Statistics of the Kolkata Paise Restaurant problem

To cite this article: Asim Ghosh *et al* 2010 *New J. Phys.* **12** 075033

View the article online for updates and enhancements.

You may also like

Statistics of the Kolkata Paise Restaurant problem

Asim Ghosh[1], Arnab Chatterjee[2], Manipushpak Mitra[3] and Bikas K Chakrabarti[1,3,4]

[1] Theoretical Condensed Matter Physics Division and Centre for Applied Mathematics and Computational Science, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700 064, India
[2] Condensed Matter and Statistical Physics Section, The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, Trieste I-34014, Italy
[3] Economic Research Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Kolkata 700 108, India
E-mail: asim.ghosh@saha.ac.in, achatter@ictp.it, mmitra@isical.ac.in and bikask.chakrabarti@saha.ac.in

Abstract. We study the dynamics of a few stochastic learning strategies for the 'Kolkata Paise Restaurant' problem, where N agents choose among N equally priced but differently ranked restaurants every evening, such that each agent tries to get dinner in the best restaurant (with each restaurant serving only one customer and the rest of the customers arriving there going without dinner that evening). We consider the learning strategies to be similar for all the agents, and assume that each follows the same probabilistic or stochastic strategy dependent on information about past successes in the game. We show that some 'naive' strategies lead to much better utilization of services than some relatively 'smarter' strategies. We also show that a service utilization fraction as high as 0.80 can result for a stochastic strategy, where each agent sticks to his past choice (independent of success achieved or not, with probability decreasing inversely in the past crowd size). The numerical results for the utilization fraction of the services in some limiting cases are analytically examined.

[4] Author to whom any correspondence should be addressed.

IOP Institute of Physics ⬤ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Contents

1. Introduction

The Kolkata Paise Restaurant (KPR) problem [1]–[3] is a repeated game, played between a large number N of agents having no interaction with each other. In the KPR problem, prospective customers (agents) choose from N restaurants each evening simultaneously (in parallel decision mode). N is fixed. Each restaurant has the same price for a meal but a different rank (agreed upon by all customers) and can serve only one customer any evening. Information regarding customer distributions for earlier evenings is available to everyone. Each customer's objective is to go to the restaurant with the highest possible rank while avoiding the crowd so as to be able to get dinner there. If more than one customer arrives at any restaurant on any evening, one of them is randomly chosen (each of them is treated anonymously) and is served. The rest do not get dinner that evening.

In Kolkata, there were very cheap and fixed-rate 'Paise Restaurants', popular among the daily laborers in the city. During lunch hours, the laborers used to walk (to save 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! Paise is the smallest Indian coin, and there were indeed some well-known rankings of these restaurants because some of them would offer tastier items compared to the others. A more general example of such a problem would be when society provides hospitals (and beds) in every locality but 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 lack of service for those people and consequently as (social) wastage of service by those unattended hospitals.

A social planner's (or dictator's) solution to the KPR problem is the following. The planner (or dictator) asks everyone to form a queue and then assigns each one a restaurant with rank matching the sequence of the person in the queue on the first evening. Then each person is told to go to the next ranked restaurant on the following evening (for the person in the last ranked

restaurant, this means going to the first ranked restaurant). This shift process then continues for successive evenings. Call this solution the *fair social norm*. This is clearly one of the most efficient solutions (with utilization fraction $\bar{f}$ of services by the restaurants equal to unity) and the system arrives at this solution immediately (from the first evening). However, in reality, this cannot be the true solution to the KPR problem, where each agent decides on his own (in parallel or democratically) every evening, based on complete information about past events. In this game, customers try to evolve a learning strategy to eventually get dinners at the best possible ranked restaurant, avoiding the crowd. It is seen that the evolution of these strategies takes considerable time to converge, and even then the eventual utilization fraction $\bar{f}$ is far below unity. The KPR problem has some basic features similar to the minority game problem [4, 5] in that diversity is encouraged (compared to herding behavior) in both. However, it differs from (two-choice) minority games in terms of the macroscopic size of the choices.

As already shown in [1], a simple random-choice algorithm, if adapted by all the agents, can lead to a reasonable value of utilization fraction ($\bar{f} \simeq 0.63$). Compared to this, several seemingly 'more intelligent' stochastic algorithms lead to lower utilization of the services. Ghosh *et al* [3] studied a few more such 'smarter' algorithms, having several attractive features (including analytical estimate possibilities), but still failed to improve the overall utilization fraction beyond its random choice value. Here we develop a stochastic strategy that maintains a naive tendency (probability decreasing with past crowd size) to stick to any agent's own past choice (successful or not), leading to a maximum, so far, value of the utilization fraction $\bar{f}$ ($\simeq 0.80$) in the KPR problem. We also estimate analytically the $\bar{f}$ values for several such strategies.

2. Stochastic learning strategies

Let the symmetric stochastic strategy chosen by each agent be such that at any time t the probability $p_k(t)$ of arriving at the kth ranked restaurant is given by

$$p_k(t) = \frac{1}{z}\left[k^\alpha \exp\left(-\frac{n_k(t-1)}{T}\right)\right], \quad z = \sum_{k=1}^{N}\left[k^\alpha \exp\left(-\frac{n_k(t-1)}{T}\right)\right], \quad (1)$$

where $n_k(t)$ denotes the number of agents arriving at the kth ranked restaurant in period t, $T > 0$ is a scaling factor and $\alpha \geqslant 0$ is an exponent. Note that under (1) the probability of selecting a particular restaurant increases with its rank and decreases with its popularity in the immediate past (given by the number $n_k(t-1)$). Certain properties of the strategies given by (1) are the following:

(i) For $\alpha = 0$ and $T \rightarrow \infty$, $p_k(t) = 1/N$ corresponds to the complete random choice case for which we know [1] that the utilization fraction is around 0.63, that is, on an average there is 63% utilization of the restaurants (see appendix A).

(ii) For $\alpha = 0$ and $T \rightarrow 0$, the agents avoid those restaurants visited last evening and choose again randomly from the remaining restaurants [1]. With appropriate simulation it was shown that the distribution of the fraction f of utilization of the restaurants is Gaussian around 0.46 (see section 2.2).

2.1. Rank-dependent strategies

For any natural number α and $T \to \infty$, an agent goes to the kth ranked restaurant with probability $p_k(t) = k^{\alpha}/\sum k^{\alpha}$, which means in the limit $T \to \infty$ in (1) gives $p_k(t) = k^{\alpha}/\sum k^{\alpha}$. Let us discuss the results for such a strategy here.

If an agent selects any restaurant with equal probability P, then the probability that a single restaurant is chosen by m agents is given by

$$\Delta(m) = \binom{N}{m} p^m (1-p)^{N-m}. \tag{2}$$

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

$$\Delta_k(m=0) = \binom{N}{0}(1-p_k)^N, \quad p_k = \frac{k^{\alpha}}{\sum k^{\alpha}}$$

$$\simeq \exp\left(\frac{-k^{\alpha}N}{\widetilde{N}}\right) \quad \text{as} \quad N \to \infty, \tag{3}$$

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

$$\Delta_k(m=0) = \exp\left(-\frac{k^{\alpha}(\alpha+1)}{N^{\alpha}}\right). \tag{4}$$

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

$$\bar{f}_k = 1 - \Delta_k(m=0) \tag{5}$$

and the numerical estimates of $\bar{f}_k$ are shown in figure 1. Naturally for $\alpha = 0$, the problem corresponds to random choice $\bar{f}_k = 1 - e^{-1}$ giving $\bar{f} = \sum \bar{f}_k/N \simeq 0.63$ and, for $\alpha = 1$, $\bar{f}_k = 1 - e^{-2k/N}$ giving $\bar{f} = \sum \bar{f}_k/N \simeq 0.58$, as already obtained analytically earlier (see appendix B).

2.2. Strict crowd-avoiding case

We consider here the case (see also [3]) where each agent chooses on any evening (t) randomly among the restaurants to which nobody had gone the last evening $(t-1)$. This corresponds to the case where $\alpha = 0$ and $T \to 0$ in equation (1). Our numerical simulation results for the distribution $D(f)$ of the fraction f of utilized restaurants are again Gaussian with a most probable value at $\bar{f} \simeq 0.46$. This can be explained in the following way. As the fraction $\bar{f}$ of restaurants visited by the agents on 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 N agents. Hence, when fitted to equation (A.1) in appendix A, $\lambda = 1/(1 - \bar{f})$. Therefore, following equation (A.1), we can write the equation for $\bar{f}$ as

$$(1 - \bar{f})\left[1 - \exp\left(-\frac{1}{1-\bar{f}}\right)\right] = \bar{f}. \tag{6}$$

The solution of this equation gives $\bar{f} \simeq 0.46$. This result agrees well with the numerical results for this limit ($\alpha = 0$, $T \to 0$).
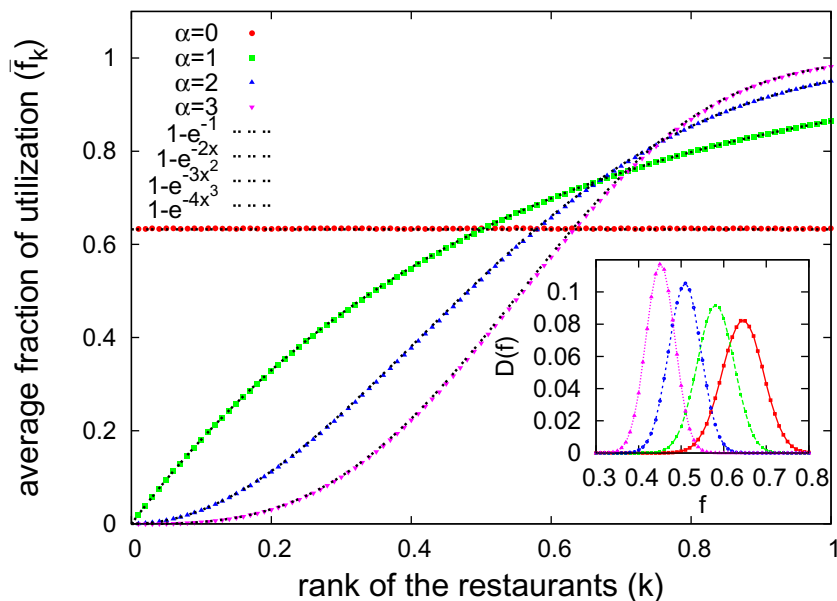
Figure 1. Average fraction of utilization ($\bar{f}_k$) versus rank of the restaurants (k) for different α values. The inset shows the distribution $D(f = \sum \bar{f}_k/N)$ of the fraction f agent getting dinner any evening for different α values.
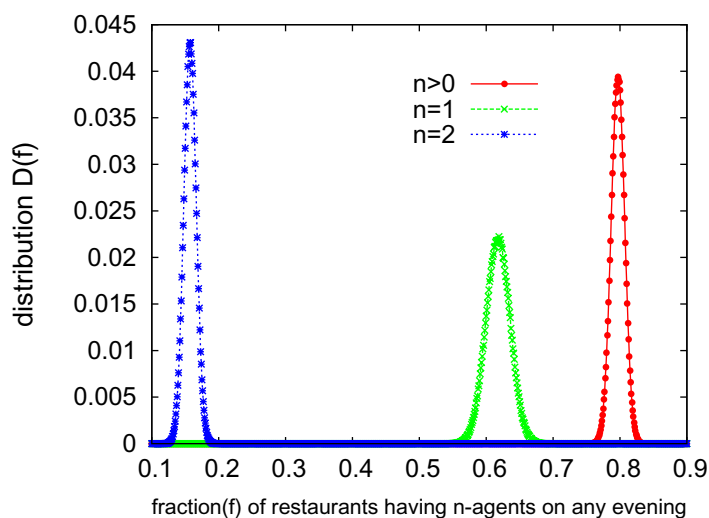

Figure 2. Numerical simulation results for a typical prospective customer distribution on any evening.

2.3. Stochastic crowd-avoiding case

In this section, we start with the following stochastic strategy. If an agent goes to restaurant k in period $(t-1)$, then the agent goes to the same restaurant in the next period with probability $p_k(t) = (1/n_k(t-1))$ and to any other restaurant $k'(\neq k)$ with probability $p_{k'}(t) = (1 - p_k(t))/(N-1)$. In this process, the average utilization fraction is $\bar{f} \simeq 0.8$ and the distribution $D(f)$ is a Gaussian around $f \simeq 0.8$ (see figure 2).
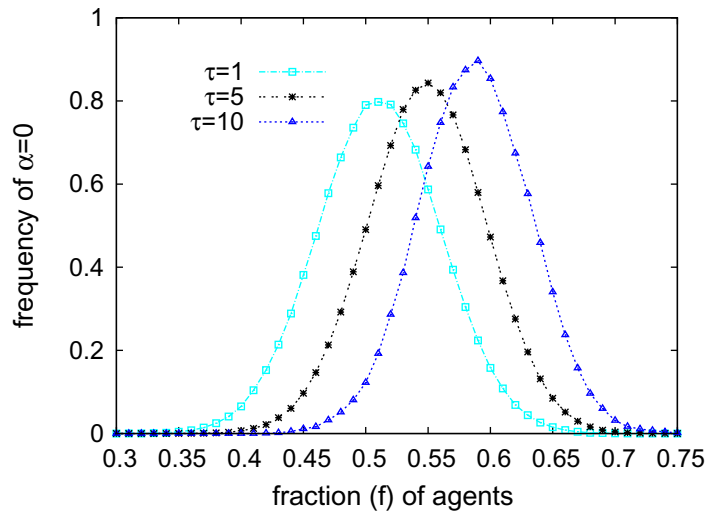
Figure 3. Steady state distribution of successful agents having $\alpha = 0$. The same for $\alpha = 1$ will be given by just a complementary function.

Here we make an approximate estimate of $\bar{f}$. 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 $i \geqslant 3$. Therefore, $a_0 + a_1 + a_2 = 1$, $a_1 + 2a_2 = 1$ and hence $a_0 = a_2$. Given the strategy, a_2 fraction of agents will make attempts to leave their respective restaurants the next evening ($t + 1$), while no intrinsic activity will occur in the restaurants where nobody came (a_0) or only one came (a_1) the previous evening (t). This 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 $a_0 - a_0 a_2$. In restaurants having exactly two agents (a_2 per cent in the last evening), some vacancies will be created due to this process, and this is equal to $(a_2/4) - a_2(a_2/4)$. Steady state implies that $a_0 - a_0 a_2 + (a_2/4) - a_2(a_2/4) = a_0$ and hence, using $a_0 = a_2$, we obtain $a_0 = a_2 = 0.2$, 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 restaurants is assumed to get more than two customers on any evening ($a_i = 0$ for $i \geqslant 3$). The advantage in assuming a_0, a_1 and a_2 only to be non-vanishing on any evening is that the activity of redistribution on the next evening starts from this a_2 fraction of 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 confirm that $a_i \leqslant 0.03$ for $i \geqslant 3$, and hence the above approximation does not lead to serious error.

3. Evolving stochastic strategy

In this section, we assume that agents have two possible exogenously given values of $\alpha : \alpha = 0$ or $\alpha = 1$. We start by taking some random allocation of α over the set of N agents. The strategy followed by each agent thereafter is the following. If an agent starts with an $\alpha = 0(1)$ and fails to get dinner for the successive τ evenings, then the next evening the agent shifts to $\alpha = 1(0)$. The steady state distribution of α values in the population of agents does not depend on the initial allocation of α values in the population (see figure 3). However, as is obvious, for large values of $\tau \geqslant N$, the stability of the distribution disappears.

4. Convergence to a fair social norm with deterministic strategies

In the KPR problem, if the rational agents interact, then a *fair social norm* that can evolve is a periodically organized state with periodicity N, where each agent in turn gets served in all the N restaurants and all agents get served every evening. Can we find deterministic strategies (in the absence of a dictator) such that society achieves this fair social norm? There is one variant of Pavlov's win shift lose stay strategy (see [6]–[8]) that can be adopted to achieve the fair social norm and another variant that can be adopted to achieve the fair social norm in an asymptotic sense. Of course, these strategies are deterministic in nature.

4.1. Fair strategy

The fair strategy works as follows:

 (i) At time (evening) $t = 0$, agents can choose any restaurant either randomly or deterministically.
 (ii) If at time t agent i was in a restaurant ranked k and was served, then at time $t + 1$ the agent moves to the restaurant ranked $k - 1$ if $k > 1$ and goes to the restaurant ranked N if $k = 1$.
(iii) If agent i was in a restaurant ranked k at time t and was not served, then at time $t + 1$ the agent goes to the same restaurant.

It is easy to verify that this strategy gives a convergence to the fair social norm in less than or equal to N periods. Moreover, after convergence is achieved, the fair social norm is retained ever after. The difficulty with this strategy is that a myopic agent will find it hard to justify the action of going to the restaurant ranked last after getting served in the best ranked restaurant. However, if the agent is not that myopic and observes the past history of strategies played by all the agents, and can figure out that this one evening's loss is a tacit commitment devised for this kind of symmetric strategy to work, then this voluntary loss is not that implausible. Therefore, one needs to run experiments before arguing for or against this kind of symmetric deterministic strategy. More importantly, the fair strategy can be modified to take care of this justification problem, provided that one wants to achieve the fair social norm in an asymptotic sense.

4.2. Asymptotically fair strategy

The asymptotically fair strategy works as follows:

 (i) At time (evening) $t = 0$, agents can choose any restaurant either randomly or deterministically.
 (ii) If at time t agent i was in a restaurant ranked k and was served, then at time $t + 1$ the agent moves to the restaurant ranked $k - 1$ if $k > 1$ and goes to the same restaurant if $k = 1$.
(iii) If agent i was in a restaurant ranked k at time t and was not served, then at time $t + 1$ the agent goes to the restaurant ranked N.

5. Summary and discussion

We consider the KPR problem where the decision made by each agent in each time period t is independent and is based on the information about the rank k of the restaurants and their

occupancy given by the numbers $n_k(t-1)\ldots n_k(0)$. We consider in section 2 several stochastic strategies where each agent chooses the kth ranked restaurant with probability $p_k(t)$ given by equation (1). The utilization fraction f_k of the kth ranked restaurants on every evening is studied and their average (over k) distributions $D(f)$ are shown in figure 1 for some special cases. From numerical studies, we find their distributions to be Gaussian with the most probable utilization fraction $\bar{f} \simeq 0.63, 0.58$ and 0.46 for the cases with $\alpha = 0, T \to \infty$; $\alpha = 1, T \to \infty$; and $\alpha = 0$, $T \to 0$, respectively. For the stochastic crowd-avoiding strategy discussed in section 2.3, we get the best utilization fraction $\bar{f} \simeq 0.8$. The analytical estimates for $\bar{f}$ in these limits are also given and they agree very well with the numerical observations.

Finally, we suggest ways of achieving the fair social norm either exactly in the presence of the incentive problem or asymptotically in the absence of such an incentive problem. Implementing or achieving such a norm in a decentralized way is impossible when $N \to \infty$. The KPR problem has similarity with the minority game problem [5], as in both the games, herding behavior is punished and diversity is encouraged. Also, both involve learning of the agents from past successes. Of course, KPR has some simple exact solution limits, a few of which are discussed here. In none of the cases considered here are learning strategies individualistic; rather, all the agents choose following the probability given by equation (1). In a few different limits of such a learning strategy, the average utilization fraction $\bar{f}$ and their distributions are obtained and compared with the analytic estimates, which are reasonably close. Needless to mention, the real challenge is to design algorithms of learning mixed strategies (e.g. from the pool discussed here) by the agents so that the fair social norm emerges eventually, even when everyone decides on the basis of their own information independently. As we have seen, some naive strategies give better values of $\bar{f}$ compared to most of the 'smarter' strategies like strict crowd-avoiding strategies (section 2.2). This observation in fact compares well with the earlier observation in minority games (see e.g. [9]).

It may be noted that all the stochastic strategies, being parallel in computational mode, have the advantage that they converge to a solution at smaller time steps ($\sim \sqrt{N}$ or weakly dependent on N), while for deterministic strategies the convergence time is typically of the order of N, which renders such strategies useless in the truly macroscopic ($N \to \infty$) limits. However, deterministic strategies are useful when N is small and rational agents can design appropriate punishment schemes for the deviators (see [6]).

In brief, the study of the KPR problem shows that a dictated solution leads to one of the best possible solutions to the problem, with each agent getting his dinner at the best ranked restaurant with a period of N evenings, and with the best possible value of $\bar{f}$ ($= 1$), starting from the first evening. The parallel decision strategies (employing evolving algorithms by the agents, and past information, e.g. of $n(t)$), which are necessarily parallel among the agents and stochastic (as in democracy), are less efficient ($\bar{f} \ll 1$; the best one discussed here in section 2.3, giving $\bar{f} \simeq 0.8$ only). We also note that most of the 'smarter' strategies lead to much lower efficiency.

Is there an upper bound for the value of utilization fraction $\bar{f}$ (less than unity; easily achieved in the dictated solution) for such stochastic strategies employed in parallel (democratically) by the agents in KPR? If so, what is this upper bound value? Also, what is the learning time required to arrive at such a solution (compared to zero waiting time to arrive at the most efficient dictated solution) in KPR? These questions are to be investigated in future.

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Appendix A. Random-choice case

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

$$\Delta(m) = \binom{\lambda N}{m} p^m (1-p)^{\lambda N - m}, \qquad p = \frac{1}{N}$$

$$= \frac{\lambda^m}{m!} \exp(-\lambda) \text{ as } N \to \infty. \tag{A.1}$$

Therefore, the fraction of restaurants not chosen by any agent is given by $\Delta(m=0) = \exp(-\lambda)$, and this implies that the average fraction of restaurants occupied on any evening is given by [1]

$$\bar{f} = 1 - \exp(-\lambda) \simeq 0.63 \quad \text{for } \lambda = 1 \tag{A.2}$$

in the KPR problem.

Appendix B. Strict-rank-dependent choice

In this case, an agent goes to the kth ranked restaurant with probability $p_k(t) = k / \sum k$; that is, $p_k(t)$ given by (1) in the limit $\alpha = 1$, $T \to \infty$. Starting with N restaurants and N agents, we make $N/2$ pairs of restaurants and each pair has restaurants ranked k and $N+1-k$, where $1 \leqslant k \leqslant N/2$. Therefore, an agent chooses any pair of restaurants with uniform probability $p = 2/N$ or N agents choose randomly from $N/2$ pairs of restaurants. Therefore, the fraction of pairs selected by the agents (from equation (A.1)) is

$$f_0 = 1 - \exp(-\lambda) \simeq 0.86 \quad \text{for } \lambda = 2. \tag{B.1}$$

Also, the expected number of restaurants occupied in a pair of restaurants with rank k and $N+1-k$ by a pair of agents is

$$E_k = 1 \times \frac{k^2}{(N+1)^2} + 1 \times \frac{(N+1-k)^2}{(N+1)^2} + 2 \times 2 \times \frac{k(N+1-k)}{(N+1)^2}. \tag{B.2}$$

Therefore, the fraction of restaurants occupied by pairs of agents is

$$f_1 = \frac{1}{N} \sum_{k=1,\ldots,N/2} E_k \simeq 0.67. \tag{B.3}$$

Hence, the actual fraction of restaurants occupied by the agents is

$$\bar{f} = f_0 \cdot f_1 \simeq 0.58. \tag{B.4}$$

References

[1] Chakrabarti A S, Chakrabarti B K, Chatterjee A and Mitra M 2009 The Kolkata Paise Restaurant problem and resource utilization *Physica* A **388** 2420–6

[2] Ghosh A and Chakrabarti B K 2009 *Kolkata Paise Restaurant (KPR) Problem* http://demonstrations.wolfram.com/KolkataPaiseRestaurantKPRProblem

[3] Ghosh A, Chakrabarti A S and Chakrabarti B K 2010 Kolkata Paise Restaurant problem in some uniform learning strategy limits *Econophysics and Economics of Games, Social Choices and Quantitative Techniques, New Economic Windows* ed B Basu, B K Chakrabarti, S R Chakravarty and K Gangopadhyay (Milan: Springer) pp 3–9

[4] Arthur W B 1994 Inductive reasoning and bounded rationality: El Farol problem *Am. Econ. Assoc. Papers Proc.* **84** 406

[5] Challet D, Marsili M and Zhang Y-C 2005 *Minority Games: Interacting Agents in Financial Markets* (Oxford: Oxford University Press)

[6] Michihiro K 2008 Repeated games *The New Palgrave Dictionary of Economics* 2nd edn vol 7 (NY: Macmillan) pp 98–105

[7] Nowak M and Sigmund K 1993 A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game *Nature* **364** 56–8

[8] Orléan A 1995 Bayesian interactions and collective dynamics of opinion: herd behavior and mimetic contagion *J. Econ. Behav. Organ.* **28** 257–4

[9] Satinover J B and Sornette D 2007 Illusion of control in minority and Parrondo games *Eur. Phys. J.* B **60** 369–84