Frédéric Abergel • Bikas K. Chakrabarti •
Anirban Chakraborti • Asim Ghosh
Editors

Econophysics of Systemic Risk and Network Dynamics

Springer

Editors
Frédéric Abergel
Chaire de Finance Quantitative
Laboratoire de Mathématiques Appliquées aux Systèmes
École Centrale Paris
Châtenay-Malabry, France

Anirban Chakraborti
Chaire de Finance Quantitative
Laboratoire de Mathématiques Appliquées aux Systèmes
École Centrale Paris
Châtenay-Malabry, France

Bikas K. Chakrabarti
Theoretical Condensed Matter Physics Division
Saha Institute of Nuclear Physics
Kolkata, India;
Economic Research Unit
Indian Statistical Institute
Kolkata, India

Asim Ghosh
Theoretical Condensed Matter Physics Division
Saha Institute of Nuclear Physics
Kolkata, India

ISSN 2039-411X ISSN 2039-4128 (electronic)
New Economic Windows
ISBN 978-88-470-2552-3 ISBN 978-88-470-2553-0 (eBook)
DOI 10.1007/978-88-470-2553-0
Springer Milan Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012939162

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Systemic risk has long been identified as a potential for financial institutions to trigger a dangerous contagion mechanism from the financial economy to the real economy itself. One of the commonly adopted definitions of systemic risk is: "risk of disruption to the flow of financial services that is

(i) caused by an impairment of all or parts of the financial system; and
(ii) has the potential to have serious negative consequences for the real economy".

Evident from this definition, or from any of its variants that one can find in the growing literature on the subject, are two characteristic aspects. The first one being that such a risk takes place at a much larger scale than that of an individual institution. The second one being that it eventually spreads to the real economy outside the financial system through various "leakage" mechanisms, of which the last crisis has given some examples: liquidity shrinkage, fire sale of assets, drop in market value of derivatives...

This type of risk, long confined to the monetary market, has spread widely in the recent past, culminating in the subprime crisis of 2008. The understanding and control of systemic risk has therefore become an extremely important societal and economic question. Such problems are now extensively being studied by people from disciplines like economics, finance and physics. The contributions by physicists are relatively new.

The Econophys-Kolkata VI conference, the 6th event in this series of international conferences, held during October 21–25 last year, was dedicated to address and discuss extensively these issues and the recent developments. Like the last event in the series, this one was also organized jointly by the École Centrale Paris and the Saha Institute of Nuclear Physics, and was held at the Saha Institute of Nuclear Physics, Kolkata.

This proceedings volume contains the written versions of most of the talks and seminars delivered by distinguished experts from all over the world, participating in the meeting, and accepted after refereeing. For some completeness in the cases of one or two important topics (like in the case Many-agent Games), some reviews, by experts who could not attend, were invited and incorporated in this volume.

These Proceedings volume is organized as follows: Part I dedicated to the study of systemic risk, network dynamics and other empirical studies. Part II devoted to model-based studies. We have also included Part III for "miscellaneous reports", to present some on-going or preliminary studies. Finally, we have summarized in a brief "discussion and comments" Appendix, some of the remarks made by the participants during the various interesting and animated exchanges that took place during the panel discussion in the conference.

We are grateful to all the participants of the conference for their participation and contributions. We are also grateful to Mauro Gallegati and the Editorial Board of the New Economic Windows series of the Springer-Verlag (Italia) for their support in getting this Proceedings volume published as well, in their esteemed series.[1]

The editors also address their thanks to the Centre for Applied Mathematics and Computational Science at Saha Institute, and École Centrale Paris for their support in organizing this conference. They would also like to thank Gayatri Tilak for providing invaluable help during the preparation of the manuscript.

Châtenay-Malabry, France — Frédéric Abergel
Kolkata, India — Bikas K. Chakrabarti
Châtenay-Malabry, France — Anirban Chakraborti
Kolkata, India — Asim Ghosh
April, 2012

[1]Past volumes:

(i) Econophysics of Order-driven Markets, Eds. F. Abergel, B.K. Chakrabarti, A. Chakraborti, M. Mitra, New Economic Windows, Springer-Verlag, Milan, 2011.
(ii) Econophysics & Economics of Games, Social Choices and Quantitative Techniques, Eds. B. Basu, B.K. Chakrabarti, S.R. Chakravarty, K. Gangopadhyay, New Economic Windows, Springer-Verlag, Milan, 2010.
(iii) Econophysics of Markets and Business Networks, Eds. A. Chatterjee, B.K. Chakrabarti, New Economic Windows, Springer-Verlag, Milan, 2007.
(iv) Econophysics of Stock and other Markets, Eds. A. Chatterjee, B.K. Chakrabarti, New Economic Windows, Springer-Verlag, Milan, 2006.
(v) Econophysics of Wealth Distributions, Eds. A. Chatterjee, S. Yarlagadda, B.K. Chakrabarti, New Economic Windows, Springer-Verlag, Milan, 2005.

Contents

Chapter 12
Kolkata Paise Restaurant Problem: An Introduction

Asim Ghosh, Soumyajyoti Biswas, Arnab Chatterjee, Anindya Sundar Chakrabarti, Tapan Naskar, Manipushpak Mitra, and Bikas K. Chakrabarti

Abstract We discuss several stochastic optimization strategies in games with many players having large number of choices (Kolkata Paise Restaurant Problem) and two choices (minority game problem). It is seen that a stochastic crowd avoiding strategy gives very efficient utilization in KPR problem. A slightly modified strategy in the minority game problem gives full utilization but the dynamics stops after reaching full efficiency, thereby making the utilization helpful for only about half of the population (those in minority). We further discuss the ways in which the dynamics may be continued and the utilization becomes effective for all the agents keeping fluctuation arbitrarily small.

A. Ghosh (✉) · S. Biswas · B.K. Chakrabarti
Theoretical Condensed Matter Physics Division, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700 064, India
e-mail: asimghosh066@gmail.com

S. Biswas
e-mail: soumyajyotibiswas@gmail.com

B.K. Chakrabarti
e-mail: bikask.chakrabarti@saha.ac.in

A. Chatterjee
Department of Biomedical Engineering and Computational Science (BECS), Aalto University School of Science, FI-00076 AALTO, Espoo, Finland
e-mail: arnab.chatterjee@aalto.fi

A.S. Chakrabarti
Economics Department, Boston University, 270 Bay State Road, Boston, MA, USA
e-mail: aschakrabarti@gmail.com

T. Naskar
Theory Division, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700 064, India
e-mail: tnaskar@gmail.com

M. Mitra · B.K. Chakrabarti
Economic Research Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700 018, India

M. Mitra
e-mail: mmitra@isical.ac.in

F. Abergel et al. (eds.), *Econophysics of Systemic Risk and Network Dynamics*, New Economic Windows,
DOI 10.1007/978-88-470-2553-0_12,

12.1 Introduction

The Kolkata Paise Restaurant (KPR) problem [1–5] is a repeated game, played between a large number (N) of agents having no interaction amongst themselves. In KPR problem, prospective customers (agents) choose from n restaurants each evening simultaneously (in parallel); N and n are both large and fixed (typically $n = N$). Each restaurant has the same price for a meal (hence no budget constraint for the agents). It is assumed that each can serve only one customer any evening (generalization to a larger value is trivial). The information regarding the customer distributions for earlier evenings is available to everyone. If more than one customer arrives at any restaurant on any evening, one of them is randomly chosen (each of them are anonymously treated) and is served, while the rest do not get dinner that evening. An alternative visualization can be one in which multiple customers arriving in a single restaurant have to share the food meant for one customer, keeping all of them unhappy. The utilization fraction $\bar{f}$ in the problem is defined as the average fraction of restaurants which were visited by people any evening in the steady state. Each agent develops its own (parallel) algorithm to choose the restaurant every evening such that he/she is alone there. Also, the times required to converge/settle to such a solution (if exists), should be low (faster than, say, $\log N$). If the restaurants have different ranks which are agreed upon by all the agents, additional complications may arise.

Paisa is the smallest monetary unit in Indian currency, and the use of the word would essentially be synonymous with anything that is very cheap. In Kolkata, there used to be very cheap and fixed rate "Paise Restaurant" which were popular among the daily labourers. During lunch hours, the labourers used to walk (to save the transport costs) to one of these restaurants and would miss lunch if they got to a restaurant where there were too many customers. Walking down to the next restaurant would mean failing to report back to work on time! There were indeed some well-known rankings of these restaurants, as some of them would offer tastier items compared to the others. A more general example of such a problem would be when the society provides hospitals (and beds) in every locality but the local patients go to hospitals of better rank (commonly perceived) elsewhere, thereby competing with the local patients of those hospitals. Unavailability of treatment in time may be considered as a lack of service for those people and consequently as (social) wastage of service by those unvisited hospitals.

A dictator's solution to the KPR problem is the following: everyone is asked to form a queue and is assigned a restaurant with rank matching the sequence of the person in the queue on the first evening. Then each person is asked to go to the next ranked restaurant in the following evening, thus for the person in the last ranked restaurant this means going to the first ranked restaurant. This shift process continues for successive evenings, thus providing clearly the most efficient solution (with utilization fraction $\bar{f}$ of the services by the restaurants equal to unity) and the system arriving at this solution trivially and immediately (from the first evening itself). However, in reality this cannot be the true solution of the KPR problem, where each agent decides on his own (in parallel and democratically) every evening, based on complete information about past events. In this game, the customers try to evolve a learning strategy to eventually get dinners at the best possible ranked restaurant, avoiding the crowd. It is seen that these strategies take considerable time to converge and even after that the eventual utilization fraction $\bar{f}$ is far below unity.

12.2 Kolkata Paise Restaurant Problem

In this review, we will talk about the KPR problem where N agents are parallelly visiting n restaurants on every day [n, $N \to \infty$; keeping n/N finite]. Each agent has been trying to get food from the best rank restaurants every day. But, each day, one agent can visit one restaurant and every restaurant has the capacity to serve food for one customer per evening. Therefore, as mentioned before, many agents go to a particular restaurant then one of the agents will be randomly chosen and will be served and the rest of the agents will not get dinner for that day, thus satisfying one of them. An alternative picture is one in which many customers have to share the food served for one customer, leaving all of them unsatisfied. Generally one can see that a few of the restaurants are not visited by any of the agents on a particular evening and that many agents crowd in other restaurants and do not get dinner for the evening. The utilization fraction $\bar{f}$ in the problem is therefore given by the average fraction of restaurants which were visited by customers on any evening in the steady state.

We discuss the case where instead of deterministic strategies, if everyone follows stochastic strategies, then one gets not only to higher values of the utilization fraction, but also attains it in very small convergence time (usually of order $\log N$ or smaller).

In general in the KPR problem $n = gN$ and $N \to \infty$ and in its primitive version, $g = 1$ ($n = N$), while for general phase transition studies (see Sect. 12.3) one considers $g \leq 1$. For the Minority Game (see Sect. 12.4) $n = 2$ (with $N \to \infty$ as usual).

12.2.1 Random-Choice Case (Stochastic)

Suppose there are N agents and n restaurants. Any agent can select any restaurant with equal probability. Therefore, the probability that a single restaurant is chosen by m agents is a Poisson distribution in the limit $N \to \infty$, $n \to \infty$:

$$\begin{aligned}\Delta(m) &= \binom{N}{m} p^m (1-p)^{N-m}; \quad p = \frac{1}{n} \\ &= \frac{(N/n)^m}{m!} \exp(-N/n) \quad \text{as } N \to \infty,\ n \to \infty. \qquad (12.1)\end{aligned}$$

Therefore the fraction of restaurants not chosen by any agent is given by $\Delta(m=0) = \exp(-(N/n))$ and that implies that average fraction of restaurants occupied on any evening is given by [2]

$$\bar{f} = 1 - \exp(-N/n) \simeq 0.63 \tag{12.2}$$

for $n = N$ in the KPR problem.

12.2.2 Rank Dependent Strategies (Stochastic)

Let us now consider that all restaurants have a well defined rank (agreed by every agent) depending upon quality of food, services, etc. although price of a meal is same for all restaurants. Thus, all agents will try to get food from best rank restaurants. But since a restaurant can serve only one customer, it means that many of the agents in crowded restaurants will remain unsatisfied. Now, assume that any kth restaurant have rank k and any agent choses that restaurant with probability $p_k(t) = k^\zeta / \sum k^\zeta$ (here ζ is any natural number). Here we discuss the results for such kind of strategy.

If an agent selects any restaurant with uniform probability p then the probability that a single restaurant is chosen by m agents is given by

$$\Delta(m) = \binom{N}{m} p^m (1-p)^{N-m}. \tag{12.3}$$

Therefore, the probability that a restaurant with rank k is not chosen by any of the agents will be given by

$$\begin{aligned} \Delta_k(m=0) &= \binom{N}{0} (1-p_k)^N; \quad p_k = \frac{k^\zeta}{\sum k^\zeta} \\ &\simeq \exp\left(\frac{-k^\zeta N}{\widetilde{N}}\right) \quad \text{as } N \to \infty, \end{aligned} \tag{12.4}$$

where $\widetilde{N} = \sum_{k=1}^{N} k^\zeta \simeq \int_0^N k^\zeta \, dk = \frac{N^{\zeta+1}}{(\zeta+1)}$. Hence

$$\Delta_k(m=0) = \exp\left(-\frac{k^\zeta(\zeta+1)}{N^\zeta}\right). \tag{12.5}$$

Therefore the average fraction of agents getting dinner in the kth ranked restaurant is given by

$$\bar{f}_k = 1 - \Delta_k(m=0) \tag{12.6}$$

and the numerical estimates of $\bar{f}_k$ is shown in Fig. 12.1. Naturally for $\zeta = 0$, the problem corresponding to random choice $\bar{f}_k = 1 - e^{-1}$, giving $\bar{f} = \sum \bar{f}_k/N \simeq 0.63$ and for $\zeta = 1$, $\bar{f}_k = 1 - e^{-2k/N}$ giving $\bar{f} = \sum \bar{f}_k/N \simeq 0.57$.

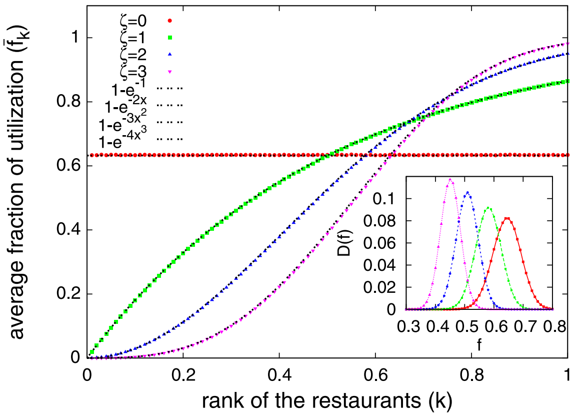

Fig. 12.1 The *main figure* shows average fraction of utilization ($\bar{f}_k$) versus rank of the restaurants (k) for different ζ values. The *inset* shows the distribution $D(f = \sum \bar{f}_k/N)$ of the fraction f agent getting dinner any evening for different ζ values. The simulations are done for $N = 10^4$ and $n = 10^4$. From [5]

12.2.3 Strict Crowd-Avoiding Case (Mixed)

We consider the case (see [4, 5]) where each agent chooses on any evening (t) randomly among the restaurants in which nobody had gone in the last evening ($t-1$). It was observed [5] that the distribution $D(f)$ of the fraction f of utilized restaurants is again Gaussian with a most probable value at $\bar{f} \simeq 0.46$. The explanation was given in the following way: As the fraction $\bar{f}$ of restaurants visited by the agents in the last evening is avoided by the agents this evening, the number of available restaurants is $N(1-\bar{f})$ for this evening and is chosen randomly by all the N agents. Hence, it fits with (12.1) by considering $(N/n) = 1/(1-\bar{f})$. Therefore, following (12.1),

$$(1-\bar{f})\left[1-\exp\left(-\frac{1}{1-\bar{f}}\right)\right] = \bar{f}. \tag{12.7}$$

The solution of this equation gives $\bar{f} \simeq 0.46$.

12.2.4 Stochastic Crowd Avoiding Case

Up to this point it is seen that indeed the random choice gives best utilization. Following a rank or strictly avoiding the crowd do not improve this fraction. While following a rank inherently prefers some restaurants and thereby making those crowed, the strict crowd avoidance on the other hand eliminates the possibility of a high utilization by not allowing repeated (successful) visits to a given restaurant.

However, in this section, we describe the following stochastic strategy: [5] If an agent goes to restaurant k on an evening $(t-1)$ then the agent goes to the same restaurant next evening with probability $p_k(t) = \frac{1}{n_k(t-1)}$ where $n_k(t-1)$ is the number of customers in kth restaurant on $t-1$ day's evening or otherwise choose any

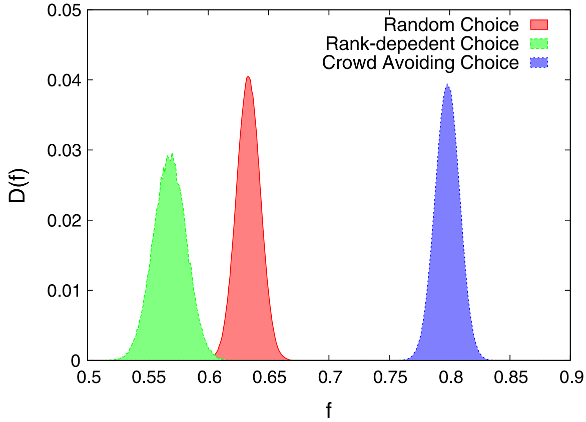

Fig. 12.2 The figure shows that distribution of utilization fraction in different condition of the KPR problem. All simulation data are shown for $N = 10^4$ and $n = 10^4$

other restaurant $k'(\neq k)$ with uniform probability. In this process, the average utilization fraction is $\bar{f} \simeq 0.8$ in the steady state and the distribution $D(f)$ is a Gaussian with peak at $f \simeq 0.8$ (see Fig. 12.2).

An approximate estimate of $\bar{f}$ can be made using the following argument: Let a_i denote the fraction of restaurants where exactly i agents ($i = 0, \ldots, N$) appeared on any evening and assume that $a_i = 0$ for $i \geq 3$. Therefore, $a_0 + a_1 + a_2 = 1$, $a_1 + 2a_2 = 1$ and hence $a_0 = a_2$. Given this strategy, a_2 fraction of agents will attempt to leave their respective restaurants in the next evening $(t + 1)$, while no intrinsic activity will occur at the restaurants where, nobody came (a_0) or only one came (a_1) in the previous evening (t). These a_2 fraction of agents will now get equally divided (each in the remaining $N - 1$ restaurants). Of these a_2, the fraction going to the vacant restaurants (a_0 in the earlier evening) is $a_0 a_2$. Hence the new fraction of vacant restaurants is now $a_0 - a_0 a_2$. In restaurants having exactly two agents (a_2 fraction in the last evening), some vacancy will be created due to this process, and this is equal to $\frac{a_2}{4} - a_2 \frac{a_2}{4}$. Steady state implies that $a_0 - a_0 a_2 + \frac{a_2}{4} - a_2 \frac{a_2}{4} = a_0$ and hence using $a_0 = a_2$ we get $a_0 = a_2 = 0.2$, giving $a_1 = 0.6$ and $\bar{f} = a_1 + a_2 = 0.8$. Of course, the above calculation is approximate as none of the restaurant is assumed to get more than two customers on any evening ($a_i = 0$ for $i \geq 3$). The advantage in assuming only a_0, a_1 and a_2 to be non vanishing on any evening is that the activity of redistribution on the next evening starts from this a_2 fraction of the restaurants. This of course affects a_0 and a_1 for the next evening and for steady state these changes must balance. The computer simulation results also conform that $a_i \leq 0.03$ for $i \geq 3$ and hence the above approximation does not lead to a serious error.

12.2.5 A General Study for Crowd Avoiding Case

The stochastic crowd avoiding case can be generalized by modifying the probability of an agent to choose the same restaurant as the previous evening as $p_i(t) =$

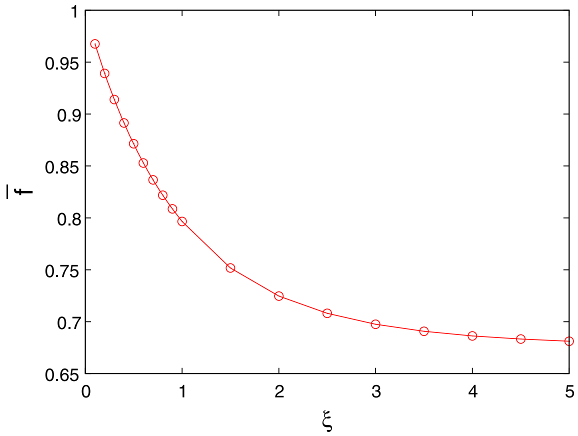

Fig. 12.3 The figure shows the average utilization fraction ($\bar{f}$) for different values of ξ. All simulation data are shown for $N = 10^4$ and $n = 10^4$

$1/n_i^{\xi}(t-1)$ where ξ is positive real number. Of course $\xi = 1$ is the case discussed in the previous section. It is observed (numerically) that the utilization fraction increases with decreasing ξ. However, the time to reach steady state value also increases. So, in this method we can reach a better utilization fraction as $\xi \to 0$ (Fig. 12.3). We observe, trivially, that the $\xi = 0$ case does not have any dynamics. On the other hand, the utilization fraction decreases to a limiting value ($\bar{f} \simeq 0.676$) for $\xi \to \infty$. The details of the critical behavior of this model will be reported elsewhere [6].

12.3 KPR and Phase Transition

Recently Ghosh et al. applied a stochastic crowd avoiding strategy in the KPR problem with considering gN agents and N number of restaurants [7]. It was observed that if the stochastic crowd avoiding strategy is applied to the problem then one can find out a particular value of $g = g_c$ below which all the agents are satisfied (and the state is called an absorbing or frozen state) and above the value of g_c, some of the agents will not be satisfied (and the state is called an active state). Therefore there is a phase transition between the an absorbing state and an active state with variation of g. The exponents of the transition in this process is well fitted with stochastic sandpile model.

12.3.1 The Models

Consider gN ($g < 1$) agents and the N restaurants. It is reminded that a restaurant can serve only one agent in an evening. Suppose in any evening a particular restaurant (ith restaurant) is visited by n_i agents and then one of the agents is chosen randomly and is served and rest $(n_i - 1)$ agents do not get any dinner for that day. Suppose all the agents are following the stochastic crowd avoiding dynamics mentioned before. Here two cases of the model are discussed (Model A & Model B). In model A, if any (ith) restaurant is visited by n_i agents in any evening then in the next evening each of the n_i agents will independently choose the same restaurant with probability $p = 1/n_i$ or a different restaurant otherwise with uniform probability. But in model B, if any (ith) restaurant is visited by n_i $(n_i > 1)$ agents in any evening then in the next evening all agents will independently choose any of the restaurants with uniformly probability $(p = 1/N)$. If, however, $n_i = 1$ then the agent will stick to his/her choice in the next evening. In both the models, one can find a value of $g = g_c$ below which all the agents will be getting food and when $g > g_c$, some of agents will not be satisfied. The order parameter is given by the steady state density of active sites ρ_a (density of sites having $n > 1$). So the absorbing phase corresponds to $\rho_a = 0$ $(g < g_c)$ whereas, for $g > g_c$ the steady state gives a non-zero value of the order parameter $(\rho_a > 0)$. Here the lattice versions (1D & 2D) models are also discussed.

12.3.2 Numerical Results

In this model one can see that below g_c the order parameter ρ_a goes to zero with time and above g_c, ρ_a goes to a stationary non zero value with time. Now, it is known that the evolution of order parameter is an exponential form and can be expressed as

$$\rho_a(t) = \rho_a^0\left[1 - e^{-t/\tau}\right] \tag{12.8}$$

for $g > g_c$, and

$$\rho_a(t) = \rho_a^0 e^{-t/\tau} \tag{12.9}$$

for $g < g_c$, where τ in the above expressions represents the relaxation time in the system. Therefore, the order parameter asymptotically goes to steady state value with time. Now, near critical point the order parameter can be scaled as $\rho_a \sim (g - g_c)^\beta$ where β is the order parameter exponent, similarly τ also scales as $\tau \sim (g - g_c)^{-\nu}$. A scaling form for ρ_a can be written as

$$\rho_a \sim t^{-\alpha} F\left(\frac{t}{\tau}\right); \qquad \tau \sim (g - g_c)^{-\nu} \sim L^z, \tag{12.10}$$

where L denotes size of the system and α, z are dynamic exponents near critical point. For time $t \to \infty$, and using (12.8), (12.9) and (12.10) we get a scaling relation $\beta = \nu\alpha$. The exponents have been obtained by numerical simulations and the scaling relations are also discussed.

12.3.3 Model A

12.3.3.1 Mean Field Case

The model in its original form (as discussed so far) is mean-field (i.e. infinite range) type, in the sense that the excess agents from a restaurant can choose from all the remaining restaurant in the next evening and the geometrical distance was not an issue. In the mean filed case, the simulations are done by taking system size $L = 10^6$ and different scaling exponents are estimated (see Fig. 12.4). The simulation results suggest that $g_c = 0.7502 \pm 0.002$ and $\beta = 0.98 \pm 0.02$. Also doing the data collapse it has been shown $z = 0.50 \pm 0.01$, $\nu = 1.00 \pm 0.01$ and $\alpha = 1.00 \pm 0.01$. Therefore, the scaling relation $\beta = \nu\alpha$ is satisfied by the estimated exponents for this case.

12.3.3.2 Lattice Cases

This model was also studied for 1-d and 2-d lattices. In 1-d, by studying the dynamics in the lattice it is meant that the excess agents can only go to the nearest neighbor sites in the next step. For 1-d, lattice size $N = L = 10^4$ have been taken and averaging over 10^3 initial conditions were permormed. For 2-d, a square lattice ($N = L^2$) with $L = 1000$ and averaging over 10^3 initial conditions were considered. Periodic boundary condition have been employed in both cases.

1. The model is defined for 1-d as follows: The agents are allowed to hop only to their nearest neighbor restaurants, and each agent can choose either left or right neighbor randomly. It is found that $g_c = 1$ and hence the phase transition is not very interesting.
2. In the 2-d version of the model, a square lattice is considered and the agents are to choose one of the 4 nearest neighbors randomly in next evening. For $N = 1000 \times 1000$, $g_c = 0.88 \pm 0.01$, $\beta = 0.68 \pm 0.01$, $z = 1.65 \pm 0.02$, $\nu = 1.24 \pm 0.01$ and $\alpha = 0.42 \pm 0.01$. It was observed that these independently estimated exponent values do not fit with the scaling relation $\beta = \nu\alpha$. However, this type of scaling violation was also observed previously in many active-absorbing transition cases [8].

12.3.4 Model B

12.3.4.1 Mean Field Case

For the mean field case, $N = 10^6$, averaging over 10^3 initial condition were taken. The phase diagram and the universality classes of the transition has been numerically investigated. In the mean field case, the phase boundary seems to be linear starting $g_c = 1/2$ for $p = 0$ and ending at $g_c = 1$ for $p = 1$ (Fig. 12.5), obeying $g_c = \frac{1}{2}(1 + p)$. In this case, for $p = 0$, $g_c = 1/2$, and this is similar to the fixed energy sandpiles [9–11]. Again the critical exponents are the same along the phase boundary and they match with those of model A.

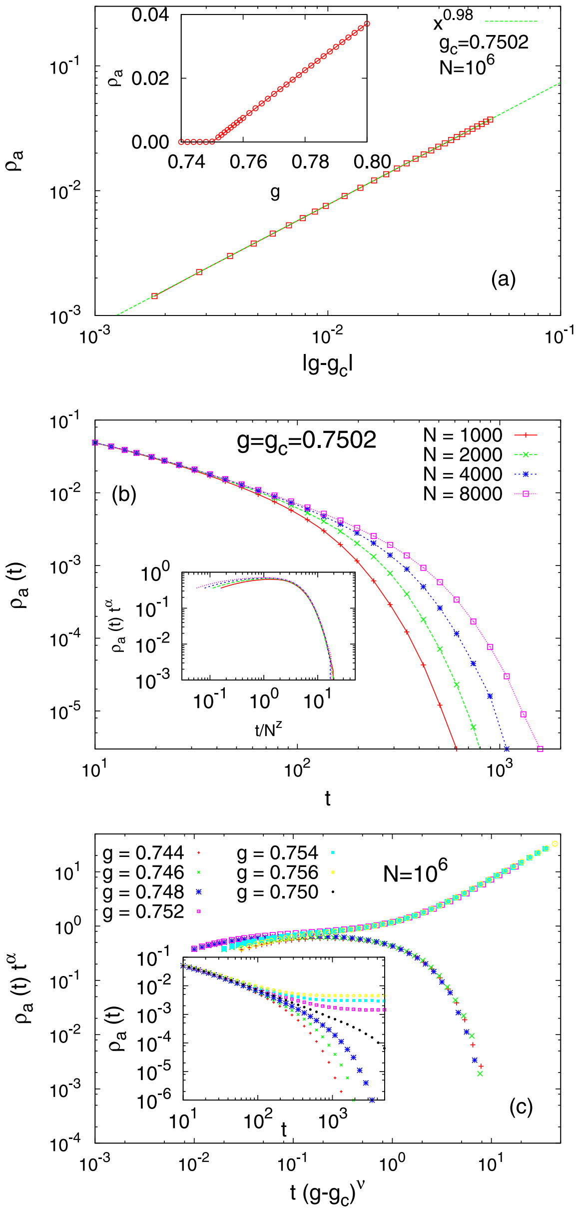

Fig. 12.4 Simulation results for mean field case, $g_c = 0.7502 \pm 0.0002$. (**a**) Variation of steady state density ρ_a of active sites versus $g - g_c$, fitting to $\beta = 0.98 \pm 0.02$. The *inset* shows the variation of ρ_a with density g. (**b**) Relaxation to absorbing state near critical point for different system sizes, the *inset* showing the scaling collapse giving estimates of critical exponents $\alpha = 1.00 \pm 0.01$ and $z = 0.50 \pm 0.01$. (**c**) Scaling collapse of $\rho_a(t)$. The *inset* shows the variation of $\rho_a(t)$ versus time t for different densities g. The estimated critical exponent is $\nu = 1.00 \pm 0.01$. The system sizes N are mentioned. From [7]

12.3.4.2 Lattice Cases

This model was also studied for 1-d and 2-d lattices. For a linear chain in 1-d, $N = L = 10^4$ and average over 10^3 initial condition were considered. For 2-d, square

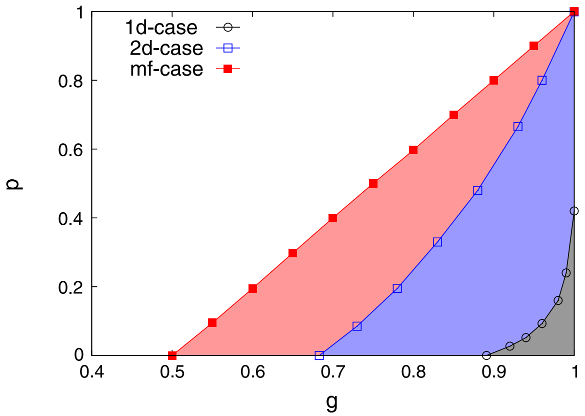

Fig. 12.5 Phase diagram for the generalized model in the (g, p) plane, showing the phase boundaries separating the active and absorbing phases in 1-d, 2-d and mean field cases. The active phases are on the right of the phase boundaries while the absorbing phases are on the left in the respective cases. The system sizes are $N = 10^5$ for mean field, 1000×1000 for 2-d, and 10^4 for 1-d. From [7]

restaurants (lattice) with $L = 1000$ and averaging over 10^3 initial conditions were considered.

1. For 1-d, for the case $p = 0$, $g_c = 0.89 \pm 0.01$, with $\beta = 0.42 \pm 0.01$, $z = 1.55 \pm 0.02$, $\nu = 1.90 \pm 0.02$ and $\alpha = 0.16 \pm 0.01$. The phase boundary in (g, p) is nonlinear: it starts from $g_c = 0.89 \pm 0.01$ at $p = 0$ to $p = 0.43 \pm 0.03$ at $g = 1$ (Fig. 12.5). Thus, one can independently define a model at unit density $(g = 1)$ and calculate the critical probability p_c for which the system goes from an active to an absorbing phase.
2. For 2-d, for the case $p = 0$, $g_c = 0.683 \pm 0.002$, with $\beta = 0.67 \pm 0.02$, $z = 1.55 \pm 0.02$, $\nu = 1.20 \pm 0.03$ and $\alpha = 0.42 \pm 0.01$. The phase boundary seems nonlinear, from $g_c = 0.683 \pm 0.002$ for $p = 0$ (Fig. 12.5) extending to $g_c = 1$ at $p = 1$.

In summary, it is shown how a crowd dynamics in a resources allocation game gives rise to a phase transition between an active and a frozen phase, as the density varies. In this respect, a class of models has been defined and studied, where gN agents compete among themselves to get the best service from N restaurants of same rank, generalizing the 'Kolkata Paise Restaurant' problem. In the original problem, where density $g = 1$, the model was far from its critical value g_c, the relaxation time τ, given by (12.10) never showed any $L = N^{1/d}$ dependence. As long as $g \leq g_c$, absorbing frozen configurations are present, and whether that can be reachable or not, depends on the underlying dynamics. The existence of a critical point g_c above which the agents are unable to find frozen configurations was found. In the case in which the agents are moving if and only if they are unsatisfied (model B) with $p = 0$, they fail to reach satisfactory configurations if the density is above $g_c = 1/2$. Strategies where agents wait longer (higher p) speed up the convergence, increasing g_c and decreasing the time to reach saturation configurations (faster-is-slower effect). The exponent values of the phase transitions in finite dimensions are in good agreement with the exponents of stochastic fixed-energy sandpile (Table. 12.1) [9–13]. Thus, it is a simple model for resource allocation, which is solvable (the MF

Table 12.1 The table shows that comparison of the critical exponents of this model with those of the conserved Manna model [13]

		Model A	Model B	Manna
β	1D		0.42 ± 0.01	0.382 ± 0.019
	2D	0.68 ± 0.01	0.67 ± 0.02	0.639 ± 0.009
	MF	0.98 ± 0.02	0.99 ± 0.01	1
z	1D		1.55 ± 0.02	1.393 ± 0.037
	2D	1.65 ± 0.02	1.55 ± 0.02	1.533 ± 0.024
	MF	0.50 ± 0.01	0.50 ± 0.01	2
α	1D		0.16 ± 0.01	0.141 ± 0.024
	2D	0.42 ± 0.01	0.42 ± 0.01	0.419 ± 0.015
	MF	1.00 ± 0.01	1.00 ± 0.01	1
ν	1D		1.90 ± 0.02	1.876 ± 0.135
	2D	1.24 ± 0.01	1.20 ± 0.03	1.225 ± 0.029
	MF	1.00 ± 0.01	1.00 ± 0.01	1

limit), and shows a variety of interesting features including phase transitions as in well known models.

12.4 KPR and Its Application on MG

So far we have dealt with the cases where the number of choices and the number of agents making those choices are of comparable magnitudes (KPR problem). However, there is another very well studied limit where the number of agents remain large but the number of choices is only two. A pay-off is given to the agents belonging to the minority group. Given there is no dictator and the agents do not communicate among themselves, how to device a strategy to extract maximum gain for maximum number of people, has been a long standing question. This problem goes by the name Minority Game (MG). This is, in fact, a particular version of the El Farol bar problem introduced by Brian Arthur [14].

In MG, the total number of agents (N) being odd, the maximum possible utilization can come when $(N-1)/2$ agents are in the minority. However, if the agents choose randomly, the utilization is far from the maximum value, in fact the deviation is of the order of $\sqrt{N}$. However, there can be deterministic strategies, where agents learn from their past experiences and in those cases this fluctuation can be considerably reduced, giving a self-organized, efficient market [15–20]. But in all those cases, the fluctuations (deviation from maximum utilization) scales with system size as $\sqrt{N}$. Only the pre-factor, depending upon the particulars of the strategy, can be reduced.

Recently, Dhar et al. [21] applied a stochastic strategy, inspired by the stochastic strategy used in KPR [2, 4, 5], to show that the fluctuations, or deviation from maximum utilization, can be reduced to be of the order of N^{ϵ} for any $\epsilon > 0$ in $\log \log N$

time. Stochastic strategy was used in MG before [22], where the fluctuation could be made of the order 1, but the time to reach that state scaled with $\sqrt{N}$. The strategy taken by Dhar et al., is the first of its kind that gives smallest fluctuation in very short time. In the following sections we discuss the main results of this strategy and its subsequent modifications.

12.5 KPR Strategy in MG: Results

As mentioned before, Minority Game deals with N (odd) agents selecting between two choices, when an incentive is associated with people belonging to minority. For example, consider the situation where there are only two restaurants in a locality and $N = 2M + 1$ agents select between these two restaurants for dinner. An agent is happy if he or she goes to the less crowded restaurant. But they cannot communicate among themselves and cannot change their choices once they fix it for a given evening. The agents, however, have in their possession the entire history of which restaurant was more crowded. This is a classic example to the MG problem. Other examples can be buying or selling of stocks and so on.

For any configuration at time (day) t, one can write the populations in the two restaurants as, $M - \Delta(t)$ and $M + \Delta(t) + 1$. In the this strategy, a deviation from the classic MG problem was made that the knowledge of $\Delta(t)$ was also available to the agents, while originally only its sign was known. In that sense, agents have more information than usual.

The strategy of the agents is as follows: At $t = 0$ the agents select randomly. Then the agents belonging to the minority stick to their choice in the next day. But the agents in the majority change their choice with a probability

$$p = \frac{\Delta(t)}{(M + \Delta(t) + 1)} \tag{12.11}$$

for $\Delta(t) > 0$ and stick to their choice with probability $1 - p$. As it is a probabilistic strategy, the number of people shifting will also have a fluctuation of the order $\sqrt{\Delta(t)}$, which is the new difference between the two populations; which leads us to the recurrence relation $\Delta(t + 1) = \sqrt{\Delta(t)}$. This shows that after $\log \log N$ time $\Delta(t)$ becomes of the order 1 and remains there.

Therefore, by following the same stochastic strategy, the difference between the populations in the two restaurants can be minimized in a very short time. This is in contrast with standard MG strategies, where the agents indeed try to differ in their strategies to maximize individual gain. However, the difference being in general the strategies were deterministic, i.e. given a history, all the subsequent steps are known. The stochasticity itself makes the agents differ. Furthermore, that the agents follow the same stochastic strategy and do not deviate from it, can be justified if it can be shown that a single individual does not gain by deviating from this strategy. Indeed it was shown that for this strategy, an individual will not gain by deviating from this strategy.

12.5.1 *Stability Against Individual Deviator*

In the above discussions it is discussed that if the agents in the MG problem follow a simple stochastic strategy, the difference between the two choices can be made of the order 1 in $\log\log N$ time. However, it is not always expected that all the agents will follow the same strategy, until it is shown that no one will gain by deviating from the strategy.

12.5.1.1 Game with One Cheater

Defining *cheater* as one who will not follow the strategy followed by rest present in the majority. Now suppose there is a cheater, say X_1 in the majority, say in restaurant A. If he want to stay, then the number of agents in the restaurant A, who will follow the conventional strategy is $M + \Delta(t)$. The probability that $\tilde{r}$ agents from $M + \Delta(t)$ agents in A will shift from A to minority, say restaurant B is

$$P(\tilde{r}) = \binom{M + \Delta(t)}{\tilde{r}} p^{\tilde{r}} (1 - p)^{M+\Delta(t)-\tilde{r}}. \tag{12.12}$$

For $M \to \infty$ the probability distribution will become Poisson with $\lambda = p(M + 1)$. So this distribution will be

$$P(r) = \frac{\lambda^r}{r!} \exp(-\lambda)(1 + Br), \tag{12.13}$$

where

$$B = \frac{\lambda}{M} - \left(\frac{\lambda^3}{2} + \lambda^2\Delta - \frac{\lambda^2}{2}\right)\frac{1}{M^2}. \tag{12.14}$$

Using the above probability distribution, it can be shown that [21] there exist a value of λ for a given $\Delta(t)$ such that existence of cheater does not effect the dynamics of the game. This λ is given by

$$\lambda - \Delta = \frac{1}{6} + \frac{\lambda^2}{2M}\sqrt{\frac{\lambda}{\Delta}}\left(1 + \frac{\Delta}{\lambda}\right). \tag{12.15}$$

Or restoring the inequality given that X_1 will gain switching as he is in majority then we get

$$\lambda < \Delta + \frac{1}{6} + \frac{\lambda^2}{2M}\sqrt{\frac{\lambda}{\Delta}}\left(1 + \frac{\Delta}{\lambda}\right). \tag{12.16}$$

As $\lambda \propto \Delta$, this means for a large difference Δ we can increase the noise safely up to $\frac{1}{6}+$ without letting the cheater to win. We have seen in Fig. 12.6 that (12.15) match the simulation result. In the simulation we took $p = \frac{\Delta+c}{M+\Delta+1}$, with vary-

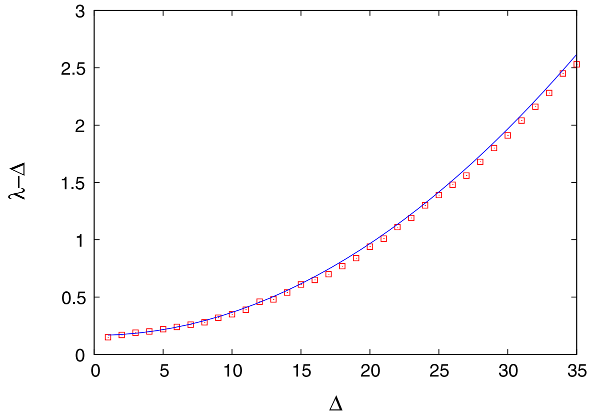

Fig. 12.6 The data points are the simulation data and the line is (12.15). The total number of player is 2001

ing the noise parameter c. Below this optimal value of λ, cheater will gain if he shift from majority to minority, above this optimal value a cheater will gain if he shift from minority to majority.

12.5.1.2 Majority Stay or Minority Flip

For a situation when one agent will stay if he finds himself in majority (in A) and will shift if he finds himself in minority (in B). Then he will win by staying in majority if r number of agent shift from majority A to minority B, given $r \geq \Delta + 1$. The total probability $P(win \,|\, stay\ in\ majority)$ that he will win, which is same as expected payoff is

$$EP(majority \,|\, stay) = \sum_{r=\Delta+1}^{\infty} P(r). \tag{12.17}$$

Now, if he is in B having total number of agent $M - \Delta$ and shift to A having total number of agent $M + \Delta + 1$, he will win if r number of agent come from A to B, with $r \geq \Delta + 2$. The total probability of his win if he flip is $P(win \,|\, flip\ from\ minority)$, which is same as his payoff given by

$$EP(minority \,|\, flip) = \sum_{r=\Delta+2}^{\infty} P(r) \tag{12.18}$$

where $P(r)$ is given by (12.13). Total probability of win or expected payoff, if he stay at majority and flip if he is in minority is sum of (12.18) and (12.17), which after little algebra is given by

$$EP_I = 2\left(1 - \frac{\Gamma(\Delta + 1, \lambda)}{\Delta!}\right) - \frac{\lambda^{\Delta+1}}{(\Delta + 1)!} \tag{12.19}$$

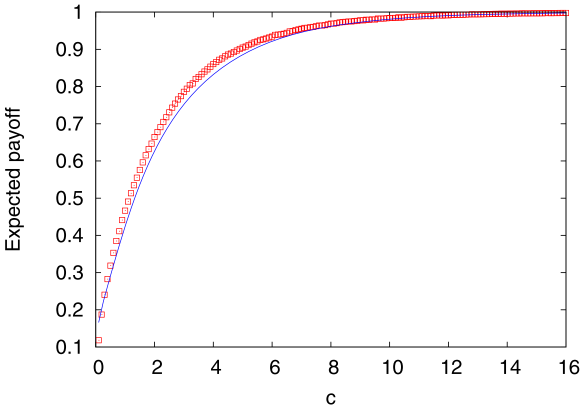

Fig. 12.7 The simulation data is compared with the solid line which is (12.19) with given in (12.21) neglecting the small correction term in the bracket, thus $\lambda = 2c$. The total number of player is 2001

where $\Gamma(s, x)$ is a *incomplete gamma function*. To get more accurate result we need to average the expected payoff (EP_I). From numerical experiment we know that fluctuation in Δ is very small. So this gives very little error when fitted with the simulated result. This error can be minimized by little adjustment of the constant terms. The best fit will come for the first argument $\Delta + a$ where $a \neq 1$, but $a = 0.65$. and the second argument $\lambda = 2\Delta$ in the $\Gamma(s, x)$. From Fig. 12.7 we find that the noise c can not be increased to very large value else a cheater will always gain the game.

We have seen that if we take $\Delta = 0$ so that $\lambda = c$, the noise parameter, then the curve have *s*ame features as the simulated curve, this is due to the fact that Δ does not become zero in the presence of non zero noise. So we need to know the average Δ in this case, which is given by

$$\langle \Delta \rangle = \frac{1}{2}\lambda\left(1 - \frac{\Gamma(\Delta_0, \lambda)}{(\Delta_0 - 1)!}\right). \tag{12.20}$$

This is the average difference if A become minority after a shift of agents. So we get

$$\lambda = 2\langle \Delta \rangle\left(1 - \frac{\Gamma(\Delta_0, \lambda)}{(\Delta_0 - 1)!}\right)^{-1}. \tag{12.21}$$

12.5.1.3 Minority Stay or Majority Flip

Let the cheater is in A having $M + \Delta + 1$ agents who will shift to B there are $M - \Delta$ agents. He will shift to B making $n_A = M + \Delta$ and $n_B = M - \Delta + 1$. Now he will win if r number of people from A shift to B with $r \leq \Delta - 1$. Then the probability that he will win is given by $P(win\,|\,\textit{flip from majority})$. Now if he is in B, then he will stay. If $r \leq \Delta$ number of people shift from A to B he will win. The probability that he will win staying in B is $P(win\,|\,\textit{stay in minority})$. The total probability of

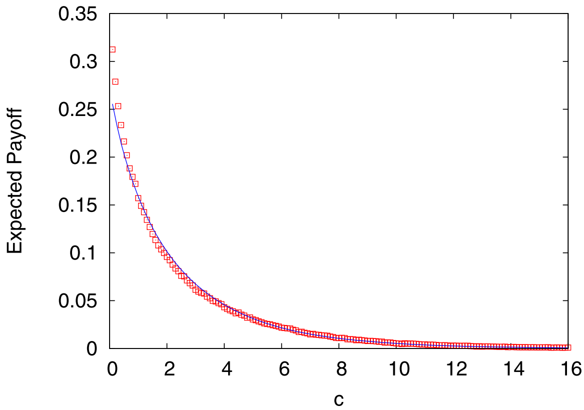

Fig. 12.8 The symbols represent the simulation data and the line is (12.22) with λ given in (12.23). In the plot average Δ is not the noise parameter c but a little less, so instead of $\lambda = 2(c+1)$, $\lambda = 2c + 1.85$ is plotted in the theoretical curve. The total number of player is 2001

winning if he always stays in minority, which is same as expected payoff E_{II} in the same case

$$EP_{II} = \sum_{r=0}^{\Delta-1} P(r) + \sum_{r=0}^{\Delta} P(r) \approx 2\frac{\Gamma(\Delta, \lambda)}{(\Delta-1)!} + \frac{\lambda^{\Delta}}{\Delta!}\exp(-\lambda). \tag{12.22}$$

We have seen that if we take $\Delta = 0$ so that $\lambda = c$, the noise parameter, then the curve have *s*ame features as the simulated curve (see Fig. 12.8), this is because Δ does not become zero in the presence of non zero noise. So we need to know the average Δ in this case which is

$$\lambda = 2\big(\langle\Delta\rangle + 1\big). \tag{12.23}$$

12.5.2 Freezing of Dynamics and Escape Routes from It

It is clear from the strategy discussed above, that once $\Delta(t) = 0$ i.e., the difference of population in the two restaurants is 1 (which is the minimum possible value as the total number is odd), the dynamics stops. This leaves the system highly asymmetric in the sense that the people in the majority (minority) will remain in the majority (minority) forever. This situation is of course socially unacceptable, although this is the most efficient division.

12.5.2.1 Resetting After a Given Time

To resolve this status quo, Dhar et al. [21] suggested that once $\Delta(t) = 0$ a major reshuffle can take place if all the agents (whether in majority or in minority) shift after waiting T time steps. This time period needs to be much smaller than the lifetime of the agents. If an inefficiency parameter is defined as follows

$$\eta = \lim_{N\to\infty} \frac{4}{N}\left\langle (r - N/2)^2 \right\rangle, \tag{12.24}$$

then for this case this would be

$$\eta \simeq \frac{K_1 N^{\epsilon-1}}{T + K_2 \log\log N} \tag{12.25}$$

where K_1, K_2 and ϵ are constant. This means that efficiency increases with T. However, large T would mean longer wait in the majority. Clearly, other parameters like overall social welfare and equality needs to be considered here. Also, as indicated in the Dhar et al. paper, it will be interesting to see what if the agents try to maximize their pay-offs for next $n > 1$ days.

12.5.2.2 Continuous Transition of Social Efficiency

In the above method, the system becomes efficient only when the agents act for overall social welfare or have a long-term gain strategy. Even then, efficiency depends upon time waiting time T, which gives rise to a competition regarding its magnitude.

Biswas et al. [23] suggested a subsequent modification in the strategy such that the fluctuation could be reduced to any arbitrarily small value by tuning a parameter. This, therefore, gives a continuous phase transition and as long as a finite fluctuation is kept in the system, the frozen condition can be avoided.

The modified strategy is the following: The agents in the majority in a given day shifts to the other choice with a probability

$$p_+(t) = \frac{\Delta'(t)}{M + \Delta'(t) + 1}, \tag{12.26}$$

(where $\Delta'(t) = q\Delta(t)$ and q is a constant) and people in the minority stick to their choices ($p_- = 0$).

Regarding the steady-state behavior, consider the following: Suppose the populations in the majority and minority are $M + \Delta(t)$ and $M - \Delta(t)$ respectively, at time t. Now, if $2\Delta(t)$ number of people can be shifted from majority to minority, then the population difference will remain same and the same process can be repeated, sustaining a steady state. Of course, this possibility can only arise when $q > 1$. If Δ_s is the steady state value for fluctuation, then

$$(M + \Delta_s + 1)\frac{q\Delta_s}{M + q\Delta_s + 1} = 2\Delta_s. \tag{12.27}$$

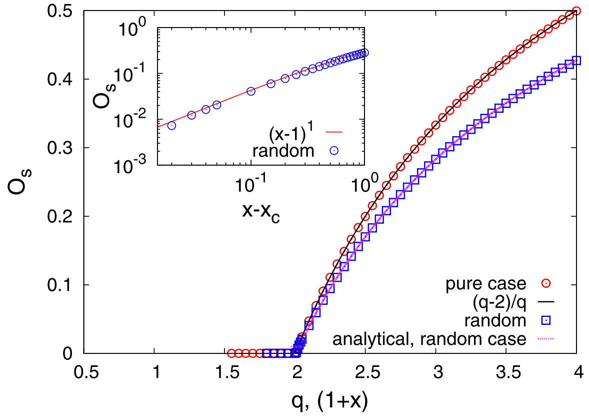

Fig. 12.9 Steady state values of the order parameter O_s for different values of q and x. The *solid lines* show the analytical results for the pure and annealed disordered cases. Both match very well with the simulation points. *Inset* shows the log–log plot near the critical point for the disordered case, confirming $\beta = 1.00 \pm 0.01$. All simulation data are shown for $M = 10^5$. From [23]

The two solutions of this quadratic equation are

$$\Delta_s = 0 \quad \text{or} \quad \frac{q-2}{q}(M+1). \tag{12.28}$$

This means that for $q < q_c = 2$, the system will reach the zero fluctuation state (although the dynamics of the system will be very much different for $q < 1$ and $q > 1$), and for $q > 2$ there will remain a residual fluctuation in the system signifying an active-absorbing type phase transition around $q = q_c = 2$.

Formally, one can define an order parameter like $O(t) = \Delta(t)/M$ and in the steady state the saturation value is $O_s = 0$ when $q < 2$ and $O_s = (q - q_c)/q$ for $q > 2$ both for $M \gg 1$, giving the order parameter exponent $\beta = 1$. Figure 12.9 shows the numerical results and its comparison with the above calculations.

Regarding the dynamics of the system in approaching this steady state, assume that at time t the populations at the two restaurants are $N_A(t)$ and $N_B(t)$ and $N_A(t) > N_B(t)$. Therefore,

$$\Delta(t) = \frac{N_A(t) - N_B(t) - 1}{2}. \tag{12.29}$$

Now, according to the strategy in (12.26), the number of people shifted from choice A to choice B will be

$$\begin{aligned} S(t) &= \frac{q\Delta(t)}{M + q\Delta(t) + 1}\big(M + \Delta(t) + 1\big) \\ &\approx q\Delta(t) \end{aligned} \tag{12.30}$$

up to leading order term, when $\Delta(t) \ll M$, i.e., when q is close to q_c, or in the long time limit if $q < q_c$ and not too close to it. With this transfer amount, in the next step $N_A(t+1) = N_A(t) - S(t)$ and $N_B(t+1) = N_B(t) + S(t)$. For $q > 1$, majority will become minority, so

$$\Delta(t+1) = \frac{N_B(t+1) - N_A(t+1) - 1}{2}$$
$$\approx q\Delta(t) - \Delta(t) - 1. \tag{12.31}$$

Subtracting $\Delta(t)$ from both sides and dividing by M, one arrives at

$$\frac{dO(t)}{dt} = -(2-q)O(t) - \frac{1}{M}. \tag{12.32}$$

The last term can be neglected for large M. The it follows

$$O(t) = O(0)\exp\bigl[-(2-q)t\bigr]. \tag{12.33}$$

So this exponential decay in the region $1 < q < 2$ gives a time scale $\tau \sim (q_c - q)^{-1}$, diverging at the critical point with exponent 1.

In (12.30), if one keeps the second order term, one gets

$$S(t) \approx q\Delta(t) - \frac{1}{M}\bigl(q^2\Delta^2(t) - q\Delta^2(t)\bigr). \tag{12.34}$$

The time evolution equation becomes

$$\frac{dO(t)}{dt} = -(2-q)O(t) - q(q-1)O^2(t). \tag{12.35}$$

Now, exactly at the critical point $q = 2$, the solution is

$$O(t) = \frac{O(0)}{2O(0)t + 1}, \tag{12.36}$$

which, in the long time limit gives $O(t) \sim t^{-1}$, giving the critical exponent value $\alpha = 1$.

A more general solution of (12.35) can be obtained (for any q) as follow: Consider the auxiliary variable $u(t) = |q-1|^t/O(t)$ and substitute it in (12.35). This gives after simplifications

$$u(t+1) = u(t) + q|q-1|^t. \tag{12.37}$$

Using this recursion relation, one can write $u(t)$ in a GP series and can perform the sum to get the following:

$$O(t) = \frac{1-|q-1|}{q} \frac{1}{(\frac{1-|q-1|}{qO(0)} + 1)|q-1|^{-t} - 1}. \tag{12.38}$$

Putting $q = 2$ in the above equation, one gets back (12.36). Also, a time scale is obtained from the above equation in the form

$$\tau \sim \frac{1}{|\ln(|q-1|)|}. \tag{12.39}$$

As $q \to q_c$, the power law divergence $(q_c - q)^{-1}$ is recovered.

Furthermore, for $q < 1$ the dynamical equation (12.38) will reduce to

$$O(t) \sim \frac{O(0)}{O(0)+1}(1-q)^t. \tag{12.40}$$

12.5.3 Reducing Fluctuation with Less Informed Agents

As is clear from the strategies discussed above, the agents in those versions of the game, posses more information than the usual minority game problem. Particularly, the agents are aware of the amount of excess population in the majority, while in the usual case then only know whether they were in the majority or minority. This extra information is important. Although it is logical that the would agents eventually come to know about this excess population, there have been studies to confirm if this extra information is essential in obtaining the maximum efficient state. It is found that this information is not essential. The system can indeed reach the maximum efficient state even when this knowledge is partially or even fully absent.

12.5.3.1 Non-uniform Guessing of the Excess Crowd: Phase Transition

It has been argued in Ref. [23] that in considering less informed agent a natural step would be the agents with different guessing abilities. This means that although the agents do not know the exact value of the excess population, they can make a guess about the value. This acts as an annealed disorder. Formally, the ith agent at time t makes a guess about $\Delta(t)$ which is

$$\Delta_i(t) = \Delta(t)(1+\epsilon_i), \tag{12.41}$$

where ϵ_i is an annealed random variable taken from a uniform distribution in the range $[0:2x]$. This means,

$$\big\langle \Delta_i(t)\big\rangle = \Delta(t)\big(1+\langle\epsilon_i\rangle\big) = \Delta(t)(1+x), \tag{12.42}$$

where the angular brackets denote average over disorder. One can generally write

$$\Delta(t+1) = \big|\Delta(t) - S(t)\big|, \tag{12.43}$$

where

$$S(t) = \left|\left\langle \frac{\Delta(t)(1+\epsilon)}{M+\Delta(t)(1+\epsilon)}\right\rangle\right|. \tag{12.44}$$

This leads to

$$O(t+1) = O(t)\left|\left\langle \frac{\epsilon}{1+(1+\epsilon)O(t)}\right\rangle\right|. \tag{12.45}$$

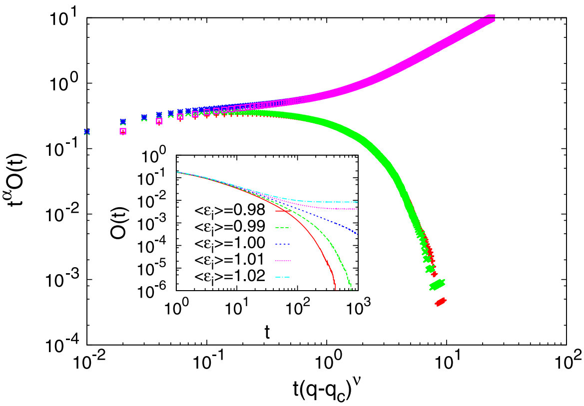

Fig. 12.10 Data collapse for finding ν in the disordered case for different x values. The estimate is $\nu = 1.00 \pm 0.01$. *Inset* shows the uncollapsed data. The *straight line* at the critical point gives $\alpha = 1.00 \pm 0.01$. Simulation data is shown for $M = 10^6$. From [23]

In the steady state $O(t+1) = O(t) = O^*$, leading to

$$\frac{(1-O^*)2xO^*}{(1+O^*)} = \ln\left[1 + \frac{2xO^*}{1+O^*}\right]. \tag{12.46}$$

A numerical solution of this self-consistent equation was found to agree with the simulation results (see Fig. 12.9). For small O^*, $O^* \sim (x-1)$ giving $\beta = 1$. Also, for small $O(t)$, the dynamical equation can be written as

$$\frac{dO(t)}{dt} = (x-1)O(t) - xO^2(t). \tag{12.47}$$

The critical point is at $x_c = 1$. So at the critical point, $O(t) \sim t^{-1}$, giving $\alpha = 1$ and above the critical point the exponential decay would give a time scale, diverging at $x = x_c$ with an exponent $\nu = 1$.

The above results were also verified using numerical simulations. A finite size scaling form was considered

$$O(t) \approx t^{-\alpha}\mathcal{F}\big(t^{1/\nu}(q-q_c), t^{d/z}/N\big), \tag{12.48}$$

where d is the spatial dimension, which was taken as 4 in this mean-field scenario. This form suggests that at the critical point the order parameter decays in a power-law, with exponent α, which was numerically found to be 1.00 ± 0.01 (see inset of Fig. 12.10). One can also plot (see Fig. 12.10) $O(t)t^{\alpha}$ against $t(q-q_c)^{\nu}$, where by knowing α, ν can be tuned to get best data collapse, giving $\nu = 1.00 \pm 0.01$. Also, $O(t)t^{\alpha}$ can be plotted against $t/N^{z/d}$, where z/d can be obtained from the data collapse (Fig. 12.11) to be 0.50 ± 0.01. Therefore, it was concluded that the analytical estimates were verified and the scaling relation $\alpha = \beta/\nu$ was satisfied.

In the above mentioned case, the non-uniform guessing power acts as an annealed disorder. When this disorder is quenched, the case slightly complicated. It is no longer possible to tackle analytically as done above. It was seen that the agents with higher ϵ are more likely to change side and be in the majority. So, if the average pay-offs are plotted against ϵ, a monotonic decay is observed (Fig. 12.12).

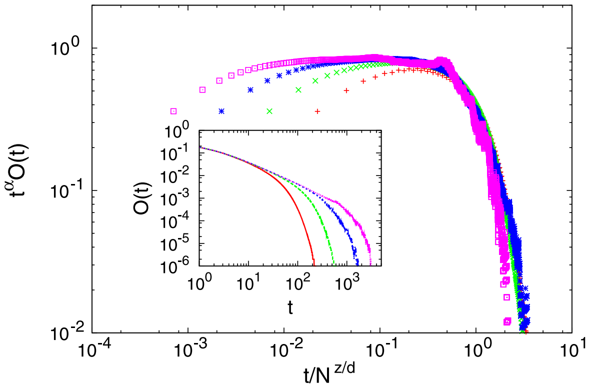

Fig. 12.11 Data collapse for finding z in the disordered case for different system sizes ($M = 10^3, 10^4, 10^5, 10^6$) at $x = 1.0$. The estimate is $z/d = 0.50 \pm 0.01$. *Inset* shows the uncollapsed data. The *linear part* in the *inset* confirms $\alpha = 1.00 \pm 0.01$. From [23]

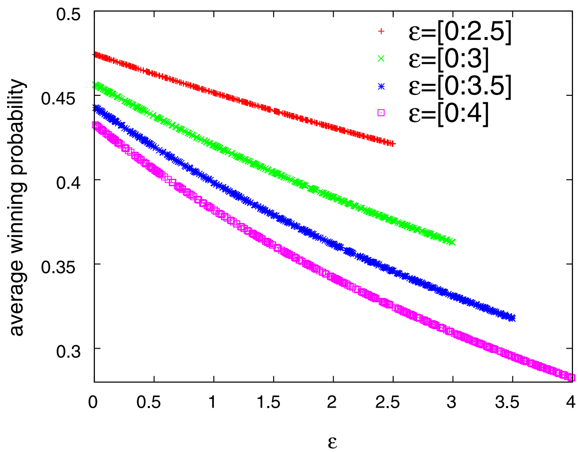

Fig. 12.12 For quenched ϵ_i the average pay-offs of the agents are plotted for different ϵ values having different ranges as indicated. The monotonic decay with increasing ϵ clearly indicates that agents with higher ϵ are more likely to be in the majority. From [23]

12.5.3.2 Following an Annealing Schedule

Usually in minority game, agents do not have any information about the amount of excess population in the majority. They are only aware whether they are in the minority or majority. All the strategies mentioned above require this information in some form (fully or partially). However, it was studied in Ref. [23] that even without this information, the system can reach the fully efficient state in $\ln N$ time.

In this case of least informed agents, the agents assume a simple time evolution for the excess population. An example can be

$$\Delta^T(t) = \Delta^T(0)\exp(-t/\tau), \tag{12.49}$$

where $\Delta^T(0)$ is close to $\sqrt{M}$, corresponding to the initial random choice. Assuming this form, one can plot the actual $\Delta(t)$ along with this trial function with time. They have a simple relation as follows:

$$2\Delta(t) = \Delta^T(t). \tag{12.50}$$

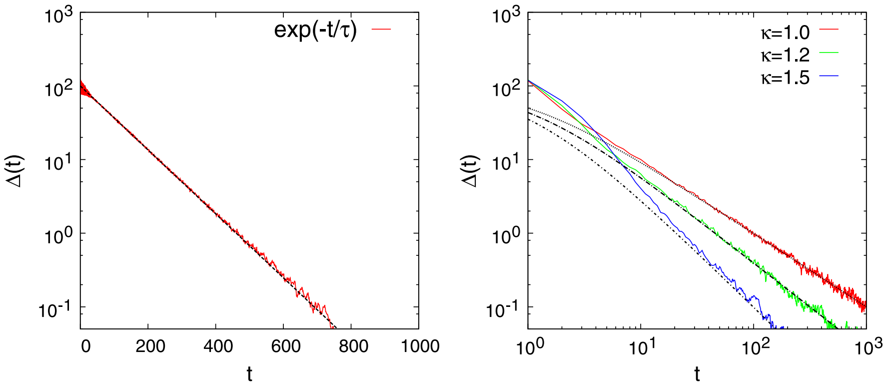

Fig. 12.13 Time variation of the excess population $\Delta(t)$ are plotted for different functional forms of $\Delta^T(t)$. *Left*: In log-linear scale the excess population are plotted for exponential decay. *Right*: For power law ($\Delta^T(0)/(1+t)^\kappa$ decay, with different values of κ). $M = 5 \times 10^3$ for the simulations. From [23]

This implies that even when the agents are completely unaware of the excess population, they can reach an efficient state ($\Delta(t) \sim 1$) in $\ln N$ time.

It was also checked in Ref. [23] if this process is specific to the functional form considered for the trial function. For this purpose a power-law decay was also considered

$$\Delta^T(t) = \frac{\Delta^T(0)}{(1+t)^\kappa}. \tag{12.51}$$

Again it was found that for different κ values, the relation in (12.50) is satisfied. It was therefore concluded that this relation is true for a wide range of the functional form (the restrictions in the functional form is discussed later).

The behavior of the order parameter when a trial function is considered, can be verified as follows: The dynamical evolution of $O(t)$ would be

$$O(t+1) = \frac{|\eta(t) - O(t)|}{1+\eta(t)}, \tag{12.52}$$

where $\eta(t) = \Delta^T(t)/M$. When $\eta(t) > O(t)$, one can obtain (by Taylor series expansion)

$$\frac{dO(t)}{dt} - \big[\eta(t) - 2\big]O(t) = \eta(t)\big[1 - \eta(t)\big]. \tag{12.53}$$

A general solution of the above equation will be of the form

$$O(t) = \frac{\int_0^t dt_1\, \eta(t_1)(1-\eta(t_1))e^{\int_0^{t_1}(2-\eta(t_2))\,dt_2}}{e^{\int_0^t (2-\eta(t_1))\,dt_1}} + C_1 e^{-\int_0^t (2-\eta(t_1))\,dt_1}, \tag{12.54}$$

where C_1 is a constant. This is valid only when $\eta(t)$ is not a fast decaying function. When $\eta(t) < 2$, the dominant terms in the above equation is

$$O(t) \approx \frac{\eta(t)(1-\eta(t))}{2-\eta(t)} \approx \frac{\eta(t)}{2}, \tag{12.55}$$

which was the numerical observation (see Fig. 12.13). If one evaluates (12.54) using $\eta(t) = \eta_0 \exp(-t/\tau)$ for $\tau > 1/2$, one gets

$$O(t) \sim \frac{\tau}{2\tau - 1}\eta(t). \tag{12.56}$$

Therefore, $O(t) \approx \eta(t)/2$ is only valid when $\tau \gg 1/2$, which limits the fastness in the trial function.

When one considers a fast decaying trail function, one would simply have

$$O(t) \sim O(t-1) - \eta(t-1) \sim O(0) - \sum_{k=0}^{t-1} \eta(k). \tag{12.57}$$

So, $O(t)$ will saturate to a finite value in this case.

12.5.4 *Effect of Random Traders*

The above mentioned strategies concern with agents following a given strategy (this does not remove their heterogeneity, since these are stochastic strategies that involve uncorrelated random numbers). However, it is often the case in real markets that there exist agents who do not follow the market signals (fluctuations) in deciding their trade strategies. Whatever might be their logic, it terms of market signals, they can be treated as random traders who decide completely randomly as opposed to the chartists who follow given strategies (deterministic or stochastic). Following discussions deal with effect of such random traders in minority games.

12.5.4.1 Single Random Trader

Consider the scenario when there is only one random trader in the system. The other agents follow some strategy mentioned before, and reach the minimum fluctuation state. After that $\Delta(t) = 0$, so no chartist will shift from his or her choice. However, the single random trader will continue to shift on average in a 2 days time period. The majority will be determined by this random trader. Therefore, that random trader will always be the loser. Although the resource utilization will be perfect in this case, it will be at the cost of one player being in the majority for ever.

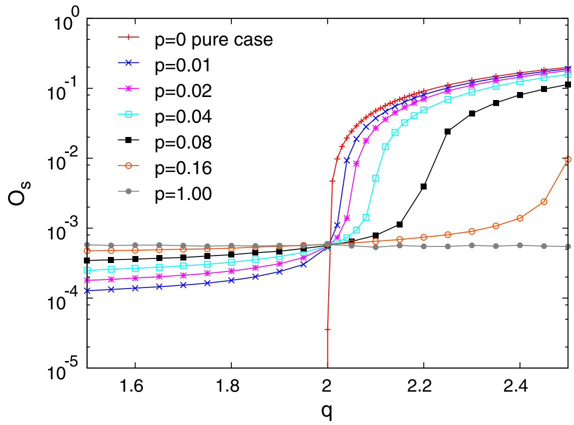

Fig. 12.14 The saturation values of O_s are plotted against q for different fractions p of the random traders. $M = 10^6$ for the simulations. From [23]

12.5.4.2 More than One Random Trader

The case of the single random trader has the problem that the random trader is always a loser. This makes the system unstable in the sense that resource allocation is unfair for that agent as long as he or she follows that strategy (random in this case). However, this problem can be avoided by considering more than one random player. In this case it is not always possible to keep all the random players in the majority, since the majority is no longer determined by a single random player. Also, as the average time period is 2 days for the random players, both the choices will become majority and minority in this time period (due to symmetry of the choices). It is true that random players would make the fluctuations to grow. If the number of random player is pN, then the fluctuation would scale as $\sqrt{N}$ (see Fig. 12.14). However, one can always keep the number of random players at a minimum value. If this number is 2, then the fluctuation would be minimum and uniform resource allocation is guaranteed.

12.6 Summary

We consider a repetitive game performed by N agents choosing every time (parallelly) one among the $n(\leq N)$ choices, such that each agent can be in minority: no one else made the same choice in the KPR case (typically $n = N$) and $N_k < N/2$ for the Minority Game ($n = 2$; $k = 1, 2$). The strategies to achieve this objective evolve with time bounded by N. Acceptable strategies are which evolve quickly (say within $\log N$ time). Also the effectiveness of a strategy is measured by the resulting utilization factor $\bar{f}$ giving the (steady state) number of occupied restaurants in any evening for the KPR, by the value of fluctuation Δ in the minority game case ($\Delta = 0$ corresponds to maximum efficiency).

The study of the KPR problem shows that a dictated solution leads to one of the best possible solution to the problem, with each agent getting his dinner at the best ranked restaurant with a period of N evenings, and with best possible value of $\bar{f}$ (=1) starting from the first evening on itself. For a democratic situation (for parallel decision strategies), the agents employ stochastic algorithms based on past occupation informations (e.g., of $N_k(t)$)). These strategies are of course less efficient ($\bar{f} \ll 1$; the best one discussed in [5], giving $\bar{f} \simeq 0.8$ only). Here the time required is very weakly dependent on N, if any. We also note that most of the "smarter" strategies lead to much lower efficiency.

Finally we note that the stochastic strategy Minority Game [21], a very efficient one: The strategy is described by (12.11), where the agents very quickly (in $\log\log N$ time; $N = 2M + 1$) get divided almost equally (M and $M + 1$) between the two choices. This strategy guarantees that a single cheater, who does not follow this strategy, will always be a loser. However, the dynamics in the system stops very quickly (leading to the absorbing state), making the resource distribution highly asymmetric (people in the majority stays there for ever) thereby making this strategy socially unacceptable. To rectify this, we noted that the presence of a single random trader (who picks between the two choices completely randomly) will avoid this absorbing state and the asymmetric distribution. However, this will always make that random trader a loser. But the presence of more than one random trader will avoid that situation too, making the average time period of switching between majority and minority for all the traders (irrespective of whether they are chartists or random traders) to be 2. Hence, the system will always evolve collectively such that only two agents will make random choices between the binary choices, while the rest $N - 2$ will follow the probabilities given by (12.11).

References

1. Chakrabarti BK, Chakraborti A, Chatterjee A (eds) (2006) Econophysics and sociophysics. Wiley-VCH, Berlin
2. Chakrabarti AS, Chakrabarti BK, Chatterjee A, Mitra M (2009) Physica A 388:2420
3. Ghosh A, Chakrabarti BK (2009) Kolkata Paise Restaurant (KPR) problem. http://demonstrations.wolfram.com/KolkataPaiseRestaurantKPRProblem
4. Ghosh A, Chakrabarti AS, Charabarti BK (2010) In: Basu B, Charabarti BK, Ghosh A, Gangopadhyay K (eds) Econophysics and economics of games. Social choices and quantitative techniques Springer, Milan
5. Ghosh A, Chatterjee A, Mitra M, Chakrabarti BK (2010) New J Phys 12:075033
6. Ghosh A, Chatterjee A, Chakrabarti BK (in preparation)
7. Ghosh A, Martino DD, Chatterjee A, Marsili M, Chakrabarti BK (2012) Phys Rev E 85:021116
8. Rossi M, Pastor-Satorras R, Vespignani A (2000) Phys Rev Lett 85:1803
9. Dickman R, Muñoz MA, Vespignani A, Zapperi S (2000) Braz J Phys 30:27
10. Vespignani A, Dickman R, Muñoz MA, Zapperi S (2000) Phys Rev E 62:4564
11. Dickman R, Alava M, Muñoz MA, Peltola J, Vespignani A, Zapperi S (2001) Phys Rev E 64:056104
12. Manna SS (1991) J Phys A 24:L363
13. Lübeck S (2004) Int J Mod Phys B 18:3977
14. Arthur BW (1994) Am Econ Assoc Pap & Proc 84:406
15. Challet D, Zhang YC (1997) Physica A 246:407

16. Challet D, Zhang YC (1998) Physica A 256:514
17. Challet D, Marsili M, Zhang Y-C (2005) Minority games: interacting agents in financial markets. Oxford University Press, Oxford
18. Moro E (2004) In: Korutcheva E, Cuerno R (eds) Advances in condensed matter and statistical mechanics. Nova Science Publishers, New York. arXiv:cond-mat/0402651v1
19. De Martino A, Marsili M (2006) J Phys A 39:R465
20. Kets W (2007) Preprint. arXiv:0706.4432v1 [q-fin.GN]
21. Dhar D, Sasidevan V, Chakrabarti BK (2011) Physica A 390:3477
22. Reents G, Metzler R, Kinzel W (2001) Physica A 299:253
23. Biswas S, Ghosh A, Chatterjee A, Naskar T, Chakrabarti BK (2012) Phys Rev E 85:031104
24. Evans MR, Hanney T (2005) J Phys A, Math Gen 38:R195

Chapter 13
Kolkata Paise Restaurant Problem and the Cyclically Fair Norm

Priyodorshi Banerjee, Manipushpak Mitra, and Conan Mukherjee

Abstract In this paper we revisit the Kolkata Paise Restaurant problem by allowing for a more general (but common) preference of the n customers defined over the set of n restaurants. This generalization allows for the possibility that each pure strategy Nash equilibrium differs from the Pareto efficient allocation. By assuming that n is small and by allowing for mutual interaction across all customers we design strategies to sustain cyclically fair norm as a sub-game perfect equilibrium of the Kolkata Paise Restaurant problem. We have a cyclically fair norm if n strategically different Pareto efficient strategies are sequentially sustained in a way such that each customer gets serviced in all the n restaurants exactly once between periods 1 and n and then again the same process is repeated between periods $(n + 1)$ and $2n$ and so on.

13.1 Introduction

The Kolkata Paise Restaurant problem [2, 3, 5–7] is a repeated game with identical stage (or one-shot) games and with the same set of n customers (or agents or players). In each stage these n customers have to simultaneously choose between n restaurants to get served. All the customers have a common and rational preference ordering over the service of these n restaurants and, to each customer, the least preferred outcome is not getting the service. Without loss of generality, we assume that the first restaurant is the most preferred followed by the second restaurant and so on and that getting served in the last restaurant is preferred to not getting the service. The price of getting the service from each restaurant is identical. Each restaurant

The authors would like to thank Tarun Kabiraj for his comments and suggestions.

P. Banerjee (✉) · M. Mitra · C. Mukherjee
Economic Research Unit, Indian Statistical Institute, Kolkata, India
e-mail: banpriyo@gmail.com

M. Mitra
e-mail: mmitra@isical.ac.in

C. Mukherjee
e-mail: conanmukherjee@gmail.com

F. Abergel et al. (eds.), *Econophysics of Systemic Risk and Network Dynamics*, New Economic Windows,
DOI 10.1007/978-88-470-2553-0_13,

can serve only one customer so that if more than one customer arrives at the same restaurant, the restaurant randomly chooses one customer to serve and the others do not get the service in that stage. Thus, given the common preferences of the customers over the set of restaurants, the stage game of the Kolkata Paise Restaurant problem is a symmetric one. Moreover, as long as the first restaurant is strictly preferred to the last restaurant, the stage game of the Kolkata Paise Restaurant problem is non-trivial. Given the restrictions on the preferences, Pareto efficiency means that each customer goes to a different restaurant and each restaurant gets exactly one customer to serve.

In the very first work on the Kolkata Paise Restaurant problem [2], it was assumed that the common preferences of the customers is such that going to any unoccupied restaurant is strictly preferred to going to any other restaurant where at least another customer is present. This restriction implied that the set of pure strategy Nash equilibria of the stage game were all Pareto efficient. Hence there are exactly $n!(= n(n-1)\dots 2.1)$ pure strategy Nash equilibria of this version of the stage game of the Kolkata Paise Restaurant problem. If customers are rational, n is small and if customers can mutually interact, then, given the fact that the set of pure strategy Nash equilibrium are also Pareto efficient, one can show that it is easy to sustain any pure strategy Nash equilibrium of the stage game of the Kolkata Paise Restaurant problem as a sub-game perfect equilibrium outcome of the Kolkata Paise Restaurant problem without designing any punishment strategy. This is because, in this context, unilateral deviation means going to a restaurant where there is already another customer which is payoff reducing. In this context it seems quite unfair to sustain exactly one pure strategy Nash equilibrium of the stage game repeatedly as a sub-game perfect Nash equilibrium of the Kolkata Paise Restaurant problem. This is because in any pure strategy Nash equilibrium of the stage game, the customer going to the first restaurant derives a strictly higher payoff than the customer going to the last restaurant. Instead it seems more natural to sustain the *cyclically fair norm* where n strategically different Pareto efficient allocations are sequentially sustained in a way such that each customer gets serviced in all the n restaurants exactly once between periods 1 and n and then again the same process is repeated from the $(n+1)$th period to period $2n$ and so on. A variant of the cyclically fair norm was proposed in [7] under the large player assumption. However, this type of cyclically fair norm can also be sustained as a sub-game perfect Nash equilibrium because unilateral deviation at any stage means going to a restaurant already occupied by another customer which is always payoff reducing. Therefore, the existing structure of the Kolkata Paise Restaurant problem is such that if the number of customers n is small and if the customers can coordinate their action then the problem becomes uninteresting as there is no need to design punishment strategies to induce customers to remain in the equilibrium path. Thus it is natural that the existing literature on Kolkata Paise Restaurant problem [2, 5–7] deals with situations where n is macroscopically large so that the agents cannot rely on the other agents' actions and therefore what matters to each agent is the past collective configuration of actions and the resulting average utilization of the restaurants.

In this paper we revisit the Kolkata Paise Restaurant problem by relaxing the assumption on preferences that ensures the Pareto efficiency of all the pure strategy Nash equilibria of the stage game. Therefore, we analyze the Kolkata Paise Restaurant problem by looking at a more general (but common across agents) rational preference structure over the restaurants such that the stage game allows for the possibility of inefficient pure strategy Nash equilibria. In this scenario we assume that n is small and that customers can take coordinated actions and then analyze the possibility of sustaining the cyclically fair norm as a sub-game perfect equilibrium of the Kolkata Paise Restaurant problem. Clearly, in this context, there is a need for designing punishment schemes in order to sustain the cyclically fair norm as a subgame perfect equilibrium since unilateral deviation from the proposed norm can be payoff enhancing as the configurations under the cyclically fair norm may not be pure strategy Nash equilibria of the stage game.

13.2 The Stage Game

We start by formally defining and analyzing the stage game associated with the Kolkata Paise Restaurant (or KPR) problem. Let $N = \{1, \ldots, n\}$ be the finite set of agents, $S = \{R1, \ldots, Rn\}$ be the set of restaurants and vector $u = (u_1, \ldots, u_n) \in \Re^n$ represent the utility (in terms of money) associated with each restaurant which is common to all customers or agents. Assume w.l.o.g. that $0 < u_n \leq \cdots \leq u_2 \leq u_1$ with $u_1 \neq u_n$. Formally, the one shot KPR game is $G(u) \equiv (N, S, \pi)$, where $S = \{R1, \ldots, Rn\}$ is the common *action* space and $\pi_i : S^n \mapsto \Re$ is the payoff function of agent i. For any agent i, $s_i = k \in S$ implies that agent i chooses the strategy of going to restaurant k. It may so happen that more than one agents goes to the same restaurant. In that case, service is provided to only one of them and this selection is completely random. Therefore, for any strategy profile $s \in S^n$, the *expected* payoff to agent i, $\pi_i(s) = \frac{u_{s_i}}{\eta_i(s)}$ where $\eta_i(s) = 1 + |\{j \in N : j \neq i, s_i = s_j\}|$ is the number of agents that have selected the same restaurant as agent i. We call a strategy profile $s = (s_1, \ldots, s_n) \in S^n$ Pareto efficient, if the sum of payoffs of the agents is maximized, that is,

$$s \in \arg\max_{s' \in S^n} \sum_{i \in N} \pi_i\left(s'\right).$$

Given the current setting, a strategy combination leads to Pareto efficiency if and only if the strategies of the agents are such that they end up in different restaurants, that is, $\forall i, j \in N$, $s_i \neq s_j$. A strategy combination $s^* = (s_1^*, \ldots, s_n^*)$ is a pure strategy *Nash equilibrium* (NE) if no agent i has incentive to deviate from the existing strategy s_i^* given the strategies $s_{-i}^* = (s_1^*, \ldots, s_{i-1}^*, s_{i+1}^*, \ldots, s_n^*)$ of the other players, that is, for each agent $i \in N$,

$$\pi_i\left(s^*\right) \geq \pi_i\left(s_i, s_{-i}^*\right) \quad \forall s_i \in S.$$

Remark 13.1 It was proved in [2] that if $u_1 < 2u_n$ then the set of all pure strategy Nash equilibria of the one-shot KPR problem coincides with the set of all Pareto efficient strategies. This result is quite intuitive since the restriction $u_1 < 2u_n$ means that going to any unoccupied restaurant is strictly preferred to going to any other restaurant where at least another agent is present. Hence for any agent $i \in N$, given the strategy of all other agents, it is always optimum for agent i to select the most preferred unoccupied restaurant. Since the number of restaurant is the same as the number of agents, it is always possible for agent i to find an unoccupied restaurant. Hence in any pure strategy Nash equilibrium all agents end up in different restaurants which is also a Pareto optimal strategy combination.

Before concluding this section we provide a discussion on symmetric mixed strategy equilibria in the following remark.

Remark 13.2 (Symmetric mixed strategy equilibria) Let $u_1 < 2u_n$ and let $A(S)$ denote the set of all mixed strategies defined over S.[1] A symmetric mixed strategy Nash equilibrium $\underline{p}^* = (\underbrace{p^*, \ldots, p^*}_{n}) \in A(S)^N$ where $p^* = (p_1^*, \ldots, p_n^*) \in [0,1]^n$ with $\sum_{i=1}^n p_i^* = 1$ is a solution to the following sets of equation:

For each $i \in N$, $\sum_{k=0}^{n-1}(1-p_i^*)^k = \frac{nc(n)}{u_i}$ for some constant $c(n)$ which is positive real.[2]

(i) For $N = \{1,2\}$, the symmetric mixed strategy Nash equilibrium is $\underline{p}^* = (p^*, p^*)$ where $p^* = (p_1^* = \frac{2u_1 - u_2}{u_1 + u_2}, p_2^* = \frac{2u_2 - u_1}{u_1 + u_2})$ and $c(2) = \frac{3u_1u_2}{2(u_1+u_2)}$.
(ii) For $N = \{1,2,3\}$, there are two symmetric mixed strategy Nash equilibria. These equilibria are characterized by $p^* = (p_1^*, p_2^*, p_3^*)$ and $c(3)$ where $p_i^* = \frac{3}{2} - \frac{1}{2}\sqrt{\frac{12c(3)}{u_i} - 3}$ for all $i \in \{1,2,3\}$, the constant $c(3)$ takes two values given by $c(3) = \sqrt{E_1E_2E_3}\Big(\frac{3(E_1+E_2+E_3) \pm \sqrt{9(E_1+E_2+E_3)^2 - 20(E_1^2+E_2^2+E_3^2)}}{(E_1^2+E_2^2+E_3^2)}\Big)$ and $E_i = u_ju_l$ for all $i \neq j \neq l \neq i$. It can be verified that given $u_3 < 2u_1$, $9(E_1 + E_2 + E_3)^2 - 20(E_1^2 + E_2^2 + E_3^2) > 0$ and hence $c(3)$ is always positive real.
(iii) In general, for $n > 3$ such symmetric mixed strategy equilibria always exists [1]. A general feature of the symmetric mixed strategy equilibria is that $0 < p_n^* \leq \cdots \leq p_1^* < 1$ and $p_1^* \neq p_n^*$.

It is quite clear from Remark 13.2 that working out the mixed strategy equilibria, in general, is difficult. Therefore, in this paper, we concentrate only on pure strategy equilibria of the stage game.

[1] A mixed strategy is a probability distribution defined on the strategy set. Therefore, in the present context, $A(S)$ is the set of all probability distributions on the set of restaurants S.

[2] For mixed strategy equilibria the required condition is $\sum_{r=0}^{n-1}\{\binom{n-1}{r}(p_i^*)^r(1-p_i^*)^{n-r-1}\frac{u_i}{r+1}\} = c(n)$ for all $i \in N$ and after simplification we get $\sum_{k=0}^{n-1}(1-p_i^*)^k = \frac{nc(n)}{u_i}$ for all $i \in N$.

13.3 The KPR Problem

The KPR problem is an infinitely repeated game where in each stage the same set of $N = \{1, \ldots, n\}$ agents play the one shot KPR game $G(u)$ defined in the previous section.[3] We represent the KPR problem as $G^{\infty}(u) = (N, (\Sigma_i)_{i\in N}, (\Pi_i)_{i\in N})$ where N is the set of agents and for any agent i, Σ_i is the set of *strategies* available to i, while Π_i is the payoff function of i. However, the concepts of strategy and payoff, have now become more complex, due to this repeated interaction setting.[4]

Let us start with the concept of strategy in $G^{\infty}(u)$. In each period t, the play of the one-shot KPR game would result in some action profile $s^t = (s_1^t, \ldots, s_n^t) \in S^n$. Given any period t, define history $h_t = (s^1, s^2, \ldots, s^{t-1})$ as the description of past play. That is, h_t is a sequence of action profiles realized through times 1 to $t-1$. For any t, h_t is assumed to be common knowledge. Let H^t denote the set of all possible histories at time t. Strategy of i in $G^{\infty}(u)$, specifies an action, that is, the restaurant that i goes to, in each period t, for each possible history h^t. Therefore, $\forall\, i \in N, \forall\, \sigma_i \in \Sigma_i, \sigma_i : H^t \mapsto S$.

For each possible sequence of action profiles over time, we get a sequence of payoffs, for each agent. To calculate the payoff of an agent we define the concept of the discount factor $\delta \in (0, 1)$. It is presumed that agents are impatient, and hence, discount future payoffs, so that *present discounted value* of a dollar to be received one period later is δ, two periods later is δ^2, and so on. In general, any payoff x to be received τ periods later, is valued at the present period as $\delta^{\tau} x$. Therefore, present discounted value of the infinite sequence of payoffs corresponding to any infinite sequence of action profiles $\{s^1, s^2, s^3, \ldots\}$, for agent i, is $\sum_{t=1}^{\infty} \delta^{t-1} \pi_i(s^t)$. We assume that each agent discounts the future payoffs at same rate.[5]

Remark 13.3 In this remark we provide two interpretations of the discount factor.

(i) The popular interpretation of discount factor δ is that it is the *time-value* of money. Suppose a person puts an amount of money x in a bank at the beginning of present period. If the bank pays interest r per period, upon withdrawal the person gets $x(1+r)$ money at the beginning of the next period. Therefore, we can say that amount x to be received in the beginning of the next period is worth only $\frac{1}{1+r}x$ money in the present period. Setting $\delta = \frac{1}{1+r}$ we get that; at present, the next period payoff x is worth δx and the next to next period payoff x is worth $\delta^2 x$. Therefore, a sequence of future payoffs $\{x^1, x^2, x^3, \ldots\}$ is worth $\sum_{t=1}^{\infty} \delta^{t-1} x^t$ at present.

(ii) The concept of δ, can also be used to view the infinitely repeated game as a *finite period* repeated game that ends after a *random* number of periods. Suppose that

[3]An infinitely repeated game like the KPR problem, where the same stage game is played repeatedly, is also referred to as a *supergame* [4].

[4]The analysis of the concepts of repeated game theory is taken from [8] and [9].

[5]It can be easily verified that the conclusions of this paper remain qualitatively same if we allow for unequal discount factors across agents.

after each period is played, a (possible weighted) coin is flipped to determine whether the game will end. If the probability that the game ends immediately is p and then, with probability $1 - p$, the game continues for a least one more period and then the payoff x, to be obtained in the next stage (if it occurs), is worth only $\frac{(1-p)x}{(1+r)}$. Similarly, a payoff x to be received two periods from now (if both periods are played) is worth only $\frac{(1-p)^2 x}{(1+r)^2}$ before this stage's coin flip occurs. Therefore, the sum $x + \delta x + \delta^2 x + \cdots$ with $\delta = \frac{1-p}{1+r}$ reflects both the time value of money and the possibility that the game may end.

For different values of δ, we get different KPR problems $G_\delta^\infty(u)$. Therefore, for any agent i, payoff function Π_i in $G_\delta^\infty(u)$ is a mapping $\Pi_i : \Sigma_1 \times \ldots \Sigma_n \mapsto \Re$ such that for any strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$, $\Pi_i(\sigma) = \sum_{t=1}^{\infty} \delta^{t-1} \pi_i(\sigma_1^t, \sigma_2^t, \ldots, \sigma_n^t)$. A strategy profile $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ is a Nash equilibrium (NE) of $G_\delta^\infty(u)$, if no agent i finds it profitable to deviate unilaterally from σ^*, that is for each $i \in N$,

$$\Pi_i\left(\sigma^*\right) \geq \Pi_i\left(\sigma_i, \sigma_{-i}^*\right) \quad \forall \sigma_i \in \Sigma_i.$$

We focus on a particular strategy profile $\bar{\sigma}$ satisfying the following conditions.

(i) Without loss of generality, in period $t = 1$ each agent $i(\in N)$ goes to restaurant i.
(ii) For any period $t > 1$, if agent i went to restaurant 1 in the last period $t - 1$, then i goes to restaurant n at period t.
(iii) For any period $t > 1$, if agent i went to restaurant $k > 1$ in the last period $t - 1$, then i goes to restaurant $k - 1$ at period t.

Note that strategy $\bar{\sigma}$ requires that action of any agent i at any period t depend only on i's action at period $t - 1$ and *not on other agents' actions* in the past. If all agents play $\bar{\sigma}$ at $G_\delta^\infty(u)$, we get the *cyclically fair norm.*

Proposition 13.1 *If $u_1 < 2u_n$, then for all $\delta \in (0, 1)$, $\bar{\sigma}$ is a Nash equilibrium of $G_\delta^\infty(u)$.*

Proof If $u_1 < 2u_n$ then we know that in any period t, going to any unoccupied restaurant is strictly preferred to going to any other restaurant where at least another agent is present. Hence it is always optimum for any agent $i \in N$, in any period t, to select the most preferred unoccupied restaurant. Since the number of restaurant is the same as the number of agents, it is always possible for agent i in any period t to find an unoccupied restaurant.

Given $\bar{\sigma}$ it is clear that any unilateral deviation from $\bar{\sigma}$ by any agent i, at any time t, would lead to i being tied with another agent at some restaurant thereby ensuring a strict reduction in payoff in that period.

Depending on the deviation strategy σ_i, in all periods after t, agent i can face a tie or he may not face a tie. If agent i faces a tie then he is strictly worse off in that period in comparison to $\bar{\sigma}_i$ and if he does not face a tie then he gets the same payoff in that period in comparison to $\bar{\sigma}_i$. The reason is that given that all other agents $j \in N \setminus \{i\}$ are continuing with the strategy $\bar{\sigma}_j$, in each period after t, for agent i, there is exactly one restaurant which is not occupied and hence, given the preference of agent i we get the result. Thus, in either case $\Pi_i(\sigma_i, \bar{\sigma}_{-i}) < \Pi_i(\bar{\sigma})$ implying that $\bar{\sigma}$ is a Nash equilibrium of $G_{\delta}^{\infty}(u)$. □

Our objective is to sustain $\bar{\sigma}$ as a sub-game perfect equilibrium in order to implement the cyclically fair norm. That is, we need to show that $\bar{\sigma}$ constitutes a sub-game perfect equilibrium of $G_{\delta}^{\infty}(u)$. Before defining the sub-game perfect equilibrium we need to define a sub-game. We call any 'piece' of game $G_{\delta}^{\infty}(u)$ following any history h^t, at any period t, a *subgame* of $G_{\delta}^{\infty}(u)$. Therefore, a sub-game is that piece of the game that remains to be played beginning at any point at which the complete history of the game thus far is common knowledge. The definition of a strategy in any infinitely repeated game is closely related to the definition of a sub-game. In particular, an agent's strategy specifies the actions the agent will take in the first period of the repeated game and the first stage of each of its sub-game. There are infinite number of sub-games of $G_{\delta}^{\infty}(u)$. Since $G_{\delta}^{\infty}(u)$ is an infinitely repeated game, each of its sub-games, beginning at period $t+1$ of $G_{\delta}^{\infty}(u)$ is identical to $G_{\delta}^{\infty}(u)$. Note that the tth period of a repeated game taken in isolation is *not* a sub-game of the repeated game. Therefore, a sub-game is a piece of the original game that not only starts at a point where the history of the game thus far is common knowledge among the agents, but also includes all the moves that follow this point in the original game. A Nash equilibrium strategy profile $\sigma^{**} = (\sigma_1^{**}, \ldots, \sigma_n^{**})$ is a *sub-game perfect equilibrium* if these strategies constitute a Nash equilibrium in every sub-game.

Corollary 13.1 *If $u_1 < 2u_n$, then for any $\delta \in (0, 1)$, $\bar{\sigma}$ is a sub-game perfect equilibrium of $G_{\delta}^{\infty}(u)$.*

Proof Given particular property of $\bar{\sigma}$ where i's behavior depends only on his own past behavior, no deviation by i in any period t (and hence in the sub-game starting from period t) can induce a change in future actions of other agents in $N \setminus \{i\}$. Thus, using the arguments from the last paragraph of the proof of Proposition 13.1, it follows that $\bar{\sigma}$ continues to be a Nash equilibrium in every sub-game of $G_{\delta}^{\infty}(u)$. □

Remark 13.4 If $u_1 = 2u_n$ then, by making minor alterations in the arguments in the proofs of Proposition 13.1 and Corollary 13.1, one can implement the cyclically fair norm as a sub-game perfect equilibrium with the same strategy $\bar{\sigma}$ and for any $\delta \in (0, 1)$. The proof is left to the reader.

Observe that the strategy profile $\bar{\sigma}$, that implements the cyclically fair norm as a sub-game perfect equilibrium when $u_1 \leq 2u_n$, is such that there is no specification of punishment in the sense that it is silent about what happens if one agent deviates from the existing strategy. However, if $u_1 > 2u_n$ then implementing the cyclically fair norm becomes non-trivial and one has to design appropriate punishment

Table 13.1 The two-agent payoff matrix

$G(u_1, u_2)$	$R1$	$R2$
$R1$	$\left(\frac{u_1}{2}, \frac{u_1}{2}\right)$	(u_1, u_2)
$R2$	(u_2, u_1)	$\left(\frac{u_2}{2}, \frac{u_2}{2}\right)$

schemes. To see this, consider the simple stage game with $n = 2$. The two-agent payoff matrix with 1 as the row player and 2 as the column player is presented in Table 13.1.

If $u_1 < 2u_2$ then there are only two pure strategy Nash equilibria ($R1$, $R2$) and ($R2$, $R1$) of the stage game. Even when $u_1 = 2u_2$, ($R1$, $R2$) and ($R2$, $R1$) continue to be pure strategy Nash equilibria of the stage game. Therefore, the Pareto efficient strategies where both agents go to different restaurants are pure strategy Nash equilibria for $u_2 \geq \frac{u_1}{2}$. It is this strong result that drives Proposition 13.1 and we can easily implement the cyclically fair norm. Note that if $u_2 = \frac{u_1}{2}$, then there are three pure strategy Nash equilibria ($R1$, $R2$), ($R2$, $R1$) and ($R1$, $R1$) of the stage game. Therefore, for $u_1 = 2u_2$, there exists a sub-game prefect equilibrium that leads to inefficiency in every period. Specifically, the strategy that specifies that each agent should go to the first restaurant in all periods is a sub-game perfect equilibrium, where sum of the stage game payoffs of the two agents is u_1 which is strictly less that sum $u_1 + u_2$ that results under any Pareto optimal strategy. The problem gets only worse if $2u_2 < u_1$, because now there is only one pure strategy Nash equilibrium ($R1$, $R1$) of the stage game which is not Pareto efficient. How to design strategies to implement the cyclically fair norm as a sub-game perfect equilibrium when $n = 2$ and $2u_2 < u_1$ is discussed in the next section.

13.4 The Two Agent Problem

In this section we show that for $N = \{1, 2\}$ and for $u_2 < \frac{u_1}{2}$, if agents are sufficiently patient (that is, if δ is sufficiently high), then, by designing an appropriate strategy one can implement the cyclically fair norm as a sub-game perfect equilibrium of $G_\delta^\infty(u_1, u_2)$. The strategy profile we propose, to implement the cyclically fair norm, is $\sigma^c = (\sigma_1^c, \sigma_2^c)$, that specifies the following.

(i) Without loss of generality, if t is odd, then agent 1 plays $R1$ and agent 2 plays $R2$.
(ii) If t is even, then agent 2 plays $R1$ and agent 1 plays $R2$.
(iii) If in any period t both agents end up in the same restaurant, then from $t + 1$ onwards both agents play $R1$.

Strategy profiles of the type $\bar{\sigma}^c$ above, are called *trigger strategies* because agents cooperate until someone fails to cooperate, which triggers a switch to non-cooperation forever. In other words, each agent is willing to settle for lower payoffs under the expectation that the other agent would do the same. If some agent breaks this cooperative arrangement, the other agent *punishes* the deal-breaker by playing certain actions (for all periods in future) that ensure lower present discounted payoffs. Thus, a deviation *triggers* a punishment play by the non-deviating agents. Such trigger strategies are sub-game perfect only if the punishment play for all future periods, induced by these strategies, are *credible*. This credibility, in turn, requires that the punishment play be the Nash equilibrium of $G_\delta^\infty(u_1, u_2)$ as a whole.

Proposition 13.2 *For all* $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$, *the strategy profile* $\bar{\sigma}^c = (\sigma_1^c, \sigma_2^c)$ *is a Nash equilibrium of* $G_\delta^\infty(u_1, u_2)$.

Proof We first assume that agent 1 plays strategy $\bar{\sigma}_1^c$. Given $\bar{\sigma}_1^c$, we show that if $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$ then $\bar{\sigma}_2^c$ is the best response of agent 2.

It is clear that at any history, if agent 1 decides to play $R1$ in all future periods, then, given $2u_2 < u_1$, the best response of agent 2 is to play $R1$ in all future periods.

Consider the other possibility, that is, agent 1 decides to alternate between restaurants 1 and 2 at each odd and even period, respectively. Then it is obvious that the best response of agent 2 at any even period is to play $R1$ (since agent 1 is playing $R2$ and $u_1 > u_2$). However, finding the best response of agent 2, at *odd* periods (when agent 1 plays $R1$), is a little more complicated. If agent 2 chooses $R1$, then as per $\bar{\sigma}^c$, agent 1 plays $R1$ at all future periods, giving 2 a present discounted payoff $\frac{u_1}{2} + \delta\frac{u_1}{2} + \delta^2\frac{u_1}{2} + \cdots = \frac{u_1}{2(1-\delta)}$. Define P to be the present discounted payoff that agent 2 gets by making the *optimal action choice* at any such odd period. Therefore, if the optimal choice of agent 2 is $R1$, then $P = \frac{u_1}{2(1-\delta)}$. If the optimal choice of agent 2 is $R2$, then in the next period, that is in period $t + 1$ which is even, agent 1 plays $R2$. As mentioned before, the best response of 2 at $t + 1$ is $R1$, and so we have the following: (i) agent 2 gets payoff u_1 at $t + 1$ and (ii) agent 2 faces the same choice problem in period $t + 2$ as in period t and, since all sub-games of $G_\delta^\infty(u_1, u_2)$ is $G_\delta^\infty(u_1, u_2)$ itself, agent 2 selects P. Therefore, if $R2$ has to be the optimal choice of agent 2 at all odd periods t then agent 2 gets $u_2 + \delta u_1 + \delta^2 P$ and hence, by definition of P and using observations (i) and (ii), it follows that P has to satisfy the condition that $P = \max\{\frac{u_1}{2(1-\delta)}, u_2 + \delta u_1 + \delta^2 P\} = u_2 + \delta u_1 + \delta^2 P$. If $P = u_2 + \delta u_1 + \delta^2 P$ then we get $P = \frac{u_1\delta + u_2}{1-\delta^2}$. Finally, for $\bar{\sigma}_2^c$ to be the best response it is both necessary and sufficient that $\frac{u_1\delta+u_2}{1-\delta^2} > \frac{u_1}{2(1-\delta)}$ which holds for all $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$. Therefore, at any odd period the best response of agent 2 is $R2$ implying that for all $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$, the strategy $\bar{\sigma}_2^c$ is the best response of agent 2 when agent 1 plays $\bar{\sigma}_1^c$. Using very similar arguments it is now quite easy to show that for all $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$, the strategy $\bar{\sigma}_1^c$ is the best response of agent 1 when agent 2 plays $\bar{\sigma}_2^c$. Hence, $\bar{\sigma}^c = (\bar{\sigma}_1^c, \bar{\sigma}_2^c)$ is a Nash equilibrium of $G_\delta^\infty(u_1, u_2)$ for all $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$. □

Corollary 13.2 *For all* $\delta \in (\frac{u_1-2u_2}{u_1}, 1)$, *the strategy profile* $\bar{\sigma}^c = (\sigma_1^c, \sigma_2^c)$ *is a sub-game perfect equilibrium of* $G_\delta^\infty(u_1, u_2)$.

Proof The set of sub-games of $G_\delta^\infty(u_1, u_2)$ can be partitioned into two classes. One class following those histories where each agent followed the cyclically fair norm and alternated between restaurants 1 and 2 in a way such that Pareto efficiency is achieved in every period. The other class following those histories where there has been a tie at some restaurant and agents have shifted to $(R1, R1)$ from the next period onwards. Recall that every sub-game of $G_\delta^\infty(u_1, u_2)$ is $G_\delta^\infty(u_1, u_2)$ itself. If agents adopt strategy $\bar{\sigma}^c$ for the game as a whole, then they end up playing strategy $\bar{\sigma}^c$ in sub-games of the first class and (ii) the punishment play $(R1, R1)$ in each period of sub-games of the second class. By Proposition 13.2, for sub-games of the first type, strategies $\bar{\sigma}^c$ constitute a Nash equilibrium. For sub-games of the second type, the punishment play of $R1$ by both agents at all periods constitutes a Nash equilibrium of $G_\delta^\infty(u_1, u_2)$ since $(R1, R1)$ is the unique Nash equilibrium of $G(u_1, u_2)$ when $2u_2 < u_1$. Hence, the punishment play is always credible and the result follows. □

From Corollary 13.2 it follows that as long as agents are sufficiently patient, the strategy profile $\bar{\sigma}^c$ implements the cyclically fair norm. Therefore, the bound on δ, obtained in Proposition 13.2 above, signifies the need for sufficiently patient agents to implement cyclically fair norm that calls for cooperative behavior. If agents feel the need to obtain high payoffs in the future (or equivalently if δ is high enough) then they are willing to make a sacrifice by going to the inferior restaurant in alternate periods in order to maximize long term individual payoff. In the next section we analyze the KPR problem with three agents. We show how using different strategy profiles one can implement the cyclically fair norm.

13.5 The Three Agent Problem

We depict the payoff matrices of $G(u_1, u_2, u_3)$ in Tables 13.2, 13.3 and 13.4.

Recall that if the one shot game $G(u_1, u_2, u_3)$ represents the one shot game of the KPR problem then $u_1 \geq u_2 \geq u_3 > 0$ and $u_1 \neq u_3$. With different types of additional conditions on u_1, u_2 and u_3, we identify and discuss the associated set of pure strategy Nash equilibria in the following cases.

(N1) If $u_2 < \frac{u_1}{3}$ then there is a unique pure strategy Nash equilibrium $(R1, R1, R1)$ of $G(u_1, u_2, u_3)$. This equilibrium is inefficient.

(N2) If $u_3 < u_2 = \frac{u_1}{3}$ then the four pure strategy Nash equilibria of $G(u_1, u_2, u_3)$ are $(R1, R1, R1)$, $(R1, R1, R2)$, $(R1, R2, R1)$ and $(R2, R1, R1)$. All these equilibria are inefficient. The equilibrium $(R1, R1, R1)$ is Pareto dominated by all the remaining equilibria since the aggregate payoff under $(R1, R1, R1)$ is u_1 which is strictly less than the aggregate payoff $(u_1 + u_2)$ that results from each of the remaining equilibria.

(N3) If we have $u_3 = u_2 = \frac{u_1}{3}$, then $(R1, R1, R3)$, $(R1, R3, R1)$ and $(R3, R1, R1)$ are also pure strategy Nash equilibria along with the other equilibria specified in (N2) and hence we have seven pure strategy Nash equilibria. Again,

Table 13.2 The payoff matrix when agent 3 plays $R1$

$G(u_1, u_2, u_3)$	$R1$	$R2$	$R3$
$R1$	$(\frac{u_1}{3}, \frac{u_1}{3}, \frac{u_1}{3})$	$(\frac{u_1}{2}, u_2, \frac{u_1}{2})$	$(\frac{u_1}{2}, u_3, \frac{u_1}{2})$
$R2$	$(u_2, \frac{u_1}{2}, \frac{u_1}{2})$	$(\frac{u_2}{2}, \frac{u_2}{2}, u_1)$	(u_2, u_3, u_1)
$R3$	$(u_3, \frac{u_1}{2}, \frac{u_1}{2})$	(u_3, u_2, u_1)	$(\frac{u_3}{2}, \frac{u_3}{2}, u_1)$

Table 13.3 The payoff matrix when agent 3 plays $R2$

$G(u_1, u_2, u_3)$	$R1$	$R2$	$R3$
$R1$	$(\frac{u_1}{2}, \frac{u_1}{2}, u_2)$	$(u_1, \frac{u_2}{2}, \frac{u_2}{2})$	(u_1, u_3, u_2)
$R2$	$(\frac{u_2}{2}, u_1, \frac{u_2}{2})$	$(\frac{u_2}{3}, \frac{u_2}{3}, \frac{u_2}{3})$	$(\frac{u_2}{2}, u_3, \frac{u_2}{2})$
$R3$	(u_3, u_1, u_2)	$(u_3, \frac{u_2}{2}, \frac{u_2}{2})$	$(\frac{u_3}{2}, \frac{u_3}{2}, u_2)$

Table 13.4 The payoff matrix when agent 3 plays $R3$

$G(u_1, u_2, u_3)$	$R1$	$R2$	$R3$
$R1$	$(\frac{u_1}{2}, \frac{u_1}{2}, u_3)$	(u_1, u_2, u_3)	$(u_1, \frac{u_3}{2}, \frac{u_3}{2})$
$R2$	(u_2, u_1, u_3)	$(\frac{u_2}{2}, \frac{u_2}{2}, u_3)$	$(u_2, \frac{u_3}{2}, \frac{u_3}{2})$
$R3$	$(\frac{u_3}{2}, u_1, \frac{u_3}{2})$	$(\frac{u_3}{2}, u_2, \frac{u_3}{2})$	$(\frac{u_3}{3}, \frac{u_3}{3}, \frac{u_3}{3})$

the equilibrium $(R1, R1, R1)$ is Pareto dominated by the other six non-comparable equilibria. The equilibria are inefficient.

(N4) If $\max\{u_3, \frac{u_1}{3}\} < u_2 < \frac{u_1}{2}$ then the three pure Nash strategy equilibria of the game $G(u_1, u_2, u_3)$ are $(R1, R1, R2)$, $(R1, R2, R1)$ and $(R2, R1, R1)$. All these equilibria lead to the same aggregate payoff and hence, are Pareto non-comparable. The equilibria are inefficient.

(N5) If $u_3 = u_2 < \frac{u_1}{2}$ then $(R1, R1, R3)$, $(R1, R3, R1)$ and $(R3, R1, R1)$ are also pure strategy Nash equilibria along with the other equilibria specified in (N4) and hence we have six pure strategy Nash equilibria of $G(u_1, u_2, u_3)$. The equilibria are Pareto non-comparable and inefficient.

(N6) If $u_3 < \frac{u_1}{2} \leq u_2 \leq u_1$ then the three pure strategy equilibria of $G(u_1, u_2, u_3)$ are $(R1, R1, R2)$, $(R1, R2, R1)$ and $(R2, R1, R1)$. All these equilibria are inefficient and Pareto non-comparable.

(N7) If $\frac{u_1}{2} \leq u_3$ then we do not identify all possible pure strategy Nash equilibria. However, what is important is that the Pareto efficient strategies, characterized by all agents going to different restaurants, are all included in the set of all pure strategy Nash equilibria.

The equilibria in case (N7) above are uninteresting as implementation of the cyclically fair norm as a sub-game perfect equilibrium of $G^{\infty}(u_1, u_2, u_3)$ is quite easy (see Corollary 13.1 and Remark 13.4). For cases (N1)–(N3), all agents going to the best restaurant, that is, the strategy profile $(R1, R1, R1)$ constitutes a pure strategy Nash equilibrium and is Pareto inefficient. Therefore, as in Proposition 13.2 and Corollary 13.2, for cases (N1)–(N3), we can implement the cyclically fair norm as a sub-game perfect equilibrium of $G_\delta^\infty(u_1, u_2, u_3)$ for δ sufficiently close to one. It can be easily shown that, by designing a strategy profile which is a natural extension of $\sigma^c = (\sigma_1^c, \sigma_2^c)$ to the three agent case and that specifies that the non-deviating agents punish the deviating agent by going to the best restaurant for all future periods, one can implement the cyclically fair norm. For the rest of the cases, that is, (N4)–(N6), strategy profile $(R1, R1, R1)$ fails to be a pure strategy Nash equilibrium of the stage game, and so, implementing the cyclically fair norm becomes more subtle. This is because the threat of punishment embodied in strategy of type σ^c, that is, going to the best restaurant for all future periods, is no longer a credible one as it is not a Nash equilibrium of the stage game.

In the rest of this section, we focus on the interesting cases (N4), (N5) and (N6). Since there is no qualitative difference between (N4) and (N5), we analyze only cases (N4) and (N6) in detail. As long as agents are sufficiently patient, we can show that for both cases we can implement the cyclically fair norm as a sub-game perfect equilibrium of the KPR problem $G_\delta^\infty(u_1, u_2, u_3)$. Interesting to note here is that, for cases (N4) and (N6), the set of pure strategy Nash equilibria is $\{(R1, R1, R2), (R1, R2, R1), (R2, R1, R1)\}$. However, while for (N4), each agent playing $R1$ gets *more expected payoff* than the agent playing $R2$, for (N6), each agent playing $R1$ gets an expected payoff which is *no more* than the payoff of the agent playing $R2$. It is precisely for this difference in payoffs for the same given pure strategy Nash equilibrium for cases (N4) and (N6) that calls for designing different punishment strategies to implement the cyclically fair path.

Consider first (N4), that is $\max\{u_3, \frac{u_1}{3}\} < u_2 < \frac{u_1}{2}$. Consider the strategy profile $\sigma^a = (\sigma_1^a, \sigma_2^a, \sigma_3^a)$ that specifies the following.

(i) Without loss of generality at $t = 1$, each agent $i \in \{1, 2, 3\}$ plays Ri.
(ii) If agent i plays $R1$ in period $t - 1$, then i plays $R3$ in period t.
(iii) If agent i plays $Rk \neq R1$ in period $t - 1$, then i plays $R(k - 1)$ in period t.
(iv) If any agent i violates either of 1, 2 or 3 in some period t then in all future periods $t + 1, t + 2, \ldots$, all the non-deviating agents $(N \setminus \{i\})$ plays $R1$.

Conditions (i)–(iii) in the strategy profile σ^a ensures that agents follow the cyclically fair norm. Condition (iv) is the punishment requirement that specifies that, if an agent deviates, then the non-deviating agents punish the deviating agent by playing $R1$, for all future periods.

Proposition 13.3 *If* $\max\{u_3, \frac{u_1}{3}\} < u_2 < \frac{u_1}{2}$*, then there exists* $\bar{\delta} \in (0, 1)$ *such that for all* $\delta \in (\bar{\delta}, 1)$*,* $\sigma^a = (\sigma_1^a, \sigma_2^a, \sigma_3^a)$ *is a sub-game perfect equilibrium of* $G_\delta^\infty(u_1, u_2, u_3)$.

Proof Fix agent 2's strategy at σ_2^a and agent 3's strategy at follows σ_3^a. We first show that, given this specification, σ_1^a is the best response for agent 1 provided agent 1 is sufficiently patient. Consider agent 1 at any history and at any time t. Given the utility restriction $\max\{u_3, \frac{u_1}{3}\} < u_2 < \frac{u_1}{2}$, agent 1 has an incentive to deviate only if at time t, agent 1 is supposed to play either $R3$ or $R2$ (otherwise agent 1 has no profitable deviation in the stage game at t when the strategy prescribes $R1$). Also, the best deviation for agent 1 is to play $R1$ and get a payoff of $\frac{u_1}{2}$. If agent 1 deviates, then, following the strategy profile σ^a, agents 2 and 3 play $R1$ for all future periods. At each of such future periods, the best response of agent 1 is to play $R2$. Therefore, the present discounted payoff of agent 1 from deviation is

$$D_1(\delta) = \frac{u_1}{2} + \left(\delta u_2 + \delta^2 u_2 + \cdots\right) = \frac{u_1(1-\delta) + 2\delta u_2}{2(1-\delta)}. \tag{13.1}$$

By not deviating in a period t where agent 1 had to play $R2$ (under conditions (i)–(iii)), agent 1's present discounted value of payoff from period t onwards is

$$E_2(\delta) = \left(u_2 + \delta u_1 + \delta^2 u_3\right) + \left(\delta^3 u_2 + \delta^4 u_1 + \delta^5 u_3\right) + \cdots = \frac{u_2 + \delta u_1 + \delta^2 u_3}{(1-\delta^3)}. \tag{13.2}$$

Similarly, by not deviating in a period t' where agent 1 had to play $R3$, agent 1's present discounted value of payoff from period t' onwards is

$$E_3(\delta) = \left(u_3 + \delta u_2 + \delta^2 u_1\right) + \left(\delta^3 u_3 + \delta^4 u_2 + \delta^5 u_1\right) + \cdots = \frac{u_3 + \delta u_2 + \delta^2 u_1}{(1-\delta^3)}. \tag{13.3}$$

A sufficient condition for σ_1^a to be a best response for agent 1 (given the strategies σ_2^a and σ_3^a of agents 2 and 3 respectively) is that $\min\{E_2(\delta), E_3(\delta)\} > D_1(\delta)$. Note that $\min\{E_2(\delta), E_3(\delta)\} = E_3(\delta)$ since $E_2(\delta) - E_3(\delta) = \frac{(u_2-u_3)(1-\delta^2)+(u_1-u_2)\delta(1-\delta)}{(1-\delta^3)} > 0$. Therefore, for any $\delta \in (0, 1)$ such that $E_3(\delta) - D_1(\delta) > 0$, σ_1^a is the best response for agent 1. Observe that $E_3(\delta) - D_1(\delta) = \frac{F(\delta)}{2(1-\delta^3)}$ where $F(\delta) = 2(u_3 + \delta u_2 + \delta^2 u_1) - (u_1(1-\delta) + 2\delta u_2)(1 + \delta + \delta^2)$. Note that $F(\delta)$ is continuous in δ and, given (N4), $F(0) = 2u_3 - u_1 < 0$ and $F(1) = 4(\frac{u_1}{2} + \frac{u_3}{2} - u_2) > 0$. Hence, there exists a $\bar{\delta} \in (0, 1)$ such that for all $\delta \in (\bar{\delta}, 1)$, $F(\delta) > 0$ and σ_1^a is the best response for agent 1. Using similar arguments it is easy to show that σ_2^a is the best response for agent 2 against σ_1^a and σ_3^a and σ_3^a is the best response for agent 3 against σ_1^a and σ_2^a. Hence σ^a is a Nash equilibrium of $G_\delta^\infty(u_1, u_2, u_3)$ for all $\delta \in (\bar{\delta}, 1)$. Finally, since the punishment play induced by σ^a is either of the three pure strategy Nash equilibria $(R1, R2, R1)$, $(R2, R1, R1)$ and $(R1, R1, R2)$; it is credible. Hence, for all $\delta \in (\bar{\delta}, 1)$, the strategy profile $\sigma^a = (\sigma_1^a, \sigma_2^a, \sigma_3^a)$ is also a sub-game perfect equilibrium. □

To implement the cyclically fair norm for the KPR problem for the case (N6), that is for the stage game $G(u_1, u_2, u_3)$ satisfying $u_3 < \frac{u_1}{2} \le u_2 \le u_1$, we consider the strategy profile $\sigma^b = (\sigma_1^b, \sigma_2^b, \sigma_3^b)$ that specifies the following conditions.

(i) Without loss of generality at $t = 1$, each agent $i \in \{1, 2, 3\}$ plays Ri.
(ii) If agent i plays $R1$ in period $t-1$, then i plays $R3$ in period t.
(iii) If agent i plays $Rk \neq R1$ in period $t-1$, then i plays $R(k-1)$ in period t.
(iv) If any agent i violates either of 1, 2 or 3 in some period t, then we have the following:

a. If the deviation is initiated by agent 1, then, for all future periods, agent 2 plays $R2$ and agent 3 plays $R1$.
b. If the deviation is initiated by agent 2, then, for all future periods, agent 1 plays $R1$ and agent 3 plays $R2$.
c. If the deviation is initiated by agent 3, then, for all future periods, agent 1 plays $R2$ and agent 2 plays $R1$.

The first three conditions of strategy profile σ^b are identical to that of the strategy profile σ^a since these three conditions are meant to induce cooperative behavior across agents in order to implement the cyclically fair norm. However, the punishment scheme (that is, condition (iv)) under the strategy profile σ^b is different and more subtle compared the punishment scheme under σ^a. Under σ^b, the two non-deviating agents punish the deviating agent by going to two different restaurant by playing $R1$ and $R2$. The best response to this behavior, at any stage game (irrespective of the identity of the deviating agent) is to go to $R1$. Thus, the punishment scheme generates any one of the three pure strategy Nash equilibria—$(R1, R2, R1)$, $(R2, R1, R1)$ and $(R1, R1, R2)$; where the deviating agent gets stage game payoff of $\frac{u_1}{2}$ for all future periods after the deviation period.

Proposition 13.4 *If* $u_3 < \frac{u_1}{2} \leq u_2 \leq u_1$, *then there exists* $\delta^* \in (0, 1)$ *such that for all* $\delta \in (\delta^*, 1)$, *the strategy profile* $\sigma^b = (\sigma_1^b, \sigma_2^b, \sigma_3^b)$ *is a sub-game perfect equilibrium of* $G_\delta^\infty(u_1, u_2, u_3)$.

Proof We first show that if agent 2 plays σ_2^b and agent 3 plays σ_3^b then playing σ_1^b is the best response for agent 1. Observe that the most profitable deviation at any time t available to 1 is to play $R1$ in that period t where the prescribed strategy under σ^b for agent 1 is $R3$. If agent 1 decides to deviate then, as per σ^b, for all future periods, agent 2 plays $R2$ and agent 3 plays $R1$. Given this punishment strategy followed by agents 2 and 3, the best response of agent 1, in all future periods, is to play $R1$. Therefore, the resultant punishment play at each period in future is $(R1, R2, R1)$ with each stage payoff of $\frac{u_1}{2}$ to 1. So, the present discounted value of the payoff sequence that results after deviation for agent 1 is $D_2(\delta) = \frac{u_1}{2} + \delta\frac{u_1}{2} + \delta^2\frac{u_1}{2} + \cdots = \frac{u_1}{2(1-\delta)}$. By not deviating in a period t (where agent 1 had to play $R3$) and following σ_1^b, agent 1 gets a present discounted value payoff that equals $E_3(\delta) = (u_3+\delta u_2+\delta^2 u_1)+\cdots = \frac{u_3+\delta u_2+\delta^2 u_1}{1-\delta^3}$. If for any $\delta \in (0, 1)$, $E_3(\delta) - D_2(\delta) > 0$, then σ_1^b is the best response for agent 1. Observe that $E_3(\delta) - D_2(\delta) = \frac{G(\delta)}{2(1-\delta^3)}$ where $G(\delta) = 2(u_3 + \delta u_2 + \delta^2 u_1) - u_1(1 + \delta + \delta^2)$. Note that $G(\delta)$ is continuous in δ and, given (N6), $G(0) = 2u_3 - u_1 < 0$ and $G(1) = 2(u_2 + u_3 - \frac{u_1}{2}) > 0$. Hence, there exists a $\delta^* \in (0, 1)$ such that for all $\delta \in (\delta^*, 1)$, $G(\delta) > 0$ and σ_1^b is the best response for agent 1. Using similar arguments it is easy to show that σ_2^b is the best response for agent 2 and σ_3^b is the best response for agent 3. Hence σ^b is a Nash equilibrium of $G_\delta^\infty(u_1, u_2, u_3)$ for all $\delta \in (\delta^*, 1)$. Finally, since the punishment play induced by σ^b is a pure strategy Nash equilibrium, it is credible. Hence, for all $\delta \in (\delta^*, 1)$, the strategy profile $\sigma^b = (\sigma_1^b, \sigma_2^b, \sigma_3^b)$ is also a sub-game perfect equilibrium. □

13.6 Summary

Using simple and basic techniques from infinitely repeated games with discounting we have established how with small number of players one can sustain the cyclically fair norm as an equilibrium in a KPR problem with general preference structures. In particular, we have highlighted how to design pure strategies, that at times requires careful designing of the punishment scheme for perpetrators, to sustain the cyclically fair norm that requires cyclical stage game sacrifices on part of the agents. We have established the following results.

(i) If for the KPR problem the set of pure strategy Nash equilibria of the stage game includes the set of all Pareto efficient strategies then there is no need to design punishment schemes to implement the cyclically fair norm as a subgame perfect equilibrium of the KPR problem.
(ii) Sufficiently high patience level of the agents and the design of appropriate punishment strategies become mandatory when, for the KPR problem with either two agents or three agents, the set of pure strategy Nash equilibria of the stage game does not include the set of all Pareto efficient strategies.
(iii) The punishment scheme that works for the two agent case is one where the deviating agent is punished by shifting to the inefficient Nash equilibrium of the stage game for all future periods after the deviation. This kind of punishment is enough to deter a rational agent with sufficiently high patience level from unilateral deviation.
(iv) For the three agent KPR problem one needs to design different types of punishment schemes as, depending on the restrictions on the (common) preferences, we have different sets of pure strategy Nash equilibria of the stage game. The restrictions on preferences that are of interest are the following—(a) $\frac{u_1}{2} > u_2 > u_3$ and (b) $u_2 \geq \frac{u_1}{2} > u_3$. For both these cases the set of pure strategy Nash equilibria of the stage game are identical and yet one needs to design different pure strategies to implement the cyclically fair norm. For both (a) and (b), the pure strategy Nash equilibrium of the stage game requires two agents going to the first restaurant and one agent going to the second restaurant. However, for case (a), the expected payoff associated with going to the first restaurant is more than the payoff obtained from going to the second restaurant, but, for case (b), the expected payoff associated with going to the first restaurant is weakly less than the payoff obtained from going to the second restaurant. Therefore, while designing the punishment scheme for the perpetrators one needs to incorporate this payoff difference across (a) and (b) and hence we require two different strategies to implement the same cyclically fair norm.

We believe that for the KPR problems with more than three agents and general preference structure, the designing of punishment schemes to implement the cyclically fair norm is an important issue that needs to be addressed in greater detail.

References

1. Becker JG, Damianov DS (2006) On the existence of symmetric mixed strategy equilibria. Econ Lett 90:84–87
2. Chakrabarti AS, Chakrabarti BK, Chatterjee A, Mitra M (2009) The kolkata paise restaurant problem and resource utilization. Physica A 388(12):2420
3. Chakrabarti BK, Chakraborti A, Chatterjee A (eds) (2006) Econophysics and sociophysics. Wiley-VCH, Berlin
4. Friedman J (1971) A non-cooperative equilibrium for supergames. Rev Econ Stud 38:1–12
5. Ghosh A, Chakrabarti AS, Chakrabarti BK (2010) Kolkata paise restaurant problem in some uniform learning strategy. In: Basu B, Chakrabarti BK, Ghosh A, Chakravarty SR, Gangopadhyay K (eds) Econophysics and economics of games, social choices and quantitative techniques. Springer, Milan
6. Ghosh A, Chakrabarti BK (2009) Kolkata paise restaurant (KPR) problem. http://demonstrations.wolfram.com/KolkataPaiseRestaurantKPRProblem
7. Ghosh A, Chatterjee A, Mitra M, Chakrabarti BK (2010) Statistics of the kolkata paise restaurant problem. New J Phys 12:075033
8. Gibbons R (1992) Game theory for applied economists. Princeton University Press, Princeton
9. Mailath GJ, Samuelson L (2006) Repeated games and reputations: long-run relationships. Oxford University Press, London

Chapter 14
An Introduction to Multi-player, Multi-choice Quantum Games: Quantum Minority Games & Kolkata Restaurant Problems

Puya Sharif and Hoshang Heydari

Abstract We give a self contained introduction to a few quantum game protocols, starting with the quantum version of the two-player two-choice game of Prisoners dilemma, followed by an n-player generalization trough the quantum minority games, and finishing with a contribution towards an n-player m-choice generalization with a quantum version of a three-player Kolkata restaurant problem. We have omitted some technical details accompanying these protocols, and instead laid the focus on presenting some general aspects of the field as a whole. This review contains an introduction to the formalism of quantum information theory, as well as to important game theoretical concepts, and is aimed to work as a review suiting economists and game theorists with limited knowledge of quantum physics as well as to physicists with limited knowledge of game theory.

14.1 Introduction

Quantum game theory is the natural intersection between three fields. Quantum mechanics, information theory and game theory. At the center of this intersection stands one of the most brilliant minds of the 20th century, John von Neumann. As one of the early pioneers of quantum theory, he made major contributions to the mathematical foundation of the field, many of them later becoming core concepts in the merger between quantum theory and information theory, giving birth to quantum computing and quantum information theory [1], today being two of the most active fields of research in both theoretic and experimental physics. Among economists may he be mostly known as the father of modern game theory [2–4], the study of rational interactions in strategic situations. A field well rooted in the influential book *Theory of Games and Economic Behavior* (1944), by Von Neumann and Oscar Morgenstern. The book offered great advances in the analysis of strategic games

P. Sharif (✉) · H. Heydari
Department of Physics, Stockholm University, 10691 Stockholm, Sweden
e-mail: ps@puyasharif.net

H. Heydari
e-mail: hoshang@fysik.su.se

F. Abergel et al. (eds.), *Econophysics of Systemic Risk and Network Dynamics*, New Economic Windows,
DOI 10.1007/978-88-470-2553-0_14, © Springer-Verlag Italia 2013

and in the axiomatization of measurable utility theory, and drew the attention of economists and other social scientists to these subjects. For the last decade or so there has been an active interdisciplinary approach aiming to extend game theoretical analysis into the framework of quantum information theory, through the study of quantum games [5–10]; offering a variety of protocols where use of quantum peculiarities like entanglement in quantum superpositions, and interference effects due to quantum operations has shown to lead to advantages compared to strategies in a classical framework. The first papers appeared in 1999. Meyer showed with a model of a penny-flip game that a player making a *quantum move* always comes out as a winner against a player making a *classical* move regardless of the classical players choice [11]. The same year Eisert et al. published a quantum protocol in which they overcame the dilemma in Prisoners dilemma [12]. In 2003 Benjamin and Hayden generalized Eisert's protocol to handle multi-player quantum games and introduced the quantum minority game together with a solution for the four player case which outperformed the classical randomization strategy [13]. These results were later generalized to the n-players by Chen et al. in 2004 [14]. Multi-player minority games has since then been extensively investigated by Flitney et al. [15–17]. An extension to multi-choice games, as the Kolkata restaurant problem was offered by the authors of this review, in 2011 [18].

14.1.1 Games as Information Processing

Information theory is largely formulated independent of the physical systems that contains and processes the information. We say that the theory is substrate independent. If you read this text on a computer screen, those bits of information now represented by pixels on your screen has traveled through the web encoded in electronic pulses through copper wires, as burst of photons trough fiber-optic cables and for all its worth maybe on a piece of paper attached to the leg of a highly motivated raven. What matters from an information theoretical perspective is the existence of a differentiation between some states of affairs. The general convention has been to keep things simple and the smallest piece of information is as we all know a *bit* $b \in \{0, 1\}$, corresponding to a binary choice: *true* or *false*, *on* or *off*, or simply *zero* or *one*. Any chunk of information can then be encoded in strings of bits: $\mathbf{b} = b_{n-1}b_{n-2}\cdots b_0 \in \{0, 1\}^n$. We can further define functions on strings of bits, $f : \{0, 1\}^n \to \{0, 1\}^k$ and call these functions computations or actions of information processing.

In a similar sense games are in their most general form independent of a physical realization. We can build up a formal structure for some strategic situation and model cooperative and competitive behavior within some constrained domain without regards to who or what these game playing agents are or what their actions actually is. No matter if we consider people, animals, cells, multinational companies or nations, simplified models of their interactions and the accompanied consequences can be formulated in a general form, within the framework of game theory.

Lets connect these two concepts with an example. We can create a one to one correspondence with between the conceptual framework of game theory and the formal structure of information processing. Let there be n agents faced with a binary choice of joining one of two teams. Each choice is represented by a binary bit $b_i \in \{0, 1\}$. The final outcome of these individual choices is then given by an n-bit output string $\mathbf{b} \in \{0, 1\}^n$. We have 2^n possible outcomes, and for each agent we have some preference relation over these outcomes $\mathbf{b}_j$. For instance, agent 1 may prefer to have agent 3 in her team over agent 4, and may prefer any configuration where agent 5 is on the other team over any where they are on the same and so on. For each agent i, we'll have a preference relation of the following form, fully determining their objectives in the given situation:

$$\mathbf{b}_{x_1} \succeq \mathbf{b}_{x_2} \succeq \cdots \succeq \mathbf{b}_{x_m}, \quad m = 2^n, \tag{14.1}$$

where $\mathbf{b}_{x_i} \succeq \mathbf{b}_{x_j}$ means that the agent in question prefers $\mathbf{b}_{x_i}$ to $\mathbf{b}_{x_j}$, or is at least indifferent between the choices. To formalize things further we assign a numerical value to each outcome $\mathbf{b}_{x_j}$ for *each* agent, calling it the *payoff* $\$_i(\mathbf{b}_{x_j})$ to agent i due to outcome $\mathbf{b}_{x_j}$. This allows us to move from the preference relations in (14.1) to a sequence of inequalities. $\mathbf{b}_{x_i} \succeq \mathbf{b}_{x_j} \iff \$(\mathbf{b}_{x_i}) \geq \$(\mathbf{b}_{x_j})$. The aforementioned binary choice situation can now be formulated in terms of functions $\$_i(\mathbf{b}_{x_j})$ of the output strings $\mathbf{b}_{x_j}$, where each entry b_i in the strings corresponds to the choice of an agent i.

So far has the discussion only regarded the output string without mentioning any input. We could without loss of generality define an input as string where all the entries are initialized as 0's, and the individual choices being encoded by letting each participant either leave their bit unchanged or performing a NOT-operation, where $\text{NOT}(0) = 1$.

More complicated situations with multiple choices could be modeled by letting each player control more than one bit or letting them manipulate strings of information bearing units with more states than two; of which we will see an example of later.

14.1.2 Quantization of Information

Before moving on to the quantum formalism of operators and quantum states, there is one intermediate step worth mentioning, the *probabilistic* bit, which has a certain probability p of being in one state and a probability of $1 - p$ of being in the other. If we represent the two states '0' and '1' of the ordinary bit by the two-dimensional vectors $(1, 0)^T$ and $(0, 1)^T$, then a probabilistic bit is given by a linear combination of those basis vectors, with real positive coefficients p_0 and p_1, where $p_0 + p_1 = 1$. In this formulation, randomization between two different choices in a strategic situation would translate to manipulating an appropriate probabilistic bit.

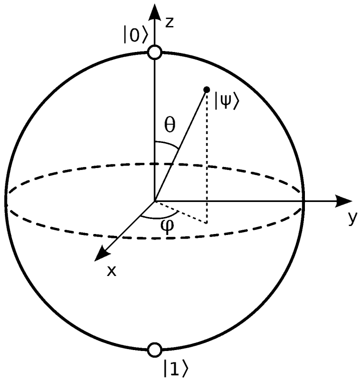

Fig. 14.1 The Bloch sphere. A geometric representation of the state space of a single qubit

The Quantum Bit Taking things a step further, we introduce the quantum bit or the *qubit*, which is a representation of a two level quantum state, such as the spin state of an electron or the polarization of a photon. A qubit lives in a two dimensional complex space spanned by two basis states denoted $|0\rangle$ and $|1\rangle$, corresponding to the two states of the classical bit.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{14.2}$$

Unlike the classical bit, the qubit can be in any superposition of $|0\rangle$ and $|1\rangle$:

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle, \tag{14.3}$$

where a_0 and a_1 are complex numbers obeying $|a_0|^2 + |a_1|^2 = 1$. $|a_i^2|$ is simply the probability to find the system in the state $|i\rangle, i \in \{0, 1\}$. Note the difference between this and the case of the probabilistic bit! We are now dealing with complex coefficients, which means that if we superpose two qubits, then some coefficients might be eliminated. This interference is one of many effects without counterpart in the classical case. The state of an arbitrary qubit can be written in the *computational basis* as:

$$|\psi\rangle = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}. \tag{14.4}$$

The state of a general qubit can be parameterized as:

$$|\psi\rangle = \cos\left(\frac{\vartheta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\vartheta}{2}\right)|1\rangle, \tag{14.5}$$

where we have factored out and omitted a global phase due to the physical equivalence between the states $e^{i\phi}|\psi\rangle$ and $|\psi\rangle$. This so called *state vector* describes a point on a spherical surface with $|0\rangle$ and $|1\rangle$ at its poles, called the Bloch-sphere, parameterized by two real numbers θ and φ, depicted in Fig. 14.1. A simple comparison between the state space of the classical, probabilistic and quantum bit is shown in Fig. 14.2.

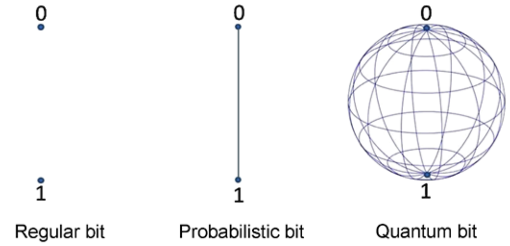

Fig. 14.2 The classical bit has only two distinct states, the probabilistic bit can be in any normalized convex combination of those states, whereas the quantum bit has a much richer state space

14.1.2.1 Hilbert Spaces and Composite Systems

The state vector of a quantum system is defined in a complex vector space called *Hilbert space* $\mathscr{H}$. Quantum states are represented in common Dirac notation as "ket's", written as the right part $|\psi\rangle$ of a bracket ("bra-ket"). Algebraically a "ket" is column vector in our state space. This leaves us to define the set of "bra's" $\langle\phi|$ on the dual space of $\mathscr{H}$, $\mathscr{H}^\star$. The dual Hilbert space $\mathscr{H}^\star$ is defined as the set of linear maps $\mathscr{H} \rightarrow \mathbf{C}$, given by

$$\langle\phi| : |\psi\rangle \mapsto \langle\phi|\psi\rangle \in \mathbb{C}, \tag{14.6}$$

where $\langle\phi|\psi\rangle$ is the inner product of the vectors $|\psi\rangle$, $|\phi\rangle \in \mathscr{H}$. We can now write down a more formal definition of a Hilbert space: It is a complex inner product space with the following properties:

(i) $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^\dagger$, where $\langle\psi|\phi\rangle^\dagger$ is the complex conjugate of $\langle\psi|\phi\rangle$.
(ii) The inner product$\langle\phi|\psi\rangle$ is linear in the first argument: $\langle a\phi_1 + b\phi_2|\psi\rangle = a^\dagger\langle\phi_1|\psi\rangle + b^\dagger\langle\phi_2|\psi\rangle$.
(iii) $\langle\psi|\psi\rangle \geq 0$.

The space of an n-qubit system is spanned by a basis of 2^n orthogonal vectors $|e_i\rangle$; one for each possible combination of the basis-states of the individual qubits, obeying the orthogonality condition:

$$\langle e_i|e_j\rangle = \delta_{ij}, \tag{14.7}$$

where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. We say that the Hilbert space of a composite system is the tensor products of the Hilbert spaces of its parts. So the space of an n-qubit system is simply the tensor product of the spaces of the n qubits.

$$\mathscr{H}_{\mathscr{Q}} = \mathscr{H}_{\mathscr{Q}_n} \otimes \mathscr{H}_{\mathscr{Q}_n-1} \otimes \mathscr{H}_{\mathscr{Q}_{n-2}} \otimes \cdots \otimes \mathscr{H}_{\mathscr{Q}_1}, \tag{14.8}$$

where $\mathscr{Q}_i$ the quantum system i is a vector in $\mathbf{C}^2$. A general n-qubit system can therefore be written

$$|\psi\rangle = \sum_{x_n,\ldots,x_1=0}^{1} a_{x_n\cdots x_1}|x_n\cdots x_1\rangle, \tag{14.9}$$

where

$$|x_n \cdots x_1\rangle = |x_n\rangle \otimes |x_{n-1}\rangle \otimes \cdots \otimes |x_1\rangle \in \mathscr{H}_{\mathscr{Q}} \tag{14.10}$$

with $x_i \in \{0, 1\}$ and complex coefficients a_{x_i}. For a two qubit system, $|x_2\rangle \otimes |x_1\rangle = |x_2\rangle|x_1\rangle = |x_2 x_1\rangle$, we have

$$|\psi\rangle = \sum_{x_2, x_1=0}^{1} a_{x_2 x_1}|x_2 x_1\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle. \tag{14.11}$$

This state space is therefore spanned by four basis vectors:

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle, \tag{14.12}$$

which are represented by the following 4-dimensional column vectors respectively:

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{14.13}$$

14.1.2.2 Operators

A linear operator on a vector space $\mathscr{H}$ is a linear transformation $\mathrm{T} : \mathscr{H} \to \mathscr{H}$, that maps vectors in $\mathscr{H}$ to vectors in the same space $\mathscr{H}$. Quantum states are normalized, and we wish to keep the normalization; we are therefore interested in transformations that can be regarded as rotations in $\mathscr{H}$. Such transformations are given by *unitary operators* U. An operator U is called unitary if $U^{-1} = U^{\dagger}$. They preserve inner products between vectors, and thereby their norm. A *projection operator* P is Hermitian i.e. $P = P^{\dagger}$ and satisfies $P^2 = P$. We can create a projector P, by taking the outer product of a vector with itself:

$$P = |\phi\rangle\langle\phi|. \tag{14.14}$$

P is a matrix with every element P_{ij} being the product of the elements i, j of the vectors in the outer product. This operator projects any vector $|\gamma\rangle$ onto the 1-dimensional subspace of $\mathscr{H}$, spanned by $|\phi\rangle$:

$$P|\gamma\rangle = |\phi\rangle\langle\phi||\gamma\rangle = \langle\phi|\gamma\rangle|\phi\rangle. \tag{14.15}$$

It simply gives the portion of $|\gamma\rangle$ along $|\phi\rangle$.

We will often deal with unitary operators $U \in \mathrm{SU}(2)$, i.e operators from the *special unitary group* of dimension 2. The group consists of 2×2 unitary matrices with determinant 1. These matrices will be operating on single qubits (often in systems of 2 or more qubits). The generators of the group are the *Pauli spin matrices* σ_x, σ_y, σ_z, shown together with the identity matrix I:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{14.16}$$

Note that σ_x is identical to a classical (bit-flip) 'NOT'-operation. General 2×2 unitary operators can be parameterized with three parameters θ, α, β, as follows:

$$U(\theta, \alpha, \beta) = \begin{pmatrix} e^{i\alpha}\cos(\theta/2) & ie^{i\beta}\sin(\theta/2) \\ ie^{-i\beta}\sin(\theta/2) & e^{-i\alpha}\cos(\theta/2) \end{pmatrix}. \tag{14.17}$$

An operation is said to be local if it only affects a part of a composite (multi-qubit) system. Connecting this to the concept of the bit-strings in the previous section; a local operation translates to just controlling one such bit. This is a crucial point in the case of modeling the effect of individual actions, since each agent in a strategic situation is naturally constrained to decisions regarding their own choices. The action of a set of local operations on a composite system is given by the tensor product of the local operators. For a general n-qubit $|\psi\rangle$ as given in (14.9) and (14.10) we get:

$$U_n \otimes U_{n-1} \otimes \cdots \otimes U_1 |\psi\rangle = \sum_{x_n,\ldots,x_1=0}^{1} a_{x_n \ldots x_1} U_n|x_n\rangle \otimes U_{n-1}|x_{n-1}\rangle \otimes \cdots \otimes U_1|x_1\rangle. \tag{14.18}$$

14.1.2.3 Mixed States and the Density Operator

We have so far only discussed *pure states*, but sometimes we encounter quantum states without a definite state vector $|\psi\rangle$, these are called *mixed states* and consists of a states that has certain probabilities of being in some number of different pure states. So for example a state that is in $|\psi_1\rangle = a_0^1|0\rangle + a_1^1|1\rangle$ with probability p_1 and in $|\psi_2\rangle = a_0^2|0\rangle + a_1^2|1\rangle$ with probability p_2 is mixed. We handle mixed states by defining a density operator ρ, which is a hermitian matrix with unit trace:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{14.19}$$

where $\sum_i p_i = 1$. A pure state in this representation is simply a state for which all probabilities, except one is zero. If we apply a unitary operator U on a pure state, we end up with $U|\psi\rangle$ which has the density operator $U\rho U^\dagger = U|\psi\rangle\langle\psi|U^\dagger$. Regardless if we are dealing with pure or mixed states, we take the expectation value of upon measurement ending up in a $|\phi\rangle$ by calculating $\mathrm{Tr}(|\phi\rangle\langle\phi|\rho)$, where $|\phi\rangle\langle\phi|$ is a so called projector. For calculating the expectation values of a state to be in *any* of a number of states $|\phi_i\rangle$, we construct a projection operator $P = \sum_i |\phi_i\rangle\langle\phi_i|$ and take the trace over P multiplied by ρ.

14.1.2.4 Entanglement

Entanglement is the resource our game-playing agents will make use of in the quantum game protocols to achieve better than classical performance. Non-classical correlations are thus introduced, by which the players can synchronize their behavior without any additional communication. An entangled state is basically a quantum system that *cannot* be written as a tensor product of its subsystems, we'll thus define two classes of quantum states. Examples below refers to two-qubit states.

Product states:

$$|\Psi_{\mathscr{Q}}\rangle = |\Psi_{\mathscr{Q}_2}\rangle \otimes |\Psi_{\mathscr{Q}_1}\rangle, \quad \text{or using density matrix} \quad \rho_{\mathscr{Q}} = \rho_{\mathscr{Q}_2} \otimes \rho_{\mathscr{Q}_1}, \tag{14.20}$$

and entangled states

$$|\Psi_{\mathscr{Q}}\rangle \neq |\Psi_{\mathscr{Q}_2}\rangle \otimes |\Psi_{\mathscr{Q}_1}\rangle, \quad \text{or using density matrix} \quad \rho_{\mathscr{Q}} \neq \rho_{\mathscr{Q}_2} \otimes \rho_{\mathscr{Q}_1}. \tag{14.21}$$

For a mixed state, the density matrix is defined as mentioned by $\rho_{\mathscr{Q}} = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i|$ and it is said to be separable, which we will denote by $\rho_{\mathscr{Q}}^{sep}$, if it can be written as

$$\rho_{\mathscr{Q}}^{sep} = \sum_i p_i \left(\rho_{\mathscr{Q}_2}^i \otimes \rho_{\mathscr{Q}_1}^i\right), \qquad \sum_i p_i = 1. \tag{14.22}$$

A set of very important two-qubit entangled states are the Bell states

$$\left|\Phi_{\mathscr{Q}}^{\pm}\right\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big), \qquad \left|\Psi_{\mathscr{Q}}^{\pm}\right\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle \pm |10\rangle\big). \tag{14.23}$$

The GHZ-type-states

$$|\mathrm{GHZ}_n\rangle = \frac{1}{\sqrt{2}}\big(|00\cdots 0\rangle + e^{i\phi}|11\cdots 1\rangle\big) \tag{14.24}$$

could be seen as an n-qubit generalization of $|\Phi_{\mathscr{Q}}^{\pm}\rangle$-states.

14.1.3 Classical Games

It is instructive to review the theory of classical games and some major solution concepts before moving on to examples of quantum games. We'll start by defining classical pure and mixed strategy games, and then move on to introducing some relevant solution concepts and finish off with a definition of quantum games.

A game is a formal model over the interactions between a number of agents (*agents, players, participants*, and *decision makers* may be used interchangeably) under some specified sets of choices (*choices, strategies, actions* and *moves*, may be used interchangeably). Each combination of choices made, or strategies chosen by the different players leads to an outcome with some certain level of desirability for each of them. The level of desirability is measured by assigning a real number, a so called *payoff* \$ for each game outcome for each player. Assuming rational players, each will choose actions that maximizes their expected payoff $E(\$)$, i.e. in an deterministic as well as in an probabilistic setting acting in a way that, based on the known information about the situation, maximizes the expectation value of their payoff. The structure of the game is fully specified by the relations between the different combinations of strategies and the payoffs received by the players. A key point is the interdependence of the payoffs with the strategies chosen by the other players. A situation where the payoff of one player is independent of the strategies of the others would be of little interest from a game theoretical point of view. It is natural to extend the notion of payoffs to *payoff functions* whose arguments are the chosen strategies of all players and ranges are the real valued outputs that assigns a level of desirability for each player to each outcome.

Pure Strategy Classical Game We have a set of n players $\{1, 2, \ldots, n\}$, n strategy sets S_i, one for each player i, with $s_i^j \in S_i$, where s_i^j is the jth strategy of player i. The strategy space $S = S_1 \times S_2 \times \cdots \times S_n$ contains all n-tuples pure strategies, one from each set. The elements $\sigma \in S$ are called strategy profiles, some of which will earn them the status of being a *solution* with regards to some solution concept.

We define a game by its payoff-functions $\$_i$, where each is a mapping from the strategy space S to a real number, the payoff or utility of player i. We have:

$$\$_i : S_1 \times S_2 \times \cdots \times S_n \to \mathbf{R}. \tag{14.25}$$

Mixed Strategy Classical Game Let $\Delta(S_i)$ be the set of convex linear combinations of the elements $s_i^j \in S_i$. A mixed strategy $s_i^{mix} \in \Delta(S_i)$ is then given by:

$$\sum_{s_i^j \in S_i} p_i^j s_i^j \quad \text{with} \sum_j p_i^j = 1, \tag{14.26}$$

where p_i^j is the probability player i assigns to the choice s_i^j. The space of mixed strategies $\Delta(S) = \Delta(S_1) \times \Delta(S_2) \times \cdots \times \Delta(S_n)$ contains all possible mixed strategy profiles σ_{mix}. We now have:

$$\$_i : \Delta(S_1) \times \Delta(S_2) \times \cdots \times \Delta(S_n) \to \mathbf{R}. \tag{14.27}$$

Note that the pure strategy games are fully confined within the definition of mixed strategy games and can be accessed by assigning all strategies except one, the probability $p^j = 0$. This class of games could be formalized in a framework using probabilistic information units, such as the probabilistic bit.

14.1.4 Solution Concepts

We will introduce two of many game theoretical solution concepts. A solution concept is a strategy profile $\sigma^* \in S$, that has some particular properties of strategic interest. It could be a strategy profile that one would expect a group of rational self-maximizing agents to arrive at in their attempt to maximize their minimum expected payoff. Strategy profiles of this form i.e. those that leads to a combination of choices where each choice is the best possible response to any possible choice made by other players tend to lead to an equilibrium, and are good predictors of game outcomes in strategic situations. To see how such equilibria can occur we'll need to develop the concept of *dominant strategies*.

Definition 14.1 (Strategic dominance) A strategy $s^{dom} \in S_i$ is said to be dominant for player i, if for any strategy profile $\sigma_{-i} \in S/S_i$, and any other strategy $s^j \neq s^{dom} \in S_i$:

$$\$_i\left(s^{dom}, \sigma_{-i}\right) \geq \$_i\left(s^j, \sigma_{-i}\right) \quad \text{for all } i = 1, 2, \ldots, n. \tag{14.28}$$

Lets look at a simple example. Say that we have two players, Alice with legal strategies s^1_{Alice}, $s^2_{Alice} \in S_{Alice}$ and Bob with s^1_{Bob}, $s^2_{Bob} \in S_{Bob}$. Now, if the payoff Alice receives when playing s^1_{Alice} against any of Bob's two strategies is higher than (or at least as high as) what she receives by playing s^2_{Alice}, then s^1_{Alice} is her dominant strategy. Her payoff can of course vary depending on Bob's move but regardless what Bob does, her dominant strategy is the *best response*. Now there is no guarantee that such dominant strategy exists in a pure strategy game, and often must the strategy space be expanded to accommodate for mixed strategies for them to exist.

If both Alice and Bob has a dominant strategy, then this strategy profile becomes a *Nash Equilibrium*, i.e. a combination of strategies for which none of them can gain by unilaterally deviating from. The Nash equilibrium profile acts as an attractor in the strategy space and forces the players into it, even though it is not always an optimal solution. Combinations can exist that can lead to better outcomes for both (all) players.

Definition 14.2 (Nash equilibrium) Let $\sigma^{NE}_{-i} \in S/S_i$ be a strategy profile containing the dominant strategies of every player except player i, and let $s^{NE}_i \in S_i$ be the dominant strategy of player i. Then for all $s^j_i \neq s^{NE}_i \in S_i$:

$$\$_i\left(s^{NE}_i, \sigma^{NE}_{-i}\right) \geq \$_i\left(s^j_i, \sigma^{NE}_{-i}\right) \quad \text{for all } i = 1, 2, \ldots, n. \tag{14.29}$$

If we have a situation where an agent can increase its payoff without decreasing any others, then this would per definition mean that nobody would mind if that agent would do so. Each such increase in payoff is called a *Pareto improvement*. When no such improvement can be done, then the strategy profile is said to be *Pareto optimal*.

Definition 14.3 (Pareto efficiency) A Pareto efficient or Pareto optimal strategy profile is one where none of the participating agents can increase their payoff without decreasing the payoff of someone else.

14.2 Quantum Games

In the quantum game protocols (*protocol* and *scheme* may be used interchangeably) presented in this paper, the m_i different choices available to a player i will be encoded in the basis states of an m_i-level quantum system, where the m_i denotes the dimensionality of the Hilbert space $\mathscr{H}_{\mathscr{Q}_i}$ associated with that subsystem. Each of the n player holds one subsystem leading to a total system with a state vector a in an $\prod_{i=1}^{n}\dim(\mathscr{H}_{\mathscr{Q}_i})$—dimensional space. The definition of a quantum game must therefore include a Hilbert space of a multipartite multilevel system $\mathscr{H}_{\mathscr{Q}} = \mathscr{H}_{\mathscr{Q}_n} \otimes \mathscr{H}_{\mathscr{Q}_{n-1}} \otimes \cdots \otimes \mathscr{H}_{\mathscr{Q}_1}$.

The different subsystems must in general be allowed to have a have a common origin to accommodate entanglement in the shared initial state $\rho_{in} \in \mathscr{H}_{\mathscr{Q}}$. This is often modeled by including a referee that prepares an initial state and distributes the subsystems among the players. Whether or not this step invokes on the non-communication criteria certain games have, is under debate. We justify it by the fact that no communication is done under the crucial step of choosing a strategy. The strategies are applied by local quantum operations on the quantum state held by each player. No player has any access to any part of the system except its own subsystem, and no information can be sent between the players with aid of the shared quantum resource. Classical strategies becomes quantum strategies by expanding the strategy sets:

$$s_i \in S_i \Rightarrow U_i \in \mathrm{S}(m_i), \tag{14.30}$$

where the set of allowed quantum operations $\mathrm{S}(m_i)$ is some subset of the special unitary group $\mathrm{SU}(m_i)$. We will later see that the nature of the game can be determined by restrictions on $\mathrm{S}(m_i)$. It is an important point to be able to show that the classical version of a game is recoverable just by restricting the set of allowed operators. At least if we want it to be a *proper quantization* [9], i.e. an extension of the classical game into the quantum realm, and not a whole new game without a classical counterpart.

We define a quantum game in two steps:

$$U_n \otimes U_{n-1} \otimes \cdots \otimes U_1 : \mathscr{H}_{\mathscr{Q}_n} \otimes \mathscr{H}_{\mathscr{Q}_{n-1}} \otimes \cdots \otimes \mathscr{H}_{\mathscr{Q}_1} \rightarrow \mathscr{H}_{\mathscr{Q}_n} \otimes \mathscr{H}_{\mathscr{Q}_{n-1}} \otimes \cdots \otimes \mathscr{H}_{\mathscr{Q}_1}, \tag{14.31}$$

$$\$_i : \mathscr{H}_{\mathscr{Q}_n} \otimes \mathscr{H}_{\mathscr{Q}_{n-1}} \otimes \cdots \otimes \mathscr{H}_{\mathscr{Q}_1} \rightarrow \mathbf{R}, \tag{14.32}$$

where the first step is a transformation of the state of the complete system by local operations, and the second is a mapping from the Hilbert space of the quantum state to a real number, the expected payoff of player i.

14.2.1 The Quantum Game Protocol

- The game begins with an entangled initial state $|\psi_{in}\rangle$. Each subsystem has a dimensionality m that equal to the number of pure strategies in each players strategy set. In the protocols covered in this paper, all players will face the same number of choices. The number of subsystems equals the number of players. One can assume that $|\psi_{in}\rangle$ has been prepared at some location by a referee that then has distributed the subsystems among the players [12, 13].
- The players then chooses an unitary operator U from a subset of SU(m), and applies it to their subsystem. The initial state ρ_{in} transforms to a final state ρ_{fin}, given by:

$$\rho_{fin} = U \otimes U \otimes \cdots \otimes U \rho_{in} U^{\dagger} \otimes U^{\dagger} \otimes \cdots \otimes U^{\dagger}. \tag{14.33}$$

  In the absence of communication, and due to the symmetry of these games, all players are expected to do the same operation.
- The players then measures their own subsystem, collapsing their quantum states to units of classical information. For the case of a two-choice protocol, each player ends up with a classical bit b_i, and the complete system has thus collapsed into a classical string $\mathbf{b}$, corresponding to a pure strategy profile $\sigma \in S$. For the quantum game to have an advantage over a classical game, the collective action of the players must have decreased the probability of the final state ρ_{fin} to collapse into such basis states (classical information strings/strategy profiles) that are undesired, i.e. leading to lower or zero payoff \$.
- To calculate the expected payoffs $E(\$)$, we define for each player i a payoff-operator P_i, which contains the sum of orthogonal projectors associated with the states for which player i receives a payoff \$. We have:

$$P_i = \sum_j \$_i^j |\chi_i^j\rangle\langle\chi_i^j|, \tag{14.34}$$

  where the states $|\chi_i^j\rangle$ are those sates that leads to a payoff for player i, and $\$_i^j$ the associated payoffs. The expected payoff $E(\$_i)$ of player i is calculated by taking the trace of the product of the final state ρ_{fin} and the payoff-operator P_i:

$$E_i(\$) = \mathrm{Tr}(P_i \rho_{fin}). \tag{14.35}$$

14.2.2 Prisoners Dilemma

The prisoners dilemma is one of the most studied game theoretical problems. It was introduced in 1950 by Merrill Flood and Melvin Dresher, and has been widely used ever since to model a variety of situations, including oligopoly pricing, auction bidding, salesman effort, political bargaining and arms races. In is in its standard form, a symmetric simultaneous game of complete information. Two players, Alice and Bob (A and B) are faced with a choice to *cooperate* or to *defect*, without any

Table 14.1 The normal-form representation of prisoners dilemma

		Bob	
		Cooperate	Defect
Alice	Cooperate	(3, 3)	(0, 5)
	Defect	(5, 0)	(1, 1)

information about the action taken by the other. The payoffs they receive due to any combination of choices is shown in Table 14.1, where the first entry in each parenthesis shows the payoff $\$_A$ of Alice and the second entry the payoff $\$_B$ of Bob. Given that Bob chooses to cooperate, Alice receives $\$_A = 3$ if she chooses to do the same, and she receives $\$_A = 5$ if she chooses to defect. If Bob instead defects, then Alice receives $\$_A = 0$ by cooperating and $\$_A = 1$ by choosing to defect. No matter what Bob does, Alice will always gain by choosing to defect, equipping her with a strictly dominant strategy! Due to the symmetry of the game, the same is true for Bob, forcing them into a Nash equilibrium strategy profile of (defect, defect), which pays out $\$_{AB} = 1$ to each. This outcome is clearly far from efficient, since there is a Pareto optimal strategy profile (cooperate, cooperate) that would have given them $\$_{AB} = 3$, and hence the dilemma.

Quantum prisoners dilemma was introduced by J. Eisert, M. Wilkens, and M. Lewenstein in 1999 [11]. Here Alice and Bob are equipped with a quantum resource, a maximally entangled Bell-type-state, and each of them are in possession of a subsystem. The Hilbert space of the game is given by: $\mathscr{H} = \mathscr{H}_B \otimes \mathscr{H}_A$, with $\mathscr{H}_A = \mathscr{H}_B = \mathbf{C}^2$. We'll identify the following relations, mapping classical outcomes with basis states of the Hilbert space: (cooperate, cooperate) $\rightarrow |00\rangle$, (cooperate, defect) $\rightarrow |01\rangle$, (defect, cooperate) $\rightarrow |10\rangle$ and (defect, cooperate) $\rightarrow |11\rangle$. The entangled initial state is created by acting with an entangling operator $J = \frac{1}{\sqrt{2}} I^{\otimes 2} + i\sigma_x^{\otimes 2}$ on a product state initialized as (cooperate, cooperate):

$$J|00\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + i|11\rangle\big). \tag{14.36}$$

Note that the entangling operator performs a global operation, i.e. an operation performed on both subsystems simultaneously. One can consider it to be performed by a referee, loyal to both parties. The game proceeds by Alice and Bob performing their local strategies U_A and U_B, and the state is turned into its final form: $|\psi_{fin}\rangle = (U_B \otimes U_A) J|00\rangle$. Before measurement is performed, an disentangling operator $J^\dagger$ is applied. The inclusion of J and $J^\dagger$ into the protocol assures that the classical game is embedded into the quantum version, whereby the classical prisoners dilemma can be accessed by restricting the set of allowed operators to $U_A, U_B \in \{I, \sigma_x\}$. It is a simple task to show that any combination of the identity operator I and the bit-flip operator σ_x commutes with J, and together with the fact that $JJ^\dagger = I$, one concludes that this restriction turns the protocol into classical (one-bit) operations on a bit string '00'. The complete protocol is shown as a circuit diagram in Fig. 14.3.

It is now left to define a set of operators U, representing allowed *quantum* strategies, and the payoff operators P_A and P_B. Eisert et al. considered a two parameter subset of SU(2) as the strategy space:

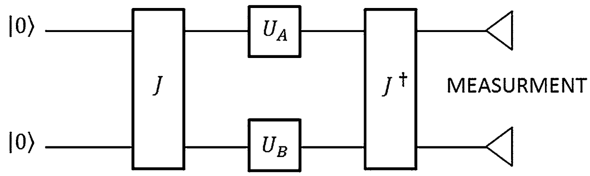

Fig. 14.3 Circuit diagram of the quantum prisoners dilemma protocol

$$U(\theta, \alpha) = \begin{pmatrix} e^{i\alpha}\cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & e^{-i\alpha}\cos(\theta/2) \end{pmatrix}. \tag{14.37}$$

The classical strategies are represented by $U(0,0) = I$ and $U(0,\pi) = \sigma_x$. We construct Alice's payoff operator P_A as defined in (14.34) with values from the payoff matrix:

$$P_A = 3|00\rangle\langle 00| + 5|01\rangle\langle 01| + 1|11\rangle\langle 11|. \tag{14.38}$$

Her expected payoff is calculated by taking the trace of the final state and the payoff operator: $E(\$_A) = \mathrm{Tr}(P_A \rho_{fin})$, where $\rho_{fin} = |\psi_{fin}\rangle\langle\psi_{fin}|$. It can be shown that when the set of strategies are expanded to allow any $U(\theta,\alpha)$, the old Nash equilibrium (defect, defect) $\rightarrow U(0,\pi) \otimes U(0,\pi)$ ceases to exist! Instead a new Nash equilibrium emerges at

$$U_A = U_B = U(0, \pi/2) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{14.39}$$

This strategy leads to an expected payoff $E(\$_A) = E(\$_A) = 3$. Thereby they both receive an expected payoff that equals the Pareto optimal solution in the classical pure strategy version, with the addition that this solution is also a Nash equilibrium. Dilemma resolved. It should be added that if the strategy sets are further expanded to include all SU(2) operations, this solution vanishes, and there is no Nash equilibrium strategy profile in pure quantum strategies, whereby one has to include mixed quantum operations to find an equilibrium [19].

14.2.3 Minority Games

We extend the previous protocol to ones with multiple agents, by introducing the minority game. The game consists of n of non-communicating players that must independently make up their mind between two choices. We could regard these players as investors on a market deciding between two equally attractive securities, as commuters choosing between two equally fast routes to a suburb, or any collection of agents facing situations where they wish to make the minority choice. The core objective of the players are thus to avoid the crowd. We encode the two choices as $|0\rangle$ and $|1\rangle$ in the computational basis like before. The players receive payoff a $\$ = 1$ if they happen to be in the smaller group. So if the number of players choosing $|0\rangle$ is less than the number of players choosing $|1\rangle$, the first group receives payoff whereas the second group is left with nothing. Would the players happen to be evenly distributed between the two choices, then they'll all go empty handed.

The Nash equilibrium solution is to randomize between $|0\rangle$ and $|1\rangle$ using a fair coin. The *one shoot* version we are considering will necessarily have a mixed strategy solution, since any deterministic strategy would lead all players to the same choice and thus a maximally undesired outcome. The expected payoff $E(\$)$ for a player is simply the number outcomes with that player in the minority group divided by the number of different possible outcomes. For a four player game, there are two minority outcomes for each player, out of sixteen possible. This gives a expected payoff of $1/8$.

A quantum version of a four player minority game was presented by Benjamin and Hayden in 2000 [13], offering a solution that significantly outperformed the classical version of the game. The advantage comes from the possibility of eliminating (or reducing the probability of) such final outcomes where the players are evenly distributed among the two choices. The collective application of local unitary operators on the subsystems of an entangled state can thus transform this initial state in such a way that a better-than-classical result is achieved. This transformation does not have a classical analogue, and the performance is due to interference effects from the local phases added to the qubits by the players local operations. We are not including the action of an entangling operator J in this section, we simply assume the initial state to be entangled at the start of the protocol, and it can again be assumed that the state has been prepared by an unbiased referee and distributed among the players. Considering the four-player case, we begin the protocol with an GHZ-type state similar to the one used in the previous two-player game, but now consisting of *four* entangled qubits.

$$|\psi_{in}\rangle = \frac{1}{\sqrt{2}}\big(|0000\rangle + |1111\rangle\big). \tag{14.40}$$

The Hilbert space of the game is sixteen dimensional, accounting for all possible game outcomes. $\mathscr{H}_{\mathscr{Q}} = \mathscr{H}_{\mathscr{Q}_4} \otimes \mathscr{H}_{\mathscr{Q}_3} \otimes \mathscr{H}_{\mathscr{Q}_2} \otimes \mathscr{H}_{\mathscr{Q}_1}$, with $\mathscr{H}_{\mathscr{Q}_i} = \mathbf{C}^2$. Each player $i = 1,2,3,4$ is permitted to manipulate its subsystem with the full machinery of local quantum operations: $U_i \in \mathrm{SU}(2)$ given in (14.17). The payoff operator P_i projects the final state onto the desired states of player i, and is given by

$$P_i = \sum_{j=1}^{k} \big|\xi_i^j\big\rangle\big\langle\xi_i^j\big|. \tag{14.41}$$

The sum is over all the k different states $|\xi_i^j\rangle$, for which player i is in the minority. Its worth to note that the sums are always over a even number k, and that they run over the states of the following form:

$$P_i = \sum_{j=1}^{k} \big|\xi_i^j\big\rangle\big\langle\xi_i^j\big| = \sum_{j=1}^{k/2} \big|\vartheta_i^j\big\rangle\big\langle\vartheta_i^j\big| + \sum_{j=1}^{k/2} \big|\overline{\vartheta_i^j}\big\rangle\big\langle\overline{\vartheta_i^j}\big|, \tag{14.42}$$

where $|\overline{\vartheta_i^j}\rangle$ is the bit-flipped version of $|\vartheta_i^j\rangle$, i.e 0's and 1's are interchanged. The payoff operator P_1 for player 1 in the four player case is given by:

$$P_1 = |0001\rangle\langle 0001| + |1110\rangle\langle 1110|. \tag{14.43}$$

By playing $U(\theta, \alpha, \beta) = U(\frac{\pi}{2}, -\frac{\pi}{8}, \frac{\pi}{8})$, the four players can completely eliminate the risk of upon measurement ending up with an outcome where none of them receives a payoff. This quantum strategy leads to an expected payoff $E(\$) = \frac{1}{4}$ that is twice as good as in the classical case $E(\$) = \frac{1}{8}$. The strategy profile is a Nash equilibrium as well as Pareto optimal. Quantum minority games has been extensively studied for cases of arbitrary n, and it can be shown that the quantum versions gives rise to better than classical payoffs for any game with an even number of players [14].

14.2.4 Kolkata Restaurant Problem

The Kolkata restaurant problem is an extension of the minority game [20–24], where the n players now has m choices. As the story goes, the choice is between m restaurants. The players receive a payoff if their choice is not too crowded, i.e the number of agents that chose the same restaurant is under some limit. We will discuss the case for which this limit is one. Just like in the minority game previously discussed, the Kolkata restaurant problem offers a way for modeling heard behavior and market dynamics, where visiting a restaurant translates to buying a security, in which case an agent wishes to be the only bidder. In our simplified model there are just three agents, Alice, Bob and Charlie. They have three possible choices: security 0, security 1 and security 2. They receive a payoff $\$ = 1$ if their choice is unique, i.e that nobody else has made the same choice, otherwise they receive $\$ = 0$. The game is so called *one shoot*, which means that it is non-iterative, and the agents have no information from previous rounds to base their decisions on. Under the constraint that they cannot communicate, there is nothing left to do other than randomizing between the choices just like in the minority games in the previous section. Given the symmetric nature of the problem, any deterministic strategy would lead all three agents to the same strategy, which in turn would mean that all three would leave empty handed. There are 27 different strategy profiles possible, i.e combinations of choices. 12 of which gives a payoff of $\$ = 1$ to each one of them. Randomization gives therefore agent i an expected payoff of $E(\$) = \frac{4}{9}$.

In the quantum version we let Alice, Bob and Charlie share a quantum resource [18]. Each has a part of a multipartite quantum state. They play their strategy by manipulating their own part of the combined system, before measuring their subsystems and choosing accordingly. Whereas classically the players would be allowed randomizing over a discrete set of choices, in the quantum version each subsystem is allowed to be transformed with arbitrary local quantum operations, just like before. In the absence of entanglement, quantum games of this type usually yield the same payoffs as their classical counterparts, whereas the combination of unitary operators (or a subset therein) and entanglement, will be shown to outperform the classical randomization strategy.

When moving from quantum game protocols with two choices into ones with three, we'll need some additional structure. Instead of qubits will we be dealing with qutrits, which are their three level versions. The local operations on qutrits are now represented by a more complicated group of matrices, the SU(3) group. Everything else will essentially be similar to that of the quantum minority game.

A qutrit is a 3-level quantum system on 3-dimensional Hilbert space $\mathscr{H}_{\mathscr{Q}} = \mathbf{C}^3$, written in the computational basis as:

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle + a_2|2\rangle \in \mathbf{C}^3, \tag{14.44}$$

with $a_0, a_1, a_2 \in \mathbf{C}$ and $|a_0|^2 + |a_1|^2 + |a_2|^2 = 1$. A general n-qutrit system $|\Psi\rangle$ is a vector on 3^n-dimensional Hilbert space, and is written as a linear combination of 3^n orthonormal basis vectors.

$$|\Psi\rangle = \sum_{x_n,\ldots,x_1=0}^{2} a_{x_n\ldots x_1}|x_n \cdots x_1\rangle, \tag{14.45}$$

where

$$|x_n \cdots x_1\rangle = |x_n\rangle \otimes |x_{n-1}\rangle \otimes \cdots \otimes |x_1\rangle \in \mathscr{H}_{\mathscr{Q}} = \overbrace{\mathbf{C}^3 \otimes \cdots \otimes \mathbf{C}^3}^{n\text{-times}}, \tag{14.46}$$

with $x_i \in \{0, 1, 2\}$ and complex coefficients a_{x_i}, obeying $\sum |a_{x_n\cdots x_1}|^2 = 1$.

Single qutrits are transformed with unitary operators $U \in \mathrm{SU}(3)$, i.e operators from the special unitary group of dimension 3, acting on $\mathscr{H}_{\mathscr{Q}}$ as $U : \mathscr{H}_{\mathscr{Q}} \to \mathscr{H}_{\mathscr{Q}}$. In a multi-qutrit system, operations on single qutrits are said to be local. They affect the state-space of the corresponding qutrit only. The SU(3) matrix is parameterized by defining three general, mutually orthogonal complex unit vectors $\bar{x}$, $\bar{y}$, $\bar{z}$, such that $\bar{x} \cdot \bar{y} = 0$ and $\bar{x}^* \times \bar{y} = \bar{z}$. We construct a SU(3) matrix by placing $\bar{x}$, $\bar{y}^*$ and $\bar{z}$ as its columns [25]. Now a general complex unit vector is given by:

$$\bar{x} = \begin{pmatrix} \sin\theta\cos\phi e^{i\alpha_1} \\ \sin\theta\sin\phi e^{i\alpha_2} \\ \cos\theta e^{i\alpha_3} \end{pmatrix}, \tag{14.47}$$

and one complex unit vector orthogonal to $\bar{x}$ is given by:

$$\bar{y} = \begin{pmatrix} \cos\chi\cos\theta\cos\phi e^{i(\beta_1-\alpha_1)} + \sin\chi\sin\phi e^{i(\beta_2-\alpha_1)} \\ \cos\chi\cos\theta\sin\phi e^{i(\beta_1-\alpha_2)} - \sin\chi\cos\phi e^{i(\beta_2-\alpha_2)} \\ -\cos\chi\sin\theta e^{i(\beta_1-\alpha_3)} \end{pmatrix}, \tag{14.48}$$

where $0 \leq \phi, \theta, \chi, \leq \pi/2$ and $0 \leq \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2 \leq 2\pi$. We have a general SU(3) matrix U, given by:

$$U = \begin{pmatrix} x_1 & y_1^* & x_2^* y_3 - y_3^* x_2 \\ x_2 & y_2^* & x_3^* y_1 - y_1^* x_3 \\ x_3 & y_3^* & x_1^* y_2 - y_2^* x_1 \end{pmatrix}, \tag{14.49}$$

and it is controlled by eight real parameters $\phi, \theta, \chi, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$.

The initial state, a maximally entangled GHZ-type state

$$|\psi_{in}\rangle = \frac{1}{\sqrt{3}}\big(|000\rangle + |111\rangle + |222\rangle\big) \in \mathscr{H}_{\mathscr{Q}} = \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3, \tag{14.50}$$

is symmetric and unbiased in regards to permutation of player position and has the property of letting us embed the classical version of the game, accessible trough restrictions on the strategy sets. To show this, we define a set of operators corresponding to classical pure strategies that gives raise to deterministic payoffs when applied to $|\psi_{in}\rangle$. The cyclic group of order three, C_3, generated by the matrix:

$$s = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{14.51}$$

where $s^3 = s^0 = I$ and $s^2 = s^T$, has the properties we are after. The set of classical strategies $S = \{s^0, s^1, s^2\}$ with $s^i \otimes s^j \otimes s^k|000\rangle = |ijk\rangle$ acts on the initial state $|\psi_{in}\rangle$ as:

$$\begin{aligned} & s^i \otimes s^j \otimes s^k \frac{1}{\sqrt{3}}\big(|000\rangle + |111\rangle + |222\rangle\big) \\ & \quad = \frac{1}{\sqrt{3}}\big(|0+i0+j0+k\rangle + |1+i1+j1+k\rangle + |2+i2+j2+k\rangle\big). \end{aligned} \tag{14.52}$$

Note that the superscripts denotes powers of the generator and that the addition is modulo 3. In the case under study, where there is no preference profile over the different choices, any combination of the operators in $S = \{s^0, s^1, s^2\}$ leads to the same payoffs when applied to $|\psi_{in}\rangle$ as to $|000\rangle$. We form a density matrix ρ_{in} out of the initial state $|\psi_{in}\rangle$ and add noise that can be controlled by the parameter f [17]. We get:

$$\rho_{in} = f|\psi_{in}\rangle\langle\psi_{in}| + \frac{1-f}{27} I_{27}, \tag{14.53}$$

where I_{27} is the 27×27 identity matrix. Alice, Bob and Charlie now applies a unitary operator U that maximizes their chances of receiving a payoff $\$ = 1$, and thereby the initial state ρ_{in} is transformed into the final state ρ_{fin}.

$$\rho_{fin} = U \otimes U \otimes U \rho_{in} U^\dagger \otimes U^\dagger \otimes U^\dagger. \tag{14.54}$$

We define for each player i a payoff-operator P_i, which contains the sum of orthogonal projectors associated with the states for which player i receives a payoff $\$ = 1$. For Alice this would correspond to

$$P_A = \left(\sum_{x_3,x_2,x_1=0}^{2} |x_3x_2x_1\rangle\langle x_3x_2x_1|, x_3 \neq x_2, x_3 \neq x_1, x_2 \neq x_1 \right) + \left(\sum_{x_3,x_2,x_1=0}^{2} |x_3x_2x_1\rangle\langle x_3x_2x_1|, x_3 = x_2 \neq x_1 \right). \tag{14.55}$$

The expected payoff $E_i(\$)$ of player i is as usual calculated by taking the trace of the product of the final state ρ_{fin} and the payoff-operator P_i:

$$E(\$_i) = \mathrm{Tr}(P_i \rho_{fin}). \tag{14.56}$$

It can be shown that if Alice, Bob and Charlie acts with a general SU(3), there exist a $U^{opt}(\phi, \theta \cdot \chi \cdot \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2) \in \mathrm{SU}(3)$, given by $U^{opt}(\frac{\pi}{4}, \cos^{-1}(\frac{1}{\sqrt{3}}), \frac{\pi}{4}, \frac{5\pi}{18}, \frac{5\pi}{18}, \frac{5\pi}{18}, \frac{\pi}{3}, \frac{11\pi}{6})$, that outperforms classical randomization. The strategy profile $U^{opt} \otimes U^{opt} \otimes U^{opt}$ leads to a payoff of $E(\$) = \frac{6}{9}$, assuming $(f = 1)$, compared to the classical $E^c(\$) = \frac{4}{9}$. Letting the payoff function depend on the fidelity parameter f, we get a payoff function $E(\$(f)) = \frac{2}{9}(f + 2)$ where we can clearly see that the expected payoff reaches the classical value as $f \to 0$.

14.3 Outlook

As the field of quantum information theory matures and information processing moves into the quantum realm, will it be increasingly important to study the broad spectrum of effects of this transition. Game theory is the study of strategic decision making under limited information. How decision making should or will change as situations are played out in a world where this information is *quantum* information, will be some of many conceptual challenges to address if classical communication and computing, is due to be replaced by systems governed by the peculiar and counter-intuitive laws of quantum mechanics.

Acknowledgements The work was supported by the Swedish Research Council (VR).

References

1. Nielsen M, Chuang I (2000) Quantum computation and quantum information. Cambridge University Press, Cambridge
2. Hargreaves Heap S, Varoufakis Y (2004) Game theory – a critical introduction. Routledge, London

3. Fudenberg D, Tirole J (1991) Game theory. MIT Press, Cambridge
4. Osbourne MJ, Rubinstein A (1994) A course in game theory. MIT Press, Cambridge
5. Flitney AP (2008) Review of quantum games. In: Haugen IN, Nilsen AS (eds) Game theory: strategies, equilibria, and theorems, p 140
6. Piotrowski EW, Sladkowski J (2003) An invitation to quantum game theory. Int J Theor Phys 42:1089
7. Khan FS, Phoenix SJD (2011) Nash equilibrium in quantum superpositions. In: Proceedings of SPIE, vol 8057 80570K-1
8. Landsburg SE (2011) Quantum game theory. Wiley encyclopedia of operations research and management science
9. Bleiler SA (2008) A formalism for quantum games and an application. Preprint http://arxiv.org/abs/0808.1389
10. Marinatto L, Weber T (2000) A quantum approach to static games of complete information. Phys Lett A 272:291–303
11. Meyer D (1999) Quantum strategies. Phys Rev Lett 82:1052–1055
12. Eisert J, Wilkens M, Lewenstein M (1999) Quantum games and quantum strategies. Phys Rev Lett 83:3077–3080
13. Benjamin S, Hayden P (2001) Multiplayer quantum games. Phys Rev A 64:030301
14. Chen Q, Wang Y (2004) N-player quantum minority game. Phys Lett A 327(98):102
15. Flitney A, Hollenberg LCL (2007) Multiplayer quantum minority game with decoherence. Quantum Inf Comput 7:111–126
16. Flitney A, Greentree A (2007) Coalitions in the quantum minority game: classical cheats and quantum bullies. Phys Lett A 362:132137
17. Schmid C, Flitney AP, Wieczorek W, Kiesel N, Weinfurter H, Hollenberg LCL (2010) Experimental implementation of a four-player quantum game. New J Phys 12:063031
18. Sharif P, Heydari H (2011) Quantum solution to a three player Kolkata restaurant problem using entangled qutrits. Preprint http://arxiv.org/abs/1111.1962
19. Benjamin S, Hayden P (2001) Comment on quantum games and quantum strategies. Preprint http://arxiv.org/abs/quant-ph/0003036
20. Chakrabarti BK (2007) Kolkata restaurant problem as a generalised El Farol Bar problem. In: Chatterjee A, Chakrabarti BK (eds) Econophysics of markets & business networks. New economic windows series. Springer, Milan, pp 239–246
21. Chakrabarti AS, Chakrabarti BK, Chatterjee A, Mitra M (2009) The Kolkata paise restaurant problem and resource utilization. Physica A 388:2420–2426
22. Ghosh A, Chakrabarti AS, Chakrabarti BK (2010) Kolkata paise restaurant problem in some uniform learning strategy limits. In: Basu B, Chakrabarti BK, Chakravarty SR, Gangopadhyay K (eds) Econophysics & economics of games, social choices & quantitative techniques, New economic windows. Springer, Milan, pp 3–9
23. Ghosh A, Chatterjee A, Mitra M, Chakrabarti BK (2010) New J Phys 12:075033
24. Arthur WB (1994) Inductive reasoning and bounded rationality: El Farol problem. Am Econ Assoc Papers & Proc 84:406
25. Mathur M, Sen D (2001) Coherent states for SU(3). J Math Phys 42:4181–4196