IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 324005 (30pp) doi:10.1088/1751-8113/41/32/324005

Generating functional analysis of minority games with inner product strategy definitions

A C C Coolen and N Shayeghi

Department of Mathematics, King's College London, The Strand, London WC2R 2LS, United Kingdom

E-mail: ton.coolen@kcl.ac.uk and nima@mth.kcl.ac.uk

Received 2 December 2007, in final form 23 January 2008
Published 30 July 2008
Online at stacks.iop.org/JPhysA/41/324005

Abstract

We use generating functional methods to solve the so-called inner product versions of the minority game (MG), with fake and/or real market histories, by generalizing the theory developed recently for look-up table MGs with real histories. The phase diagrams of the look-up table and inner product MG versions are generally found to be identical, with the exception of inner product MGs where histories are sampled linearly, which are found to be structurally critical. However, we encounter interesting differences both in the theory (where the role of the history frequency distribution in look-up table MGs is taken over by the eigenvalue spectrum of a history covariance matrix in inner product MGs) and in the static and dynamic phenomenology of the models. Our theoretical predictions are supported by numerical simulations.

PACS numbers: 02.50.Le, 87.23.Ge, 05.70.Ln, 64.60.Ht

1. Introduction

Minority games (MG) [1, 2] are relatively simple and transparent models that were designed to increase our understanding of the complex collective processes which result from inductive decision making by large numbers of interacting agents in simplified 'markets'. They can be seen as mathematical implementations of the so-called El Farol bar problem [3]. Many versions of the MG have by now been studied in the literature. They differ in the type of microscopic dynamics used (e.g. batch versus on-line, stochastic versus deterministic), in the definition of the information provided to the agents (real-valued versus discrete, true versus fake market history) and the agents' decision making strategies, and also in the specific recipe defined for the conversion of the external information into a trading action (inner products versus look-up tables). In MG models with 'fake' market memory, proposed first in [4], at each point in time all agents are given random data upon which to base their decisions, rather than the actual market history. They have the advantage of being Markovian and were

therefore the first to be studied and solved in the theoretical physics literature using techniques from equilibrium [5, 6] and non-equilibrium [7–9] statistical mechanics, as first developed for and applied to spin-glasses. Models with true market history are non-Markovian and therefore more demanding, but also here theoretical progress is being made [10, 11]. For a more extensive introduction to MGs and their statistical mechanical theory we refer to the recent textbooks [12, 13].

To our knowledge, however, there has not yet been any such exact solution for MG models in which the decision making strategies are defined via inner products rather than look-up tables [14, 15]; not even for the simplest Markovian case, where the market histories are fake[1]. Inner product MGs have so far tended to be studied either numerically, or at the level of investigating properties of the stochastic microscopic laws, in spite of the fact that one might well argue that the definitions of the 'inner product MGs' are possibly closer to how financial time series tend to be predicted by practitioners [21] (as these usually involve generalized linear market prediction models). Moreover, in fact not a single study has so far been published of inner product MGs with real as opposed to fake histories, there are not even papers based on numerical simulations only.

In this paper we first generalize the existing generating functional theory [11] that was developed for look-up table MGs with real and/or fake histories to a larger class of MG versions, which includes the familiar look-up table MGs and inner product MGs as special cases. We then focus on the analysis of the dynamics and phase transitions of inner product MG versions with and without real histories, and show that in the infinite system size limit: (i) for inconsistent fake histories the inner product and look-up table MG versions behave fully identically, including short time correlation and response functions and the volatility (apart from a trivial transformation related to time-scales), (ii) similar to what was found in [11] for look-up table MGs, for inner product MGs the introduction of real histories affects the observables in the ergodic region, but does not generally change the phase diagram, except when the histories are sampled by unbounded functions (see below), (iii) the key observable to measure the differences between real and fake histories in our generalized MG models is the spectrum of the history covariance matrix in the stationary state, which for inner product MGs reduces to the history frequency distribution, (iv) for inner product MGs an approximate calculation of this spectrum leads to a fully closed stationary state theory, the predictions of which are supported by numerical simulations, and (v) inner product MGs with predominantly real histories where the inner product is taken between strategy entries and the actual values of the past market bids (or, more generally, by an unbounded function of these bids rather than e.g. a saturating function) are permanently critical, giving a volatility that is either zero or infinity (separated by a novel critical value for the control parameter α, unrelated to the standard ergodicity breaking transition point).

2. Definitions

We imagine having a system of N agents, labeled by $i = 1, \ldots, N$. At each step $\ell \in \{0, 1, 2, \ldots\}$ of the game each agent submits a 'bid' $b_i(\ell) \in \mathbb{R}$ to the market. The (re-scaled) cumulative bid at stage ℓ is defined as

$$A(\ell) = \frac{1}{\sqrt{N}} \sum_i b_i(\ell) + A_e(\ell). \tag{1}$$

[1] At least qualitative equivalence between inner product and look-up table MGs was anticipated for models with fake histories on the basis of an approximate analysis of the processes involved [16, 17].

Here $A_e(\ell)$ denotes an external contribution that represents e.g. random market perturbations or actions by market regulators. Each agent i determines his bid $b_i(\ell)$ at each step ℓ on the basis of external information, which is to represent historic market data, using his preferred decision making strategy at that point in time. Each agent has S such strategies, labeled by $a = 1, \ldots, S$. Profit is made by those agents who find themselves subsequently in the minority group, i.e. in the case of $A(\ell) > 0$ by those with $b_i(\ell) < 0$ and in the case of $A(\ell) < 0$ by those with $b_i(\ell) > 0$.

MG versions differ in their definition of the information provided to the agents and in how this information is converted into trading decisions. In the original MG, the information consisted of the signs of the overall market bids $A(\ell)$ over a number of time steps in the past. In inner product MGs, in contrast, the external information at stage ℓ will consist of a p-dimensional vector $\mathbf{I}(\ell)$ (generalizing the 'information vector' of [14, 15]), where $\alpha = p/N$ remains finite as $N \to \infty$, with entries

$$I_\lambda(\ell) = \frac{1}{\sqrt{p}} f[(1-\zeta)A(\ell-\lambda) + \zeta Z(\ell,\lambda)] \tag{2}$$

and with $\lambda \in \{1, \ldots, p\}$. The function $f[.]$ (which we choose to be anti-symmetric) allows the information given to the agents to involve more general properties of the overall bids than just their sign, on which the original model [1, 2] was based. The $Z(\ell, \lambda)$ are zero-average Gaussian random variables, which represent the 'fake' alternative to the true market data $A(\ell - \lambda)$, and p is the number of iteration steps in the past for which market information is made available. For reasons that will become clear we write $I_\lambda(\ell) = p^{-\frac{1}{2}} \mathcal{F}_\lambda[\ell, A, Z]$, so

$$\mathcal{F}_\lambda[\ell, A, Z] = f[(1-\zeta)A(\ell-\lambda) + \zeta Z(\ell,\lambda)]. \tag{3}$$

If we ensure that $f[.]$ remains finite, one has $\lim_{N\to\infty} \sum_\lambda [I_\lambda(\ell)]^2 = \mathcal{O}(1)$ for all ℓ. Definition (2) implies that for $\zeta = 1$ the external information is fully random, whereas for $\zeta = 0$ it represents true historic market data. Following [11] we distinguish between two types of definitions for the statistical properties of the 'fake memory' variables:

$$\text{consistent}: \quad Z(\ell,\lambda) = Z(\ell-\lambda), \qquad \langle Z(\ell)Z(\ell')\rangle = S^2\delta_{\ell\ell'} \tag{4}$$

$$\text{inconsistent}: \quad Z(\ell,\lambda)\ \text{all independent}, \qquad \langle Z(\ell,\lambda)Z(\ell',\lambda')\rangle = S^2\delta_{\ell\ell'}\delta_{\lambda\lambda'}. \tag{5}$$

We take the agents' strategies to be represented by vectors $\mathbf{R}^{ia} \in \mathbb{R}^p$ with components R^{ia}_λ. In accordance with standard definitions [14, 15] we choose these components to be independent zero-average and unit-variance Gaussian random variables, assigned before the start of the game, and remaining fixed throughout. They represent quenched disorder. The arrival of information $\mathbf{I}(\ell)$ at step ℓ prompts each agent i to submit the bid $b_i(\ell) = \sum_\lambda R^{ia_i(\ell)}_\lambda I_\lambda(\ell)$, where $a_i(\ell)$ denotes the preferred strategy at stage ℓ of agent i. In order to detect which of their S private strategies to use, all agents keep track of valuations $p_{ia}(\ell)$ of their strategies. These measure to what extent each strategy would have led to minority decisions if it had been used all the time:

$$p_{ia}(\ell+1) = p_{ia}(\ell) - \frac{\tilde{\eta}}{\sqrt{N}} A(\ell) \sum_\lambda R^{ia}_\lambda I_\lambda(\ell). \tag{6}$$

The factor $\tilde{\eta} > 0$ represents a learning rate. The active strategy $a_i(\ell)$ of trader i at stage ℓ is chosen to be the one with the highest valuation $p_{ia}(\ell)$ at that moment, i.e. $a_i(\ell) = \arg\max_a\{p_{ia}(\ell)\}$. Our full equations now become, in more explicit form and with the notation (3):

$$p_{ia}(\ell+1) = p_{ia}(\ell) - \frac{\tilde{\eta}}{N\sqrt{\alpha}} \sum_\lambda R^{ia}_\lambda \mathcal{F}_\lambda[\ell, A, Z] A(\ell) \tag{7}$$

$$A(\ell) = \frac{1}{\sqrt{\alpha}N} \sum_i \sum_\lambda R_\lambda^{i a_i(\ell)} \mathcal{F}_\lambda[\ell, A, Z]. \tag{8}$$

Henceforth we restrict ourselves to the case $S = 2$, two strategies per agent so $a \in \{1, 2\}$, and introduce the conventional definitions

$$q_i(\ell) = \tfrac{1}{2}[p_{i1}(\ell) - p_{i2}(\ell)], \qquad \boldsymbol{\omega}^i = \tfrac{1}{2}(\mathbf{R}^{i1} + \mathbf{R}^{i2}), \qquad \boldsymbol{\xi}^i = \tfrac{1}{2}(\mathbf{R}^{i1} - \mathbf{R}^{i2})$$

as well as $\boldsymbol{\Omega} = N^{-1/2} \sum_i \boldsymbol{\omega}^i$. This allows us to write $\mathbf{R}^{i a(\ell)} = \boldsymbol{\omega}^i + \mathrm{sgn}[q_i(\ell)]\boldsymbol{\xi}^i$. We then proceed to include decision noise in the familiar way via the substitution $\mathrm{sgn}[q_i(\ell)] \to \sigma[q_i(\ell), z_i(\ell)]$, in which the $\{z_i(\ell)\}$ are independent zero average random numbers, described by a symmetric and unit-variance distribution $P(z)$. The function $\sigma[q, z]$ is taken to be non-decreasing in q for any z, and parametrized by a parameter $T \geqslant 0$ such that $\sigma[q, z] \in \{-1, 1\}$, with $\lim_{T\to 0} \sigma[q, z] = \mathrm{sgn}[q]$ and $\lim_{T\to\infty} \int \mathrm{d}z P(z)\sigma[q, z] = 0$. The two main examples are additive and multiplicative noise:

$$\text{additive :} \quad \sigma[q, z] = \mathrm{sgn}[q + Tz] \tag{9}$$

$$\text{multiplicative :} \quad \sigma[q, z] = \mathrm{sgn}[q]\,\mathrm{sgn}[1 + Tz] \tag{10}$$

T measures the degree of randomness in the agents' decision making, with $T = 0$ bringing us back to $a_i(\ell) = \arg\max_a\{p_{ia}(\ell)\}$, and with random strategy selection for $T = \infty$. In on-line MG theories one will in practice only need the decision noise average

$$\sigma[q] = \int \mathrm{d}z P(z)\sigma[q, z]. \tag{11}$$

Upon also adding the usual external perturbation fields $\{\theta_i(\ell)\}$ to define response functions, our microscopic equations (7), (8) are then replaced by

$$q_i(\ell + 1) = q_i(\ell) + \theta_i(\ell) - \frac{\tilde{\eta}}{N\sqrt{\alpha}} \sum_{\lambda=1}^{p} \xi_\lambda^i \mathcal{F}_\lambda[\ell, A, Z] A(\ell) \tag{12}$$

$$A(\ell) = A_e(\ell) + \frac{1}{\sqrt{\alpha N}} \sum_{\lambda=1}^{p} \left\{ \Omega_\lambda + \frac{1}{\sqrt{N}} \sum_i \sigma[q_i(\ell), z_i(\ell)]\xi_\lambda^i \right\} \mathcal{F}_\lambda[\ell, A, Z]. \tag{13}$$

The values of $A(\ell)$ for $\ell \leqslant 0$ and of the $q_i(0)$ play the role of initial conditions.

Equations (12), (13) are identical to those studied in [11] for look-up table MGs; the only difference between look-up table and inner product definitions for the information-to-decision conversion in the MG is in the definition of the function $\mathcal{F}_\lambda[\ell, A, Z]$. We may thus take over from [11] the derivation of the effective single agent problem, and simply make the appropriate substitutions in the final result[2]:

$$\text{look-up table :} \quad \boldsymbol{\lambda} \in \{-1, 1\}^M, \qquad 2^M = p, \qquad \mathcal{F}_{\boldsymbol{\lambda}}[\ell, A, Z] = \sqrt{\alpha N}\delta_{\boldsymbol{\lambda}, \boldsymbol{\lambda}(\ell, A, Z)}$$

$$\lambda_\mu(\ell, A, Z) = \mathrm{sgn}[(1 - \zeta)A(\ell - \mu) + \zeta Z(\ell, \mu)] \tag{14}$$

$$\text{inner product :} \quad \lambda \in \{1, \ldots, p\}, \qquad \mathcal{F}_\lambda[\ell, A, Z] = f[(1 - \zeta)A(\ell - \lambda) + \zeta Z(\ell, \lambda)]. \tag{15}$$

The inner product models of [14, 15] correspond to choosing recipe (15), with $\zeta = 1$ (fake history) and $f[A] = A$, giving $\mathcal{F}_\lambda[\ell, A, Z] = Z(\ell, \lambda)$.

[2] In fact, the formulation of the derivation in [11] was chosen with the present generalization in mind.

3. Generating functional analysis of generalized MGs

3.1. The effective single agent equation

We choose the real-time duration of individual iterations of the MG equations to be $\delta_N = \tilde{\eta}/2p$. This was proposed within the replica approach [5, 6, 10], and later confirmed to be the canonical scaling via generating functional analysis studies [9, 11]. The generating functional analysis in [11], which via (15) can be made to apply also to inner-product models, leads as always to closed deterministic dynamical order parameter equations which are fully exact in the limit $N \to \infty$ and on times which do not scale with N. The order parameters are the disorder-averaged single-site correlation and response function, which in terms of the original discrete iterations $\ell = 0, 1, 2, \ldots$ are defined as

$$C(\ell, \ell') = \lim_{N\to\infty} \frac{1}{N} \sum_i \overline{\langle \sigma[q_i(\ell), z_i(\ell)]\sigma[q_i(\ell'), z_i(\ell')]\rangle} \tag{16}$$

$$G(\ell, \ell') = \lim_{N\to\infty} \frac{1}{N} \sum_i \frac{\partial}{\partial\theta_i(\ell')} \overline{\langle \sigma[q_i(\ell), z_i(\ell)]\rangle}. \tag{17}$$

The brackets $\langle \ldots \rangle$ refer to averaging over all realizations of both the decision noise variables $\{z_i(\ell)\}$ and the fake bid variables $\{Z(\ell, \lambda)\}$, at all times. It was shown in [11] (to which we refer for details) that for $N \to \infty$ and in terms of new time variable $t = \ell\delta_N$ (which will become real-valued in this limit) the order parameters (16), (17) are to be extracted self-consistently from the following disorder-free effective single agent process with a retarded self-interaction and zero-average Gaussian noise $\eta(t)$ with non-trivial temporal correlations $\langle \eta(t)\eta(t')\rangle = \Sigma(t, t')$:

$$\frac{\mathrm{d}}{\mathrm{d}t} q(t) = \theta(t) - \alpha \int_0^t \mathrm{d}t' R(t, t')\sigma[q(t')] + \sqrt{\alpha}\eta(t). \tag{18}$$

As a result of the limit $N \to \infty$ we must also write $C(\ell, \ell') \to C(t, t')$ and $G(\ell, \ell') \to G(t, t')$. One always has $G(t, t') = 0$ for $t \leqslant t'$ (due to causality), and $C(t, t') = C(t', t)$ with $C(t, t) = 1$. If we write averages over realizations of the non-Markovian process (18) as $\langle \ldots \rangle_\star$, the kernels $\{C, G\}$ must satisfy for $t > t'$:

$$C(t, t') = \langle \sigma[q(t)]\sigma[q(t')]\rangle_\star \qquad G(t, t') = \frac{\delta}{\delta\theta(t')} \langle \sigma[q(t)]\rangle_\star. \tag{19}$$

The relation between the auxiliary kernels $\{R, \Sigma\}$ and the order parameters $\{C, G\}$ was found to be defined via an effective process for the overall bid $A(\ell)$, namely by

$$R(t, t') = \frac{\delta}{\delta A_e(t')} \lim_{\delta_N\to 0} \langle\langle \overline{W}[\ell', \ell; \{A, Z\}]A(\ell)\rangle\rangle_{\{A,Z\}}|_{\ell=t/\delta_N, \ell'=t'/\delta_N} \tag{20}$$

$$\Sigma(t, t') = \tilde{\eta} \lim_{\delta_N\to 0} {\delta_N}^{-1} \langle\langle \overline{W}[\ell, \ell'; \{A, Z\}]A(\ell)A(\ell')\rangle\rangle_{\{A,Z\}}|_{\ell=t/\delta_N, \ell'=t'/\delta_N}. \tag{21}$$

Here $\langle\langle \ldots \rangle\rangle_{A,Z}$ refers to an average over this effective stochastic process for the bids $\{A\}$ and the pseudo-history $\{Z\}$, a process that was found [11] to take the form

$$A(\ell) = A_e(\ell) + \phi_\ell - \frac{1}{2}\tilde{\eta} \sum_{\ell'<\ell} G(\ell, \ell')\overline{W}[\ell, \ell'; \{A, Z\}]A(\ell') \tag{22}$$

with zero-average Gaussian random fields $\{\phi\}$, characterized by

$$\langle \phi_\ell \phi_{\ell'} \rangle_{\{\phi|A,Z\}} = \tfrac{1}{2}[1 + C(\ell, \ell')]\overline{W}[\ell, \ell'; \{A, Z\}]. \tag{23}$$

The function $\overline{W}[.;.]$ in these formulae is defined as

$$\overline{W}[\ell,\ell';\{A,Z\}]=\frac{1}{p}\sum_{\lambda}\mathcal{F}_{\lambda}[\ell,A,Z]\mathcal{F}_{\lambda}[\ell',A,Z]. \tag{24}$$

The macroscopic theory is closed. The remaining problem is to (i) express the kernels $\{R,\Sigma\}$ in terms of $\{C,G\}$ using (20), (21), followed by (ii) calculating the order parameters $\{C,G\}$ from (19).

It is quite satisfactory to find that the theory of [11] can be made to apply to a much broader class of MG models, which saves us from having to redo the generating functional analysis here. Mathematically, the differences between the two main MG families (look-up table versus inner product strategies) are within the macroscopic theory found to be limited to which expression to substitute in equations (20)–(22) for the function $\overline{W}[\ell,\ell';\{A,Z\}]$, which measures the similarity between the market 'histories' (whether fake or real) as observed at stages ℓ and ℓ' of the process:

$$\text{look-up table:}\quad \overline{W}[\ell,\ell';\{A,Z\}]=\delta_{\boldsymbol{\lambda}(\ell,A,Z),\boldsymbol{\lambda}(\ell',A,Z)} \tag{25}$$

$$\text{inner product :}\quad \overline{W}[\ell,\ell';\{A,Z\}]=\frac{1}{\alpha N}\sum_{\lambda}f[(1-\zeta)A(\ell-\lambda)+\zeta Z(\ell,\lambda)]$$
$$\times f[(1-\zeta)A(\ell'-\lambda)+\zeta Z(\ell',\lambda)]. \tag{26}$$

Here $\boldsymbol{\lambda}(\ell,A,Z)\in\{-1,1\}^M$ (with $2^M=p$) denotes the 'history string' as observed by agents in the look-up table MGs at time ℓ, with entries $\lambda_k(\ell,A,Z)=\text{sgn}[(1-\zeta)A(\ell-k)]+\zeta Z(\ell,k)]$. Similarly, the differences between the two 'fake memory' definitions (4), (5) are limited to the details of the zero-average Gaussian variables $\{Z\}$ in (25), (26) and the associated averaging process $\langle\ldots\rangle_Z$.

3.2. Time translation invariant stationary states

In the special case of fully ergodic and time-translation invariant states (TTI) without anomalous response, where $C(t,t')=C(t-t')$, $G(t,t')=G(t-t')$, $R(t,t')=R(t-t')$ and $\Sigma(t,t')=\Sigma(t-t')$, one can derive from the effective single agent equation relatively simple and familiar expressions for persistent order parameters [11]. The relevant scalar quantities are $\chi=\int_0^\infty \mathrm{d}t\,G(t)$, $\chi_R=\int_0^\infty \mathrm{d}t\,R(t)$, $c=\lim_{t\to\infty}C(t)$, and the fraction ϕ of 'frozen' agents in the game. They were found [11] to obey:

$$\phi=1-\text{Erf}[u] \tag{27}$$

$$c=\sigma^2[\infty]\left\{1-\text{Erf}[u]+\frac{1}{2u^2}\text{Erf}[u]-\frac{1}{u\sqrt{\pi}}\,\mathrm{e}^{-u^2}\right\} \tag{28}$$

$$\chi=\text{Erf}[u]/\alpha\chi_{\text{R}} \tag{29}$$

with the short-hands $u=\sqrt{\alpha}\chi_{\text{R}}\sigma[\infty]/S_0\sqrt{2}$ and $S_0^2=\Sigma(\infty)$. In order to find the TTI stationary solution $\{\phi,c,\chi\}$ and the phase transition point (defined by $\chi\to\infty$), we therefore do not need to solve for our order parameter kernels in full but just need to extract expressions for χ_{R} and S_0 from the stochastic overall bid process (22). These latter two quantities can be written as

$$\chi_{\text{R}}=\lim_{\delta_N\to 0}\left\{\overline{W}[0,0;\{A,Z\}]+\sum_{\ell=1}^{\infty}\frac{\partial}{\partial A(0)}\langle\langle\overline{W}[\ell,0;\{A,Z\}]A(\ell)\rangle\rangle_{\{A,Z\}}\right\} \tag{30}$$

$$S_0^2 = \lim_{\delta_N \to 0} \lim_{L \to \infty} \frac{\tilde{\eta}}{L^2 \delta_N} \sum_{\ell,\ell'=1}^{L} \langle\langle \overline{W}[\ell, \ell'; \{A, Z\}] A(\ell) A(\ell') \rangle\rangle_{\{A,Z\}}. \tag{31}$$

Note that $A(0)$ and $\overline{W}[0, 0; \{A, Z\}]$ express the initial conditions; for the purpose of evaluating the stationary state we may now drop the global bid perturbations $\{A_e(t)\}$.

3.3. Expressions for the kernels R and Σ

In order to work out the theory also away from TTI stationary states and compare it to that of look-up table MGs, knowledge of just χ_R and S_0 is no longer sufficient; one needs to calculate the kernels R and Σ in full. Here it is no longer clear to what extent the analysis of [11] can be adapted. The differences between the two model families start to manifest themselves. We first define

$$\Delta_{r+1}(\ell_0, \ldots, \ell_r) = p^r \left\langle\!\!\left\langle \overline{W}[\ell_0, \ell_r; \{A, Z\}] \prod_{i=1}^{r} \overline{W}[\ell_{i-1}, \ell_i; \{A, Z\}] \right\rangle\!\!\right\rangle_{\{A,Z\}} \tag{32}$$

$$\begin{aligned} \tilde{\Delta}_{r+r'+2}(\ell_0, \ldots, \ell_r; \ell'_0, \ldots, \ell'_{r'}) = p^{r+r'+1} \Big\langle\!\!\Big\langle & \overline{W}[\ell_0, \ell'_0; \{A, Z\}] \overline{W}[\ell_r, \ell'_{r'}; \{A, Z\}] \\ & \times \left[\prod_{i=1}^{r} \overline{W}[\ell_{i-1}, \ell_i; \{A, Z\}] \right] \left[\prod_{j=1}^{r'} \overline{W}[\ell'_{j-1}, \ell'_j; \{A, Z\}] \right] \Big\rangle\!\!\Big\rangle_{\{A,Z\}}. \end{aligned} \tag{33}$$

As was done in in [11] we re-write the global bid equation (22) as

$$\sum_{\ell' \leqslant \ell} \left\{ \delta_{\ell\ell'} + \frac{1}{2} \tilde{\eta} G(\ell, \ell') \overline{W}[\ell, \ell'; \{A, Z\}] \right\} A(\ell') = A_e(\ell) + \phi_\ell$$

and we invert the operator on the left-hand side, using $\delta_N = \tilde{\eta}/2p$:

$$\begin{aligned} A(\ell) = A_e(\ell) + \phi_\ell + \sum_{r>0} (-\delta_N p)^r \sum_{\ell_1 \ldots \ell_r} & G(\ell, \ell_1) G(\ell_1, \ell_2) \ldots G(\ell_{r-1}, \ell_r) \\ & \times \overline{W}[\ell, \ell_1; \{A, Z\}] \overline{W}[\ell_1, \ell_2; \{A, Z\}] \ldots \overline{W}[\ell_{r-1}, \ell_r; \{A, Z\}] \big[A_e(\ell_r) + \phi_{\ell_r} \big]. \end{aligned} \tag{34}$$

We insert (34) into (20), and consider only infinitesimal external bid perturbations A_e:

$$\begin{aligned} R(t, t') &= \delta(t - t') + \lim_{\delta_N \to 0} \frac{1}{\delta_N} \left\{ \sum_{r>0} (-\delta_N)^r \sum_{\ell_1 \ldots \ell_{r-1}} G(\ell_0, \ell_1) G(\ell_1, \ell_2) \ldots G(\ell_{r-1}, \ell_r) \right. \\ &\qquad \left. \times p^r \left\langle\!\!\left\langle \overline{W}[\ell_0, \ell_r; \{A, Z\}] \prod_{i=1}^{r} \overline{W}[\ell_{i-1}, \ell_i; \{A, Z\}] \right\rangle\!\!\right\rangle_{\{A,Z\}} \right\} \Bigg|_{\ell_0 = \frac{t}{\delta_N}, \ell_r = \frac{t'}{\delta_N}} \\ &= \delta(t - t') + \lim_{\delta_N \to 0} \frac{1}{\delta_N} \left\{ \sum_{r>0} (-\delta_N)^r \sum_{\ell_1 \ldots \ell_{r-1}} G(\ell_0, \ell_1) \ldots G(\ell_{r-1}, \ell_r) \right. \\ &\qquad \left. \times \Delta_{r+1}(\ell_0, \ldots, \ell_r) \right\} \Bigg|_{\ell_0 = \frac{t}{\delta_N}, \ell_r = \frac{t'}{\delta_N}}. \end{aligned} \tag{35}$$

This is clearly identical to the corresponding expression found in [11] (although now the meaning of the kernels $\Delta_k(\ldots)$ is allowed to be different, dependent on which model family

we choose to apply the generalized theory to). Similarly we can insert (34) into (21), again with $A_e \to 0$, and find

$$\Sigma(t,t') = \tilde{\eta} \lim_{\delta_N \to 0} \frac{1}{\delta_N} \left\{ \sum_{r,r' \geqslant 0} (-\delta_N)^{r+r'} \sum_{\ell_1 \ldots \ell_r} G(\ell_0,\ell_1) \ldots G(\ell_{r-1},\ell_r) \right.$$
$$\times \sum_{\ell'_1 \ldots \ell'_{r'}} G(\ell'_0,\ell'_1) \ldots G(\ell'_{r'-1},\ell'_{r'}) p^{r+r'} \left\langle\!\!\left\langle \langle \phi_{\ell_r} \phi_{\ell'_{r'}} \rangle_{\{\phi|A,Z\}} \overline{W}[\ell_0,\ell'_0;\{A,Z\}] \right.\right.$$
$$\left.\left. \times \left[\prod_{i=1}^{r} \overline{W}[\ell_{i-1},\ell_i;\{A,Z\}] \right] \left[\prod_{j=1}^{r'} \overline{W}[\ell'_{j-1},\ell'_j;\{A,Z\}] \right] \right\rangle\!\!\right\rangle_{\{A,Z\}} \Bigg\} \Bigg|_{\ell_0 = \frac{t}{\delta_N}, \ell'_0 = \frac{t'}{\delta_N}}$$
$$= \lim_{\delta_N \to 0} \left\{ \sum_{r,r' \geqslant 0} (-\delta_N)^{r+r'} \sum_{\ell_1 \ldots \ell_r} G(\ell_0,\ell_1) \ldots G(\ell_{r-1},\ell_r) \right.$$
$$\times \sum_{\ell'_1 \ldots \ell'_{r'}} G(\ell'_0,\ell'_1) \ldots G(\ell'_{r'-1},\ell'_{r'}) [1 + C(\ell_r,\ell'_{r'})]$$
$$\left. \times \tilde{\Delta}_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) \right\} \Bigg|_{\ell_0 = \frac{t}{\delta_N}, \ell'_0 = \frac{t'}{\delta_N}} . \tag{36}$$

The limits $\delta_N \to 0$ in (35), (36) are well-defined. Each time summation combines with a factor δ_N to generate an integral, whereas pairwise identical times in (36) leave a 'bare' factor δ_N but can also be anticipated to cause $\tilde{\Delta}_{r+r'+2}(\ldots)$ gaining a factor $p = \tilde{\eta}/2\delta_N$ in compensation.

Equations (35), (36) show, upon inserting the respective definitions (25, 26) of the function $\overline{W}[.;.]$, that replacing look-up table strategy definitions by inner product ones does have implications. For instance, the following relation holds for look-up table MGs but appears to be no longer valid for inner product ones:

$$\tilde{\Delta}_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) = \Delta_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) \tag{37}$$

4. Inner product MGs with inconsistent fake market history

The simplest instance of our inner product MG model is the case $\zeta = 1$, i.e. fake market history, of the inconsistent type. Here the function $\overline{W}[\ell,\ell';\{A,Z\}] = p^{-1} \sum_\lambda f[Z(\ell,\lambda)] f[Z(\ell',\lambda)]$ is no longer dependent on the real market bids A, and $\langle Z(\ell,\lambda) Z(\ell',\lambda') \rangle = S^2 \delta_{\ell\ell'} \delta_{\lambda\lambda'}$. Even for this case no exact macroscopic solution has so far been published. We define the short-hands $\tilde{Z}_{\ell,\lambda} = f[Z(\ell,\lambda)]$ and $\kappa_n = \int Dz\, f^n[Sz]$ with the Gaussian measure $Dz = (2\pi)^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}z^2} \mathrm{d}z$ (since $f[A]$ is anti-symmetric, $\kappa_n = 0$ for all odd n).

4.1. The retarded self-interaction kernel R

To proceed we have to calculate the two functions (32), (33). They occur only in expressions (35) and (36), where causality of the response function G enforces a helpful ordering of the time arguments. The kernel $\Delta_{r+1}(\ell_0,\ldots,\ell_r)$ as occurring in (35) is easy to evaluate, since here we may use $\ell_0 > \ell_1 > \cdots > \ell_r$. This property, in combination with $\kappa_1 = \langle \tilde{Z}_{\ell,\lambda} \rangle = 0$ for all (ℓ,λ), implies that the only non-zero contributions are those where the Gaussian variables $\{\tilde{Z}\}$ have pair-wise identical indices:

$$\Delta_{r+1}(\ell_0,\ldots,\ell_r) = \frac{1}{p} \sum_{\lambda_0 \ldots \lambda_r = 1}^{p} \left\langle \tilde{Z}_{\ell_0,\lambda_0} \tilde{Z}_{\ell_r,\lambda_0} \prod_{i=1}^{r} \tilde{Z}_{\ell_{i-1},\lambda_i} \tilde{Z}_{\ell_i,\lambda_i} \right\rangle_{\{Z\}}$$

$$= \frac{1}{p}\sum_{\lambda_0\ldots\lambda_r=1}^{p}\langle\tilde{Z}_{\ell_0,\lambda_0}\tilde{Z}_{\ell_0,\lambda_1}\rangle\left(\prod_{i=1}^{r-1}\langle\tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\rangle\right)\langle\tilde{Z}_{\ell_r,\lambda_r}\tilde{Z}_{\ell_r,\lambda_0}\rangle$$

$$= \frac{1}{p}\kappa_2^{r+1}\sum_{\lambda_0\ldots\lambda_r=1}^{p}\delta_{\lambda_0,\lambda_1}\left(\prod_{i=1}^{r-1}\delta_{\lambda_i,\lambda_{i+1}}\right)\delta_{\lambda_r,\lambda_0} = \kappa_2^{r+1}. \tag{38}$$

Insertion into (35) then gives us the fully explicit form

$$R(t,t') = \kappa_2\left\{\delta(t-t') + \lim_{\delta_N\to 0}\frac{1}{\delta_N}\sum_{r>0}(-\delta_N)^r\kappa_2^r\sum_{\ell_1\ldots\ell_{r-1}}G(\ell_0,\ell_1)\ldots G(\ell_{r-1},\ell_r)\right\}$$

$$= \kappa_2\left\{\delta(t-t') + \sum_{r>0}(-\kappa_2)^r G^r(t,t')\right\}$$

$$= \kappa_2(\mathbb{1}+\kappa_2 G)^{-1}(t,t'). \tag{39}$$

This result for inner product MGs depends on the function $f[.]$ and the variance S only via a single parameter $\kappa_2 = \int Dz\, f^2[Sz]$, and reduces to the corresponding expression found earlier [9] for look-up table MGs with fake histories when $\kappa_2 = 1$.

4.2. The effective noise covariance kernel Σ

We now turn to the function (33). This is needed in (36) but only for time combinations with $\ell_0 > \ell_1 > \cdots > \ell_r$ and $\ell'_0 > \ell'_1 > \cdots > \ell'_{r'}$. Again we first group together the various terms that have identical time labels, and we also introduce two auxiliary summation indices λ that allow us to separate the terms with times of the type $\{\ell_i\}$ from those of the type $\{\ell'_j\}$ in a clean way:

$$\tilde{\Delta}_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) = \frac{1}{p}\sum_{\lambda_0\ldots\lambda_r=1}^{p}\sum_{\lambda'_1\ldots\lambda'_{r'+1}=1}^{p}\left\langle\tilde{Z}_{\ell_0,\lambda_0}\tilde{Z}_{\ell'_0,\lambda_0}\tilde{Z}_{\ell_r,\lambda'_{r'+1}}\tilde{Z}_{\ell'_{r'},\lambda'_{r'+1}}\right.$$
$$\left.\times\left[\prod_{i=1}^{r}\tilde{Z}_{\ell_{i-1},\lambda_i}\tilde{Z}_{\ell_i,\lambda_i}\right]\left[\prod_{j=1}^{r'}\tilde{Z}_{\ell'_{j-1},\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_j}\right]\right\rangle_{\{Z\}}$$
$$= \frac{1}{p}\sum_{\lambda_0\ldots\lambda_{r+1}=1}^{p}\sum_{\lambda'_0\ldots\lambda'_{r'+1}=1}^{p}\delta_{\lambda_{r+1},\lambda'_{r'+1}}\delta_{\lambda'_0,\lambda_0}\left\langle\tilde{Z}_{\ell_0,\lambda_0}\tilde{Z}_{\ell'_0,\lambda'_0}\tilde{Z}_{\ell_r,\lambda_{r+1}}\tilde{Z}_{\ell'_{r'},\lambda'_{r'+1}}\right.$$
$$\left.\times\left[\prod_{i=1}^{r}\tilde{Z}_{\ell_{i-1},\lambda_i}\tilde{Z}_{\ell_i,\lambda_i}\right]\left[\prod_{j=1}^{r'}\tilde{Z}_{\ell'_{j-1},\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_j}\right]\right\rangle_{\{Z\}}$$
$$= \frac{1}{p}\sum_{\lambda_0\ldots\lambda_{r+1}=1}^{p}\sum_{\lambda'_0\ldots\lambda'_{r'+1}=1}^{p}\delta_{\lambda_{r+1},\lambda'_{r'+1}}\delta_{\lambda'_0,\lambda_0}$$
$$\times\left\langle\left[\prod_{i=0}^{r}\tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\right]\left[\prod_{j=0}^{r'}\tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\right]\right\rangle_{\{Z\}}. \tag{40}$$

Due to the time ordering relations of the problem, there can never be time coincidences amongst members of the set $\{\ell_0,\ldots,\ell_r\}$, nor amongst members of the set $\{\ell'_0,\ldots,\ell'_{r'}\}$. Thus the various factors within the first pair of round brackets $\left(\prod_{i=0}^{r}\ldots\right)$ in (40) are all independent, and so are those within the second pair of round brackets $\left(\prod_{j=0}^{r'}\ldots\right)$. The only possible time

coincidences are *pairings* between un-primed times $\{\ell_i\}$ and primed times $\{\ell'_j\}$. Thus we have to consider the following cases:

- If there are no time coincidences, i.e. $\ell_i \neq \ell'_j$ for all (i, j), the evaluation of (40) proceeds similarly to that of the previous functions $\Delta_{r+1}(\ldots)$:

$$\tilde{\Delta}_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) = \frac{1}{p}\kappa_2^{r+r'+2}\sum_{\lambda_0\ldots\lambda_{r+1}=1}^{p}\sum_{\lambda'_0\ldots\lambda'_{r'+1}=1}^{p}\delta_{\lambda_{r+1},\lambda'_{r'+1}}\delta_{\lambda'_0,\lambda_0}\left[\prod_{i=0}^{r}\delta_{\lambda_i,\lambda_{i+1}}\right]\left[\prod_{j=0}^{r'}\delta_{\lambda'_j,\lambda'_{j+1}}\right]$$
$$= \frac{1}{p}\kappa_2^{r+r'+2}\sum_{\lambda_0=1}^{p}\sum_{\lambda'_0=1}^{p}\delta_{\lambda'_0,\lambda_0} = \kappa_2^{r+r'+2} \tag{41}$$

- Now consider the effect of a time pairing, where $\ell_i = \ell'_j$ (note: there can be multiple pairings, but the number of coinciding times is two at most, due to the built-in time ordering). The contribution of the Gaussian averages with times (ℓ_i, ℓ'_j) to the sum $\sum_{\lambda_0\ldots\lambda_{r+1}}\sum_{\lambda'_0\ldots\lambda'_{r'+1}}$ in (40), which in the absence of pairing equalled simply

$$\text{no pairing}: \quad \langle\tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\rangle = \langle\tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\rangle\langle\tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\rangle$$
$$= \kappa_2^2\delta_{\lambda_i,\lambda_{i+1}}\delta_{\lambda'_j,\lambda'_{j+1}}$$

now becomes:

$$\ell_i = \ell'_j: \quad \langle\tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\rangle = (\kappa_4 - 3\kappa_2^2)\delta_{\lambda_i,\lambda_{i+1}}\delta_{\lambda'_j,\lambda'_{j+1}}\delta_{\lambda_i,\lambda'_j}$$
$$+\kappa_2^2\left[\delta_{\lambda_i,\lambda_{i+1}}\delta_{\lambda'_j,\lambda'_{j+1}} + \delta_{\lambda_i,\lambda'_j}\delta_{\lambda_{i+1},\lambda'_{j+1}} + \delta_{\lambda_i,\lambda'_{j+1}}\delta_{\lambda_{i+1},\lambda'_j}\right]$$

The effect of time pairings can be represented diagrammatically; see figure 1.

We now make the crucial observation that in expression (36) any time pairing would inevitably generate a 'bare' factor δ_N that would no longer be absorbed into an integration via $\sum_{\ell_k}\delta_N \to \int \mathrm{d}t_k$. Hence the only contributions to $\tilde{\Delta}_{r+r'+2}$ to survive the limit $\delta_N \to 0$ in (36) are those where each time pairing also generates an extra $\mathcal{O}(p)$ factor to compensate for the emerging δ_N. Let us inspect the various diagrams, see figure 1 (and the higher order versions generated due to multiple time coincidences) and their contributions to $\tilde{\Delta}_{r+r'+2}$. Each *connected* diagram implies that ultimately in the summation $\sum_{\lambda_0\ldots\lambda_r}\sum_{\lambda'_0\ldots\lambda'_{r'}}$ we are restricted to $\lambda_0 = \lambda_1 = \ldots\lambda_r = \lambda'_0 = \cdots = \lambda'_{r'}$. This leaves a sum over λ_0 only, and a final $\mathcal{O}(1)$ contribution to $\tilde{\Delta}_{r+r'+2}$. On the other hand, in the case of a *disconnected* diagram, each internally connected sub-diagram will give an $\mathcal{O}(p)$ factor. Here the order of the final contribution to $\tilde{\Delta}_{r+r'+2}$ is p^L where L denotes the total number of time pairings that gave rise to the diagram cuts. Hence the only relevant diagrams in figure 1 are the top one (when there are no pairings) and the third from the top (when there are pairings, but where these also generate compensating $\mathcal{O}(p)$ factors), including the higher order diagrams with multiple vertical connections (in the case of multiple time pairings).

It follows, in combination with the contribution coming from the unpaired terms and from each possible time pairing as calculated earlier, and taking into account the crucial time orderings in (36), that

$$\tilde{\Delta}_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\cdots,\ell'_{r'}) = \kappa_2^{r+r'+2}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\delta_{\ell_i,\ell'_j}(p+\mathcal{O}(p^0)\right]$$
$$= \kappa_2^{r+r'+2}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{\tilde{\eta}}{2\delta_N}\delta_{\ell_i,\ell'_j}(1+\mathcal{O}(\delta_N)\right]. \tag{42}$$

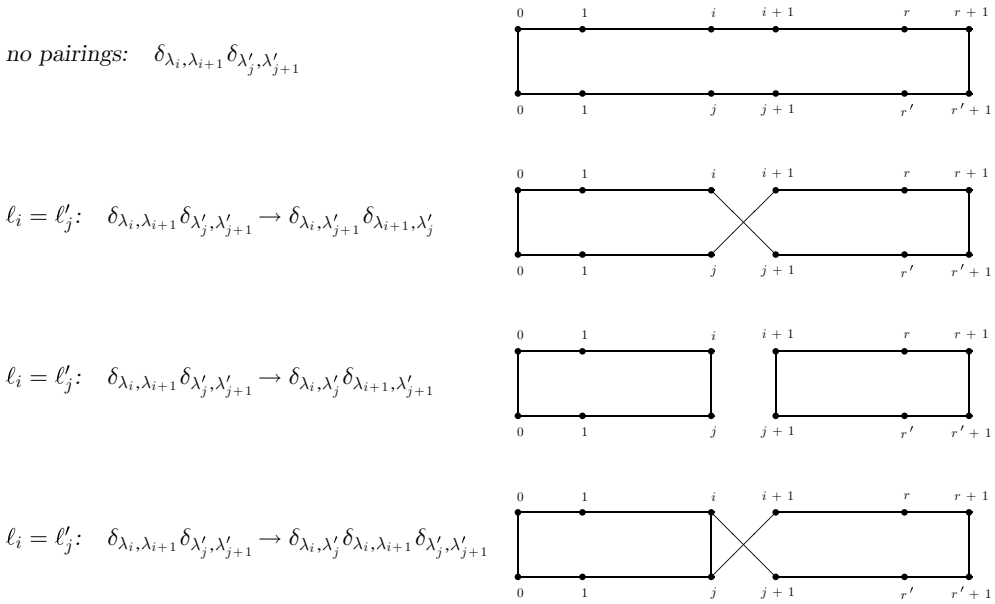

Figure 1. Diagrammatical representation of the different contributions of the Gaussian averages to the function $\tilde{\Delta}_{r+r'+2}$. Each label $i \in \{0,\ldots,r+1\}$ and each label $j \in \{0,\ldots,r'+1\}$ is drawn as a distinct vertex of a graph. Each factor $\delta_{\lambda_i,\lambda'_j}$ is drawn as a line segment connecting the vertices i and j. Top graph: the case where there are no time coincidences. Bottom three graphs: the three different new contributions that are generated by the occurrence of a time pairing where $\ell_i = \ell'_j$.

Insertion into (36) gives

$$\Sigma(t_0,t'_0) = \kappa_2^2 \sum_{r,r'\geqslant 0} (-\kappa_2)^{r+r'} \int_0^\infty \mathrm{d}t_1 \ldots \mathrm{d}t_r\, \mathrm{d}t'_1 \ldots \mathrm{d}t'_{r'} \prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{1}{2}\tilde{\eta}\delta(t_i-t'_j)\right]$$
$$\times [1+C(t_r,t'_{r'})]G(t_0,t_1)\ldots G(t_{r-1},t_r)G(t'_0,t'_1)\ldots G(t'_{r'-1},t'_{r'}). \tag{43}$$

We see again, as with the retarded self-interaction kernel R, that also this result for inner product MGs reduces to the corresponding expression found earlier [9] for look-up table MGs with fake histories when $\kappa_2 = 1$.

4.3. Summary, TTI stationary state and phase diagram

The full and closed dynamical equations for the inner product MGs with inconsistent fake market information are thus found to be given by

$$C(t,t') = \langle \sigma[q(t)]\sigma[q(t')]\rangle_\star \qquad G(t,t') = \frac{\delta}{\delta\theta(t')}\langle\sigma[q(t)]\rangle_\star. \tag{44}$$

Averages are defined with respect to the effective single agent process with a retarded self-interaction and an effective zero-average Gaussian noise $\eta(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}q(t) = \theta(t) - \alpha\kappa_2 \int_0^t \mathrm{d}t'(\mathbb{1}+\kappa_2 G)^{-1}(t,t')\sigma[q(t')] + \sqrt{\alpha}\eta(t) \tag{45}$$

$$\langle\eta(t_0)\eta(t'_0)\rangle = \kappa_2^2 \sum_{r,r'\geqslant 0} (-\kappa_2)^{r+r'} \int_0^\infty \mathrm{d}t_1 \ldots \mathrm{d}t_r\, \mathrm{d}t'_1 \ldots \mathrm{d}t'_{r'} \prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{1}{2}\tilde{\eta}\delta(t_i-t'_j)\right]$$
$$\times [1+C(t_r,t'_{r'})]G(t_0,t_1)\ldots G(t_{r-1},t_r)G(t'_0,t'_1)\ldots G(t'_{r'-1},t'_{r'}). \tag{46}$$

We notice that the role of the remaining parameter κ_2 is to define a characteristic time scale for the process, and that apart from this our equations reduce in the limit of infinitesimal perturbation fields exactly to the theory of the fake history look-up table MGs as derived in [9]. To see this we re-define $G(t,t') = \kappa_2^{-1}\hat{G}(t,t'), \theta(t) = \kappa_2\theta\hat{(t)}$ and $\eta(t) = \kappa_2\eta\hat{(t)}$, and find

$$C(t,t') = \langle\sigma[q(t)]\sigma[q(t')]\rangle_\star \qquad \hat{G}(t,t') = \frac{\delta}{\delta\hat{\theta}(t')}\langle\sigma[q(t)]\rangle_\star \tag{47}$$

$$\frac{1}{\kappa_2}\frac{\mathrm{d}}{\mathrm{d}t}q(t) = \hat{\theta}(t) - \alpha\int_0^t \mathrm{d}t'(\mathbb{1}+\hat{G})^{-1}(t,t')\sigma[q(t')] + \sqrt{\alpha}\hat{\eta}(t) \tag{48}$$

$$\begin{aligned}\langle\hat{\eta}(t_0)\hat{\eta}(t_0')\rangle = \sum_{r,r'\geqslant 0}(-1)^{r+r'}\int_0^\infty \mathrm{d}t_1\ldots\mathrm{d}t_r\,\mathrm{d}t_1'\ldots\mathrm{d}t_{r'}'\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{1}{2}\tilde{\eta}\delta(t_i-t_j')\right]\\ \times[1+C(t_r,t_{r'}')]\hat{G}(t_0,t_1)\ldots\hat{G}(t_{r-1},t_r)\hat{G}(t_0',t_1')\ldots\hat{G}(t_{r'-1}',t_{r'}').\end{aligned} \tag{49}$$

It will no longer come as a surprise that upon using (39), (43) to calculate in time-translation invariant stationary states the persistent order parameters $\chi_{\mathrm{R}} = \int \mathrm{d}t\,R(t)$ and $S_0^2 = \Sigma(\infty)$, we find that the only difference between the stationary state order parameter equations of look-up table MGs and the present inner product MGs is a re-scaling of the static susceptibility χ:

$$u = \sigma[\infty]\sqrt{\alpha/2(1+c)} \tag{50}$$

$$\phi = 1 - \mathrm{Erf}[u] \tag{51}$$

$$c = \sigma^2[\infty]\left\{1 - \mathrm{Erf}[u] + \frac{1}{2u^2}\,\mathrm{Erf}[u] - \frac{1}{u\sqrt{\pi}}\,\mathrm{e}^{-u^2}\right\} \tag{52}$$

$$\chi = \mathrm{Erf}[u]/\kappa_2(\alpha - \mathrm{Erf}[u]). \tag{53}$$

Since phase transitions in the present type of MGs are defined by a divergence of χ, we conclude that not only the values of the static observables $\{\phi, c\}$ but also the phase diagrams of the two fake history MG model families are identical.

4.4. The volatility

Finally we calculate the volatility. Since the average overall bid is zero, the volatility is here defined by $\sigma^2 = \lim_{t\to\infty}\lim_{\delta_N\to 0}\langle\langle A^2(\ell)\rangle\rangle_{\{A,Z\}}|_{\ell=t/\delta_N}$. Inserting (34) into the definition of σ, with $A_e \to 0$, gives

$$\begin{aligned}\sigma^2 = \lim_{t\to\infty}\lim_{\delta_N\to 0}\Bigg\{&\sum_{r,r'\geqslant 0}(-\delta_N)^{r+r'}\sum_{\ell_0\ldots\ell_r}G(\ell_0,\ell_1)\ldots G(\ell_{r-1},\ell_r)\\ &\times\sum_{\ell_0'\ldots\ell_{r'}'}G(\ell_0',\ell_1')\ldots G(\ell_{r'-1}',\ell_{r'}')\delta_{\ell,\ell_0}\delta_{\ell,\ell_0'}p^{r+r'}\Bigg\langle\Bigg\langle\langle\phi_{\ell_r}\phi_{\ell_{r'}'}\rangle_{\{\phi|A,Z\}}\\ &\times\left[\prod_{i=1}^{r}\overline{W}[\ell_{i-1},\ell_i;\{A,Z\}]\right]\left[\prod_{j=1}^{r'}\overline{W}[\ell_{j-1}',\ell_j';\{A,Z\}]\right]\Bigg\rangle\Bigg\rangle_{\{A,Z\}}\Bigg\}\Bigg|_{\ell=\frac{t}{\delta_N}}\\ = \lim_{t\to\infty}\lim_{\delta_N\to 0}\Bigg\{&\frac{1}{2}\sum_{r,r'\geqslant 0}(-\delta_N)^{r+r'}\sum_{\ell_0\ldots\ell_r}G(\ell_0,\ell_1)\ldots G(\ell_{r-1},\ell_r)\end{aligned}$$

$$\times \sum_{\ell'_0 \ldots \ell'_{r'}} G(\ell'_0, \ell'_1) \ldots G(\ell'_{r'-1}, \ell'_{r'})[1 + C(\ell_r, \ell'_{r'})]\delta_{\ell,\ell_0}\delta_{\ell,\ell'_0}$$

$$\times \tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0, \ldots, \ell_r; \ell'_0, \ldots, \ell'_{r'})\Bigg\}\Bigg|_{\ell=t/\delta_N} \tag{54}$$

with

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0, \ldots, \ell_r; \ell'_0, \ldots, \ell'_{r'})$$
$$= p^{r+r'}\left\langle\left\langle \overline{W}[\ell_r, \ell'_{r'}; \{A, Z\}]\left[\prod_{i=1}^{r} \overline{W}[\ell_{i-1}, \ell_i; \{A, Z\}]\right]\left[\prod_{j=1}^{r'} \overline{W}[\ell'_{j-1}, \ell'_j; \{A, Z\}]\right]\right\rangle\right\rangle_{\{A,Z\}}. \tag{55}$$

We need this function (55) only for times with $\ell_0 > \ell_1 > \cdots > \ell_r$ and $\ell'_0 > \ell'_1 > \cdots > \ell'_{r'}$. For the case of inner product MG, where $\overline{W}[\ell, \ell'; \{A, Z\}] = p^{-1}\sum_{\lambda=1}^{p} \tilde{Z}_{\ell,\lambda}\tilde{Z}_{\ell',\lambda}$, its evaluation is very similar to that of (33):

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0, \cdots, \ell_r; \ell'_0, \ldots, \ell'_{r'}) = \frac{1}{p}\sum_{\lambda_1 \ldots \lambda_{r+1}=1}^{p}\sum_{\lambda'_1 \ldots \lambda'_{r'+1}=1}^{p} \delta_{\lambda_{r+1},\lambda'_{r'+1}}$$
$$\times\left\langle \tilde{Z}_{\ell_0,\lambda_1}\tilde{Z}_{\ell'_0,\lambda'_1}\left[\prod_{i=1}^{r} \tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\right]\left[\prod_{j=1}^{r'} \tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\right]\right\rangle_{\{Z\}} \tag{56}$$

Due to the time ordering, the factor $\tilde{Z}_{\ell_0,\lambda_1}\tilde{Z}_{\ell'_0,\lambda'_1}$ is statistically independent of all others, so that

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0, \ldots, \ell_r; \ell'_0, \ldots, \ell'_{r'}) = \frac{\kappa_2}{p}\sum_{\lambda_1 \ldots \lambda_{r+1}=1}^{p}\sum_{\lambda'_1 \ldots \lambda'_{r'+1}=1}^{p} \delta_{\lambda_{r+1},\lambda'_{r'+1}}$$
$$\times\left\langle\left[\prod_{i=1}^{r} \tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\right]\left[\prod_{j=1}^{r'} \tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\right]\right\rangle_{\{Z\}}. \tag{57}$$

Again we inspect the effect of time pairings:

- If there are no time coincidences, i.e. $\ell_i \neq \ell'_j$ for all (i, j), we simply find

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0, \ldots, \ell_r; \ell'_0, \ldots, \ell'_{r'})$$
$$= \frac{1}{p}\kappa_2^{r+r'+1}\sum_{\lambda_1 \ldots \lambda_{r+1}=1}^{p}\sum_{\lambda'_1 \ldots \lambda'_{r'+1}=1}^{p} \delta_{\lambda_{r+1},\lambda'_{r'+1}}\left[\prod_{i=1}^{r} \delta_{\lambda_i,\lambda_{i+1}}\right]\left[\prod_{j=1}^{r'} \delta_{\lambda'_j,\lambda'_{j+1}}\right]$$
$$= \frac{1}{p}\kappa_2^{r+r'+1}\sum_{\lambda_1=1}^{p}\sum_{\lambda'_1=1}^{p} \delta_{\lambda'_1,\lambda_1} = \kappa_2^{r+r'+1} \tag{58}$$

- The effect of a time pairing $\ell_i = \ell'_j$ on the contribution of the Gaussian averages with times (ℓ_i, ℓ'_j) is exactly as before:

$$\text{no pairing}: \quad \left\langle \tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\right\rangle = \left\langle \tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\right\rangle\left\langle \tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\right\rangle$$
$$= \kappa_2^2\delta_{\lambda_i,\lambda_{i+1}}\delta_{\lambda'_j,\lambda'_{j+1}}$$
$$\ell_i = \ell'_j: \quad \left\langle \tilde{Z}_{\ell_i,\lambda_i}\tilde{Z}_{\ell_i,\lambda_{i+1}}\tilde{Z}_{\ell'_j,\lambda'_j}\tilde{Z}_{\ell'_j,\lambda'_{j+1}}\right\rangle = \left(\kappa_4 - 3\kappa_2^2\right)\delta_{\lambda_i,\lambda_{i+1}}\delta_{\lambda'_j,\lambda'_{j+1}}\delta_{\lambda_i,\lambda'_j}$$
$$+\kappa_2^2\left(\delta_{\lambda_i,\lambda_{i+1}}\delta_{\lambda'_j,\lambda'_{j+1}} + \delta_{\lambda_i,\lambda'_j}\delta_{\lambda_{i+1},\lambda'_{j+1}} + \delta_{\lambda_i,\lambda'_{j+1}}\delta_{\lambda_{i+1},\lambda'_j}\right)$$

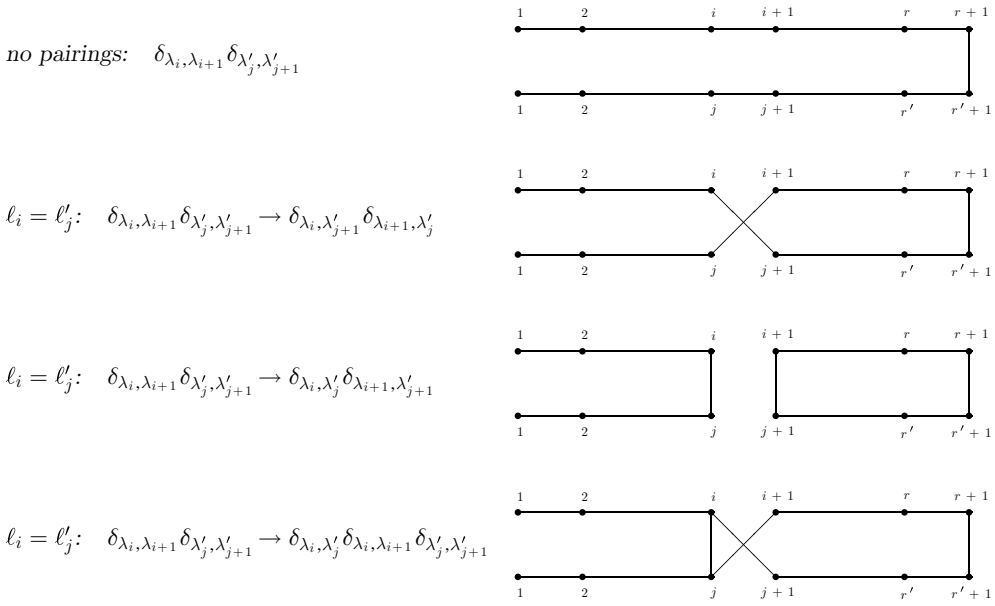

Figure 2. Diagrammatical representation of the different contributions of the Gaussian averages to the function $\tilde{\tilde{\Delta}}_{r+r'+1}$. Each label $i \in \{1,\ldots,r+1\}$ and each label $j \in \{1,\ldots,r'+1\}$ is drawn as a distinct vertex of a graph. Each factor $\delta_{\lambda_i,\lambda'_j}$ is drawn as a line segment connecting the vertices i and j. Top graph: the case where there are no time coincidences. Bottom three graphs: the three different new contributions that are generated by the occurrence of a time pairing where $\ell_i = \ell'_j$.

The corresponding diagrammatic representation is shown in figure 2. Since each time pairing will have to be compensated by an $\mathcal{O}(p)$ factor to retain relevance in the limit $\delta_N \to 0$, the only relevant diagrams in figure 2 are again the top one (no pairings) and the third from the top (a time pairing with vertical connections that causes a diagram cut), including the higher order diagrams with multiple vertical connections (in the case of multiple time pairings). We conclude that

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) = \kappa_2^{r+r'+1}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\delta_{\ell_i,\ell'_j}(p+\mathcal{O}(p^0)\right]$$
$$= \kappa_2^{r+r'+1}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{\tilde{\eta}}{2\delta_N}\delta_{\ell_i,\ell'_j}(1+\mathcal{O}(\delta_N)\right] \tag{59}$$

and hence

$$\sigma^2 = \frac{1}{2}\kappa_2\lim_{t\to\infty}\sum_{r,r'\geqslant 0}(-\kappa_2)^{r+r'}\int_0^\infty \mathrm{d}t_1\ldots\mathrm{d}t_r\,\mathrm{d}t'_1\ldots\mathrm{d}t'_{r'}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{1}{2}\tilde{\eta}\delta(t_i-t'_j)\right]$$
$$\times[1+C(t_r,t'_{r'})]G(t,t_1)G(t_1,t_2)\ldots G(t_{r-1},t_r)G(t,t'_1)G(t'_1,t'_2)\ldots G(t'_{r'-1},t'_{r'})$$
$$= \lim_{\tau\to\infty}\frac{\kappa_2}{2\tau}\int_0^\tau \mathrm{d}t\sum_{r,r'\geqslant 0}(-\kappa_2)^{r+r'}\int_0^\infty \mathrm{d}t_1\ldots\mathrm{d}t_r\,\mathrm{d}t'_1\ldots\mathrm{d}t'_{r'}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{1}{2}\tilde{\eta}\delta(t_i-t'_j)\right]$$
$$\times G(t,t_1)G(t_1,t_2)\ldots G(t_{r-1},t_r)G(t,t'_1)G(t'_1,t'_2)\ldots G(t'_{r'-1},t'_{r'})$$
$$\times[1+C(t_r,t'_{r'})]. \tag{60}$$

Comparison with the corresponding expression in [9] for look-up table MGs shows that the only difference is in the various occurrences of κ_2. If we implement the transformation that maps the effective single agent exactly onto that of the look-up-table models, viz. $G(t,t') = \kappa_2^{-1}\hat{G}(t,t')$, we find that the inner-product volatility differs from that of the look-up table models by a factor κ_2:

$$\sigma^2 = \lim_{\tau\to\infty}\frac{\kappa_2}{2\tau}\int_0^\tau \mathrm{d}t \sum_{r,r'\geqslant 0}(-1)^{r+r'}\int_0^\infty \mathrm{d}t_1\ldots\mathrm{d}t_r\,\mathrm{d}t'_1\ldots\mathrm{d}t'_{r'}\prod_{i=1}^{r}\prod_{j=1}^{r'}\left[1+\frac{1}{2}\tilde{\eta}\delta(t_i-t'_j)\right]$$
$$\times\hat{G}(t,t_1)\hat{G}(t_1,t_2)\ldots\hat{G}(t_{r-1},t_r)\hat{G}(t,t'_1)\hat{G}(t'_1,t'_2)\ldots\hat{G}(t'_{r'-1},t'_{r'})$$
$$\times[1+C(t_r,t'_{r'})]. \tag{61}$$

5. Inner product MGs with real market history

The previous cases could be solved exactly and in full, due to the absence of real history. We now turn to the more demanding situation where $\zeta < 1$ in definition (26), so that the information vector truly depends on the past market history. Here it will turn out advantageous to define the stochastic and time dependent symmetric $p\times p$ matrices $\mathbf{B}(\ell)$ with entries

$$B_{\lambda\lambda'}(\ell) = \mathcal{F}_\lambda[\ell,A,Z]\mathcal{F}_{\lambda'}[\ell,A,Z] \tag{62}$$

with $\mathcal{F}_\lambda[\ldots]$ as defined in (3). This allows us to write the relevant functions $\Delta_{r+1}(\ldots)$, $\tilde{\Delta}_{r+r'+2}(\ldots)$ and $\tilde{\tilde{\Delta}}_{r+r'+1}(\ldots)$ (that occur in the kernels R and Σ and in the volatility) as averages over a trace:

$$\Delta_{r+1}(\ell_0,\ldots,\ell_r) = \frac{1}{p}\mathrm{Tr}\langle\langle\mathbf{B}(\ell_0)\mathbf{B}(\ell_1)\ldots\mathbf{B}(\ell_r)\rangle\rangle_{\{A,Z\}} \tag{63}$$

$$\tilde{\Delta}_{r+r'+2}(\ell_0,\ldots,\ell_r;\ell'_0,\ldots,\ell'_{r'}) = \frac{1}{p}\mathrm{Tr}\langle\langle\mathbf{B}(\ell_0)\mathbf{B}(\ell_1)\ldots\mathbf{B}(\ell_r)\mathbf{B}(\ell'_{r'})\ldots\mathbf{B}(\ell'_1)\mathbf{B}(\ell'_0)\rangle\rangle_{\{A,Z\}} \tag{64}$$

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0,\ell_1\ldots,\ell_r;\ell_0,\ell'_1,\ldots,\ell'_{r'}) = \frac{1}{p}\mathrm{Tr}\langle\langle\mathbf{B}(\ell_0)\mathbf{B}(\ell_1)\ldots\mathbf{B}(\ell_r)\mathbf{B}(\ell'_{r'})\ldots\mathbf{B}(\ell'_1)\rangle\rangle_{\{A,Z\}} \tag{65}$$

These expressions are of course still exact, but not easy to evaluate.

5.1. Short history correlation times in TTI stationary states

We now make an approximation for the TTI stationary state solution in the spirit of the short history correlation times ansatz made in [11] for look-up table MGs; in fact we will find below that for look-up table MGs the two are identical. In working out our expressions for $\chi_R = \int \mathrm{d}t\,R(t)$, $\Sigma(\infty) = \lim_{\tau\to\infty}\Sigma(t+\tau,t)$ and the volatility σ^2 we will replace the kernels (63), (64), (65) by the values they will take for times which are sufficiently widely separated to de-correlate the random matrices $\mathbf{B}(\ell)$, i.e. we replace

$$\prod_{i=1}^{r}\mathbf{B}(\ell_i)\to\mathbf{B}^r(A,Z) \qquad \mathbf{B}(A,Z) = \lim_{L\to\infty}\frac{1}{L}\sum_{\ell\leqslant L}\mathbf{B}(\ell). \tag{66}$$

This is equivalent to assuming ergodicity in the space of the bid history strings on time-scales of $\mathcal{O}(N)$ discrete iteration steps of the MG. For TTI states it allows us to express everything in

terms of the average eigenvalue distribution $\varrho(\mu)$ of the now time independent random $p \times p$ matrix $\mathbf{B}(A, Z)$:

$$\Delta_{r+1}(\ell_0, \ldots, \ell_r) = \frac{1}{p} \mathrm{Tr}\langle\langle \mathbf{B}^{r+1}(A, Z)\rangle\rangle_{\{A,Z\}} = \int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\mu^{r+1}. \tag{67}$$

$$\tilde{\Delta}_{r+r'+2}(\ell_0, \ldots, \ell_r; \ell'_0, \ldots, \ell'_{r'}) = \frac{1}{p} \mathrm{Tr}\langle\langle \mathbf{B}^{r+r'+2}(A, Z)\rangle\rangle_{\{A,Z\}} = \int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\mu^{r+r'+2} \tag{68}$$

$$\tilde{\tilde{\Delta}}_{r+r'+1}(\ell_0, \ell_1 \ldots, \ell_r; \ell_0, \ell'_1, \ldots, \ell'_{r'}) = \frac{1}{p} \mathrm{Tr}\langle\langle \mathbf{B}^{r+r'+1}(A, Z)\rangle\rangle_{\{A,Z\}} = \int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\mu^{r+r'+1}. \tag{69}$$

Since $\mathbf{B}(A, Z)$ is non-negative definite one always has $\mu \geqslant 0$. Insertion of (67), (68), (69) into expressions (35), (36) for the kernels R and Σ followed by appropriate integration and re-summation of the series leads us to the following expressions for χ_R and $S_0^2 = \Sigma(\infty)$:

$$\chi_R = \int_0^\infty \mathrm{d}\mu\, \varrho(\mu) \frac{\mu}{1 + \mu\chi} S_0^2 = (1 + c) \int_0^\infty \mathrm{d}\mu\, \varrho(\mu) \frac{\mu^2}{(1 + \mu\chi)^2}. \tag{70}$$

The effects of having real memory are concentrated fully in the eigenvalue distribution $\varrho(\mu)$. Once this distribution has been calculated in terms of $\{c, \phi, \chi\}$, the persistent order parameters of the MG will be given by a closed set of equations. Upon introducing a convenient parameter $\omega \in [0, 1]$ via

$$\omega = \frac{\int \mathrm{d}\mu\, \varrho(\mu)\mu\chi/(1 + \mu\chi)}{\sqrt{\int \mathrm{d}\mu\, \varrho(\mu)[\mu\chi/(1 + \mu\chi)]^2}}. \tag{71}$$

our closed order parameter equations take the form

$$u = \frac{\omega\sigma[\infty]\sqrt{\alpha}}{\sqrt{2(1 + c)}}, \qquad \frac{1 - \phi}{\alpha} = \int_0^\infty \frac{\mathrm{d}\mu\, \varrho(\mu)}{1 + (\mu\chi)^{-1}}, \qquad \phi = 1 - \mathrm{Erf}[u] \tag{72}$$

$$c = \sigma^2[\infty]\left\{1 - \mathrm{Erf}[u] + \frac{1}{2u^2}\mathrm{Erf}[u] - \frac{1}{u\sqrt{\pi}}\mathrm{e}^{-u^2}\right\}. \tag{73}$$

From this we can already extract several results, and explain observations made in the past for MG types other than look-up table ones on the basis of numerical simulations:

- Irrespective of the extent to which the histories are real, and irrespective of the nature of the function $f[A]$, the ergodicity-breaking phase transition point marked by a divergence of χ occurs *always* at the value $\alpha_c(T)$ that was found originally for fake history look-up table MG models [5–8], as solved from the closed set

$$u = \sigma[\infty]\sqrt{\alpha/2(1 + c)}, \qquad \mathrm{Erf}[u] = \alpha, \tag{74}$$

$$c = \sigma^2[\infty]\left\{1 - \mathrm{Erf}[u] + \frac{1}{2u^2}\mathrm{Erf}[u] - \frac{1}{u\sqrt{\pi}}\mathrm{e}^{-u^2}\right\} \tag{75}$$

- Equations (71)–(73) return to those derived earlier for inconsistent random histories when $\omega = 1$, whereas real histories affect the persistent observables as soon as $\omega < 1$. For finite χ we thus arrive at the conclusion, via (71), that our observables are affected by having real histories if and only if $\varrho(\mu)$ is *not* a δ-distribution. A nonzero width of $\varrho(\mu)$ is the fingerprint of real histories influencing the behaviour of the MG.

We note that the combined equations (70)–(73) are identical to those found for the look-up table MGs with true market histories [11], with the eigenvalue distribution $\varrho(\mu)$ taking over the role of the relative history frequency distribution in [11]. This can be understood. Our choice of notation allows us to apply the present derivation also to the look-up table models, simply by making in (62) the substitutions $\lambda \in \{1, \ldots, p\} \to \boldsymbol{\lambda} \in \{-1, 1\}^M$ and $\mathcal{F}_\lambda[\ell, A, Z] \to \sqrt{p}\delta_{\lambda,\lambda(\ell,A,Z)}$, where $\boldsymbol{\lambda}(\ell, A, Z)$ is the history string (potentially partly fake) as observed by the agents at step ℓ. In look-up table MGs one would therefore have $B_{\lambda\lambda'}(\ell) = p\delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}(\ell,A,Z)}\delta_{\boldsymbol{\lambda}',\boldsymbol{\lambda}(\ell,A,Z)}$ and hence

$$B_{\lambda\lambda'}(A, Z) = p\delta_{\lambda\lambda'} \lim_{L\to\infty} \frac{1}{L} \sum_{\ell\leqslant L} \delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}(\ell,A,Z)} = p\delta_{\lambda\lambda'}\pi_{\boldsymbol{\lambda}}(A, Z) \tag{76}$$

Here $\pi_{\boldsymbol{\lambda}}(A, Z)$ denotes the stationary state frequency of occurrence of history string $\boldsymbol{\lambda}$. Since this matrix $\mathbf{B}(A, Z)$ is of a diagonal form, its eigenvalues are simply the diagonal entries $p\pi_{\boldsymbol{\lambda}}(A, Z)$, and the asymptotic eigenvalue distribution becomes

$$\varrho(\mu) = \lim_{p\to\infty} \frac{1}{p} \sum_{\boldsymbol{\lambda}} \langle\langle \delta\left[\mu - p\pi_{\boldsymbol{\lambda}}(A, Z)\right]\rangle\rangle_{\{A,Z\}} \tag{77}$$

Thus, for look-up table MGs the eigenvalue distribution $\varrho(\mu)$ is indeed identical to the relative history frequency distribution of [11].

The question of whether and to what extent having real as opposed to fake market histories affects the MG thus boils down to assessing the shape of the eigenvalue distribution $\varrho(\mu)$. For look-up table MGs this distribution is known to become nontrivial above the critical point [10, 11]. Numerical simulations show that the same is true for inner product MGs, so also here models with real history will behave differently from those with fake histories.

5.2. *Calculation of eigenvalue spectrum $\varrho(\mu)$ for weakly correlated bids*

In the remainder of this paper we will deal exclusively with the inner product version of the MG with real history and inconsistent bid noise, i.e. from now on we will have

$$B_{\lambda\lambda'}(A, Z) = \frac{1}{L} \sum_{\ell\leqslant L} f[(1-\zeta)A(\ell-\lambda) + \zeta Z(\ell, \lambda)] f[(1-\zeta)A(\ell-\lambda') + \zeta Z(\ell, \lambda')]. \tag{78}$$

where $\langle Z(\ell, \lambda)Z(\ell', \lambda')\rangle = S^2\delta_{\ell\ell'}\delta_{\lambda\lambda'}$, and where we focus on the regime where $1 \ll p \ll L$. We try to understand the origin of the nontrivial eigenvalue distribution $\varrho(\mu)$ in inner product MGs and its impact on the order parameter equations. Numerical simulations show that, even when the histories are real, the bid correlations are still very weak, see e.g. figure 3; this we will exploit in our analysis. The remaining complications arise from the fact that, although the matrix $\mathbf{B}(A, Z)$ is an average over a diverging number of at most weakly correlated variables, the number of entries of $\mathbf{B}(A, Z)$ also diverges. To prepare the stage we first write the time-translation invariant bid covariances as $\Xi(\lambda) = \langle A(\ell)A(\ell+\lambda)\rangle_{\{A,Z\}}$, and introduce the short-hands

$$Q_0 = \int Dz\, f^2[z\sqrt{(1-\zeta)^2\sigma^2 + \zeta^2 S^2}] \tag{79}$$

$$Q_1 = \int Dy \left[\int Dz\, f[(1-\zeta)\sigma y + \zeta S z]\right]^2. \tag{80}$$

with $\sigma^2 = \lim_{\ell\to\infty} \langle\langle A^2(\ell)\rangle\rangle_{\{A,Z\}} = \Xi(0)$, and $0 \leqslant Q_1 \leqslant Q_0$. As for the look-up table MGs, one expects also for inner product MGs the distribution $\varrho(\mu)$ to be well-defined as $p \to \infty$, since $\int \mathrm{d}\mu\, \varrho(\mu)\mu = p^{-1} \sum_{\lambda\leqslant p} \langle\langle B_{\lambda\lambda}(A, Z)\rangle\rangle_{A,Z} = Q_0$.

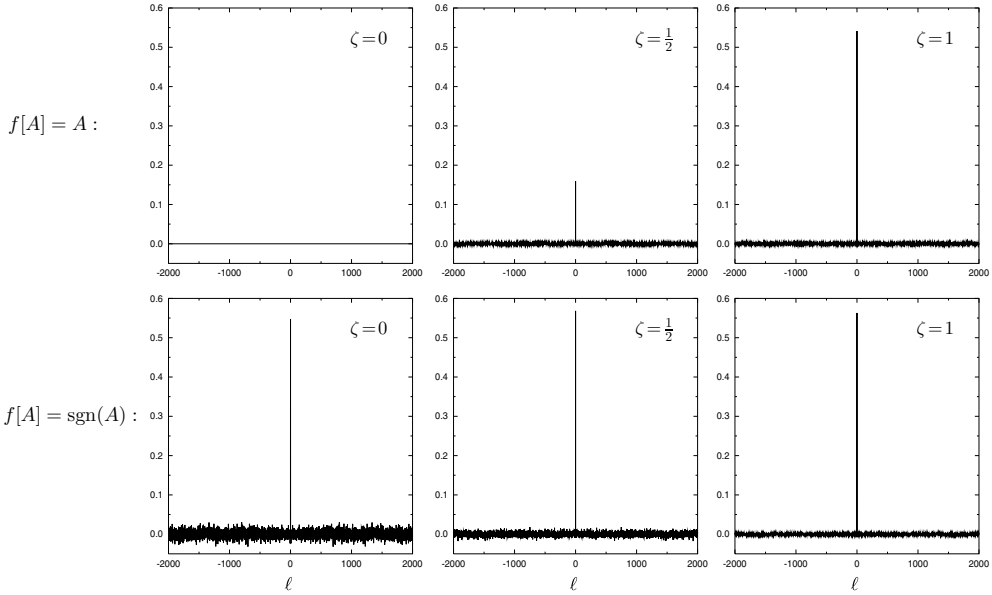

Figure 3. Time-translation invariant bid covariances $\Xi(\ell) = \langle A(t+\ell)A(t)\rangle$ of inner product MGs (78) measured in numerical simulations, with $N = 1025$ and $\alpha = 2$. Covariances are calculated over $200.N$ time steps, following equilibration. Top row: $f[A] = A$, with $\zeta \in \{0, \frac{1}{2}, 1\}$; bottom row: $f[A] = \mathrm{sgn}(A)$, with $\zeta \in \{0, \frac{1}{2}, 1\}$. The left pictures refer to strictly real histories, the right pictures to strictly fake histories, with the middle ones representing a mixing of the two. These data are representative of the typical behaviour throughout the ergodic regime. The bid correlations are *always* seen to be weak, which justifies our approximate calculation of the spectrum $\varrho(\mu)$. We also observe qualitative differences between $f[A] = A$ and $f[A] = \mathrm{sgn}(A)$, e.g. vanishing volatility as $\zeta \to 0$ (strictly real histories) for $f[A] = A$, which will be addressed later.

In the absence of bid correlations, as for fully fake histories, $\mathbf{B}(A, Z)$ will be self-averaging for $L \to \infty$ and fixed p; this implies (due to the time-translation invariance of the bid correlations) that $\mathbf{B}$ will be of the Toeplitz form. One expects time translation invariance to hold also if the correlations between bids at different times are sufficiently small, and also $\mathbf{B}(A, Z)$ to remain self-averaging at least with respect to the inconsistent noise variables, although possibly not with respect to the bids. Here we explore the consequences for the spectrum $\varrho(\mu)$ of assuming the bid correlations $\Xi(\lambda)$ for $\lambda \neq 0$ to be small, and of taking $\mathbf{B}(A, Z)$ to be Toeplitz and self-averaging with respect to the $\{Z(\ell, \lambda)\}$, i.e. $B_{\lambda\lambda'}(A, Z) = B(\lambda - \lambda'|A)$. If $\Xi(\lambda)$ decays to zero or to small random values sufficiently fast, we may continue $\mathbf{B}(A)L$-periodically, so that it becomes a circular Toeplitz matrix. Its L eigenvalues $\mu_r (r = 0, \ldots, L-1)$ will now be given by

$$\mu_r = \sum_{\ell'=0}^{L-1} B(\lambda|A)\, \mathrm{e}^{-2\pi\, \mathrm{i}\ell' r/L} = Q_0 + S_r(A) \tag{81}$$

with Q_0 as given in (79) and, upon replacing $\sum_{\ell \leqslant L}$ by $\sum_{\ell < L}$ in (78) (which is allowed since $L \to \infty$):

$$S_r(A) = \frac{1}{L} \sum_{\ell\ell'=1}^{L-1} \mathrm{e}^{-2\pi\, \mathrm{i}\ell' r/L} \left[\int Dz\, f[(1-\zeta)A(\ell - \ell') + \zeta Sz] \right] \left[\int Dz\, f[(1-\zeta)A(\ell) + \zeta Sz] \right]$$

$$= \frac{1}{\sqrt{L}} \sum_{\ell=1}^{L-1} \int Dz\, f[(1-\zeta)A(\ell) + \zeta Sz]\, \mathrm{e}^{-2\pi\,\mathrm{i}\ell r/L}$$
$$\times \frac{1}{\sqrt{L}} \sum_{\ell'=\ell-L+1}^{\ell-1} \int Dz\, f[(1-\zeta)A(\ell') + \zeta Sz]\, \mathrm{e}^{2\pi\,\mathrm{i}\ell' r/L}. \tag{82}$$

The first and second line are both sums over a large number of nearly independent random variables. If the correlations between the bids at different times are sufficiently small, we may take the second sum to be independent of ℓ, which allows us to put $\ell = L$ in this term, provided we correct for the inappropriate generation in doing so of terms where the bids involve identical times (i.e. where $\ell = \ell'$, as such terms were absent initially):

$$S_r(A) = |z_r(A)|^2 - Q_1 \tag{83}$$

$$z_r(A) = \frac{1}{\sqrt{L}} \sum_{\ell=1}^{L-1} \int Dz\, f[(1-\zeta)A(\ell) + \zeta Sz]\, \mathrm{e}^{2\pi\,\mathrm{i}\ell r/L}. \tag{84}$$

Clearly $z_{L-r} = \overline{z}_r$. Given our assumption of weakly correlated bids and given that $f[A]$ is by definition anti-symmetric, the L complex random variables z_r (where $r = 0 \ldots L-1$) will for $L \to \infty$ acquire zero-average Gaussian statistics (central limit theorem), with

$$\langle z_r(A) z_{r'}(A)\rangle = \frac{Q_1}{L-1} \sum_{\ell=1}^{L-1} \mathrm{e}^{2\pi\,\mathrm{i}\ell(r+r')/L} = \frac{Q_1 L}{L-1}\delta_{r+r',\mathrm{mod}L} - \frac{Q_1}{L-1} \tag{85}$$

$$\langle \overline{z}_r(A) \overline{z}_{r'}(A)\rangle = \frac{Q_1}{L-1} \sum_{\ell=1}^{L-1} \mathrm{e}^{-2\pi\,\mathrm{i}\ell(r+r')/L} = \frac{Q_1 L}{L-1}\delta_{r+r',\mathrm{mod}L} - \frac{Q_1}{L-1} \tag{86}$$

$$\langle z_r(A) \overline{z}_{r'}(A)\rangle = \frac{Q_1}{L-1} \sum_{\ell=1}^{L-1} \mathrm{e}^{2\pi\,\mathrm{i}\ell(r-r')/L} = \left(Q_1 + \frac{Q_1}{L-1}\right)\delta_{r-r',\mathrm{mod}L} - \frac{Q_1}{L-1} \tag{87}$$

If the spectrum is well-defined as $L \to \infty$, we may take L to be odd and write the spectrum as an average over $\{z_1, \ldots, z_{L/2-1/2}\}$, using $z_{L-r} = \overline{z}_r$ to deal with the remaining $\{z_{L/2-1/2}, \ldots, z_{L-1}\}$. We may disregard $r = 0$, since it gives only a vanishing contribution to the spectrum. The remaining $\frac{L-1}{2}$ complex Gaussian variables are written in terms of their real and imaginary parts as $z_r = x_r + \mathrm{i}y_r$, for which one finds

$$\langle x_r x_{r'}\rangle = \frac{Q_1 L}{2(L-1)}\delta_{rr'} - \frac{Q_1}{L-1}, \qquad \langle y_r y_{r'}\rangle = \frac{Q_1 L}{2(L-1)}\delta_{rr'},\ \langle x_r y_{r'}\rangle = 0 \tag{88}$$

If for $L \to \infty$ we are allowed to neglect the $\mathcal{O}(1/L)$ correlations between the Gaussian x_r, we may use the law of large numbers and write the spectrum $\varrho(\mu)$ as an average over their distribution, giving (with $|z_r|^2 = x_r^2 + y_r^2$):

$$\varrho(\mu) = \int Du\, Dv\, \delta\left[\mu - Q_0 + Q_1 - \frac{1}{2}Q_1(u^2+v^2)\right]$$
$$= Q_1^{-1}\theta[\mu - Q_0 + Q_1]\, \mathrm{e}^{-(\mu - Q_0 + Q_1)/Q_1}. \tag{89}$$

It is encouraging to see that this simple expression passes the two tests at our disposal. Firstly, it obeys the exact equation $\int \mathrm{d}\mu\, \mu\varrho(\mu) = Q_0$. Second, since $\lim_{\zeta\to 1} Q_1 = 0$ one always has $\lim_{\zeta\to 1}\varrho(\mu) = \delta[\mu - Q_0]$, so in the fake history limit it leads us correctly to the standard stationary state equations of the fake history MGs. Let us inspect what our general results imply

for the two canonical choices for $f[A]$. For $f[A] = A$ one obtains $Q_0 = (1-\zeta)^2\sigma^2 + \zeta^2 S^2$ and $Q_1 = (1-\zeta)^2\sigma^2$, so that

$$\varrho(\mu) = \frac{\theta[\mu - \zeta^2 S^2]}{(1-\zeta)^2\sigma^2}\, \mathrm{e}^{-(\mu-\zeta^2 S^2)/(1-\zeta)^2\sigma^2}. \tag{90}$$

Fully real history would give $\lim_{\zeta\to 0}\varrho(\mu) = \theta[\mu]\sigma^{-2}\,\mathrm{e}^{-\mu/\sigma^2}$. For $f[A] = \mathrm{sgn}(A)$, in contrast, one has $Q_0 = 1$ and

$$Q_1 = \int Dy\, \mathrm{Erf}^2\left[\frac{(1-\zeta)\sigma y}{\zeta S\sqrt{2}}\right] = \frac{4}{\pi}\arctan\left[\sqrt{1 + 2\frac{(1-\zeta)^2\sigma^2}{\zeta^2 S^2}}\right] - 1 \tag{91}$$

and so

$$\varrho(\mu) = \frac{\theta\left[\mu - 2 + \frac{4}{\pi}\arctan\left[\sqrt{1+2\frac{(1-\zeta)^2\sigma^2}{\zeta^2 S^2}}\right]\right]}{\frac{4}{\pi}\arctan\left[\sqrt{1+2\frac{(1-\zeta)^2\sigma^2}{\zeta^2 S^2}}\right] - 1} \times \mathrm{e}^{-\left[\mu-2+\frac{4}{\pi}\arctan\left[\sqrt{1+2\frac{(1-\zeta)^2\sigma^2}{\zeta^2 S^2}}\right]\right]\Big/\left[\frac{4}{\pi}\arctan\left[\sqrt{1+2\frac{(1-\zeta)^2\sigma^2}{\zeta^2 S^2}}\right]-1\right]}. \tag{92}$$

Here the limit of fully real history gives $\lim_{\zeta\to 0} Q_1 = 1$ and $\lim_{\zeta\to 0}\varrho(\mu) = \theta[\mu]\,\mathrm{e}^{-\mu}$. The persistent order parameter equations (71), (72), (73) do not involve the volatility, but it turns out that we generally need to know the volatility in order to use (92) and (90) for closure, unless $\zeta \in \{0, 1\}$, i.e. unless the histories are strictly real or strictly fake. In figure 4 we compare the above predictions with the eigenvalue spectra as measured in numerical simulations of the inner product MG with $f[A] = \mathrm{sgn}(A)$, where (for $\zeta \notin \{0, 1\}$) we inserted in our formulae the value of σ as measured. Given the severe finite size and finite time effects in such simulations, the agreement is surprisingly acceptable. The remaining deviations are very likely to be finite size effects. Having a finite value of L in measurements removes the self-averaging of the matrix $B_{\lambda\lambda'}(A, Z)$ with respect to the noise variables $Z(\ell, \lambda)$, which our $L \to \infty$ calculation assumed (and thereby deforms the matrix in the direction of the random matrices as studied in e.g. [19–21]), and furthermore we have in these simulation experiments only been able to sample in equilibrium states matrices of limited dimension $p \in \{1025, 2049\}$. Both these limitations result in a smoothening of the spectrum as measured in practice, compared to the asymptotic shape. The effect of retaining the $\mathcal{O}(L^{-1})$ correlations between the x_r in the above calculation turns out to be an order L^{-1} reduction of the width of $\varrho(\mu)$; furthermore, when including $\mathcal{O}(L^{-1})$ terms finds that the condition $\int \mathrm{d}\mu\, \mu\varrho(\mu) = Q_0$ is violated at order L^{-1}, which implies (as expected) that in that order we can no longer treat the z_r as Gaussian variables.

We note that the above reasoning would also apply for $\Xi(\lambda) = \sigma^2\delta_{\lambda 0}$, a situation indistinguishable from replacing the actual bids in the history signal by consistent fake information, so we must conclude that the *origin* of nontrivial eigenvalue distributions $\varrho(\mu)$ is the *consistency* of the histories (although of course the precise shape of $\varrho(\mu)$ will depend on whether or not the histories are real). We note also that allowing for $\mathbf{B}(A, Z)$ not being self-averaging with respect to the bids has been crucial in our argument. Had we set the bid correlations to zero and averaged $\mathbf{B}(A, Z)$ over the bid process, we would have found the trivial spectrum $\varrho(\mu) = \delta[\mu - Q_0]$.

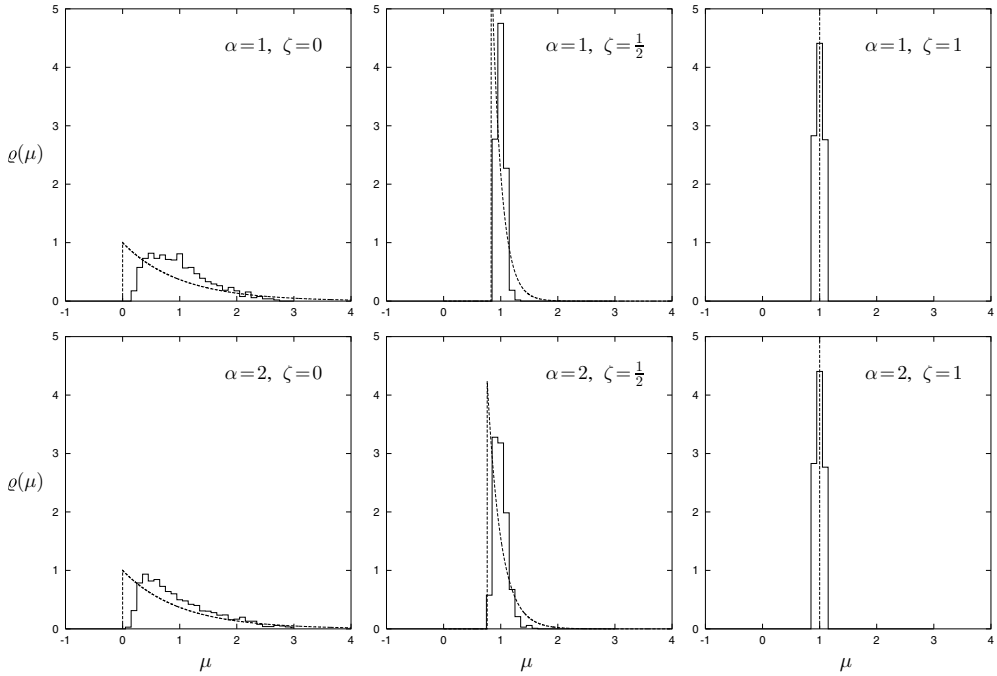

Figure 4. Comparison between eigenvalue spectra $\varrho(\mu)$ of the stochastic matrix $\mathbf{B}(A, Z)$ of inner product MGs (78) as measured in numerical simulations (histograms, for $N = 1025$) versus the approximate prediction (92) based on assuming weak bid correlations, for $f[A] = \mathrm{sgn}(A)$. Top row: $\alpha = 1$, with $\zeta \in \{0, \frac{1}{2}, 1\}$; bottom row: $\alpha = 2$, with $\zeta \in \{0, \frac{1}{2}, 1\}$. All matrix entries in simulations were calculated over $L = 400 \cdot N$ time steps, following equilibration. For $\zeta = \frac{1}{2}$ we used in (92) the value for σ as measured in the simulation; for $\zeta \in \{0, 1\}$ the volatility drops out of the formula. Note that there are at least two obvious sources of deviation between prediction and experiment (apart from the crude approximation used in the theory): finite size effects ($N = 1025$ rather than $N \to \infty$) and finite observation time effects ($L \leqslant 400 \cdot N$ rather than $L \to \infty$). Given these limitations, the qualitative agreement between prediction and experiment is surprisingly acceptable. Spectra of random matrices are indeed known to have finite size effects which decay much slower than $N^{-1/2}$, however, we will find later that at the level of the MG's conventional order parameters the finite size deviations between theory and simulations are of the expected order $N^{-1/2}$.

5.3. The volatility

We turn to the calculation of the volatility, which measures the fluctuations in the bids $A(\ell)$. It turns out that these bids are always Gaussian random variables, see figure 5, which is not obvious since the bids obey a generally nonlinear stochastic equation. The first step is to substitute (65) (obtained upon assuming short history correlation times in ergodic stationary states) into the as yet exact formula (54) for the volatility, giving

$$\sigma^2 = \lim_{t\to\infty}\lim_{\delta_N\to 0}\left\{\frac{1}{2}\sum_{r,r'\geqslant 0}(-\delta_N)^{r+r'}\int_0^\infty \mathrm{d}\mu\,\varrho(\mu)\mu^{r+r'+1}\sum_{\ell_0\ldots\ell_r}G(\ell_0,\ell_1)\ldots G(\ell_{r-1},\ell_r)\right.$$
$$\left.\times\sum_{\ell'_0\ldots\ell'_{r'}}G(\ell'_0,\ell'_1)\ldots G(\ell'_{r'-1},\ell'_{r'})[1+C(\ell_r,\ell'_{r'})]\delta_{\ell,\ell_0}\delta_{\ell,\ell'_0}\right\}\Bigg|_{\ell=t/\delta_N}. \qquad (93)$$

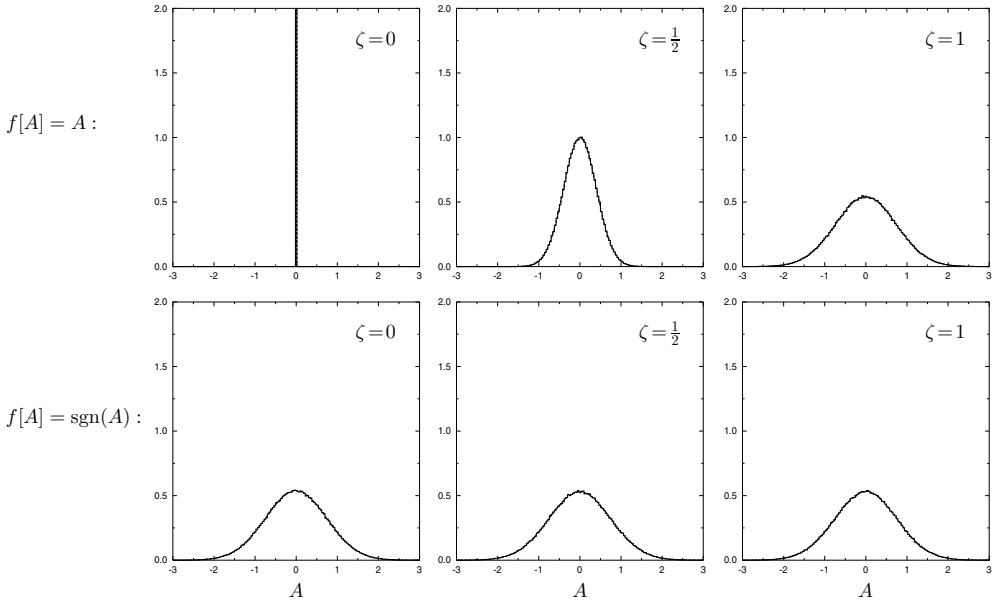

Figure 5. The bid distribution $P(A) = L^{-1}\sum_{\ell\leqslant L}\langle\delta[A - A(\ell)]\rangle$ of inner product MGs (78), as measured in numerical simulations (histograms), with $N = 1025$ and $\alpha = 2$. Statistics are calculated over $200.N$ time steps, following equilibration. The solid lines (barely visible, since they are virtually coinciding with the histograms) are Gaussian distributions with average and variance identical to those observed, for comparison. Top row: $f[A] = A$, with $\zeta \in \{0, \frac{1}{2}, 1\}$; bottom row: $f[A] = \mathrm{sgn}(A)$, with $\zeta \in \{0, \frac{1}{2}, 1\}$. The left pictures refer to strictly real histories, the right pictures to strictly fake histories, with the middle ones representing a mixing of the two. These data (corresponding to the same experiments as those of figure 3) are representative of the typical behaviour throughout the ergodic regime. The bids are *always* indistinguishable from Gaussian variables.

As always we next have to approximate non-persistent dynamical order parameters by expressions involving only persistent ones. Following procedures similar to those used in the past [13] we assume the agent correlation function $C(\ell)$ to decay very fast, so that we may replace $C(\ell) \to c + \delta_{\ell 0}(1 - c)$. Substitution into (93), followed by carrying out various sums analytically (using $\delta_N \sum_\ell G(\ell) = \chi$) leads us to the approximation

$$\sigma^2 = \frac{1}{2}(1+c)\int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\frac{\mu}{(1+\mu\chi)^2} + \frac{1}{2}(1-c)\lim_{t\to\infty}\lim_{\delta_N\to 0}\left\{\sum_{r,r'\geqslant 0}(-\delta_N)^{r+r'}\int_0^\infty \mathrm{d}\mu\,\varrho(\mu)\mu^{r+r'+1} \times \sum_{\ell_0\ldots\ell_r} G(\ell_0,\ell_1)\ldots G(\ell_{r-1},\ell_r)\sum_{\ell'_0\ldots\ell'_{r'}} G(\ell'_0,\ell'_1)\ldots G(\ell'_{r'-1},\ell'_{r'})\delta_{\ell_r,\ell'_{r'}}\delta_{\ell,\ell_0}\delta_{\ell,\ell'_0}\right\}\Bigg|_{\ell=t/\delta_N}. \tag{94}$$

In the second term we see that, unless $r = r' = 0$ (where $\delta_{\ell_r,\ell'_{r'}}$ reduces directly to $\delta_{\ell_0,\ell'_0} = \delta_{\ell\ell} = 1$), the factor $\delta_{\ell_r,\ell'_{r'}}$ will always leave us with one residual factor δ_N that is no longer compensated by a time summation, resulting in the removal of the corresponding term for

$\delta_N \to 0$. Hence one retains only

$$\sigma^2 = \frac{1}{2}(1+c)\int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\frac{\mu}{(1+\mu\chi)^2} + \frac{1}{2}(1-c)\int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\mu$$
$$= \frac{1}{2}(1+c)\int_0^\infty \mathrm{d}\mu\, \varrho(\mu)\frac{\mu}{(1+\mu\chi)^2} + \frac{1}{2}(1-c)Q_0. \tag{95}$$

For $\zeta \to 1$ (strictly inconsistently fake histories), where $\varrho(\mu) \to \delta[\mu - Q_0]$ and $Q_0 \to \kappa_2$, this expression reduces to the familiar approximation formula known from inner product MGs (apart from various factors κ_2, which appear exactly as predicted by (53) and (61):

$$\lim_{\zeta\to 1}\sigma^2 = \kappa_2\left\{\frac{1}{2}\frac{1+c}{(1+\kappa_2\chi)^2} + \frac{1}{2}(1-c)\right\} \tag{96}$$

5.4. *Compactification of the theory and universality for $\zeta \to 0$*

The final equations which describe the static limit of our theory, describing time-translation invariant stationary states, can be compactified further. The simple exponential shape of the spectrum (89) allows us to express the two relevant averages in (71), (72) in terms of the exponential integral $E_1(z) = \int_z^\infty \mathrm{d}t\, t^{-1}\,\mathrm{e}^{-t}$ [22], resulting in the following set of coupled equations for the static order parameters $\{c, \phi, \chi\}$:

$$\omega = \frac{\chi Q_1 - \mathrm{e}^{[1+\chi(Q_0-Q_1)]/\chi Q_1} E_1\left(\frac{1+\chi(Q_0-Q_1)}{\chi Q_1}\right)}{\sqrt{(\chi Q_1)^2 - \mathrm{e}^{[1+\chi(Q_0-Q_1)]/\chi Q_1} E_1\left(\frac{1+\chi(Q_0-Q_1)}{\chi Q_1}\right)(2\chi Q_1 + 1) + \frac{\chi Q_1}{1+\chi(Q_0-Q_1)}}} \tag{97}$$

$$u = \omega\sigma[\infty]\sqrt{\alpha}/\sqrt{2(1+c)} \tag{98}$$

$$\phi = 1 - \mathrm{Erf}[u] \tag{99}$$

$$\mathrm{Erf}[u] = \alpha\left\{1 - \frac{1}{\chi Q_1}\,\mathrm{e}^{[1+\chi(Q_0-Q_1)]/\chi Q_1} E_1\left(\frac{1+\chi(Q_0-Q_1)}{\chi Q_1}\right)\right\} \tag{100}$$

$$c = \sigma^2[\infty]\left\{1 - \mathrm{Erf}[u] + \frac{1}{2u^2}\,\mathrm{Erf}[u] - \frac{1}{u\sqrt{\pi}}\,\mathrm{e}^{-u^2}\right\}. \tag{101}$$

Equations (97), (98) simply define ω and u as short-hands for complicated functions of $\{c, \phi, \chi\}$. The volatility and the factors $Q_{0,1}$ are calculated from

$$\sigma^2 = \frac{1+c}{2Q_1\chi^2}\left\{\frac{1+Q_1\chi}{Q_1\chi}\,\mathrm{e}^{[1+\chi(Q_0-Q_1)]/\chi Q_1} E_1\left(\frac{1+\chi(Q_0-Q_1)}{\chi Q_1}\right) - \frac{1}{1+\chi(Q_0-Q_1)}\right\}$$
$$+ \frac{1}{2}(1-c)Q_0$$
$$= \frac{1+c}{2Q_1\chi^2}\left\{(1+Q_1\chi)\frac{\alpha - \mathrm{Erf}[u]}{\alpha} - \frac{1}{1+\chi(Q_0-Q_1)}\right\} + \frac{1}{2}(1-c)Q_0 \tag{102}$$

$$Q_0 = \int Dz\, f^2[z\sqrt{(1-\zeta)^2\sigma^2 + \zeta^2 S^2}] \tag{103}$$

$$Q_1 = \int Dy \left[\int Dz\, f[(1-\zeta)\sigma y + \zeta S z] \right]^2. \tag{104}$$

All the details regarding the history definitions (namely the sampling function $f[A]$ and the degree of 'fakeness' ζ) are fully contained within the factors $Q_{0,1}$. Our stationary state order parameter equations are now fully closed, and can be solved numerically.

Of special interest is the limit $\zeta \to 0$, where the histories are strictly real. In this limit one can show that the persistent order parameters $\{\phi, c\}$ become functions of α that no longer depend on the choice made for $f[A]$. The reason for this universality is that according to (79), (80) for $\zeta \to 0$ one always has $\lim_{\zeta\to 0} Q_1 = \lim_{\zeta\to 0} Q_0 = \int Dz\, f^2[\sigma z]$. We may now define $\tilde{\chi} = Q_0 \chi$ and find for $\zeta \to 0$ our order parameter equations for $\{c, \phi, \tilde{\chi}\}$ reducing to

$$\omega = \frac{\tilde{\chi} - \mathrm{e}^{1/\tilde{\chi}} E_1(1/\tilde{\chi})}{\sqrt{\tilde{\chi}^2 - \mathrm{e}^{1/\tilde{\chi}} E_1(1/\tilde{\chi})(2\tilde{\chi}+1) + \tilde{\chi}}} \tag{105}$$

$$u = \omega\sigma[\infty]\sqrt{\alpha}/\sqrt{2(1+c)} \tag{106}$$

$$\phi = 1 - \mathrm{Erf}[u] \tag{107}$$

$$\mathrm{Erf}[u] = \alpha\{1 - \tilde{\chi}^{-1}\, \mathrm{e}^{1/\tilde{\chi}} E_1(1/\tilde{\chi})\} \tag{108}$$

$$c = \sigma^2[\infty]\left\{ 1 - \mathrm{Erf}[u] + \frac{1}{2u^2}\,\mathrm{Erf}[u] - \frac{1}{u\sqrt{\pi}}\,\mathrm{e}^{-u^2} \right\}. \tag{109}$$

The values of $\{c, \phi, \tilde{\chi}\}$ as solved from the above closed set can only depend on the remaining control parameter α. Since $\lim_{z\to 0} z\,\mathrm{e}^z E_1(z) = 0$, the ergodicity-breaking transition point where $\tilde{\chi} \to \infty$, corresponds as always to $\phi = 1 - \alpha$. We would only need to know the choice made for $f[A]$ if we wished to disentangle χ and Q_0 from $\tilde{\chi}$, or to calculate the volatility which would be the solution of

$$\sigma^2 = \left\{ \frac{1+c}{2\tilde{\chi}^2}\left[(1+\tilde{\chi})\frac{\alpha - 1 + \phi}{\alpha} - 1 \right] + \frac{1}{2}(1-c) \right\} \int Dz\, f^2[\sigma z]. \tag{110}$$

For $f[A] = A$, however, we have a degenerate situation at $\zeta = 0$. Here we find (110) converting into

$$\sigma^2 \left\{ 1 - \frac{1}{\tilde{\chi}^2}\left[(1+\tilde{\chi})\frac{\alpha - 1 + \phi}{\alpha} - 1 \right] \right\} = 0. \tag{111}$$

We find that (possibly apart from one degenerate case) that the only stationary solution has $\sigma = 0$, so also $Q_0 = \sigma^2 = 0$. This, in turn, tells us that $\chi = \infty$ for all α, so the $f[A] = A$ system is always critical and the ergodicity assumption is always inconsistent.

6. Theory versus numerical simulations of inner product MGs

Below we test the predictions of our theory further against numerical simulations of inner product MGs without agents' decision noise, so $T = 0$ and $\sigma[\infty] = 1$. Our equations allow for such noise, but it represents a simple generalization that is well understood and delays considerably the equilibration in simulations. In contrast to look-up table MGs, where the non-Markovian nature of the microscopic process involves delayed forces going back over a time interval $\Delta t = \mathcal{O}(N^{-1}\log N)$, in inner product MGs this time interval scales as $\Delta t = \mathcal{O}(N^0)$. This already limits severely the scope of numerical simulations, and, given the

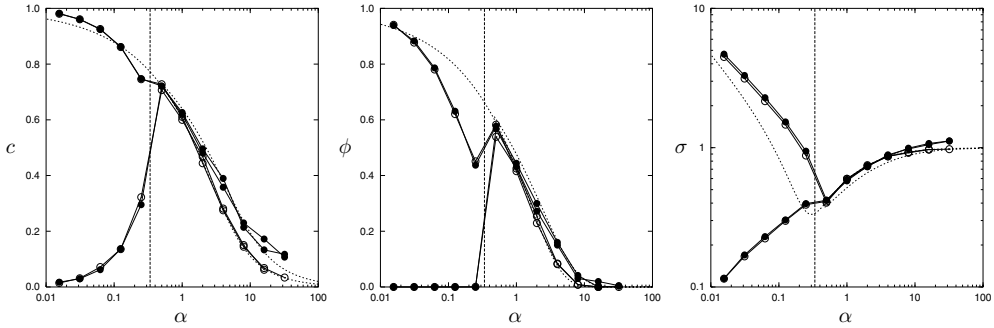

Figure 6. Observables c, ϕ and σ as functions of $\alpha = p/N$, measured in numerical simulations for $f[A] = \mathrm{sgn}(A)$ and $S = 1$. Measurements are taken over $200 \cdot p$ time steps, after an equilibration of $200 \cdot p$ time steps; system sizes: $pN^2 = 2^{30}$ (except for $\alpha = 16$ and $\alpha = 32$, where $N = 1025$ and $N = 513$, respectively). Each panel shows the result of four simulation runs: $\zeta = 0$ (strictly real histories, connected full circles) with biased and unbiased initial conditions, and $\zeta = 1$ (strictly fake histories, connected open circles) with biased and unbiased initial conditions. Biased initial conditions: $q_i(0) = \mathcal{O}(1)$ (zero average random); unbiased initial conditions: $q_i(0) = \mathcal{O}(10^{-5})$ (zero average random). Dotted curves give the theoretical predictions for $\zeta = 0$ and $\zeta = 1$, upon assuming ergodic stationary states, with the vertical dashed line marking the transition value α_c below which the theory no longer applies.

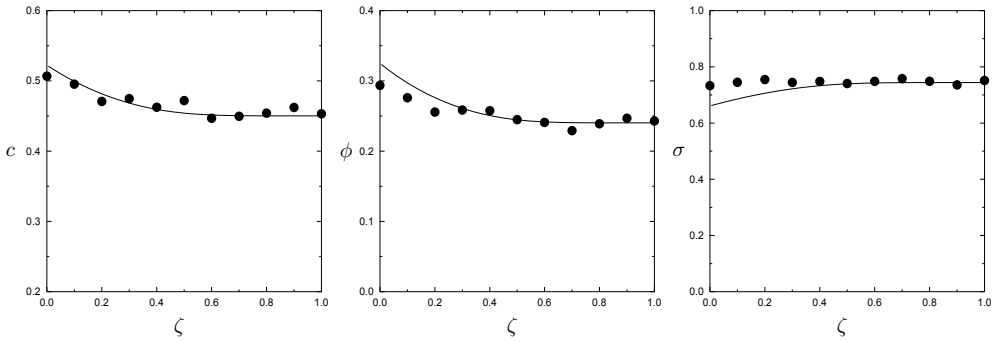

Figure 7. Observables c, ϕ and σ as functions of ζ, measured in numerical simulations for $f[A] = \mathrm{sgn}(A), \alpha = 2$, and $S = 1$. Measurements are taken over $400 \cdot N$ time steps, after an equilibration of $400 \cdot N$ time steps; system sizes: $N = 1025$. Circles give the simulation results. Curves give the theoretical predictions, upon assuming ergodic stationary states. Strictly real histories and stricly fake histories correspond to $\zeta = 0$ and $\zeta = 1$, respectively. The random deviations between theory and simulations are compatible with the expected finite size effects of order $N^{-1/2} \approx 0.03$, except for the volatility where fluctuations are small and the weak predicted trend is not confirmed.

multitude of control parameters to be varied, system sizes will have to remain modest, here mostly $N = \mathcal{O}(10^3)$. Even then, finite size and finite time effects remain a serious problem, especially for large α.

6.1. Inner product MG with $f[A] = \mathrm{sgn}(A)$

In an earlier section we have already compared the observed spectra of the history covariance matrix to our theoretical predictions, see figure 4, showing quite reasonable agreement. In figures 6 and 7 we compare the predicted values of the observables $\{c, \phi, \sigma\}$ (obtained by solving our order closed equations (97–104) numerically) to numerical simulation data, for

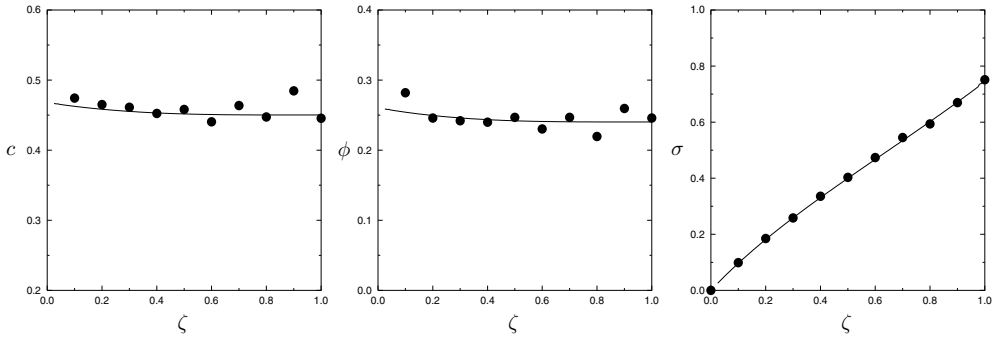

Figure 8. Observables c, ϕ and σ as functions of ζ, measured in numerical simulations for $f[A] = A, \alpha = 2$, and $S = 1$. Measurements are taken over $400 \cdot N$ time steps, after an equilibration of $400 \cdot N$ time steps; system sizes: $N = 1025$. Circles give the simulation results. Curves give the theoretical predictions, upon assuming ergodic stationary states. Strictly real histories and stricly fake histories correspond to $\zeta = 0$ and $\zeta = 1$, respectively. The deviations between theory and simulations are compatible with the expected finite size effects of order $N^{-1/2} \approx 0.03$.

$\zeta \in \{0, 1\}$, $S = 1$, and different values of α. We see our prediction that the transition point is independent of ζ, i.e. of whether histories are real or fake, is confirmed quite convincingly. In the ergodic regime $\alpha > \alpha_c$, where our theory applies, we also observe a good agreement in terms of the values of c and ϕ (except for large values of α, where finite size and finite time effects become severe). In terms of the volatility σ, for which we had to make further approximations (to express non-persistent terms in persistent ones) theory and experiment exhibit deviations: the experimentally observed value for σ appears roughly independent of ζ in the regime where the data are still reliable (if not a higher value for $\zeta = 0$ compared to $\zeta = 1$), whereas the theory predicts a slight volatility *reduction* due to having real histories compared to the fake ones. This suggests that, at least for $f[A] = \mathrm{sgn}(A)$, the usual crude step to replace $C(\ell) \to c + (1 - c)\delta_{\ell\ell'}$ in the true formula (93) for the volatility, resulting in the approximation (95), is less appropriate for real than for fake histories. A careful analysis of the short-time behaviour of the kernels C and G will be required to understand why this is so. Fortunately, the volatility drops out of our order parameter equations for $\zeta \in \{0, 1\}$. Upon measuring order parameters as a function of ζ, in order to probe those cases $0 < \zeta < 1$ where the volatility does play a vital role in closing our equations, one still observes excellent agreement in those regimes where simulation data are reliable.

6.2. Inner product MG with $f[A] = A$

In contrast to the inner product MG with $f[A] = \mathrm{sgn}(A)$, for $f[A] = A$ the theory predicts a strong dependence of the volatility on the parameter ζ, with a degeneracy at $\zeta = 0$. This prediction was already confirmed qualitatively by the data in figure 3. In figure 8 we inspect this dependence in more detail for $\alpha = 2$, and find an excellent agreement between theory and simulations, even for the volatility (in contrast to the situation for $f[A] = \mathrm{sgn}(A)$, where the predicted volatility was somewhat lower than the observed one). According to (111) we should expect that the only stationary solution for $\zeta = 0$ will have $\sigma = 0$. In the limit $\zeta \to 0$ (where solving our closed equations becomes nontrivial as a result of having $\sigma \to 0$ and $\chi \to \infty$) figure 8 indeed confirms the predicted stationary state with vanishing volatility, and the numerical solution of the order parameter equations confirms that $\chi \to \infty$. However, if

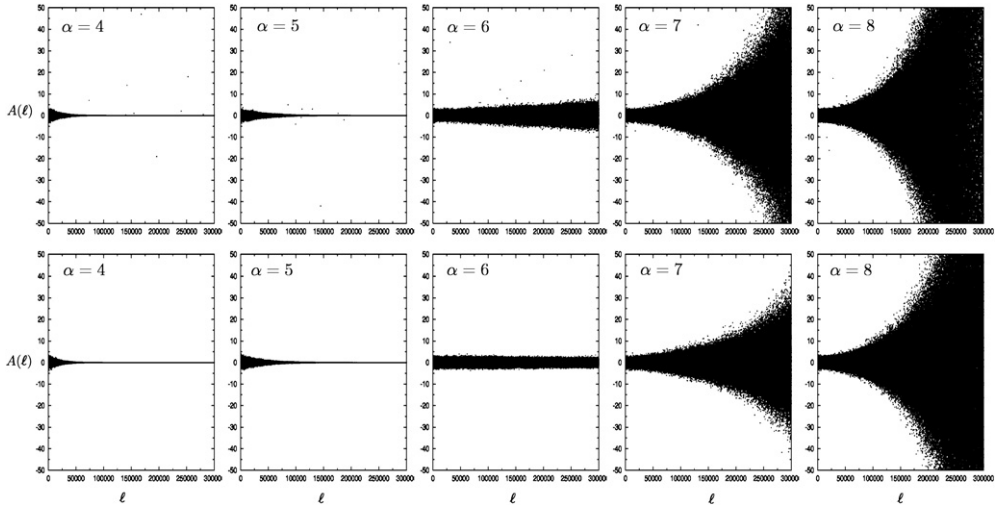

Figure 9. Overall bids $A(\ell)$ versus time $\ell = 1 \ldots 300,000$, as measured in numerical simulations of the inner product MG with $f[A] = A$ and $\zeta = 0$ (fully real histories), in systems of size $N = 1025$. Top row: unbiased initial conditions, $q_i(0) = \mathcal{O}(10^{-5})$ (zero average random); bottom row: biased initial conditions, $q_i(0) = \mathcal{O}(1)$ (zero average random). The data indicate that there exists a critical value α_c^+ such that for $\alpha > \alpha_c^+$, rather than evolve towards $\sigma = 0$, the system is unstable and the bid fluctuations diverge. If the (weak) dependence on initial conditions in the graphs for $\alpha = 6$ is not a finite size effect, then there is a parameter regime where the zero volatility state and the 'runaway' solution coexist.

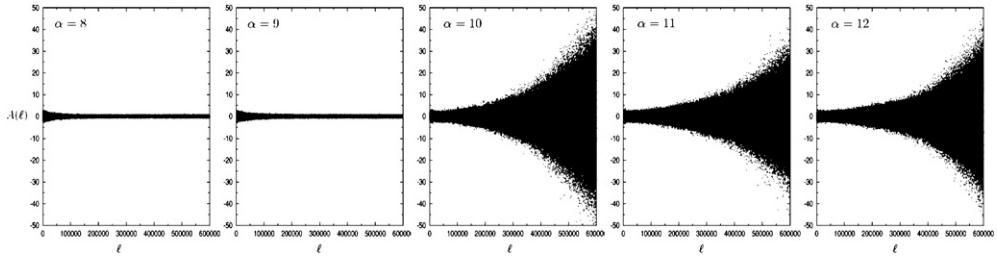

Figure 10. Overall bids $A(\ell)$ versus time $\ell = 1 \ldots 300,000$, as measured in numerical simulations of the inner product MG with $f[A] = A$, in systems of size $N = 1025$ and following unbiased initial conditions. Here $\zeta = 0.1$. The unstable state with diverging volatility exists also for $\zeta > 0$, with the instability setting in at a critical value $\alpha < \alpha_c^+(\zeta)$ that increases with ζ. Also the characteristic time scale for the instability to manifest itself appears to increase with ζ (note the different time scales of figures 9 and 10).

the experiment is repeated for larger values of α one finds that beyond a critical value $\alpha_c^+ \approx 6$ this zero volatility stationary state is no longer stable, and the fluctuations diverge. This is illustrated in figure 9. In fact, this runaway solution is found to exist also for $\zeta > 0$, see figure 10, suggesting a ζ-dependent criticality $\alpha_c^+(\zeta)$. Also at the lower end of the α-scale, below the conventional transition point, one finds the system generally entering the runaway state, see e.g. the volatility data in figure 11, below a second critical point $\alpha_c^-(\zeta)$, provided one chooses unbiased initial conditions. For intermediate α values, where the volatility remains finite, our theory is found to predict the volatility σ correctly (within the accuracy limitations

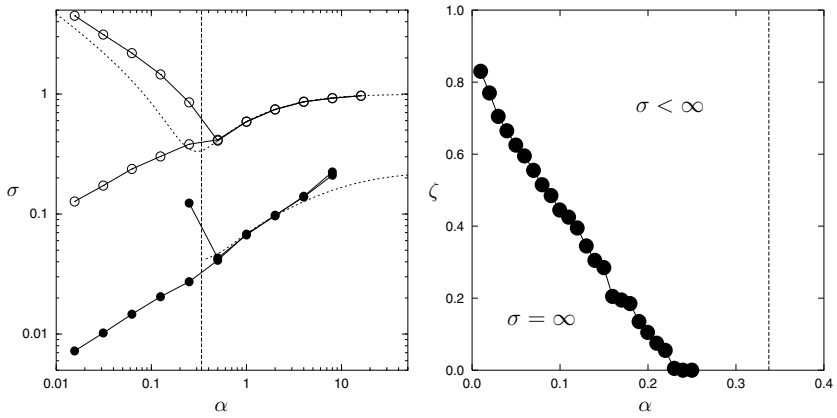

Figure 11. Left: volatility as a function of $\alpha = p/N$, measured in numerical simulations for $f[A] = A$ and $S = 1$. Measurements are taken over $200 \cdot p$ time steps, after $200 \cdot p$ equilibration time steps; system sizes: $pN^2 = 2^{30}$ (except for $\alpha = 16$ and $\alpha = 32$, where $N = 1025$ and $N = 513$, respectively). We show the results of four runs: $\zeta = 0.1$ (predominantly real histories, connected full circles) with biased and unbiased initial conditions, and $\zeta = 1$ (strictly fake histories, connected open circles) with biased and unbiased initial conditions; initial conditions are as in previous figures. Dotted curves give the theoretical predictions for $\zeta = 0.1$ and $\zeta = 1$, upon assuming ergodicity. The system destabilizes ($\sigma = \infty$) for $\alpha > \alpha_c^+(\zeta)$ (any initial conditions, terminating branches at $\alpha \approx 9$) and for $\alpha < \alpha_c^-$ (unbiased initial conditions only, terminating branch at $\alpha \approx 0.2$). Right: estimated critical line $\alpha_c^-(\zeta)$ in the low α region of the (α, ζ) plane, marking the onset of the $\sigma = \infty$ instability, obtained by numerical simulations with $N = 2049$ ($\zeta < 0.77$) and $N = 4097$ ($\zeta > 0.77$). The marker sizes indicate the error bars. The vertical dashed lines in both figures mark the $\chi \to \infty$ transition value α_c below which the ergodic theory no longer applies.

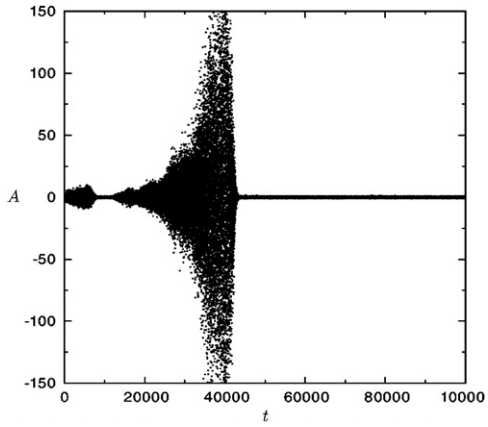

Figure 12. Typical example of the bid evolution close to the critical line that marks the onset of the $\sigma = \infty$ instability, for $f[A] = A$ (see previous figure for the estimated location of this line). The large intermittent fluctuations exhibited by the system limit the reliability with which the transition can be located via numerical simulations. The present example corresponds to $(\alpha, \zeta) = (0.14, 0.32)$, with $N = 2048$ and unbiased initial conditions.

imposed by finite size effects). In the right panel of figure 11 we show the result of estimating from numerical simulations the location of $\alpha_c^-(\zeta)$ in the (α, ζ) plane, which appears to

approach $\zeta = 1$ for $\alpha \to 0$; determining this line experimentally is not entirely trivial in view of the (expected) nontrivial bid fluctuations close to this line, see eg. figure 12. Carrying out a similar exercise for the location of the instability $\alpha_c^+(\zeta)$ in the high α region is unfortunately ruled out, due to the extreme relation times required for large α. In summary: for sufficiently small ζ (i.e. sufficiently 'real' histories), the usual high volatility state of conventional MGs in the non-ergodic regime becomes a runaway solution in inner product MGs with linearly sampled histories (a true instability with infinite volatility), which for $\zeta = 0$ vanishes slightly below the standard MG's $\chi = \infty$ transition $\alpha = \alpha_c$ if one increases α, but then re-appears at a larger value $\alpha_c^+(\zeta)$ and there becomes the *only* solution. It must in principle be possible to calculate the critical values $\alpha_c^{\pm}(\zeta)$ from our order parameter equations; in practice, however, this requires solving our equations for the correlation and response functions (formulated in terms of both the effective agent process and the effective overall bid process) at finite times, which at present we are unable to do.

7. Discussion

In this paper we have generalized the theory of look-up table Minority games with real market history [11] to a larger family of models, and applied it to MGs with so-called inner product strategy definitions. The latter MG versions had, surprisingly, never been solved; not even in their simplest Markovian version where the histories are fake. At a mathematical level it was not a priori clear which form the inner-product MG theory would take: in [11] it was found that the key object in the theory of look-up table MGs with real market history was the so-called history frequency distribution, but already on simple scaling grounds it is clear that this object cannot be defined for inner product MGs. It is satisfactory to find in the present study how the generating functional formalism resolves the issue: a more general quantity takes over the role of the history frequency distribution, namely the eigenvalue spectrum of the history covariance matrix, which reduces to the former for look-up table MGs but is also well-defined for inner product MGs. This resolution involves a generalization of the short history correlation time ansatz that led in [11] to closed stationary state order parameter equations, resulting (after a random matrix spectrum calculation) also for the present inner-product MGs in closed equations for observables such as the persistent correlations, the fraction of 'frozen' agents, the integrated response, and the market volatility (although the formula for the latter as always involves further approximations).

We find, provided all the relevant integrals and averages exists, that the phase diagrams of the two MG versions are identical. However, we encounter interesting differences in terms of the models' static and dynamic phenomenology, dependent upon which functions $f[A]$ of the past overall market bids A are being sampled in the inner product MGs. For bounded functions such as $f[A] = \mathrm{sgn}(A)$ one continues to find behaviour that, although quantitatively different, is qualitatively similar to that of look-up table MGs. If the histories are strictly real, the theory even predicts complete universality of the static observables, independent of the choice made for $f[A]$. In contrast, for unbounded sampling functions such as $f[A] = A$ the situation is quite different. There can now be a profound dependence of the volatility on the degree to which the histories are real. Especially when the histories are strictly real and $f[A] = A$ we find that the inner product MG has only one stationary solution, where the volatility σ is zero and the model is structurally critical (i.e. $\chi = \infty$). In terms of the usual control parameter $\alpha = p/N$, we find that in the regime $\alpha > \alpha_c$ (the standard non-ergodicity transition of the MG) this state $(\sigma, \chi) = (0, \infty)$ can only be reached for $\alpha < \alpha_c^+(\zeta)$, with $\alpha_c^+(0) \approx 6$ and $\alpha_c^+(\zeta)$ increasing with ζ. For $\alpha > \alpha_c^+(\zeta)$ the system destabilizes and the fluctuations diverge. Whether this divergence represents a structural instability that is independent of the system

size N, or marks system states where the bid fluctuations scale differently with N than in look-up-table MGs, is as yet unclear. In the low α regime one finds a similar critical line $\alpha_c^-(\zeta)$ marking a destabilization transition as one reduces α further, with $\sigma = \infty$ for $\alpha < \alpha_c^-(\zeta)$, but here this destabilization occurs only for unbiased initial conditions.

Those technical questions that remain open in the present study are the familiar ones that are similarly answered for all previous MG models. They relate mostly to our inability so far to extract from our order parameter equations exact solutions for time-dependent observables (beyond just a few time steps), or exact stationary state solutions in the non-ergodic regime. In addition one would like to explore further the possibility of finding solutions for MG versions in which the effects of having real histories are more dramatic; both the present study and its predecessor [11] rely on expansions that require history correlations to be weak. Finally, we would like to point out that the observed instability of the $f[A] = A$ inner product MG is qualitatively different from what has been observed in previous MG versions. Although in its present form this specific model is ill-constructed, it could perhaps form the basis of new models aimed at explaining more robustly the so-called stylized facts in real markets, since its structural instability is *not* a finite size effect and therefore requires no careful tuning of control parameters. The phenomenology observed for $f[A] = A$ is novel, should be expected to emerge also for other unbounded choices for the sampling function $f[A]$ (it would be very interesting to determine the conditions on $f[A]$ for which the MG exhibits structural instability), and emphasizes even more the need for progress in solving the GFA order parameter equations for finite times.

References

[1] Challet D and Zhang Y-C 1997 *Physica* A **246** 407–18
[2] Challet D and Zhang Y-C 1998 *Physica* A **256** 514–32
[3] Arthur W B 1994 *Am. Econ. Assoc. Papers and Proc.* **84** 406–11
[4] Cavagna A 1999 *Phys. Rev.* E **59** R3783–6
[5] Marsili M, Challet D and Zecchina R 2000 *Physica* A **280** 522–53
[6] Marsili M and Challet D 2001 *Phys. Rev.* E **64** 056138
[7] Heimel J A F and Coolen A C C 2001 *Phys. Rev.* E **63** 056121
[8] Coolen A C C, Heimel J A F and Sherrington D 2001 *Phys. Rev.* E **65** 016126
[9] Coolen A C C and Heimel J A F 2001 *J. Phys A: Math. Gen.* **34** 10783–804
[10] Challet D and Marsili M 2000 *Phys. Rev.* E **62** 1862–68
[11] Coolen A C C 2005 *J. Phys A: Math. Gen.* **38** 2311–47
[12] Challet D, Marsili M and Zhang Y C 2004 *Minority Games — Interacting Agents in Financial Markets* (Oxford: Oxford University Press)
[13] Coolen A C C 2005 *The Mathematical Theory of Minority Games — Statistical Mechanics of Interacting Agents* (Oxford: Oxford University Press)
[14] Cavagna A, Garrahan J P, Giardina I and Sherrington D 1999 *Phys. Rev. Lett.* **83** 4429–32
[15] Garrahan J P, Moro E and Sherrington D 2000 *Phys. Rev* E **62** R9–12
[16] Challet D, Marsili M and Zecchina R 2000 *Phys. Rev. Lett.* **85** 5008
[17] Cavagna A, Garrahan J P, Giardina I and Sherrington D 2000 *Phys. Rev. Lett.* **85** 5009
[18] Heimel J A F and De Martino A 2001 *J. Phys. A: Math. Gen.* **34** 2525–37
[19] Hertz J A, Krogh A and Thorbergsson G I 1989 *J. Phys A: Math. Gen.* **22** 2133–50
[20] Rosenow B, Amaral L A N, Guhr T, Stanley H E, Plerou V and Gopikrishnan P 2002 *Phys. Rev.* E **65** 066126
[21] Bouchaud JP and Potters M 2003 *Theory of Financial Risk and Derivative Pricing* 2nd (edn) (Cambridge: Cambridge University Press)
[22] Abramowitz M and Stegun I A 1972 *Handbook of Mathematical Functions* (New York: Dover)