send cancellations of existing orders, and each maintains its own local private record of assignments received, orders sent to the exchange, and transaction details received from the exchange: this is conventionally referred to as the trader’s *blotter*.

For further details of BSE, see Cliff (2012, 2018).

References

- Alexandrov, N. (2021). Competitive arms-races among autonomous trading agents: Exploring the co-adaptive dynamics. Master’s thesis, University of Bristol.
- Alexandrov, N., Cliff, D., and Figuero, C. (2022). Exploring coevolutionary dynamics of competitive arms-races between infinitely diverse heterogenous adaptive automated trading agents. In Czupryna, M. and Kaminski, B., editors, *Advances in Social Simulation: Proceedings of the 16th Annual Social Simulation Conference (SSC2021)*, pages 93–104. Springer Proceedings in Complexity. SSRN: 3901889.
- Axtell, R. and Farmer, J. D. (2018). Agent-based modeling in economics and finance: Past, present, and future. Technical report.
- Bilal, Pant, M., Zaheer, H., Garcia-Hernandez, L., and Abraham, A. (2020). Differential evolution: A review of more than two decades of research. *Engineering Appli-*

- cations of Artificial Intelligence*, 90:103479.
- Chasani, P. and Likas, A. (2022). The UU-Test for statistical modeling of unimodal data. *Pattern Recognition*, 122(108272):1–19.
- Cliff, D. (1997). Minimal-intelligence agents for bargaining behaviours in market-based environments. Technical Report HPL-97-91, HP Labs Technical Report.
- Cliff, D. (2012). *Bristol Stock Exchange: open-source financial exchange simulator*. <https://github.com/davecliff/BristolStockExchange>.
- Cliff, D. (2018). BSE : A Minimal Simulation of a Limit-Order-Book Stock Exchange. In Bruzzone, F., editor, *Proc. 30th Euro. Modeling and Simulation Symposium (EMSS2018)*, pages 194–203.
- Cliff, D. (2021). Parameterized-Response Zero-Intelligence Traders. SSRN:3823317.
- Cliff, D. (2022). Metapopulation differential co-evolution of trading strategies in a model financial market. SSRN: 4153519.
- Cliff, D. and Rollins, M. (2020). Methods matter: A trading algorithm with no intelligence routinely outperforms AI-based traders. In *Proceedings of IEEE Symposium on Computational Intelligence in Financial Engineering (CIFER2020)*.
- Das, R., Hanson, J., Kephart, J., and Tesauro, G. (2001). Agent-human interactions in the continuous double auction. In *Proc. IJCAI-2001*, pages 1169–1176.
- De Luca, M. and Cliff, D. (2011a). Agent-human interactions in the continuous double auction, redux: Using the OpEx lab-in-a-box to explore ZIP and GDX. In *Proceedings of the 2011 International Conference on Agents and Artificial Intelligence (ICAART2011)*.
- De Luca, M. and Cliff, D. (2011b). Human-agent auction interactions: Adaptive-Aggressive agents dominate. In *Proceedings IJCAI-2011*, pages 178–185.
- De Luca, M., Szostek, C., Carlidge, J., and Cliff, D. (2011). Studies of interaction between human traders and algorithmic trading systems. Technical report, UK Government Office for Science, London.
- Eckmann, J.-P., Oliffson Kamphorst, S., and Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters*, 5:973–977.
- Farmer, J. D., Patelli, P., and Zovko, I. (2005). The Predictive Power of Zero Intelligence in Financial Markets. *Proceedings of the National Academy of Sciences*, 102(6):2254–2259.
- Figuro, C. (2021). Evolving trader-agents via strongly typed genetic programming. Master’s thesis, University of Bristol Department of Computer Science.
- Gittins, J., Glazebrook, K., and Weber, R. (2011). *Multi-Armed Bandit Allocation Indices*. Wiley, 2 edition.
- Gjerstad, S. (2003). The impact of pace in double auction bargaining. Technical report, Department of Economics, University of Arizona.
- Gjerstad, S. and Dickhaut, J. (1998). Price formation in double auctions. *Games & Economic Behavior*, 22(1):1–29.
- Gode, D. and Sunder, S. (1993). Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. *Journal of Political Economy*, 101(1):119–137.
- Hartigan, J. and Hartigan, P. (1985). The dip test of unimodality. *The Annals of Statistics*, 13(1):70–84.
- Ladley, D. (2012). Zero Intelligence in Economics and Finance. *The Knowledge Engineering Review*, 27(2):273–286.
- Lattimore, T. and Szepesvari, C. (2020). *Bandit Algorithms*. Cambridge University Press.
- Marwan, N., Carmen Romano, M., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438:237–329.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.
- Montgomery, D. (2019). *Design and Analysis of Experiments*. Wiley, sixth edition.
- Myles White, J. (2012). *Bandit Algorithms for Website Optimization: Developing, Deploying, and Debugging*. O’Reilly.
- Rollins, M. and Cliff, D. (2020). Which trading agent is best? using a threaded parallel simulation of a financial market changes the pecking-order. In *Proceedings of the 32nd European Modeling and Simulation Symposium (EMSS2020)*.
- Rust, J., Miller, J., and Palmer, R. (1992). Behavior of trading automata in a computerized double auction market. In Friedman, D. and Rust, J., editors, *The Double Auction Market: Institutions, Theories, and Evidence*, pages 155–198. Addison-Wesley.
- Snashall, D. and Cliff, D. (2019). Adaptive-Aggressive traders don’t dominate. In van Herik, J., Rocha, A., and Steels, L., editors, *Agents and Artificial Intelligence: Selected papers from ICAART2019*. Springer.
- Storn, R. and Price, K. (1997). Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359.
- Tesauro, G. and Bredin, J. (2002). Sequential strategic bidding in auctions using dynamic programming. In *Proceedings AAMAS 2002*.
- Tesauro, G. and Das, R. (2001). High-performance bidding agents for the continuous double auction. In *Proc. 3rd ACM Conference on E-Commerce*, pages 206–209.
- Vach, D. (2015). Comparison of double auction bidding strategies for automated trading agents. Master’s thesis, Charles University in Prague.
- Vytelingum, P., Cliff, D., and Jennings, N. (2008). Strategic bidding in continuous double auctions. *Artificial Intelligence*, 172(14):1700–1729.
- Walsh, W., Das, R., Tesauro, G., and Kephart, J. (2002). Analyzing complex strategic interactions in multi-agent systems. In *Proc. of the AAAI Workshop on Game-Theoretic and Decision-Theoretic Agents*.
- Webber, C. and Marwan, N., editors (2015). *Recurrence Quantification Analysis: Theory and Best Practice*. Springer.