SCIENTIFIC REPORTS

OPEN

Stochastic Evolution Dynamic of the Rock–Scissors–Paper Game Based on a Quasi Birth and Death Process

Qian Yu[1], Debin Fang[2], Xiaoling Zhang[3], Chen Jin[4] & Qiyu Ren[2]

Stochasticity plays an important role in the evolutionary dynamic of cyclic dominance within a finite population. To investigate the stochastic evolution process of the behaviour of bounded rational individuals, we model the Rock-Scissors-Paper (RSP) game as a finite, state dependent Quasi Birth and Death (QBD) process. We assume that bounded rational players can adjust their strategies by imitating the successful strategy according to the payoffs of the last round of the game, and then analyse the limiting distribution of the QBD process for the game stochastic evolutionary dynamic. The numerical experiments results are exhibited as pseudo colour ternary heat maps. Comparisons of these diagrams shows that the convergence property of long run equilibrium of the RSP game in populations depends on population size and the parameter of the payoff matrix and noise factor. The long run equilibrium is asymptotically stable, neutrally stable and unstable respectively according to the normalised parameters in the payoff matrix. Moreover, the results show that the distribution probability becomes more concentrated with a larger population size. This indicates that increasing the population size also increases the convergence speed of the stochastic evolution process while simultaneously reducing the influence of the noise factor.

Cyclic competition is a very common phenomenon in nature and society, with the simplest and most studied model being the Rock-Scissors-Paper (RSP) game[1]. This can be characterised by three strategies: R (rock), S (scissors) and P (paper), where R excludes S, S excludes P and P excludes R. Because of their widespread existence, RSP games have attracted the attention of many researchers in related fields – being widely applied in biological systems, such as biological diversity[2,3] and species interaction[4].

Schreiber and Killingback[5], for example, have studied the effect of a general metacommunity structure on the coexistence of the strategies in the RSP game to show that dispersal in spatially heterogeneous environments can alter dynamical outcomes; Jeppe[6] quantitatively describes how spatial clustering slows down the dynamics and stabilizes the spatial system of the game; and Kerr and Riley[7] and Kirkup and Riley[8] show that three strands of E. colibacteria compete in a RSP fashion to reach an evolutionary stable distribution. The RSP game can also be applied to economic and social systems as well as biological systems. In the Edgeworth price cycle, for example, when not all consumers are fully informed with pricing information, oligopolistic pricing can be regarded as a RSP game[9]. Also, the RSP cycle arises in public goods games[10,11], where the populations are separated into three groups: co-operators, defectors and loners.

Various evolution properties of population behaviour for the RSP game have been analysed based on evolutionary dynamics such as the replicator dynamic and the Smith dynamic, the BNN dynamic and the BR dynamic, etc. after evolutionary game theory was developed. Around the 1990 s, Hofbauer[12] studied the evolutionary stability of the Nash equilibrium under the replicate dynamic and proved that the Nash equilibrium of the RSP game is just the fixed point of the replicate dynamic. Meanwhile, Weibull's[13] analysis showed that any evolutionary stable profile of the RSP game is not only a fixed point under the replicator dynamic but is also asymptotically stable. Furthermore, Gaunersdorfer and Hofbauer[14] found that the limiting behaviour of fictitious play and the

[1]School of Economics, Wuhan University of Technology, Wuhan, 430070, China. [2]School of Economics and Management, Wuhan University, Wuhan, 430072, China. [3]Department of Public Policy, City University of Hong Kong, Hong Kong, China. [4]Department of Economics, The George Washington University, Washington DC, 20052, United States. Correspondence and requests for materials should be addressed to D.F. (email: dbfang@whu.edu.cn)

time-averaged replicator dynamics coincide in the RSP game, while Loertscher[15] proved it has an evolutionary stable strategy (ESS) when the discount factor of the game is less than 1.

However, deterministic models can no longer be relied upon for finite populations. In this case, the random fluctuations due to sampling effects, for instance, have to be taken into account and stochastic processes, such as the Moran process and the Wright-Fisher process, must be employed to analyse the equilibrium of RSP[16,17] instead of ordinary differential equations.

Studies of the impact of stochasticity on the RSP game indicate various important factors influencing the cyclic behaviour of populations, such as noise, alliance-specific heterogeneous invasion rate, mutations, group interaction and protection spillovers except evolutionary dynamics. Perc and Szolnoki[18,19], for example, have studied noise-guided evolution in a six-species predator-prey model within cyclical interaction and cyclical interaction with alliance-specific heterogeneous invasion rates; Mobilia[20] studied the oscillatory dynamics in generic three-species RSP games with mutations; and Szolnoki[21,22] considers group interaction in RSP games and studied the vortices to determine the dynamics of biodiversity in cyclical interactions with protection spillovers. Szabó and Fáth[23] discuss the evolution dynamic of the RSP game in a lattice grid and depict heat maps of the evolutionary dynamic based on the Monte-Carlo simulation method. The limit cycles have also been found for the RSP game by Ochea[24] under logit dynamics.

Although these studies have been very successful and well verified in the evolutionary biology field, the RSP game played by the individuals with bounded rationality in a social and economic system still needs to be investigated, because the learning dynamic in such situations may differ from that based on genetics. Motivated by this question, research attention turned to focus on the evolutionary RSP game under learning dynamics, such as imitation[25] and best-response[26]. Sato et al.[27], for instance, investigated the RSP game based on the reinforcement learning dynamic, finding the learning trajectory can be simple or complex depending on initial conditions, and that the learning process displays Hamiltonian chaos for the zero sum game; while Platkowski and Zakrzewski[28] found an asymptotically stable polymorphism for the standard RSP game and the limit cycles for the general RSP game under complex personality profiles of the players and the imitation dynamics. In 2014, Szolnoki[21] reviewed the main results in well-mixed populations with the focus on oscillatory dynamics and firstly on the important role of the topology of the interaction networks[29]. Hoffman et al.[30] experimentally investigated three RSP games and found population distributions further from the centre are more frequent when the NE section is not stable, which is consistent with the predictions of non-deterministic evolutionary dynamics models.

To analyse an evolutionary model with a finite number of players and with noise or mutations, Kandori[31] firstly considered stochastic characteristics in the learning dynamics and introduced the Markov process model to describe the learning dynamic in evolutionary games. Amir and Berninghaus[32] extended this to a continuous time model[33] and modelled the evolutionary process of the evolutionary normal-form game with a finite number of players as a homogeneous Markov process with finite state space. Tadj and Touzene[34] then proposed the state dependent quasi-birth-and-death process model to describe the learning dynamic for more complicated stochastic evolution. Compared with general 3^*3 game described in Tadj and Touzene[34] that emphasises computational efficiency, this paper is more focused on the impact of stochasticity on the evolutionary dynamic of the RSP game, which considers population size, the payoff matrix and noise factor. In addition, according to the unique characteristics of the RSP game, we consider that the states above the diagonal can be transferred to each other.

To study the stochastic evolutionary dynamic of the RSP game in finite populations and analyse the effect factors of the stochastic evolutionary dynamic, we assume that the bounded rational individuals in a finite population can adjust their strategies by imitating the successful strategy according to payoffs from the last round of the game, and model the stochastic evolutionary process involved as a finite, state dependent, quasi birth and death process. The long run equilibrium of the RSP game played by bounded individuals in a finite population can be explained by the limiting distribution of the QBD process of the game evolutionary dynamic. With the limiting distribution, the probability of a stable state of the RSP game evolution dynamic can be more profoundly predicted and interpreted. Furthermore, comparisons of the numerical experimental results plotted as pseudo colour ternary heat maps shows the effect of the population size and parameters in the RSP payoff matrix and noise factor.

Results

Nash equilibrium of the general RSP game. A symmetrical Rock-Scissors-Paper game can be characterised by three strategies R, S and P. Each strategy dominates the next one in a cyclical way.

The standard RSP game can be depicted by the constant-sum payoff matrix

$$G = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \tag{1}$$

It is easily verified that this game has no Nash equilibrium in pure strategies and precisely one Nash equilibrium in mixed strategies, namely the strategy pair in which both players randomise uniformly, i.e., $(x_1^*, x_2^*, x_3^*) = (y_1^*, y_2^*, y_3^*) = (1/3, \ 1/3, 1/3)$.

Furthermore, the payoff matrix of a general Rock-Scissors-Paper game can be normalised as:

$$G' = \begin{pmatrix} 0 & -a_2 & b_3 \\ b_1 & 0 & -a_3 \\ -a_1 & b_2 & 0 \end{pmatrix}, \text{ where } a_i, b_i > 0. \tag{2}$$

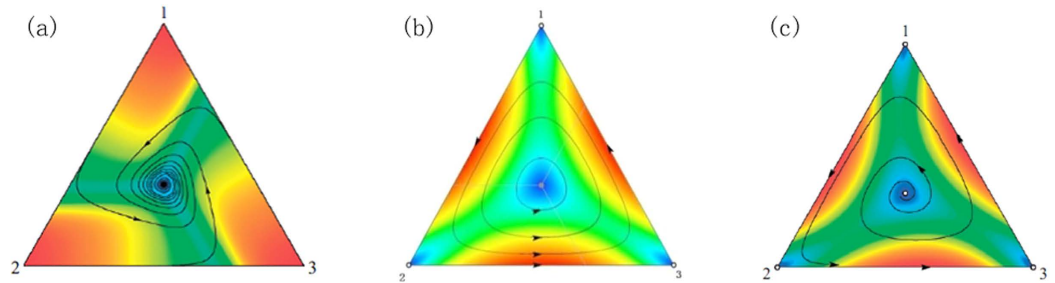It is easy to see that there always exists a unique Nash equilibrium that is

Figure 1. Phase portraits of the RSP game in replicator dynamics[24]. In replicator dynamics, the Nash equilibrium (**a**) will be asymptotically stable and the dynamical trajectory will converge to the centre of the simplex if and only if $\delta > 0$. While as (**b**), the dynamical trajectory will converge to a periodic orbit and obtain the state of neutrally stable if $\delta = 0$. For (**c**), the trajectory will diverge from the centre point and converge to the boundary of the simplex if $\delta < 0$.

$$\mathrm{p} = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \end{pmatrix} = \frac{1}{\Sigma} \begin{pmatrix} a_2 a_3 + a_3 b_2 + b_2 b_3 \\ a_1 a_3 + a_1 b_3 + b_1 b_3 \\ a_1 a_2 + a_2 b_1 + b_1 b_2 \end{pmatrix} \tag{3}$$

where Σ is a normalising constant.

Under the replicator dynamic, the unique Nash equilibrium p, is asymptotically stable when $a_1 a_2 a_3 < b_1 b_2 b_3$, the Nash equilibrium is neutrally stable when $a_1 a_2 a_3 = b_1 b_2 b_3$ and the Nash equilibrium is unstable when $a_1 a_2 a_3 > b_1 b_2 b_3$[35].

With a special transformation, we can obtain a simple payoff matrix for the general RSP game G′ as:

$$G'' = \begin{pmatrix} 0 & -1 & 1+\delta \\ 1+\delta & 0 & 1 \\ -1 & 1+\delta & 0 \end{pmatrix} \tag{4}$$

Then, the Nash equilibrium becomes $(x_1^*, x_2^*, x_3^*) = (y_1^*, y_2^*, y_3^*) = (1/3, 1/3, 1/3)$. Coding with Matlab 7, the phase portraits of the RSP game are shown in Fig. 1.

Numerical experiments for the symmetrical RSP game.

We can model the stochastic evolutionary dynamic of the RSP game as a QBD process and carry out numerical simulations to verify the dynamic stability of long run equilibrium for the RSP game in populations. The limiting distributions of the QBD process can be solved for the RSP game stochastic evolutionary dynamic (see Methods section below).

For example, by setting $N = 20$, $\delta = 5$ and $\varepsilon = 0.01$, the limiting distributions of the QBD process of the general RSP game G'' can be obtained. Then the profiles, whose appearance probability is larger than 0.01, are listed in Table 1. For instance, there are 7, 6 and 7 individuals in the population in row 1 who chose strategy R, S and P respectively and 0.1461 is the appearance probability of this state in the QBD process. As shown in Fig. 2, the main probability in the limiting distribution of the QBD process is concentrated in the states around (7,6,7), (7,7,6) and (6,7,7). The sum of these probabilities is 0.8876, indicating that (1/3, 1/3, 1/3) can be regarded as the long run equilibrium of the stochastic evolutionary RSP game.

To analyse the results more intuitively and compare the effect of population size and the payoff on the QBD process, we carried out numerical experiments with sizes 10, 20 and 100, and set the payoff parameter δ as −0.5, 0 and 1 respectively. The noise factor was set at 0.01. The limiting distribution of QBD was then obtained and plotted as pseudo colour ternary heat maps (Fig. 3), where the vertical distances of each point to the three edges in the simplex represents the probability of choosing either the R, S or P strategy respectively. Therefore, the top, bottom left and bottom right vertices of the simplex correspond to the three pure strategies (1, 0, 0), (0, 1, 0) and (0, 0, 1) respectively, and other interior points correspond to other mixed strategies. Figure 3 shows the results, where the hue value of each point, ranging from red to blue, indicates the probability of the corresponding strategy in the limiting distribution of the QBD process.

Comparison of these results indicates that:

(1) Observing Fig. (3a,3d,3g), we can see when $\delta = 1$ or $\delta > 0$, the heat maps appear as a target, the red points representing the states with higher distribution probability gathered around the centre of this simplex regardless of population size. Most of the remaining points appear as blue, which means the distribution probabilities of the states corresponding to those points approach zero. These features confirm that the mixed strategy profile (1/3, 1/3, 1/3) is a long run Nash equilibrium of the RSP game and also an asymptotically stable equilibrium when $\delta > 0$.

(2) Observing Fig. (3b,3e,3h), the colour images appear as a nebula with three spiral arms, with the colour of the unscrewing arms from the centre to the edges of the simplex gradually turning red. The impression is that the interior points in the simplex are simultaneously attracted by forces from the centre point and centrifugal

R	S	P	Limiting probability
7	6	7	0.1461
7	7	6	0.1461
6	7	7	0.1461
6	6	8	0.0646
8	6	6	0.0646
6	8	6	0.0646
7	5	8	0.0325
8	7	5	0.0325
5	8	7	0.0325
8	5	7	0.0287
5	7	8	0.0287
7	8	5	0.0287
5	6	9	0.0122
6	9	5	0.0122
9	5	6	0.0122
6	5	9	0.0117
5	9	6	0.0117
9	6	5	0.0117
…	…	…	….

Table 1. Main probabilities in the limiting distribution.
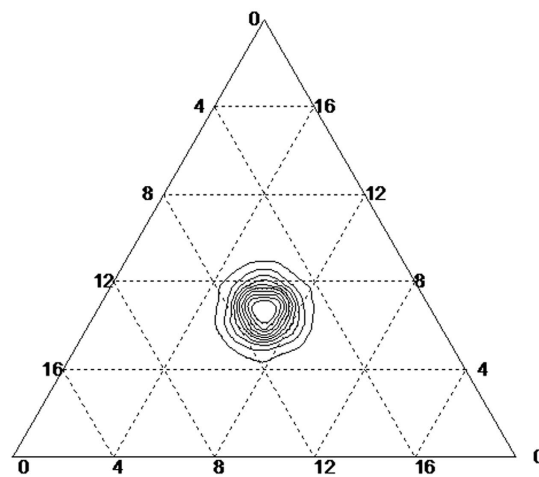
Figure 2. Contour diagram of the limiting distribution of the QBD process. The contour diagram of the QBD process causes the limiting distribution of the general RSP game G'' when $N = 20$, $\delta = 5$ and $\varepsilon = 0.01$, where the contour lines mean the states have the same probability in the limiting distributions. The diagram shows that the states in the centre of the simplex have the highest probabilities.

force from the edges of simplex. The change of colour shows that the distribution probabilities of states corresponding to the points are increasing along the spiral arms of the nebula. This corroborates that, when $\delta = 0$, the Nash equilibrium (1/3, 1/3, 1/3) is merely neutrally stable, and the states corresponding to the points on the central sections of the edges occur with higher distribution probability.

(3) Observing Fig. (3c,f,i) when $\delta = -0.5$ or $\delta < 0$, the red points representing those states with higher distribution probability are gathered around the area near the centre sections of the edges in the simplex. The rest of the parts in the simplex are almost blue, which means the states corresponding to these points have distribution probabilities approaching zero. This indicates an unstable evolutionary Nash equilibrium with condition $\delta < 0$. The long run equilibrium will be a mixed-strategy on the boundary of simplex, such as (2/5, 3/5, 0), (0, 2/5, 3/5) and (3/5, 0, 2/5) in the simulation experiment.

(4) Vertically comparing diagrams from (3a) to (3g), (3b) to (3h) and (3c) to (3i), indicates larger populations have more concentrated distribution probabilities. This indicates, along with increasing population size, the convergence speed of stochastic evolution process will be faster and the influence of the perturbations will be weaker.
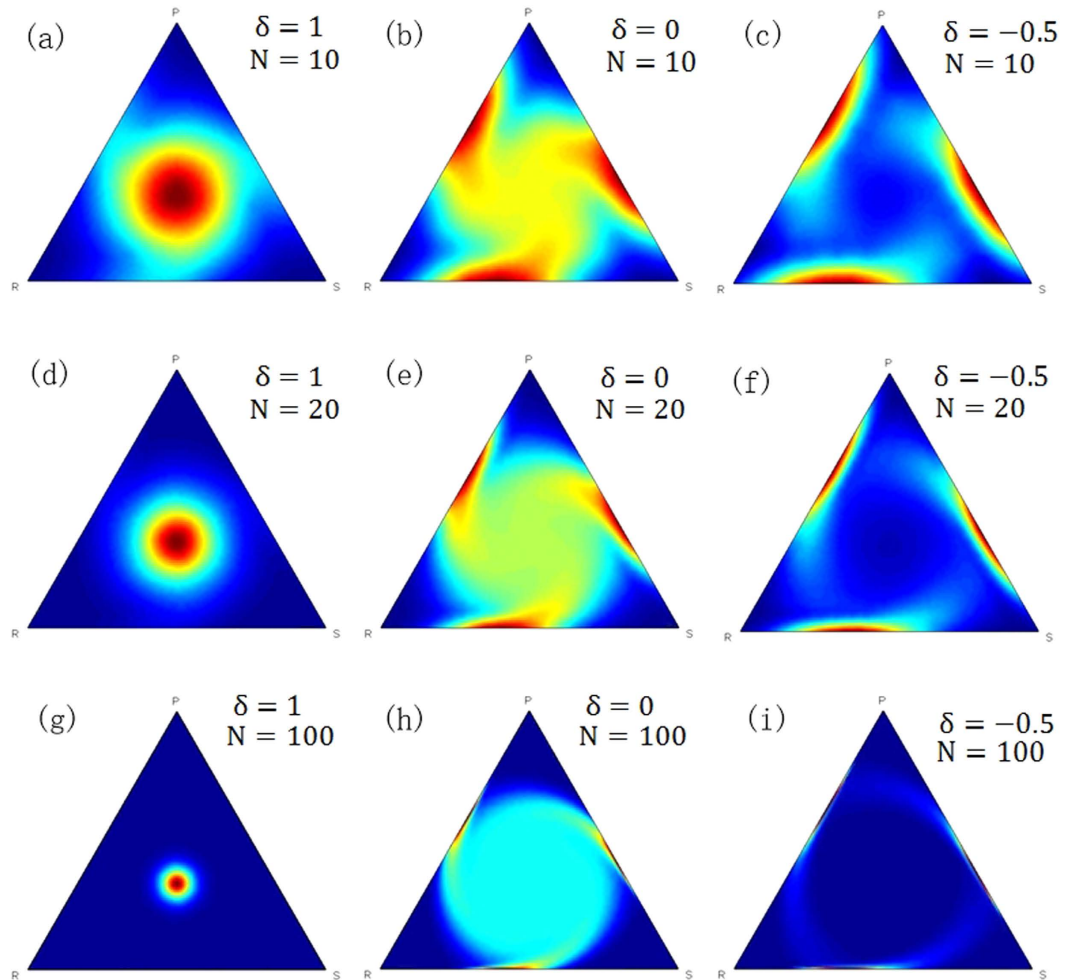
Figure 3. Pseudo colour ternary heat maps of simulations for *G‴*. The numerical experiments based on the QBD process were carried out for simulating the stochastic evolutionary dynamic of RSP game, with the different population size of 10, 20, 100, and different payoff parameter of −0.5, 0, 1, respectively. In these pseudo color ternary heat maps, each point in the simplex corresponds to mixed strategies of RSP game. The hue value of each point, ranging from red to blue, indicates the probability of the corresponding strategy in the limiting distribution of QBD process. By comparing (3a) to (3i), we can see (A): The heat map appeared as a target when $\delta > 0$ (3a, 3d, 3g). In the diagram, the red points gathered around the center indicate that the point (1/3, 1/3, 1/3) has the highest probability in the limiting distribution which confirm the Nash equilibrium is an asymptotically stable equilibrium. The heat map appeared as a nebula with three spiral arms when $\delta = 0$ (3b, 3e, 3h), which corroborate that the center is merely neutrally stable. Moreover, the spiral arms far from the center when $\delta < 0$ (3c, 3f, 3i) demonstrates that the long run equilibrium will be a mixed-strategy equilibrium on the boundary of the simplex. (B): The states corresponding to the points have more concentrated distribution probability when the size of the population is100 rather than 20 and 10, which indicates that the larger the size of the population, the convergence speed of the stochastic evolution process will be faster and the influence of perturbation will be weaker.

Numerical experiments for the asymmetrical RSP game. To avoid artificial symmetries structure, we also investigated the asymmetrical RSP game as[36]:

$$G''' = \begin{pmatrix} 0 & -1 & \gamma/2 \\ 2+\gamma & 0 & -1 \\ -1 & (1+\gamma)/3 & 0 \end{pmatrix}$$

(5)

Under the replicator dynamic, G''' has a saddle node fixed point at the vertices of the simplex and an interior fixed point at (1/2, 1/3, 1/6), independent of γ. However, the stability of the fixed point will be affected by γ. For $\gamma > 1$, (1/2, 1/3, 1/6) is a stable focus, and an unstable focus for $\gamma < 1$. The figure will exhibit closed orbits when $\gamma = 1$.
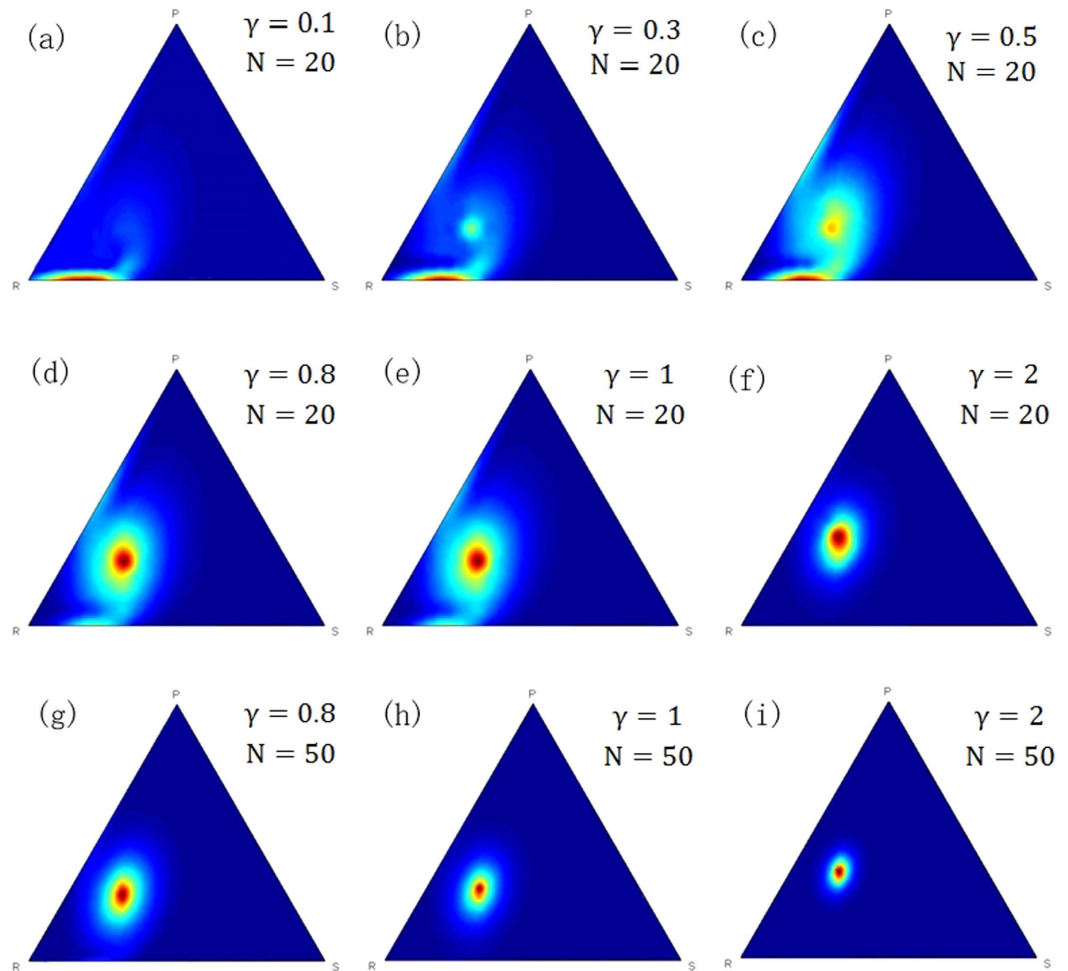
Figure 4. Pseudo colour ternary heat maps of simulations for *G'''*. The asymmetrical RSP game was played in the population with sizes of 20 and 50. The payoff parameters were set as 0.1, 0.3, 0.5, 0.8, 1 and 2 respectively. Fig. (4a–f) show the red points gathered together at the left bottom edge in the simplex at the beginning, and the vortex appear gradually and screwed to the point near (1/2,1/3,1/6) with the increasing of γ. Under the QBD process evolution dynamic, the long run equilibrium will be a mixed-strategy equilibrium on the boundary near the left bottom corner in the simplex when γ is small. The Nash equilibrium will became the long run equilibrium eventually if $\gamma \geq 1$. Figure 4d–i show that the larger the size of the population, the stronger the vortex will be. This indicates, along with the increasing population size, the convergence speed of stochastic evolutionary process will be faster and the influence of noise will be weaker.

Based on the QBD model of RSP game, we set the population to 20 and 50, and the parameter γ in the payoff matrix as 0.1, 0.3, 0.5, 0.8, 1 and 2 to inspect the effect of γ. The perturbation factor was set to $\varepsilon = 0.01$ to represent noise. The resulting pseudo colour ternary heat maps are shown in Fig. 4.

From these, we find that:

(1) Observing (4a–f), when $\gamma < 1$, the red points representing the states with higher distribution probability are gathered at the left bottom edge in the simplex at the beginning, and the vortex appears gradually and screwed to the point near (1/2,1/3,1/6) with increasing γ. The rest of the parts in the simplex are almost blue, indicating the states corresponding to the points have distribution probabilities approaching zero. Therefore, the long run equilibrium will be a mixed-strategy equilibrium on the boundary near the left bottom corner in the simplex when γ is small and the Nash equilibrium will become the long run equilibrium eventually when $\gamma \geq 1$. The Nash equilibrium at (1/2,1/3,1/6) with the condition $\gamma < 1$ is an unstable evolutionary equilibrium, will become neutrally stable when $\gamma = 1$ and will be an asymptotically stable equilibrium finally when $\gamma > 1$.

(2) Vertically comparing from (4d) to (4i), increasing the population size strengthens the vortex. This indicates that, as with the symmetric situation in Fig. 3, as the population size increases, the convergence speed of stochastic evolutionary process is faster and the influence of noise is weaker.

Discussion
In this paper, we established a QBD process for the stochastic evolutionary dynamic for the RSP game. For the general RSP game *G''*, the pseudo colour ternary heat maps of numerical experiments show that the limiting

distribution will converge to the centre of the simplex when $\delta > 0$, diverge as spiral nebula when $\delta = 0$ and diverge to the boundary of the simplex when $\delta < 0$. For the asymmetrical RSP game G''', the limiting distribution will converge to the Nash equilibrium (1/2, 1/3, 1/6) gradually when $\gamma > 1$, become neutrally stable when $\gamma = 1$ and diverge to the boundary of the simplex when $\gamma < 1$. The numerical experimental results are consistent with the conclusions proposed in the literature[12,13] based on replicator dynamics.

Moreover, the study provides a deeper understanding of the RSP game in three aspects. Firstly, the stochastic evolutionary dynamic of the RSP game here was built based on the assumption of bounded rationality. This assumption makes the model more suitable for describing an individual's behaviour in social environments and gives the study more practical significance. In fact, as Zhijian Wang[37] recently revealed in a laboratory experiment of the discrete-time iterated RSP game, a microscopic model of a win-lose-tie conditional response can offer higher payoffs to individual players in comparison with the NE mixed strategy, suggesting that high social efficiency is achievable through an optimised conditional response. Based on the assumption of bounded rationality, the numerical experiment results in Fig. 3 are very similar to the laboratory experimental evidence observed by Moshe[30].

Secondly, the probability interpretation for the long run equilibrium of the RSP game can be provided by the limiting distribution of the QBD process. The limiting probabilities of the states in the QBD process can be used to predict how likely the equilibrium of the RSP game will be.

Finally, the noise factor introduced in our model can be used for detecting the stochastic stability of the long run equilibrium in evolutionary dynamics, which could not be done by previous research with replicator dynamics.

In summary, our study provides a new means of interpretation for the long-term equilibrium of the RSP game played by bounded rational individuals, based on learning dynamics. Further research is needed to generalise these results. For example, RSP games that take place between players from heterogeneous populations can be taken into account. The learning dynamic contained in the QBD process of our model can also be further generalised.

Methods

Without loss of generality, we supposed that the general RSP games are repeatedly played between individuals in a finite population with a size of N and payoff matrix G'.

First, denote the three strategies R, S and P in the RSP game as 1, 2 and 3. Then we can define a stochastic process as: the state at time, t, is a two-dimensional stochastic variable $\omega(t) = (\omega_1(t), \omega_2(t))$, where $\omega_i(t)(i = 1, 2)$ are the number of players in the population who choose strategies R and S, while the number of players in the population who choose strategy P is $N - [\omega_1(t) + \omega_2(t)]$. Hence, the profile and the average payoffs of players are decided by $\omega_i(t)(i = 1, 2)$ and $N - [\omega_1(t) + \omega_2(t)]$ at time t.

At time t, if there are $\omega_1(t)$ individuals choosing strategy R, $\omega_2(t)$ individuals choosing strategy S and $N - [\omega_1(t) + \omega_2(t)]$ individuals choosing strategy P, then any individual in the population adopting strategy i will obtain an average payoff $\pi_i(\omega)(i = 1, 2, 3)$, where $\omega(t) = (\omega_1(t), \omega_2(t))$:

if the individual chooses strategy R,

$$\pi_1(\omega) = \frac{0 \cdot \omega_1 + (-a_2) \cdot \omega_2 + (b_3) \cdot (N - \omega_1 - \omega_2)}{N} \tag{6}$$

if the individual chooses strategy S,

$$\pi_2(\omega) = \frac{b_1 \cdot \omega_1 + 0 \cdot \omega_2 + (-a_3) \cdot (N - \omega_1 - \omega_2)}{N} \tag{7}$$

and if the individual chooses strategy P,

$$\pi_3(\omega) = \frac{-a_1 \cdot \omega_1 + b_2 \cdot \omega_2 + 0 \cdot (N - \omega_1 - \omega_2)}{N}; \tag{8}$$

The next time, $t + 1$, the individuals in the population will switch from one strategy to another at a rate depending on the average payoff of each strategy at time t.

According to the three hypotheses of bounded rationality proposed by Kandori[31], the individuals in the population are inertia, myopia and mutation. That means that not all players react instantaneously to their environment; the players are not taking into account the long run implication of their strategy choices; and the players will change their strategies at random with a small probability.

According to the profile and the average payoffs at time t we can suppose only one individual can change his strategy within a short time. That is, at time $t + 1$, $\omega(t) = (\omega_1(t), \omega_2(t))$ can only shift to an adjacent state. In this case, the strategy distribution of the population satisfies the inertia hypothesis. The myopia hypothesis will also be satisfied because only the profile and the average payoffs at time t will be considered by the individuals when they make the choice at time $t + 1$. Furthermore, the mutation feature is introduced into the model by a perturbation factor $\varepsilon \geq 0$, which allows the players to deviate from the selected strategy with a small probability.

Assume $\omega_1(t) = i$ and $\omega_2(t) = j$, then the individuals in the population will switch from strategy 1 to strategy k ($l, k = 1, 2, 3; l \neq k$) at the rate $\tau_{kl}(i, j) = \varepsilon + [\pi_k(i, j) - \pi_l(i, j)]^+$, where $(i, j) \neq \{(0, N), (N, 0), (0, 0)\}$, and $\tau_{kl}(0, N) = \tau_{kl}(N, 0) = \tau_{kl}(0, 0)$, where $f^+ = \begin{cases} f, & f \geq 0 \\ 0, & f < 0 \end{cases}$ is defined for an arbitrary real function f.
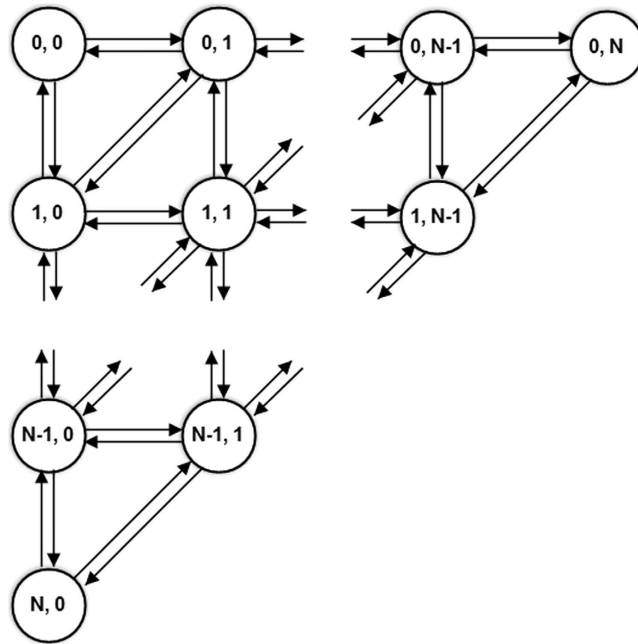
Figure 5. Switches between states in the GQBD.

This means the individuals choosing strategy l will switch to strategy k at rate $\tau_{kl}(i, j)$, if $\pi_k(i, j) - \pi_l(i, j) \geq 0$, and the transition rate will be 0, if the number of individuals adopting the same strategy in the population reach N, the size of the population.

The stochastic evolutionary process of the strategy distribution of the RSP game then becomes a two-dimension Markov stochastic process that can be described as a QBD process.

Denote the state space of $\omega(t)$ as $S = \{\omega = (\omega_1, \omega_2)\}$, where ω_i denotes the number of individuals choosing strategy i in the population (i.e. strategy distribution at time t). Obviously the number of states is $|S| = \frac{(N+1)(N+2)}{2}$. Given $\lambda_{kl}(i, j) = \tau_{kl}(i, j)$, $k < 1$ and $\mu_{kl}(i, j) = \tau_{kl}(i, j)$, $l < k$, the evolutionary process of strategy distribution is a finite, state dependent QBD process, better known as a generalised QBD process (GQBD).

The switches between states in this GQBD can be as depicted in Fig. 5.

The Q-matrix of the GQBD is then:

$$Q = \begin{bmatrix} A_{00} & A_{01} & 0 & \dots & 0 & 0 & 0 \\ A_{10} & A_{11} & A_{12} & \dots & 0 & 0 & 0 \\ 0 & A_{21} & A_{22} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_{N-1,N-2} & A_{N-1,N-1} & A_{N-1,N} \\ 0 & 0 & 0 & \dots & 0 & A_{N,N-1} & A_{N,N} \end{bmatrix} \tag{9}$$

Using a numerical algorithm based on the block Gauss-Seidel iteration method, proposed by Stewart[38], the steady-state probability of the GQBD can be obtained if the GQBD is recurrent. The solution is the steady distribution of the GQBD in the long run, which is a limiting distribution of strategies of the RSP game. In fact, it can be considered as the long run equilibrium of the stochastic evolution RSP game in the population.

References

1. McCannon, B. C. Rock paper scissors. *Journal of Economics* **92(1),** 67–68 (2007).
2. Frean, M. & Abraham, E. R. Rock-scissors-paper and the survival of the weakest. *Proceedings of the Royal Society B Biological Sciences* **268,** 1323–1327 (2001).
3. Reichenbach, T., Mobilia, M. & Frey, E. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448(7157),** 1046–1049 (2007).
4. Cheng, H. *et al.* Mesoscopic interactions and species coexistence in evolutionary game dynamics of cyclic competitions. *Scientific Reports* **4,** 7486 (2014).
5. Schreiber, S. J. & Killingback, T. P. Spatial heterogeneity promotes coexistence of rock–paper–scissors metacommunities. *Theoretical Population Biology* **86,** 1–11 (2013).
6. Jeppe, J., Kim, S. & Joachim, M. Labyrinthine clustering in a spatial rock-paper-scissors ecosystem. *Physical Review E Statistical Nonlinear & Soft Matter Physics* **87,** 187 (2013).
7. Benjamin, K., Riley, M. A., Feldman, M. W. & Bohannan, B. J. M. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418,** 171 (2002).
8. Kirkup, B. C. & Riley, M. A. Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors *in vivo. Nature* **428,** 412 (2004).

9. Maskin, E. & Tirole, J. A theory of dynamic oligopoly, II: price competition, kinked demand curves, and Edgeworth cycles. *Econometrica* **56,** 571 (1988).
10. Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science* **296,** 1129 (2002).
11. Szolnoki, A. *et al.* Cyclic dominance in evolutionary games: a review. *Journal of the Royal Society Interface* **11,** 20140735 (2014).
12. Hauer, J. & Sigmund K., *Evolutionary games and population dynamics*. Vol. 2 Ch. 7, 67–85 (Cambridge University Press, 1998).
13. Weibull, J. W. Evolutionary game theory. *Southern Economic Journal* **22,** 43 (1995).
14. Gaunersdorfer, A. & Hofbauer, J. Fictitious play, Shapley polygons, and the replicator equation. *Games & Economic Behavior* **11,** 279 (1995).
15. Loertscher, S. Rock–Scissors–Paper and evolutionarily stable strategies. *Economics Letters* **118,** 473 (2013).
16. Sigmund, K., *Calculus of Selfishness* Ch. 2, 25–45 (Princeton University. Press, 2010).
17. Taylor, C., Fudenberg, D., Sasaki, A. & Nowak, M. A. Evolutionary game dynamics in finite populations. *Bulletin of Mathematical Biology* **66,** 1621 (2004).
18. Perc, M. & Szolnoki, A. Noise-guided evolution within cyclical interactions. *New Journal of Physics* **9,** 126 (2007).
19. Perc, M., Szolnoki, A. & Gyrgy, S. Cyclical interactions with alliance-specific heterogeneous invasion rates. *Physical Review E Statistical Nonlinear & Soft Matter Physics* **75,** 490 (2007).
20. Mobilia, M. Oscillatory dynamics in rock-paper-scissors games with mutations. *Journal of Theoretical Biology* **264,** 1 (2010).
21. Szolnoki, A., Vukov, J. & Perc, M. From pairwise to group interactions in games of cyclic dominance. *Physical Review E Statistical Nonlinear & Soft Matter Physics* **89,** 062125 (2014).
22. Szolnoki, A. & Perc, M. Vortices determine the dynamics of biodiversity in cyclical interactions with protection spillovers. *New Journal of Physics* **17,** 113033 (2015).
23. Szabó, G. & Fáth, G. Evolutionary games on graphs. *Physics Reports* **446,** 97 (2007).
24. Ochea, M. I. Limit cycles and multiple attractors in logit dynamics. *Econometrics* **58,** 4 (2008).
25. Duersch P., Oechssler, J. & Schipper, B. C. Unbeatable imitation. *Games & Economic Behavior* **76,** 88 (2012).
26. Banerjee, B. & Peng J. Strategic best-response learning in multiagent systems. *Journal of Experimental & Theoretical Artificial Intelligence* **24,** 139 (2012).
27. Yuzuru, S., Eizo, A. & Doyne, F. J. Chaos in learning a simple two-person game. *Proceedings of the National Academy of Sciences of the United States of America* **99,** 4748 (2002).
28. Platkowski, T. & Zakrzewski, J. Asymptotically stable equilibrium and limit cycles in the Rock–Paper–Scissors game in a population of players with complex personalities. *Physica A Statistical Mechanics & Its Applications* **390,** 4219 (2011).
29. Szolnoki, A. & György, S. Phase transitions for rock-scissors-paper game on different networks. *Physical Physical Review E.* **70,** 037102 (2004).
30. Hoffman, M., Suetens, S., Gneezy, U. & Nowak, M. A. An experimental investigation of evolutionary dynamics in the Rock-Paper-Scissors game. *Scientific Reports* **5** (2015).
31. Kandori, M. & Rob, R. Learning, mutation, and long run equilibria in games. *Econometrica* **61,** 29 (1993).
32. Amir, M. & Berninghaus, S. K. Another approach to mutation and learning in games. *Games & Economic Behavior* **14,** 19 (1996).
33. Cason, T. N., Friedman, D. & Hopkins, E. D. Cycles and instability in a Rock-Paper-Scissors population game: a continuous time experiment. *Review of Economic Studies* **81,** 112 (2012).
34. Tadj, L. & Touzene, A. A QBD approach to evolutionary game theory. *Applied Mathematical Modelling* **27,** 913 (2003).
35. Zeeman, E. C. Population dynamics from game theory. Vol. 819 Ch. 4, *Global Theory of Dynamical System* 471–497 (Springer Berlin Heidelberg, 1980).
36. Traulsen, A., Claussen, J. C. & Hauert, C. Stochastic differential equations for evolutionary dynamics with demographic noise and mutations. *Physical Review E Statistical Nonlinear & Soft Matter Physics* **85,** 154 (2012).
37. Wang, Z., Xu, B. & Zhou, H. J. Social cycling and conditional responses in the Rock-Paper-Scissors game. *Scientific Reports* **4,** 5830 (2014).
38. Stewart, W. J. Introduction to the Numerical Solution of Markov Chains. Vol. 3, *Iterative Methods* 121–176 (Princeton University Press, 2015).

Acknowledgements

Author Contributions

Q.Y. and D.F. designed the model, simulation experiment and the main manuscript text. C.J. and Q.R. prepared the computation and figures. X.Z. wrote the part manuscript text. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Yu, Q. *et al.* Stochastic Evolution Dynamic of the Rock–Scissors–Paper Game Based on a Quasi Birth and Death Process. *Sci. Rep.* **6**, 28585; doi: 10.1038/srep28585 (2016).