Biol. Chem., Vol. 382, pp. 1343 – 1363, September 2001 

# The Stochastic Evolution of Catalysts in Spatially Resolved Molecular Systems

**John S. McCaskill*, Rudolf M. Füchslin and Stephan Altmeyer**

GMD – National Research Center for Information Technology, Schloß Birlinghoven, D-53754 St. Augustin, Germany

* Corresponding author

**A fully stochastic chemical modelling technique is derived which describes the influence of spatial separation and discrete population size on the evolutionary stability of coupled amplification in biopolymers. The model is analytically tractable for an ∞-dimensional space (simplex geometry), which also provides insight into evolution in normal Euclidean space. The results are compared with stochastic simulations describing the co-evolution of combinatorial families of molecular sequences both in the simplex geometry and in lower (one, two and three) space dimensions. They demonstrate analytically the generic limits which exploitation place on co-evolving multi-component amplification systems. In particular, there is an optimal diffusion (or migration) coefficient for co-operative amplification and minimal and maximal threshold values for stable cooperation. Over a bounded range of diffusion rates, the model also exhibits stable limit cycles. Furthermore, the co-operatively coupled system has a maximum tolerable error rate at intermediate rates of diffusion. A tractable model is thereby established which demonstrates that spatial effects can stabilize catalytic biological information. The analytic behaviour in ∞-dimensional simplex space is seen to provide a reasonable guide to the spatial dependence of the error threshold in physical space. Nanoscale possibilities for the evolution of catalysis on the basis of the model are outlined. We denote the modelling technique by PRESS, Probability Reduced Evolution of Spatially-discrete Species.**

*Key words:* Catalysis / Cooperation / Molecular evolution / Reaction-diffusion / Spatial resolution / Stochastic Kinetics.

## Introduction

The problem of exploitation appears to be a major driving force in determining evolutionary stable organizational forms in biology (Buss, 1987). This is also true at the level of viral infections and for *in vitro* biochemical systems where experimental models of evolutionary ecologies have been created (Ellinger *et al.,* 1988; Kirner *et al.,* 1999). The quasispecies theory provides a physically grounded kinetic model for the evolution of biological information (Eigen, 1971) in non-interacting self-replicating molecular systems. However, the hierarchical extension to hypercyclic coupled systems of self-replicators does not generically stabilize sequence information in the absence of spatial effects (Bresch *et al.,* 1980; Eigen, 1987). Simple alternatives to hypercyclic coupling have been proposed for spatially resolved systems in which a non-specific replicase species copies all others independently of sequence (Füchslin and McCaskill, 2001). The necessary correlations between genetic sequences encoding the replicase function and their preferential proliferation by replicases is achieved spatially. Sequence-specific (but generically only partial) recognition (binding) between replicases and their encoding is not sufficient to stabilize the evolution of replicase function in the absence of spatial effects. The key parameter characterizing the magnitude of spatial isolation effects is the diffusion (or molecular exchance or migration rate) coefficient. In this article, we develop an analytically tractable model of the spatial inhibition of functional exploitation which enables cooperative evolutionary solutions.

The simplest model in which a *trans*-acting (hetero-functional) replicase activity dictates the cooperation of two molecules for proliferation is the 2X→3X reaction. This reaction has been shown both by deterministic kinetic modeling (with resource limitation and a distinct higher diffusion coefficient for the resource; Böddeker, 1995 and by stochastic modeling with equal diffusion coefficients (McCaskill, 1997) to create self-replicating spot patterns which are favorable to the evolutionary stabilization of the cooperative replicase function. In an evolutionary context, one must distinguish between specific sequences X with the replicase function and non- or less-functional sequences Y which can still be copied by the species of type X. Furthermore, it is essential to consider the effects of erroneous copying of species. The simplest model with these properties is:

$$\begin{aligned} &2X \xrightarrow{Q} 3X \\ &2X \xrightarrow{1-Q} 2X + Y \\ &X + Y \xrightarrow{2r} X + 2Y \\ &X, Y \xrightarrow{d} \varnothing \end{aligned} \qquad (1)$$

where we have normalized the time scale such that the overall rate of synthesis by the replicase X acting on a copy of itself as template is unity. We include a uniform death rate $d$ for all species since this will prove an important parameter in the coupled evolution. $Q \leq 1$ is the fidelity of replication of the class X, where in the simplest case a single isolated sequence X is functional as a replicase, with all mutations resulting in non-functional species of class Y.

It should be stressed that while this two species simplification is used in the analytical work below, the simulations deal with the large combinatorial family of molecular sequences explicitly, with the conventional hamming distance dependent mutation probabilities between sequences as used in the quasispecies model:

$$V_{ij} = Q\left(\frac{q^{-1}-1}{\kappa-1}\right)^{d(i,j)} \tag{2}$$

where $\kappa$ is the number of monomer types, $d(i,j)$ the hamming distance between the two sequences $i$ and $j$ of length $\nu$, and $Q = q^{\nu}$ the fidelity *i. e.* probability of accurately copying a sequence. Such two valued, single peak landscapes have frequently been employed to elucidate central features of evolution in the non-interacting case. Under this two sequence classification, the quasispecies model (for comparison) reads

$$\begin{aligned} X &\xrightarrow{Q} 2X \\ X &\xrightarrow{1-Q} X + Y \\ Y &\xrightarrow{1/\sigma} 2Y \end{aligned} \tag{3}$$

where $\sigma$ is the superiority of the template X relative to the exploiting species Y. Note that Y represents any member of the large class of sequences with background functionality and so the fidelity does not play a role in the replication of Y. The very small rate of back mutation from Y to X can be neglected.

As discussed in the spatially resolved quasispecies model (Altmeyer and McCaskill, 2001), local fluctuations in the composition of finite populations of molecules are more important than fluctuations in a globally homogeneous population. A stochastic formulation is adopted below (for a treatment of stochastic effects on the error threshold in the uniformly-mixed quasispecies model see McCaskill, 1984; Nowak and Schuster, 1989). The rate coefficients above will enter into the transition probabilities (proportional to $dt$) between different copy numbers of the different species at a local site. The infinite dimensional geometry is based upon Wright's island model, in which exchange takes place between any pair from a large number of local sites with maximum population size $n$. We do not however employ the conventional large population size limit which results in a diffusion equation, since we are interested in the fluctuating effects on cooperation of small local populations. On the contrary, we seek to investigate exactly the smallest local population size consistent with the amplification reactions of the coupled system given by eq (1) above. This turns out to be $n = 3$.

**Exactly Solvable Evolution Model of Spatially Induced Cooperation**

The total population is distributed amongst a large number S of local sites, with rapid population mixing within a site and slower migration between sites. A convenient topology for separating the effects of population subdivision *per se* from those of induced spatial patterns is the simplex topology, in which all pairs of sites are equidistant. Since all subsets of sites are equivalent in such a topology, as opposed to the situation in Euclidean space where some sites are closer to one another than others, pattern formation is limited. The major effect of spatial structure in partially isolating subpopulations from each other is preserved in the simplex topology, which was first employed by Wright (1931) and later Latter (1973) to model the biological evolution of species migrating between islands. In previous work (Altmeyer and McCaskill, 2001) we have examined the relationship between the simplex topology and Euclidean space for non-interacting populations of molecules.

We consider a fixed maximum population size (carrying capacity) per site of $n$ molecules, which undergo diffusive exchange with molecules in other sites. Local reactions creating new molecules, such as catalyzed replication, can take place only if the local carrying capacity is not exhausted: *i. e.* if the local population size is less than $n$. There are of course other methods to deal with local saturation effects in population growth, including models involving newly created individuals replacing others (Moran models), but the current method will turn out to enhance the tractability of the model. Since the only interaction between identical sites is by diffusion, which involves particle exchanges between pairs of sites and not higher order correlations, the model can be characterized by the interaction between the probabilistic properties of the population of a single site and the average properties of all the other sites. The interaction between such a chosen local population and the remaining population can be described exactly, employing the mean occupance probability distribution for the other sites, which, for large S and homogeneous initial conditions, should be identical to the occupancy probability distribution of the site under consideration. For moderate numbers of sites S, this would entail a mean field approximation, but here the approximation becomes exact in the large S limit. It is possible to formulate such a stochastic simplex model for any desired molecular reaction scheme or homogeneous evolutionary model and so the modeling framework provides a tractable stochastic alternative to deterministic chemical kinetic modeling for finely subdivided populations. The number of local population states grows quadratically with $n$ [*i. e.* as $(n+1)(n+2)/2$] for two different molecule types, in general exponentially, but the models remain only quadratically nonlinear in contrast with general chemical kinetics.

The article is organized as follows. The basic stochastic simplex model is formulated in the Methods section as a dynamical model along with the equations for its steady states. It is shown that the self-consistent simplex model yields a system of ordinary differential equations with quadratic non-linearity for the local population probabilities. The equations for the stationary states of this model are solved exactly by a reduction to cubics. In the Results section the solutions are illustrated as a function of the four model parameters. Analytical results for the parameter domain giving rise to positive steady state solutions are presented and the stability of these analytical solutions is investigated. The dynamics of the model is shown to also admit limit cycle solutions in a parameter domain where the stationary solutions are unstable. The final Results section compares the predictions of this model with simulations in the simplex topology and in Euclidean space. The paper concludes with a discussion of the implications of these results and its relation to other work.

## Methods

### Mathematical Structure of the Model

The model will first be outlined in more detail. Let $x$ and $y$ be the number of X and Y type molecules in a local site, and $\bar{x}$ and $\bar{y}$ their mean values averaged over all sites. While much of the development can be made for arbitrary $n$ (see above), the explicit analytic solutions are developed for $n$=3. The basic co-operative amplification model of eq (1) is then characterized by four parameters: $r$, $Q$ (or $R=1-Q$), $d$ and $m$, where $m$ is the exchange rate coefficient of molecules between neighbouring sites.

Stochastically, the kinetics of the system are described by the (birth and death) master equations for $P_{x,y}(t)$ for $0 \le x,y \le n$ (supplemented with $P_{-1,y}=P_{x,-1}=P_{x+y>n}=0$ to give a uniform formulation). Let $z=n-x-y$ be the number of holes or vacant locations (type Z) in the site under consideration (like $x$ and $y$, $z$ is a stochastic variable with mean $\bar{z}=n-\bar{x}-\bar{y}$). Then the master equation for the probability dynamics at any site is

$$\begin{aligned}\frac{dP_{x,y}}{dt} = {} & w_{x+1\to x,y}P_{x+1,y}+w_{x-1\to x,y}P_{x-1,y}+w_{x,y+1\to y}P_{x,y+1}+w_{x,y-1\to y}P_{x,y-1}+w_{x+1\to x,y-1\to y}P_{x+1,y-1}+w_{x-1\to x,y+1\to y}P_{x-1,y+1} \\ & -(w_{x\to x-1,y}+w_{x\to x+1,y}+w_{x,y\to y-1}+w_{x,y\to y+1}+w_{x\to x-1,y\to y+1}+w_{x\to x+1,y\to y-1})P_{x,y}\end{aligned} \tag{4}$$

where the $w_{x\to x',y\to y'}$ are the transition rate coefficients for the conversion of $(x,y)$ molecules of type (X,Y) respectively to $(x',y')$ molecules. Actually, only single interconversions occur in the model. For instance, $w_{x\to x-1,y\to y+1}\equiv a_{YX}(x,y)$ is the transition probability for conversion of a molecule of type X to one of type Y in time $dt$ at a site with $x$ and $y$ molecules of types X and Y initially. Similarly, $w_{x\to x-1,y}\equiv a_{ZX}(x,y)$ is the transition probability for converting a molecule of type X to type Z (*i. e.* destruction) and so on for the transition probability functions $a_{XY}, a_{XZ} a_{YZ}, a_{ZY}$:

$$\begin{aligned}
a_{XZ}(x,y) &= m\bar{x}z + \frac{Qx(x-1)z}{(n-1)(n-2)} \\
a_{YZ}(x,y) &= m\bar{y}z + \frac{(1-Q)x(x-1)z}{(n-1)(n-2)} + \frac{2rxyz}{(n-1)(n-2)} \\
a_{YX}(x,y) &= m\bar{y}x \\
a_{XY}(x,y) &= m\bar{x}y \\
a_{ZX}(x,y) &= m\bar{z}x + d\,x = \beta\,x \\
a_{ZY}(x,y) &= m\bar{z}y + d\,y = \beta\,y
\end{aligned} \tag{5}$$

Note that the transition probabilities are set to zero if $x + y > n$ and $\beta = d+m(n-\bar{x}-\bar{y})$. The first terms on the right hand side come from diffusive exchange with other sites, with interchange rate $m$. We shall relate $m$ to the diffusion coefficient $D$ in lower dimensional space below. The other terms come from the reaction scheme of eq (1) and involve exact X catalyzed replication in the first line, erroneous catalyzed replication and exploitation by Y in the second line and destruction of X and Y with rate coefficient $d$ in the final two lines.

Defining the total transition loss rate as

$$a_s(x,y)=a_{XZ}+a_{YZ}+a_{YX}+a_{XY}+a_{ZX}+a_{ZY} \tag{6}$$

we may then rewrite eq (4) for the ten non trivial probabilities

$$\underline{P}=(P(3,0), P(2,0), P(1,0), P(0,0), P(0,1), P(0,2), P(0,3), P(1,2), P(2,1), P(1,1))^T \tag{7}$$

as

$$\frac{d\underline{P}(t)}{dt} = \underline{\underline{W}}_{10}\underline{P}(t) \tag{8}$$

where the 10 by 10 matrix of transition probabilities $\underline{\underline{W}}_{10}$ is given explicitly in the appendix.

There are two further equations which determine the mean values for x and y over all sites, the latter appearing as parameters in the matrix of eq (8):

$$\bar{x} = \sum_{x,y=0}^{n} xP(x,y)$$
$$\bar{y} = \sum_{x,y=0}^{n} yP(x,y) \tag{9}$$

and making the system of eq (8) quadratically non-linear in the probabilities.
In stationary states, the time derivatives are zero, but the resulting homogeneous matrix equation is singular, since

$$\sum_{x,y=0}^{n} P(x,y)=1 \tag{10}$$

so that the stationary probabilities are best determined by first solving for the non-zero probabilities in terms of $P(0,0)$ and then using eq (10) to determine $P(0,0)$. The equation for the vector of non-zero probabilities $\underline{P}$ = ($P(3,0)$, $P(2,0)$, $P(1,0)$, $P(0,1)$, $P(0,2)$, $P(0,3)$, $P(1,2)$, $P(2,1)$, $P(1,1)$)$^T$
takes the form

$$\underline{\underline{W}}' \underline{P} = \underline{P}'_0 \tag{11}$$

where the matrix $\underline{\underline{W}}'$ =

$$\begin{pmatrix}
3(d+3m-m\bar{x}) & -Q-m\bar{x} & 0 & 0 & 0 & 0 & 0 & -m\bar{x} & 0 \\
-3(d+3m-m\bar{x}-m\bar{y}) & 1+2d+6m-m\bar{x}+m\bar{y} & -2m\bar{x} & 0 & 0 & 0 & 0 & -d-3m+m\bar{x}+m\bar{y} & -m\bar{x} \\
0 & -2d-6m+2m\bar{x}+2m\bar{y} & d+3m+m\bar{x}+2m\bar{y} & -m\bar{x} & 0 & 0 & 0 & 0 & -d-3m+m\bar{x}+m\bar{y} \\
0 & 0 & -m\bar{y} & d+3m+2m\bar{x}+m\bar{y} & -2d-6m+2m\bar{x}+2m\bar{y} & 0 & 0 & 0 & -d-3m+m\bar{x}+m\bar{y} \\
0 & 0 & 0 & -2m\bar{y} & 2d+6m+m\bar{x}-m\bar{y} & -3(d+3m-m\bar{x}-m\bar{y}) & -d-3m+m\bar{x}+m\bar{y} & 0 & -m\bar{y} \\
0 & 0 & 0 & 0 & -m\bar{y} & 3(d+3m-m\bar{y}) & -m\bar{y} & 0 & 0 \\
0 & 0 & 0 & 0 & -m\bar{x} & -3m\bar{x} & 3(d+3m)-m\bar{x}-2m\bar{y} & -2m\bar{y} & -r-m\bar{y} \\
-3m\bar{y} & -1+Q-m\bar{y} & 0 & 0 & 0 & 0 & -2m\bar{x} & 3(d+3m)-2m\bar{x}-m\bar{y} & -m\bar{x} \\
0 & -2m\bar{y} & -2m\bar{y} & -2m\bar{x} & -2m\bar{x} & 0 & -2d-6m+2m\bar{x}+2m\bar{y} & -2d-6m+2m\bar{x}+2m\bar{y} & 2d+6m+r
\end{pmatrix} \tag{12}$$

and $\underline{P}'_0$ = (0, 0, $3m\bar{x}P(0,0)$, $3m\bar{y}P(0,0)$, 0, 0, 0, 0, 0)$^T$. Later we check that the solution is consistent with the omitted equation for *P(0,0),* as it should be, since the kinetic equations conserve total probability.

### Stationary Solutions of Spatially Resolved Stochastic Kinetics

Our solution strategy will be first to solve the system of eq (11) treating $\bar{x}$, $\bar{y}$ and $P(0,0)$ as parameters and then to use eq (10) and eq (9) as three equations for these remaining variables. Eq (11) may be inverted explicitly and algebraically to find the non-zero probabilities as a function of $\bar{x}$, $\bar{y}$ and $P(0,0)$ provided the determinant $D_n$ of the (9x9) matrix on the left is non-zero. The inverse of the matrix is best calculated first as the transposed matrix of signed minors, leaving the normalization by the determinant $D_n$ until a leter stage. The non zero-state probabilities can then be written in the form

$$P(x,y)= \frac{p(x,y)}{D_n} P(0,0) \tag{13}$$

involving the polynomials $p(x,y)$ and the normalization condition eq (10) used to obtain

$$P(0,0) = \frac{D_n}{D_n + \sum_{(x,y)\neq(0,0)} p(x,y)} \tag{14}$$

so that

$$P(x,y) = \frac{p(x,y)}{D_n + \sum_{(x,y)\neq(0,0)} p(x,y)} \tag{15}$$

### Stability Analysis of Stationary Solutions

The self-consistent dynamical eq (8) may be rewritten as a homogeneous vector equation in nine independent variables (the probabilities for non-zero, $(x,y) \neq (0,0)$, population states at each site, $P(x,y;t)$):

$$\frac{d\underline{P}(t)}{dt} = \underline{\underline{W}}\, \underline{P}(t) \tag{16}$$

where $\underline{P}(t)$ is given by deleting $P(0,0)$ from eq (7) and the matrix $\underline{\underline{W}} = \underline{\underline{W}}\,(\bar{x}(t),\bar{y}(t))$ is obtained by re-expressing the terms of eq (8) involving $P(0,0)$ in terms of $\underline{P}(t)$. $\underline{\underline{W}}$ is given explicitly in the appendix. (17)

The linear stability analysis of stationary states $\underline{P}_0$ of eq (16) involves two contributions, since changing the probabilities from $\underline{P}_0$ *to* $\underline{P}_0 + \underline{P}_1$ also changes the means $\bar{x}$, $\bar{y}$ in the matrix $\underline{\underline{W}}$ (yielding an according change of $\underline{\underline{W}}$ into $\underline{\underline{W}}_0 + \underline{\underline{W}}_1$):

$$\frac{d\underline{P}_1(t)}{dt} = \underline{\underline{W}}_0\underline{P}_1(t)+(\underline{\underline{W}}_1-\underline{\underline{W}}_0)\underline{P}_0(t) \tag{18}$$

Writing the resulting perturbation in the means as $\bar{x} = \bar{x}_0 + \xi$ and $\bar{y} = \bar{x}_0 + \eta$, we obtain

$$\underline{\underline{W}}_1 - \underline{\underline{W}}_0 =$$

$$\begin{pmatrix}
3m\xi & m\xi & 0 & 0 & 0 & 0 & 0 & m\xi & 0 \\
-3m(\eta+\xi) & m(-\eta+\xi) & 2m\xi & 0 & 0 & 0 & 0 & -m(\eta+\xi) & m\xi \\
-3m\xi & -m(2\eta+5\xi) & -2m(\eta+2\xi) & -2m\xi & -3m\xi & -3m\xi & -3m\xi & -3m\xi & -m(\eta+4\xi) \\
-3m\eta & -3m\eta & -2m\eta & -2m(2\eta+\xi) & -m(5\eta+2\xi) & -3m\eta & -3m\eta & -3m\eta & -m(4\eta+\xi) \\
0 & 0 & 0 & 2m\eta & m(\eta-\xi) & -3m(\eta+\xi) & -m(\eta+\xi) & 0 & m\eta \\
0 & 0 & 0 & 0 & m\eta & 3m\eta & m\eta & 0 & 0 \\
0 & 0 & 0 & 0 & m\xi & 3m\xi & m(2\eta+\xi) & 2m\eta & m\eta \\
3m\eta & m\eta & 0 & 0 & 0 & 0 & 2m\xi & m(\eta+2\xi) & m\xi \\
0 & 2m\eta & 2m\eta & 2m\xi & 2m\xi & 0 & -2m(\eta+\xi) & -2m(\eta+\xi) & 0
\end{pmatrix} \tag{19}$$

and since $\xi$ and $\eta$ are linear combinations of the probabilities in $\underline{P}_1$ as in eq (9)

$$\xi = \sum_{(x,y)\neq(0,0)} xP_1(x,y) \tag{20}$$
$$\eta = \sum_{(x,y)\neq(0,0)} yP_1(x,y)$$

we may write eq (18) as

$$\frac{d\underline{P}_1(t)}{dt} = (\underline{\underline{W}}_0 + \underline{\underline{V}}_0)\underline{P}_1(t) \tag{21}$$

where the matrix $V_0$ satisfies

$$\underline{\underline{V}}_0\underline{P}_1(t) = (\underline{\underline{W}}_1-\underline{\underline{W}}_0)\underline{P}_0 \tag{22}$$

and may be expressed explicitly in terms of the components of $\underline{P}_0$.

Substituting the stationary expressions from the previous section for $\underline{P}_0$ into $\underline{\underline{V}}_0$, we may derive the explicit stability criterion for a stationary state that the largest eigenvalue of the matrix $\underline{\underline{W}}_0 + \underline{\underline{V}}_0$ have a negative real part. While the matrix $\underline{\underline{W}}_0 + \underline{\underline{V}}_0$ has been computed explicitly, the characteristic equation does not factor in general, so that the eigenvalues must be computed numerically.

## Results

The above model is not only a probabilistic one, dealing with discrete numbers of molecules and their fluctuations, but is also non-linear in the probabilities. Its explicit analytical solution presented in this section is a non-trivial result providing insight into the structure of spatial cooperation in chemical systems.

### Stationary Solutions of Spatially Induced Cooperation

The determinant $D_n$ of eq (15) may be factored into two polynomials cubic in $\bar{x}$ and $\bar{y}$

$$D_n = 36\, d\, (d-\tilde{x}-\tilde{y})^3[6d^2\,(d_{00}+d_{10}R) + (d_{01}+d_{11}R)r] \tag{23}$$

with

$$\begin{aligned}
d_{00} &= (2\tilde{x}+d)\tilde{y}+3d^3 \\
d_{01} &= -2\tilde{x}\tilde{y}\,(\tilde{x}-\tilde{y}) + d\tilde{y}\,(\tilde{x}+2\tilde{y}) + d^2\,(\tilde{y}+12\tilde{x}(\tilde{x}-\tilde{y})) - 6d^3\,(\tilde{x}-\tilde{y})+3d^4 \\
d_{10} &= (d+2\tilde{x})\,(d-\tilde{x}-\tilde{y}) \\
d_{11} &= 2\tilde{x}\,(\tilde{x}+\tilde{y})^2 - d\,(3\tilde{x}^2 + 5\tilde{x}\tilde{y} + 2\,\tilde{y}^2) + d^2\,\tilde{y} + d^3
\end{aligned} \tag{24}$$

where

$$d = d+3m \qquad R = 1-Q \qquad \tilde{x} = m\bar{x} \qquad \tilde{y} = m\bar{y} \tag{25}$$

has been used to simplify the expressions. Later we shall also use the variable $\tilde{m} = 3m$.

The denominator of eq (15), $C_n \equiv D_n + \sum_{(x,y)\neq(0,0)} p(x,y)$, which normalizes the $p(x,y)$, may also be written explicitly as

$$C_n = 36\, d^{\,3}\,[6\, d^{\,3}\,(f_{00} + f_{10}\,R) + (f_{01} + f_{11}\,R)\,r] \tag{26}$$

with

$$f_{00} = \tilde{x}\,(3\tilde{x} + 2\tilde{y}) + d\,\tilde{y} + 3\,d^3$$
$$f_{10} = (d + 2\,\tilde{x})\,(d - \tilde{x} - \tilde{y})$$
$$f_{01} = 12\tilde{x}^2\,(\tilde{x} + \tilde{y})^2 + 2\,d\,\tilde{x}\,(-3\tilde{x}^2 + 2\,\tilde{x}\tilde{y} + 5\,\tilde{y}^2) + d^2\,(3\tilde{x}^2+\tilde{x}\tilde{y}+2\tilde{y}^2) + d^3\,(12\tilde{x}^2+\tilde{y}+24\tilde{x}\tilde{y}) - 6d^4\,(\tilde{x}-\tilde{y}) + 3\,d^5$$
$$f_{11} = (d + 2\tilde{x})\,(-6\tilde{x}(\tilde{x}+\tilde{y})^2 + d\,(7\tilde{x}^2 + 5\tilde{x}\tilde{y} - 2\tilde{y}^2) + d^2\,(-2\tilde{x}+\tilde{y}) + d^3) \tag{27}$$

Note that, like $D_n$, $C_n$ only depends explicitly on the three parameters ($r$, $R$ and d).

The probability factors $p(x,y)$ have the form

$$p(x,y) = 36d\,[6(c_{00} + c_{10}R)d^2 + (c_{01} + c_{11}R)r] \tag{28}$$

where the coefficients $c_{ij} = c_{ij}(x,y)$ depend on the three parameters above and the normed means $(\tilde{x},\tilde{y}) = (m\bar{x}, m\bar{y})$. Their values for the 9 different non-zero states $(x,y)$, as in eq (7), are listed in Figure 1 below. Note that the limits $R\to 0$ (no errors), $r\to 0$ (no exploitation) and $r\to\infty$ (strong exploitation) can all be evaluated simply for $P(x,y)$ from the expressions above. We postpone a physical analysis until later.

Finally, the self-consistency conditions eq (9) determine the self-consistent mean values $\bar{x}$ and $\bar{y}$. These equations can be rewritten as two coupled polynomial equations in $(\tilde{x},\tilde{y})$, of the form

$$c_{00} = \begin{pmatrix} \tilde{x}^2\,(3\,d^3\,(1+\tilde{x}) - 3\,d^2\,\tilde{y} - 2\,d\,\tilde{x}\,\tilde{y} - \tilde{x}^2\,\tilde{y}) \\ 9\,d^3\,\tilde{x}^2\,(d - \tilde{x} - \tilde{y}) \\ 3\,\tilde{x}\,(-d+\tilde{x}+\tilde{y})^2\,(3\,d^3 + d\,\tilde{y} + \tilde{x}\,\tilde{y}) \\ 3\,\tilde{y}\,(-d+\tilde{x}+\tilde{y})^2\,(3\,d^3 + d\,\tilde{y} + \tilde{x}\,(\tilde{x}+2\,\tilde{y})) \\ 3\,(d - \tilde{x} - \tilde{y})\,\tilde{y}^2\,(3\,d^3 + d\,\tilde{y} + 2\,\tilde{x}\,(\tilde{x}+\tilde{y})) \\ \tilde{y}^3\,(3\,d^3 + d\,\tilde{y} + \tilde{x}\,(3\,\tilde{x}+2\,\tilde{y})) \\ 3\,\tilde{x}\,\tilde{y}^2\,(3\,d^3 + d\,(\tilde{x}+\tilde{y}) + \tilde{x}\,(2\,\tilde{x}+\tilde{y})) \\ 3\,\tilde{x}^2\,(d^2 + 3\,d^3 + d\,\tilde{x} + \tilde{x}^2)\,\tilde{y} \\ 3\,\tilde{x}\,(d - \tilde{x} - \tilde{y})\,\tilde{y}\,(6\,d^3 + 2\,\tilde{x}\,(\tilde{x}+\tilde{y}) + d\,(\tilde{x}+2\,\tilde{y})) \end{pmatrix}$$

$$c_{01} = \begin{pmatrix} \tilde{x}^2\,(3\,d^4\,(1+\tilde{x}) - 2\,\tilde{x}^2\,\tilde{y}\,(\tilde{x}+\tilde{y}) - 3\,d^3\,(2\,\tilde{x}^2 - 2\,\tilde{x}\,(-1+\tilde{y}) + \tilde{y}) - 2\,d\,\tilde{x}\,\tilde{y}\,(4\,\tilde{x}+5\,\tilde{y}) + 4\,d^2\,\tilde{x}\,(3\,\tilde{x}^2 + 4\,\tilde{y} + \tilde{x}\,(3+6\,\tilde{y}))) \\ 9\,d^2\,\tilde{x}^2\,(d - \tilde{x} - \tilde{y})\,(d^2 - 2\,d\,\tilde{x} + 4\,\tilde{x}\,(\tilde{x}+\tilde{y})) \\ 3\,d\,\tilde{x}\,(-d+\tilde{x}+\tilde{y})^2\,(3\,d^3 - 6\,d^2\,\tilde{x} - \tilde{x}\,\tilde{y} + d\,(12\,\tilde{x}^2 + \tilde{y})) \\ 3\,\tilde{y}\,(-d+\tilde{x}+\tilde{y})^2\,(3\,d^4 + 6\,d^3\,\tilde{y} - 2\,\tilde{x}\,(\tilde{x}+\tilde{y})^2 + d^2\,(\tilde{y} - 12\,\tilde{x}\,\tilde{y}) + d\,(2\,\tilde{x}^2 + 3\,\tilde{x}\,\tilde{y} + 2\,\tilde{y}^2)) \\ 3\,(d - \tilde{x} - \tilde{y})\,\tilde{y}\,(-2\,\tilde{x}\,(\tilde{x}+\tilde{y})^3 + 3\,d^4\,(2\,\tilde{x}+\tilde{y}) + d\,(\tilde{x}+\tilde{y})^2\,(\tilde{x}+2\,\tilde{y}) + 6\,d^3\,(-\tilde{x}^2 + \tilde{x}\,\tilde{y} + \tilde{y}^2) + d^2\,(\tilde{x}^2\,(1 - 12\,\tilde{y}) + 2\,\tilde{x}\,(1 - 6\,\tilde{y})\,\tilde{y} + \tilde{y}^2)) \\ \tilde{y}^2\,(3\,d^4\,(6\,\tilde{x}+\tilde{y}) + 6\,d^3\,(-3\,\tilde{x}^2 + 2\,\tilde{x}\,\tilde{y} + \tilde{y}^2) + d^2\,(\tilde{x}^2\,(3 - 24\,\tilde{y}) + 6\,\tilde{x}\,(1 - 2\,\tilde{y})\,\tilde{y} + \tilde{y}^2) - 2\,\tilde{x}\,(3\,\tilde{x}^3 + 6\,\tilde{x}^2\,\tilde{y} + 4\,\tilde{x}\,\tilde{y}^2 + \tilde{y}^3) + d\,(3\,\tilde{x}^3 + 6\,\tilde{x}^2\,\tilde{y} + 7\,\tilde{x}\,\tilde{y}^2 + 2\,\tilde{y}^3))) \\ 3\,\tilde{x}\,\tilde{y}\,(12\,d^5 - 2\,\tilde{x}^3\,(\tilde{x}+\tilde{y}) - 3\,d^4\,(2\,\tilde{x}+\tilde{y}) + d\,\tilde{x}\,(-3\,\tilde{x}^2 - 4\,\tilde{x}\,\tilde{y} + \tilde{y}^2) - d^2\,(-2\,\tilde{x}\,\tilde{y} + \tilde{y}^2 + 3\,\tilde{x}^2\,(-1+4\,\tilde{y})) + d^3\,(-6\,\tilde{x}^2 + 4\,\tilde{y} + \tilde{x}\,(2+6\,\tilde{y}))) \\ 3\,\tilde{x}^2\,\tilde{y}\,(d^3 + 9\,d^4 + 2\,\tilde{x}^2\,(\tilde{x}+\tilde{y}) + 2\,d\,\tilde{x}\,(3\,\tilde{x}+4\,\tilde{y}) + 2\,d^2\,(\tilde{y}+6\,\tilde{x}\,\tilde{y})) \\ 0 \end{pmatrix}$$

$$c_{10} = \begin{pmatrix} -\tilde{x}^2\,(3\,d^2 + 2\,d\,\tilde{x} + \tilde{x}^2)\,(d - \tilde{x} - \tilde{y}) \\ 0 \\ 3\,\tilde{x}\,(d + \tilde{x})\,(d - \tilde{x} - \tilde{y})^3 \\ 3\,(d - \tilde{x} - \tilde{y})^3\,(\tilde{x}^2 + d\,\tilde{y} + 2\,\tilde{x}\,\tilde{y}) \\ 3\,\tilde{y}\,(-d+\tilde{x}+\tilde{y})^2\,(2\,\tilde{x}^2 + d\,\tilde{y} + 2\,\tilde{x}\,\tilde{y}) \\ (d - \tilde{x} - \tilde{y})\,\tilde{y}^2\,(3\,\tilde{x}^2 + d\,\tilde{y} + 2\,\tilde{x}\,\tilde{y}) \\ 3\,\tilde{x}\,(d - \tilde{x} - \tilde{y})\,\tilde{y}\,(d\,(\tilde{x}+\tilde{y}) + \tilde{x}\,(2\,\tilde{x}+\tilde{y})) \\ 3\,\tilde{x}^2\,(d^2 + d\,\tilde{x} + \tilde{x}^2)\,(d - \tilde{x} - \tilde{y}) \\ 3\,\tilde{x}\,(d - \tilde{x} - \tilde{y})\,(-2\,\tilde{x}\,(\tilde{x}+\tilde{y})^2 + d^2\,(\tilde{x}+2\,\tilde{y}) + d\,(\tilde{x}^2 - \tilde{x}\,\tilde{y} - 2\,\tilde{y}^2)) \end{pmatrix}$$

$$c_{11} = \begin{pmatrix} -\tilde{x}^2\,(d - \tilde{x} - \tilde{y})\,(3\,d^3 - 4\,d^2\,\tilde{x} + 2\,\tilde{x}^2\,(\tilde{x}+\tilde{y}) + 2\,d\,\tilde{x}\,(4\,\tilde{x}+5\,\tilde{y})) \\ 0 \\ 3\,d\,(d - \tilde{x})\,\tilde{x}\,(d - \tilde{x} - \tilde{y})^3 \\ 3\,(d - \tilde{x} - \tilde{y})\,(-d+\tilde{x}+\tilde{y})^2\,(d^2\,\tilde{y} - 2\,\tilde{x}\,(\tilde{x}+\tilde{y})^2 + d\,(2\,\tilde{x}^2 + 3\,\tilde{x}\,\tilde{y} + 2\,\tilde{y}^2)) \\ 3\,(d - \tilde{x})\,(\tilde{x}+\tilde{y})^2\,(-d+\tilde{x}+\tilde{y})^2\,(d + 2\,(\tilde{x}+\tilde{y})) \\ (d - \tilde{x} - \tilde{y})\,\tilde{y}\,(d^2\,(3\,\tilde{x}^2 + 6\,\tilde{x}\,\tilde{y} + \tilde{y}^2) - 2\,\tilde{x}\,(3\,\tilde{x}^3 + 6\,\tilde{x}^2\,\tilde{y} + 4\,\tilde{x}\,\tilde{y}^2 + \tilde{y}^3) + d\,(3\,\tilde{x}^3 + 6\,\tilde{x}^2\,\tilde{y} + 7\,\tilde{x}\,\tilde{y}^2 + 2\,\tilde{y}^3)) \\ 3\,(d - \tilde{x})\,\tilde{x}\,(d - \tilde{x} - \tilde{y})\,(2\,\tilde{x}^2\,(\tilde{x}+\tilde{y}) + 2\,d^2\,(\tilde{x}+2\,\tilde{y}) + d\,(5\,\tilde{x}^2 + 6\,\tilde{x}\,\tilde{y} - \tilde{y}^2)) \\ 3\,\tilde{x}^2\,(d - \tilde{x} - \tilde{y})\,(d^3 + 2\,d^2\,\tilde{y} + 2\,\tilde{x}^2\,(\tilde{x}+\tilde{y}) + 2\,d\,\tilde{x}\,(3\,\tilde{x}+4\,\tilde{y})) \\ 0 \end{pmatrix}$$

**Fig. 1** Coefficients in Denominator of Expressions for the Probabilities p(x,y).
The four terms of the form $c_{ij}$ in the expressions eq (28) for the nine non-zero occupation probability factors p(x,y): p(3,0), p(2,0), p(1,0), p(0,1), p(0,2), p(0,3), p(1,2), p(2,1), p(1,1).

$$\tilde{x}\, C_n = m \sum_{(x,y)\neq(0,0)} x\, p(x,y)$$
$$\tilde{y}\, C_n = m \sum_{(x,y)\neq(0,0)} y\, p(x,y) \tag{29}$$

where $C_n$ is given by eq (26) and the $p(x,y)$ for the site occupation values $(x,y)$ as shown in eq (7) are given by eq (28) with the coefficients obtained in Figure 1.

The $\tilde{x}$ equation in eq (29) has one solution $\tilde{x}=0$. This also implies $\tilde{y}=0$, as is clear since exploiters in the model cannot survive in the absence of catalysts. The remaining solutions are the roots of the quartic in $\tilde{x}$ and quadratic in $\tilde{y}$ with coefficients (in terms of $\tilde{m}=3m$ and $\delta=d+3m$) arranged as the matrix (increasing powers of $\tilde{x}$ down each column, and of $\tilde{y}$ across each row):

$$\begin{pmatrix} 27\delta^3(1-Q+3\delta)\,(r+6\delta)\,(\delta-\tilde{m}) & 9\delta^2(6Q\delta+r(2-Q+6\delta))\,(\delta-\tilde{m}) & 6Qr\delta\,(\delta-\tilde{m}) \\ -27\delta^2\,(2\,(-1+Q+r)\,\delta^2+(-2\,(-1+r)\,\delta+Q\,(r+4\delta))\,\tilde{m}) & 3\delta\,(\delta\,(12Q\delta+r\,(8-7Q+24\delta))+2\,(r\,(-4+5Q-12\delta)+3Q\delta)\tilde{m}) & 10Qr\,(\delta-\tilde{m}) \\ 9\delta\,(\delta\,(2\,(1+2Q)\,\delta+r\,(2-Q+4\delta))+2\,(r(-1+2Q-2\delta)+(-1+Q)\delta)\tilde{m}) & -2r\,(-(1+Q)\,\delta+(1+10Q)\tilde{m}) & 4Qr \\ -2r\,(-\delta+4Q\delta+\tilde{m}+5Q\tilde{m}) & 8Qr & 0 \\ 4Qr & 0 & 0 \end{pmatrix} \tag{30}$$

This matrix notation is chosen in order to point out similiarities in the structure of eqs (29), which, as shown below, allow for further simplification.

The $\tilde{y}$ equation in eqs (29) also has a solution with $\tilde{x}=0$, $\tilde{y}=0$. The non zero solutions are the roots of the quartic in $\tilde{x}$ and cubic in $\tilde{y}$ with coefficients

$$\begin{pmatrix} 0 & 27\delta^3(1+Q+3\delta)(r+6\delta)\,(\delta-\tilde{m}) & 9\delta^2(6Q\delta+r(2-Q+6\delta))(\delta-\tilde{m}) & 6Qr\delta(\delta-\tilde{m}) \\ 0 & -54\delta^2((-1+Q-r)\delta^2+(\delta-Q\delta+r(2-2Q+5\delta))\tilde{m}) & 3\delta(\delta(12Q\delta+r(8-7Q+24\delta))+(-12Q\delta+r(4-17Q+12\delta))\tilde{m}) & 2Qr(5\delta+\tilde{m}) \\ 81(-1+Q)\delta^2(r+2\delta)\tilde{m} & 9\delta(\delta(2(1+2Q)\delta+r(2-Q+4\delta))+(2(2-5Q)\delta+r(5-8Q+8\delta))\tilde{m}) & 2r((1+Q)\delta+(-1+2Q)\tilde{m}) & 4Qr \\ -27(-1+Q)\delta(r+2\delta)\tilde{m}^2 & 2r(\delta-4Q\delta+(-1+Q)\tilde{m}) & 8Qr & 0 \\ 0 & 4Qr & 0 & 0 \end{pmatrix} \tag{31}$$

Since the highest order coefficients in eq (30) and eq (31) have the same structure, the first of the two eqs (29) (denoted by $c_x = 0$ and $c_y = 0$) can be replaced by the linear combination $c_y - \tilde{y}\, c_x = 0$, that has a non trivial polynomial factor, linear in $\tilde{x}$ and quadratic in $\tilde{y}$ which can be solved to find

$$\tilde{x} = \tilde{y}\,\frac{(r(1-5R)+6(1-R-2r)\delta-4r(1-R)\tilde{y}}{3R\delta(r+2\delta)+4r(1-R)\tilde{y}} \tag{32}$$

This explicit relation between the mean populations of catalyst and exploiter (divided by m) exhibits a dependence on only three variables $r$, $R$ and $\delta$. Note that the denominator is positive for $\tilde{y} \neq 0$ and $R<1$.

Substituting this expression for $\tilde{x}$ into the polynomial with coefficients eq (31) gives a polynomial of order 6 in $\tilde{y}$ whose roots determine the mean exploiter concentration $\tilde{y}$. This sixth degree polynomial for $\tilde{y}$ factorizes into two cubic polynomials. These two cubics have coefficients (in increasing powers of $\tilde{y}$)

$$\begin{pmatrix} 9R^2\delta^3(r+2\delta)^2(r+6\delta) \\ 6rR\delta^2(r+2\delta)(18(1-R)\delta+r(3+R+12\delta)) \\ 4r\delta(36(1-R)\delta^2+12r\delta(2+R^2-12\delta+R(-5+6\delta))+r^2(3-2R^2+144\delta^2+R(3+60\delta))) \\ 16r^2(1-R)(r-6\delta+12r\delta) \end{pmatrix} \tag{33}$$

and

$$\begin{pmatrix} 9R^2\delta^3(r+2\delta)^2(R+3\delta)(\delta-\tilde{m}) \\ -3R\delta^2(r+2\delta)(6(-1+R)\delta(-3\tilde{m}+R(2\delta+\tilde{m}))+r(15R^2\delta+3(-8\delta^2+\tilde{m}-4\delta\tilde{m})+R(-13\delta+30\delta^2-5\tilde{m}+6\delta\tilde{m}))) \\ (1-R)\delta(108(1-R)\delta^3+12r\delta((3-5R-5R^2)\delta+18(-2+R)\delta^2+(-6+11R-4R^2)\tilde{m}) \\ +r^2((24-180R)\delta^2-432\delta^3+2(6-27R+14R^2)\tilde{m}+\delta(-3-33R+38R^2-96\tilde{m}+72R\tilde{m}))) \\ -4r(1-R)^2(18\delta^2+r(-3\delta+2R\delta-36\delta^2-6\tilde{m}+4R\tilde{m})) \end{pmatrix} \tag{34}$$

The roots of these polynomials are determinable analytically by the Cardano formulae. Having obtained explicit formulae for the possible stationary solutions $\tilde{y}$ and hence the mean exploiter concentration $\bar{y}=\tilde{y}/m$, the mean catalyst concentration $\bar{x}=\tilde{x}/m$ is determined by eq (32). This completes the explicit analytical solution of the stationary states of the model, with the probabilities given previously explicitly in terms of the solutions $\bar{x}$ and $\bar{y}$. The symbolic algebra package Mathematica® has been employed to assist in the calculations above.

The solutions of these equations are plotted for $r=1/2$ and a range of significant values of the parameters $R$, and $\tilde{m}=3m$ in Figure 2. Only solutions where both $0\leq\bar{x}, \bar{y}\leq 3$ can be physical stationary states. As we shall see in the next section, not all of these solutions are stable, but the plots demonstrate the systematic behaviour as a function of the parameters.

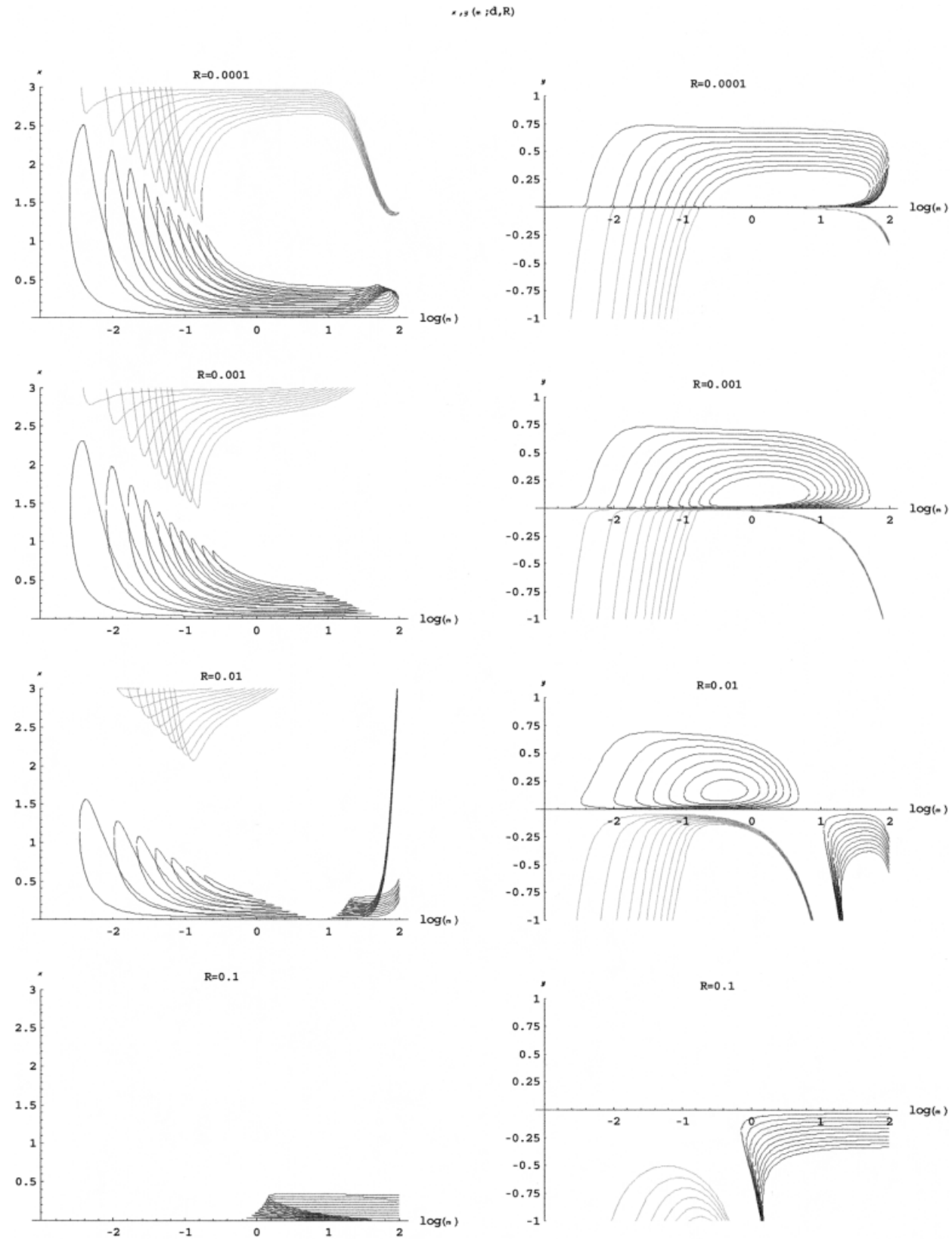


**Fig. 2** The Three Stationary Solutions for the Means $\bar{x}$ and $\bar{y}$ for r=1/2.
The solutions are shown as pairs of plots (left $\bar{x}$ and right $\bar{y}$) for four error rates $R$=$10^{-4}$, $10^{-3}$, $10^{-2}$, $10^{-1}$. The three solutions to the cubic eq (33) for $\bar{y}$ are shaded light, intermediate and dark to allow the corresponding solution for $\bar{x}$ (inverse order, from top to bottom) from eq (32) to be identified. A series of solutions with linearly increasing destruction rates d=0.01, 0.02 ... 0.1 is shown on each plot as a function of log $\tilde{m}$. Note the existence of a maximum in the curves as a function of the normalized site diffusion coefficient $\tilde{m}$=3$m$. The negative, non-physical solutions are also shown, for completeness. Only the intermediate shaded solution will turn out to be stable, and only over the lower part of the range of $\tilde{m}$ shown, see text and Figure 3.

### Stability Analysis of Steady-State Solutions

Figure 3 shows the results of the stability analysis (developed in the Methods section) applied to the stationary solutions to the mean site populations $\bar{x}$ and $\bar{y}$. The viable range of diffusion rates ($\tilde{m}$) is bounded above and below and the population of catalysts has a maximum for intermediate $\tilde{m}$. Furthermore, the stationary population of X jumps discontinuously from zero (as m increases) to a finite viable population size (at the critical value of $\tilde{m}$) and then jumps back to zero at a second, higher critical value. In Figure 4 we show all the calculated, individual, stationary probabilities for the range of parameters for which the stationary solutions are dynamically stable. This confirms that the discontinuous jumps in the mean values are as a result of discontinuous transitions in the underlying site occupation probabilities as the diffusion coefficient m is varied.

The conclusion is that over a limited range of diffusion or migration coefficients, the catalytic species can survive, despite the presence of neutral exploiters and in the absence of any direct group selection effects or compartment level proliferation or selection. As the fidelity of replication decreases or the death rate increases, this viability region shrinks and eventually disappears altogether when the exploiting species are produced too rapidly by mutation. In section 6 below, we shall see that the model also exhibits viable non-stationary solutions in the form of limit cycles for values of m above the threshold for stability of the stationary solutions.

### Evaluation of Stationary State Parameter Dependence, Limiting Cases

In the limit of large diffusion, the model is expected to behave as in the homogeneous, well-mixed case. In fact, the homogeneous deterministic kinetics of X and Y with the current mechanism of eq (1), is given by the ODEs for the continuous concentrations $x$=[X], $y$=[Y]

$$\begin{aligned}\frac{dx}{dt} &= Qx^2(1-x-y)-dx \\ \frac{dy}{dt} &= ((1-Q)x^2+2rxy)(1-x-y)-dy\end{aligned} \tag{35}$$

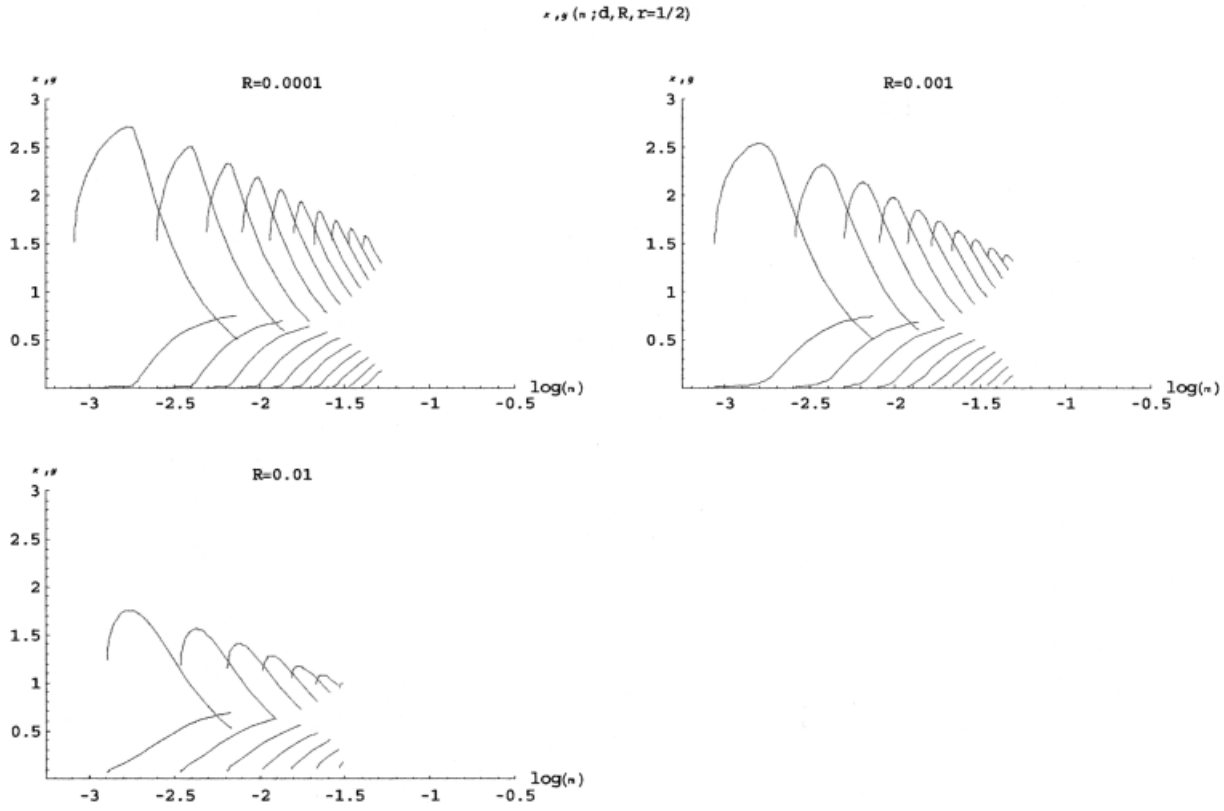


**Fig. 3** Stable Stationary Solutions for the Means $\bar{x}$ and $\bar{y}$ for r=1/2.
The stable solutions are shown for $\bar{x}$ (upper curves) and $\bar{y}$ (lower curves), only for the three error rates $R$=$10^{-4}$, $10^{-3}$, $10^{-2}$ since for $R$=$10^{-1}$ no stable solutions are found. These are the portions of the solutions shown in Figure 2 for which the dominant eigenvalue is negative. A series of solutions with linearly increasing destruction rates d=0.005, 0.01 ... 0.05 is shown on each plot as a function of log $\tilde{m}$. Note the retention of a maximum in the curves as a function of the normalized site diffusion coefficient $\tilde{m}$=3m. Only the second solution (intermediate) in Figure 2 turns out to be stable according to the calculation of eq (21). Stable solutions commence at the lower limit in $\tilde{m}$ to the viable range of $\bar{x}$ solutions, and the stationary solutions cease to be stable at an upper limit with non-zero values of $\bar{x}$ and $\bar{y}$.

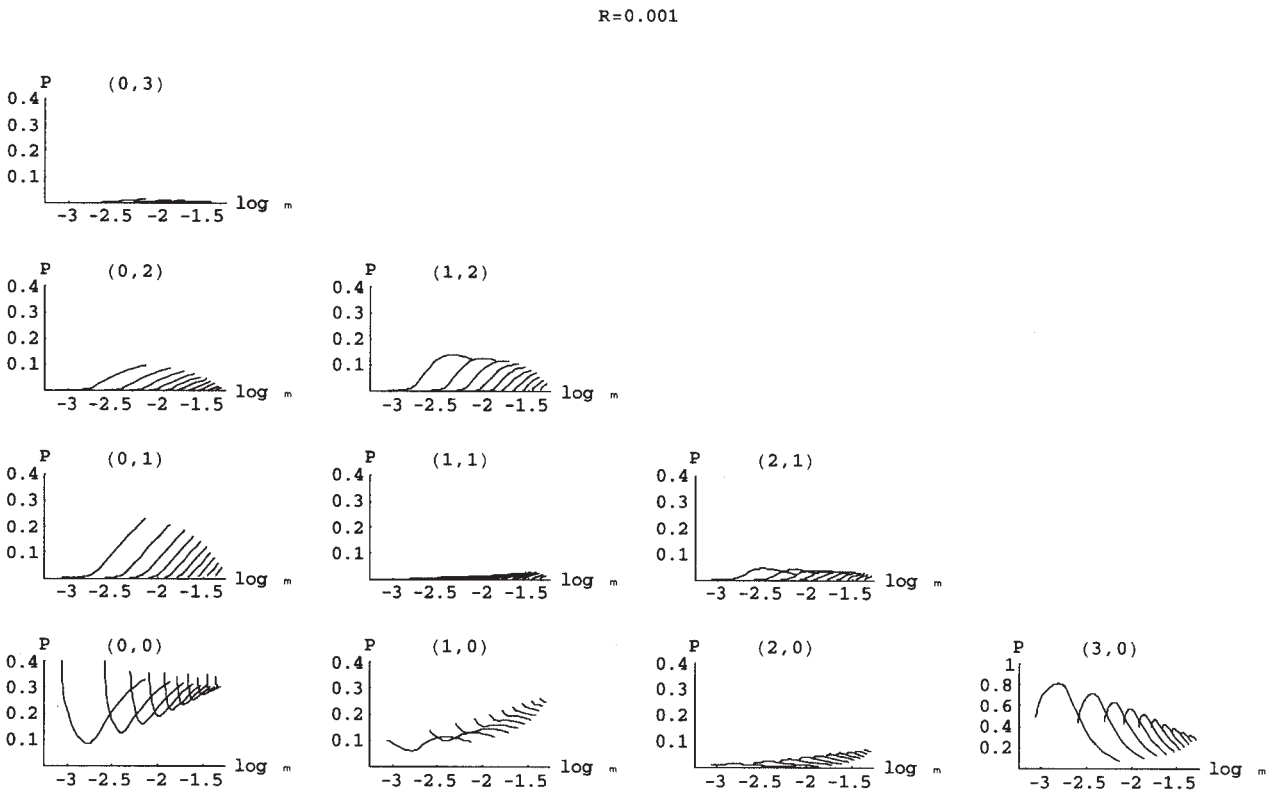


**Fig. 4** Stationary Probabilities $P(x,y)$ for $r$=1/2.
Solution of discrete probabilities $P(x,y)$ (for formulas see Section 3) arranged tabularly for $(x,y)$ as in the standard $x$–$y$ plane. For $R=10^{-3}$ there are shown with up to ten curves for d=0.01,0.02 ... 0.1 as a function of log m̃. The plots show a viability range for the catalysts X bounded above and below in the diffusion rate. The probability distributions are multipeaked. They confirm the discrete jump to a finite viable population and then back to extinction at higher m̃.

with three stationary solutions

$$\left( \begin{array}{cc} x=0 & y=0 \\ x=\dfrac{(Q-2r)\left(1-\sqrt{1-\dfrac{4d\,(1-2r)}{Q\,(Q-2r)}}\right)}{2\,(1-2r)} & y=\dfrac{(1-Q)\left(1-\sqrt{1-\dfrac{4d\,(1-2r)}{Q\,(Q-2r)}}\right)}{2\,(1-2r)} \\ x=\dfrac{(Q-2r)\left(1+\sqrt{1-\dfrac{4d\,(1-2r)}{Q\,(Q-2r)}}\right)}{2\,(1-2r)} & y=\dfrac{(1-Q)\left(1+\sqrt{1-\dfrac{4d\,(1-2r)}{Q\,(Q-2r)}}\right)}{2\,(1-2r)} \end{array} \right)$$

the first and last being stable. There are no positive solutions $x$ and $y$ for $2r \geq Q$, and in particular, the case of neutral exploitation admits no positive solutions in well mixed kinetics.

The homogeneous formulae are regained from the full individual stochastic treatment (up to the factor $n$=3 and a normalized d) in the limit of m→∞, confirming that the model gives the correct mixing limit and that space is necessary to stabilize catalysis in the presence of near neutral or strong exploitation. Not surprisingly, the opposite limit of strong localization, m→0, also admits no positive stationary solutions. The absorbing probability state (0,0) causes the then finite local site population to go to extinction for all values of the parameters. Figure 3 confirms that a non-zero stable stationary solution is possible for neutral recognition ($r$=1/2) in the case of intermediate m. As we shall see, the site model also admits positive stable solutions for arbitrarily large $r$!

We expect an error threshold result to occur as the fidelity $Q$ decreases, as with the simpler case of independent replicators in the quasispecies model, but here the limit is posed by exploitation of the catalysts X by Y, rather than by simple competition between X and Y. In the homogeneous limit, no stationary solutions exist for $Q \leq 2r$, which is always true for $r$=1/2. The error threshold for survival depends in spatial models on migration and death parameters in a non-trivial way, which turns out to be accessible to analytic study.

The recognition parameter $r$ is crucial in defining the ability of catalysts to distinguish between catalyst and non-catalysts as targets for replication (see Table 1). Of special interest is the case $r$=1/2, which corresponds to non discrimination. Based on a detailed investigation of the polynomials using Descartes rule of sign change, we distinguish the following cases for $r$ in the analysis below and observed three qualitatively different behaviours (see Table 1):

**Table 1** Dependence of Exploitation on Recognition Coefficient *r.*

| Regime | Recognition (2*r*) | Behaviour |
|---|---|---|
| 1. strong recognition | $2r<Q$ | solution for $m\to\infty$ (large diffusion) also possible |
| 2a. weak recognition | $Q<2r<1$ | stationary solutions for finite range of *m, d, R* |
| b. neutral | $2r=1$ | stationary solutions for finite range of *m, d, R* |
| c. exploiter advantage | $1<2r<2r_{max}$* | stationary solutions for certain small *m, d, R* |
| 3. extreme exploitation | $r>r_{max}(m,d,Q)$ | no stationary solutions (complete extinction) |

Three basic regimes of recognition are shown in column one, for the parameter values in column two. The associated behaviour of the stationary solutions to the model are noted in column three.
*Note that $r_{max}$ can become infinite for certain values of the parameters *m,d,R*; with stationary solutions possible even for infinite *r* over a finite range of *m, d* and *R*.

Investigation of the two cubic equations for $\bar{y}$, with coefficients eqs (33) and (34), again using Descartes rule of sign on the polynomials as a function of $\bar{y}$ and $3-\bar{y}$, reveals that the first cubic has no solutions in the range $0<\bar{y}\leq 3$. The physical realizable solutions come from the second cubic eq (34). Once $\bar{y}$ has been calculated, we may proceed to calculate $\bar{x}$ and other probabilities of interest as detailed above, which also must be verified to obey the physical limitations $\bar{x}, \bar{y} \in \mathbb{R}$; $0 \leq \bar{x}, \bar{y} \leq 3$ and $0 \leq P(x,y) \leq 1$.

Perhaps the most interesting feature of the model is the possibility of dealing with an arbitrarily large strength of exploitation. One expects exploitation specialists to be able to evolve a specialization to being amplified by a catalyst, without the constraint of having to be catalytically active themselves. Viewed as information sequences, the exploiters can devote their entire information to optimising proliferation rate by the catalyst. Catalysts are generically incapable of completely distinguishing target molecules from sequence alternatives because molecular recognition is based on local interactions. For this reason, one expects *r* values considerably greater than $^1/_2$ to also occur generically in the evolution of catalysis. We now show analytically, that a parameter domain exists where an arbitrarily large exploitive rate coefficient *r* can be tolerated without extinction of the catalyst. At the same time, we show that certain features of the neutral case ($r=1/2$) already resemble this high *r* limit.

In order to simplify the analysis of the second cubic eq (34), we focus on the parameter regime of small d and $\tilde{m}$. Analysis of the sign changes of the coefficients in the this cubic, highlighted the importance of the relative magnitude of these parameters. We write $\tilde{m}=\alpha d$, and take the limit of the coefficients as $r\to\infty$ (multiplying with a common non-zero factor) obtaining the polynomial with coefficients (using the previously introduced matrix notation for polynomials):

$$\begin{pmatrix} -243\, d^4R^3\,(1+\alpha)^3\,(R+3d(1+\alpha)) \\ 27\, d^4R^2\alpha(1+\alpha)^2(3\alpha+15R^2\,(1+\alpha)-R(13+18\alpha)+6d(1+\alpha)(-4-6\alpha+R(5+6\alpha))) \\ 3\, d^4(1-R)R\alpha^2(1+\alpha)(-3+9\alpha-432\, d^2(1+\alpha)^3-3R(11+29\alpha)+2R^2(19+33\alpha)-12d(1+\alpha)(-2+6\alpha+3R(5+3\alpha))) \\ -4\, d^4(-1+R)^2R\alpha^3(36d(1+\alpha)^2-(-3+2R)(1+3\alpha)) \end{pmatrix} \tag{36}$$

Taking the limit $d\to 0$ one obtains

$$\begin{pmatrix} -243\, R^3(1+\alpha)^3 \\ 27\, R\alpha(1+\alpha)^2(3\alpha+15R^2(1+\alpha)-R(13+18\alpha)) \\ -3(-1+R)\alpha^2(1+\alpha)(-3+9\alpha-3R(11+29\alpha)+2R^2(19+33\alpha)) \\ 4(-1+R)^2(-3+2R)\alpha^3(1+3\alpha) \end{pmatrix} \tag{37}$$

and finally, looking for finite solutions $\bar{y}$ in the limit as $R\to 0$, the cubic reduces to

$$-3\alpha^2\bar{y}^2(3-6\alpha-9\alpha^2+(4\alpha+12\alpha^2)\bar{y})=0 \tag{38}$$

with non-zero solution for $\bar{y}$ and then $\bar{x}$ *via* eq (32) in the same limit

$$\bar{x}=\frac{3+3\alpha}{2\alpha+6\alpha^2} \qquad \bar{y}=\frac{3(-1+2\alpha+3\alpha^2)}{4\alpha(1+3\alpha)} \tag{39}$$

There are only solutions in the physical range [0,3] if $\alpha\geq 1/3$, *i. e.* $m>d/9$. In fact, $\bar{x}=3$ and $\bar{y}=0$ at $\alpha=1/3$. Plotting the larger *r*, low d, m and *R* limit of eq (39) for the values of d and m as in Figure 3, we can see in Figure 5 that the lowest values of the diffusion coefficient where stationary solutions occur is captured quite well by these simple limiting expressions.

A wide range of other analytical results can be obtained from the model using similar analysis of the polynomials, but these are best studied in relation to specific applications.

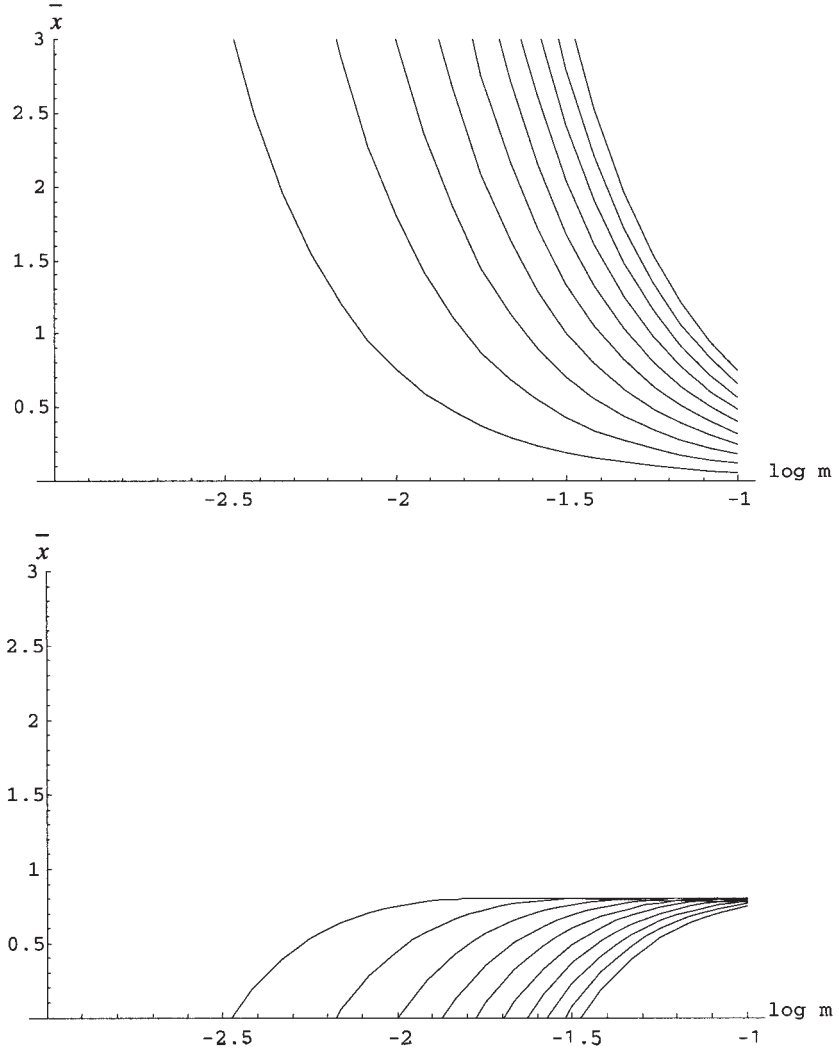


**Fig. 5** Limiting Mean Site Populations for Large Exploitation ($r \to \infty$).
The limiting solutions for $\bar{x}$ and $\bar{y}$ are plotted as in for d=0.01,0.02 ... 0.1 as a function of log $\tilde{m}$. The simple limiting formula of eq (39) captures basic features of the dependence of the stationary solutions at low diffusion rates, exhibiting the finite starting diffusion coefficient at correct locations for $\tilde{m}$ small, the bunching *d* dependence as *d* increases and the correct qualitative behaviour of the curves.

## Dynamical Solutions of the Model

The dynamical probability eq (8) may be integrated numerically directly. We shall show here that, for a range of diffusion rates above the range having stable stationary states, the dynamical equations admit limit cycle solutions. We shall also integrate the dynamics for several related models, including those with larger site populations $n$=4,5,...10, to check the robustness of our results. The dynamical solutions are of interest in their own right, but it is important to note that it is the infinite dimensional simplex space which allows temporal correlations between the probability vectors on different sites to be ignored. In lower dimensional space, such correlations, inducing spatial pattern formation, can play a decisive role. We shall compare the results of this and previous sections with three and lower dimensional calculations in the following section.

The oscillatory dynamics of the mean site populations $\bar{x}$ and $\bar{y}$ for appropriate parameters is shown in Figure 6. These solutions are somewhat similar to the limit cycles in predator prey systems, but structurally stable, unlike the cycles in the standard Lotka-Volterra system.

The simplex stochastic model allows all the individual occupation probabilities to be calculated explicitly as a function of time, not just the mean occupation numbers $\bar{x}$ and $\bar{y}$. Figure 7 displays all the probabilities as a function of time, showing the succession of probability states in the limit cycle. This analysis reveals the strongly bimodal probability distribution and gives insight into the succession of stages of amplification and exploitation.

The domain of diffusion coefficients exhibiting limit cycles can be seen in Figure 7. Note that limit cycles occur for diffusion coefficients $m$ above the limit for stable stationary solutions and up to a maximum value. For increasing $m$, the lim-

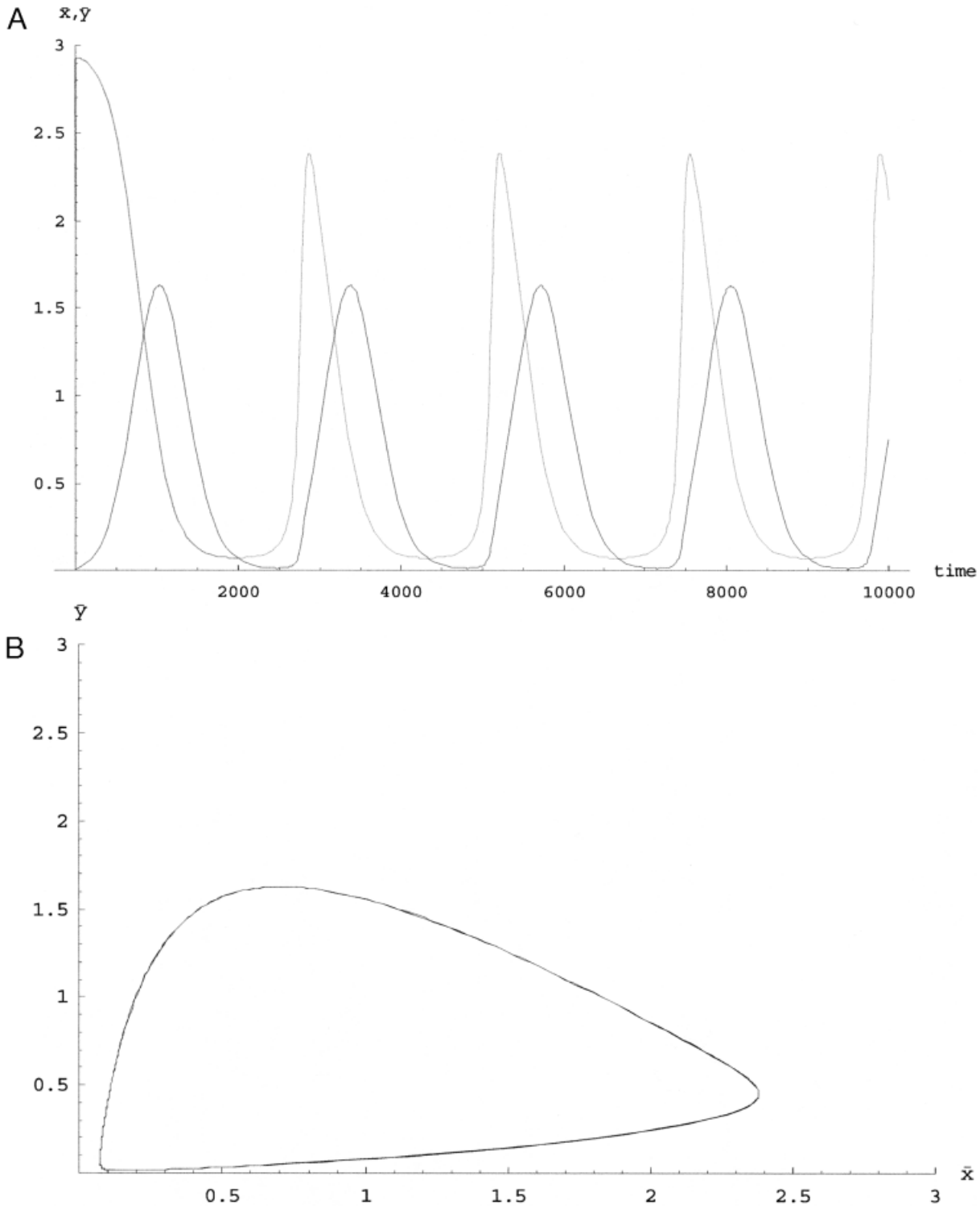


**Fig. 6** Oscillations in the Mean Populations $\bar{x}$ and $\bar{y}$ and Phase Portrait $\bar{y}$ ($\bar{x}$).
(A) The mean site population of autocatalyst X (upper curve) and exploiter Y (lower curve) are plotted as a function of time for the neutral recognition case ($r$=1/2) and representative model parameters $R$=0.005, d=0.01, m=0.01. Waves of exploitation follow the catalyst. (B) Phase plot showing limit cycle behavior (independent of initial conditions over a finite range) for the same parameter values. The populations oscillate between three states: (i) high catalyst, (ii) exploiter dominated and (iii) near extinction.

it cycles grow continuously, reaching sizes at which the trajectory passes close to (0,0) before passing over into trajectories converging to the stable attractor (0,0). The phase plot shows an extended domain where limit cycles occur and these were checked to be independent of the initial conditions over a finite range.

Finally, the model can be extended readily to site sizes of $n$=4,5 ... and the detailed probability kinetics integrated nu-

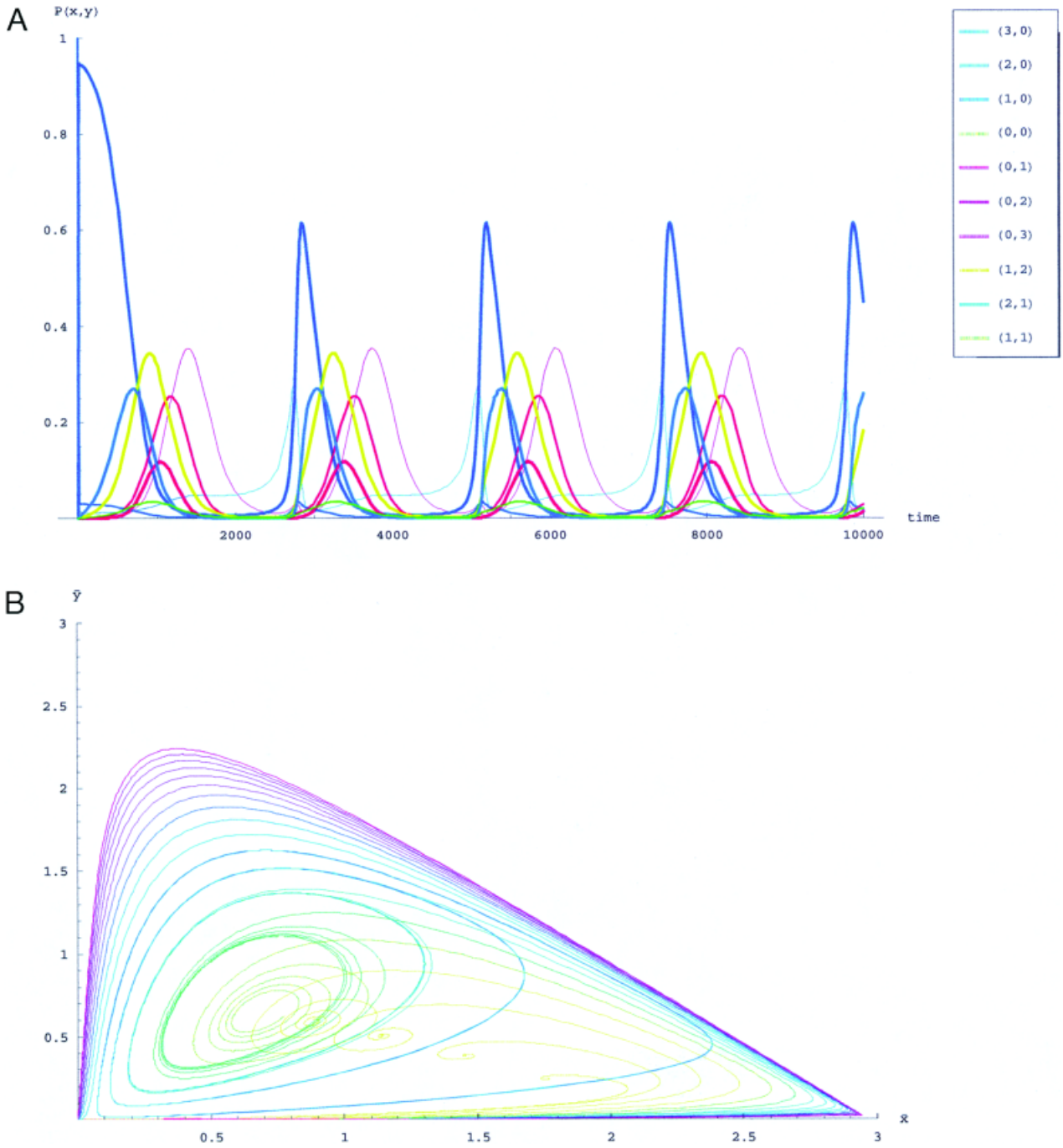


**Fig. 7** Probabilities of all Non-Zero Site Occupation States as a Function of Time.
(A) The individual probabilities are colour and line thickness coded. The total population $x+y$ is given by the line thickness and the exploitation fraction is given by the curve colour: from blue (pure catalyst) to red (pure exploiter). A peak of pure catalyst sites [$P(3,0)$] in thick blue is followed in rapid succession by singly, doubly and triply exploited sites, with the catalyst population dropping rapidly. Only when the residual population of sites occupied with single exploiters decays sufficiently, can the next catalytic wave take off. For parameters see Figure 5.
(B) The phase portrait of the dynamics of mean populations $\bar{x}$ (t) and $\bar{y}$ (t), $\bar{y}$ ($\bar{x}$), is shown for increasing diffusion rates m (from $m=10^{-3}$ to $10^{-1}$ in equal exponent steps of 0.1) color codes from orange through green, blue and crimson. Four parameter regimes are observed with increasing m: (i) extinction, (ii) stable stationary state, (iii) limit cycle, (iv) extinction. Initial conditions were chosen to start from (2,0) and altered to check that the final outcome did not depend strongly on initial conditions (not shown). The remaining parameter values are as in Figures 5 and 6 ($r=1/2$, $R=0.005$, $d=0.01$).

merically for site populations up to 10 or more. The dynamics for $n=4$ are shown in Figure 8, and demonstrate that although the catalyst concentration steadily decreases, there is a continuing extended domain of stationary limit cycle behavior (shifting to lower diffusion coefficients). This confirms that the properties analysed in this paper are not just arte-

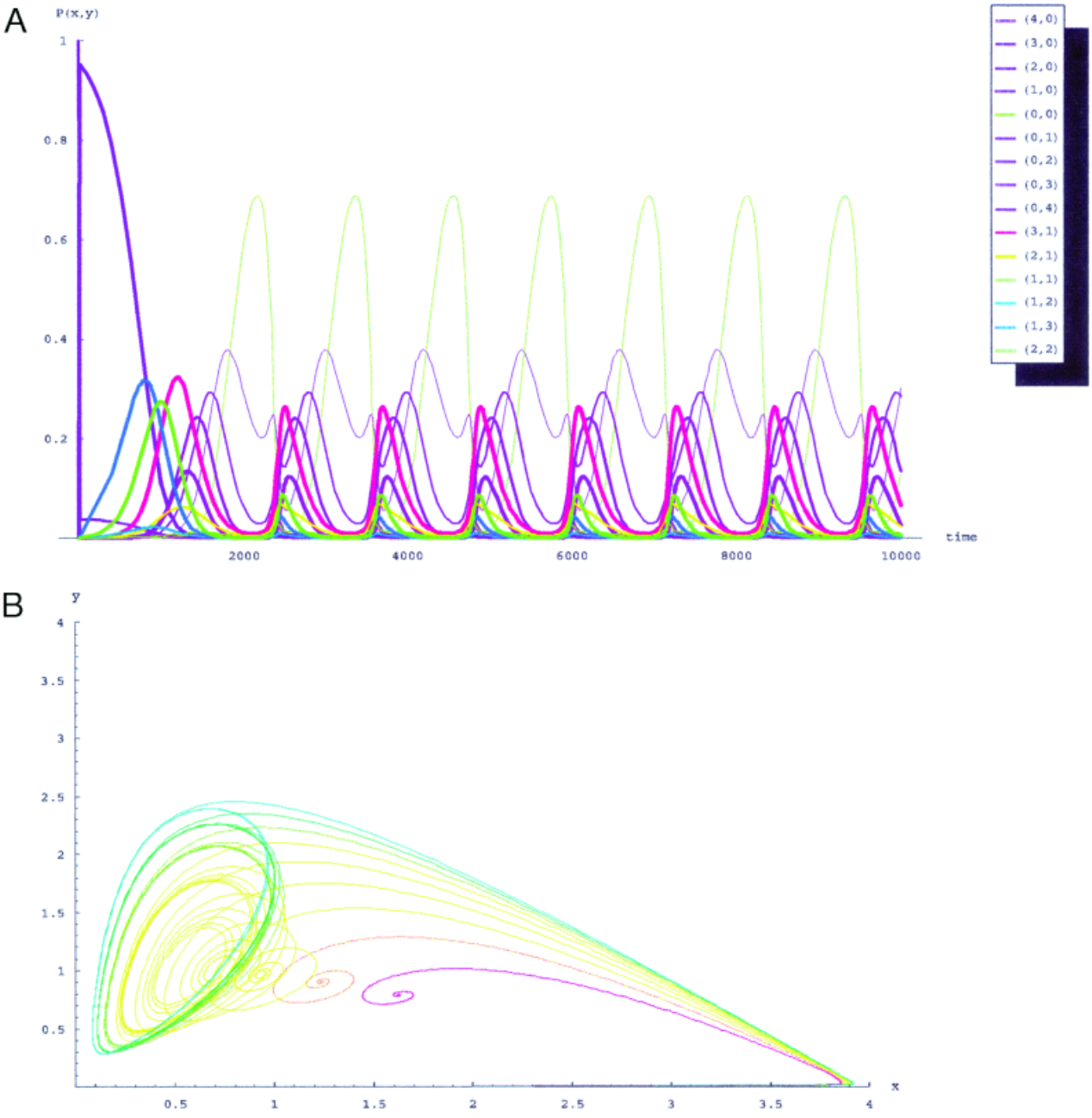


**Fig. 8** Solutions of Probability Dynamics for Larger Site Capacity n=4.
(A) Individual probabilities $P(x,y)$ The parameter values are as in Figure 5 (m = 0.01; d = 0.01; $R = 0.005$, $r = 1/2$) and the coding of the probability states in thickness and color is done analogously: the curve thickness is proportional to total site occupation $x+y$ and the colour is a spectrally proportional blend between the blue catalyst species and the pink exploiting species. Green curves have an equal number of catalysts and exploiters (including the zero occupation state, which reaches the highest probabilities between catalytic bursts). (B) Phase portrait as a function of diffusion coefficient. The colour coding and parameter values are as in Figure 7. Note the continued existence of a diffusion regime for stationary states and limit cycles for this large site size.

facts of the minimal site size $n$=3. In the next section, we shall show that the phenomena reported here also carry over to lower dimensional physical space.

### Prediction of Simulation Results in Euclidean Space

All of the previous results have dealt with the infinite dimensional simplex topology, which allows a neat exact self-consistent closure of the probability equations, since with an infinite number of neighbors the probabilities of any single site have a negligible impact on the probability of any other. We now seek to determine to what extent the features of the solutions carry over to lower dimensional physical space.

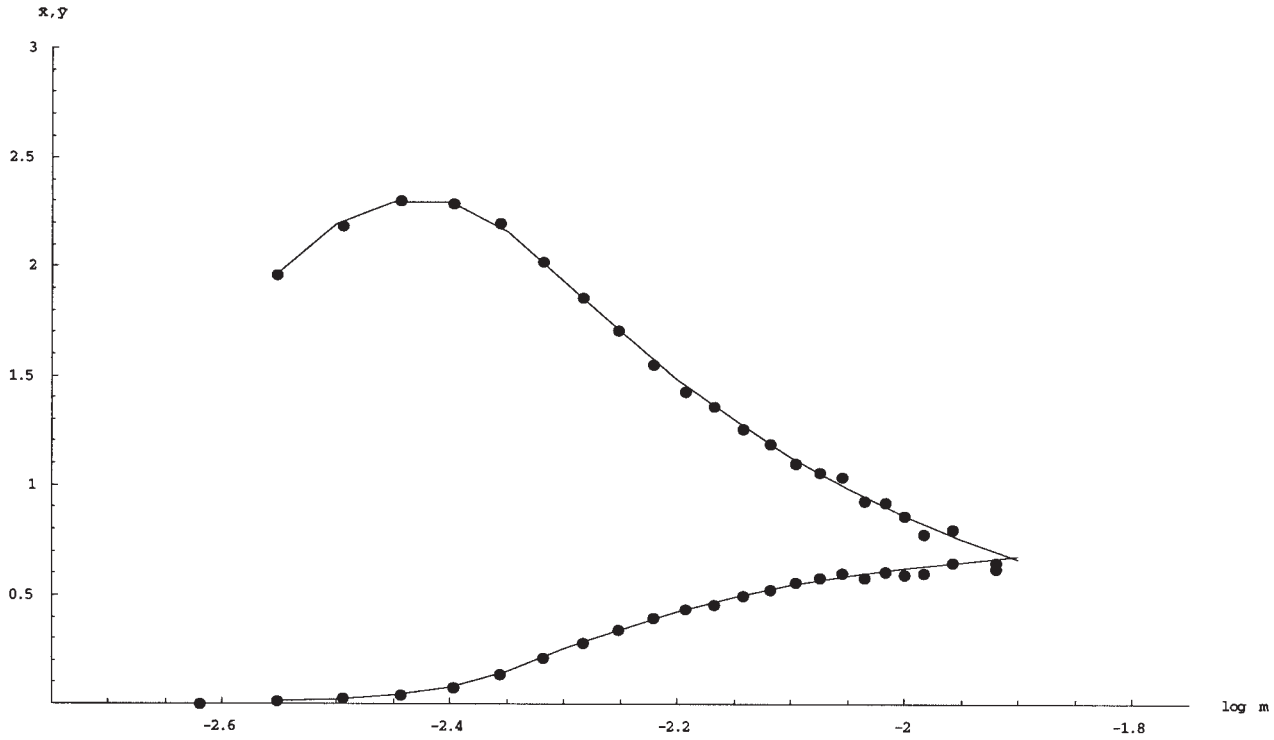


**Fig. 9** Comparison of the Results of the PRESS Model with Simulations on a Simplex Topology.
The simulation data are shown as points, the analytical theory (cf. Figure 3) as the lines for $\bar{x}$ (upper curve) and $\bar{y}$ (lower curve). The parameters were $r$=1/2, d=0.01, $R$=0.001. The simulations were run on 60 000 sites for 100 000 time units, to ensure the stationary states were reached.

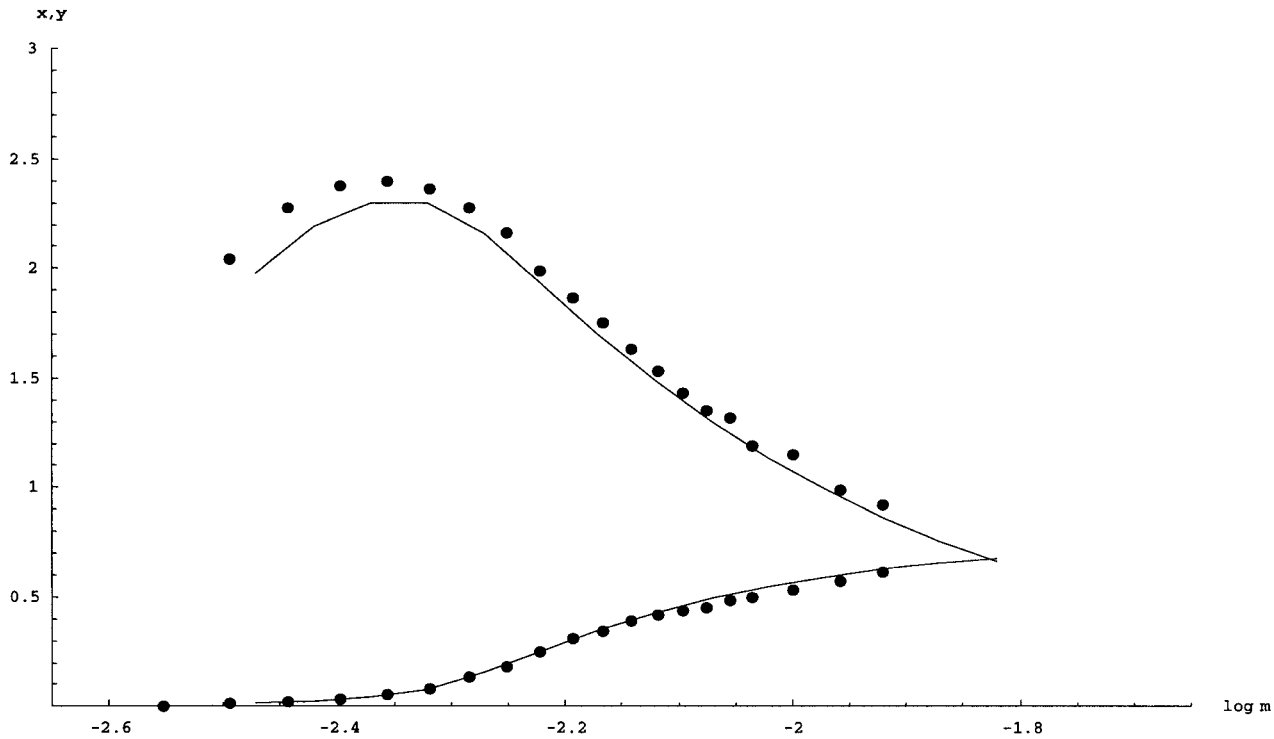


**Fig. 10** Comparison of the Results of the PRESS Model with Simulations in Three-Dimensional Euclidean Space.
The simulation data are shown as points, the analytical theory [using eq (40)] as the lines. The latter is rescaled by the dimension correction factor 5/6, see eq (40), to predict the self consistent behaviour in three dimensions for $\bar{x}$ (upper curve) and $\bar{y}$ (lower curve). The parameters were $r$=1/2, d=0.01, $R$=0.001 as in Figure 9. The simulations were run on 60 000 sites for 100 000 time units, to ensure that the stationary states were reached.

The only differences to be expected in the transition probabilities concern the treatment of diffusion between neighboring sites. The exact self-consistency condition determining the probability distribution of neighboring sites in the simplex geometry limit carries over to a mean field approximation for the probability dynamics in Euclidean space. With an appropriately scaled diffusion coefficient, it turns out that the results of the simplex model provides a good approximate description of the behaviour in Euclidean Space. Aspects of the dynamics that depend on specific correlations in occupation probabilities between neighboring sites which are different from the average will be ignored in a mean field treatment. Notice however, that the mean field treatment is in terms of individual site probabilities, rather than specific random variables.

Because the mean field assumption is only an approximation in lower dimensional space, we sought to compare the predictions based on the dimensionality-scaled diffusion of the simplex model to simulations in both simplex and Euclidean space. The simulations involved a large number (60 000) of sites in three dimensional space, with a Monte Carlo (*i. e.* random number generated) simulation of the dynamics according to the above model. The simulations were performed for various diffusion coefficients and the existence of stationary states and limit cycles monitored by following the mean populations of X and Y per site, $\bar{x}$ (t) and $\bar{y}$ (t), over time.

The comparison with simplex simulations in Figure 9 confirms the fact that the self-consistent theory is exact in the case simplex geometry. The assumption of probabilistic independence of fluctuations in individual sites is near perfect in the large site number simulations and so the agreement is excellent. The projection of the predictions of the self consistent model from the simplex geometry to three dimensions requires the diffusion coefficient m to be rescaled. The naïve rescaling, based on the fraction of new neighbours which a diffusing particle perceives, is

$$m_{\mathrm{K}} \frac{2\mathrm{K}-1}{2\mathrm{K}} = m_{\infty} \tag{40}$$

where K ist the dimensionality, and with this scaling, the analytical theory appears to predict the stationary states of the cooperation model in three dimensions rather well (Figure 10). The self consistent probability theory is only an approximation in Euclidean space. In particular, correlations between neighbouring sites which are different from correlations between randomly chosen sites can cause departures from the predictions of the model. However, these correlations do not appear to be so important, and the main features of the model appear to carry over to physical space.

Indeed the simulations in both the simplex case and three-dimensional space also demonstrate that the system goes over into limit cycle oscillations at high diffusion rates when the stationary states become unstable (Hopf bifurcation) as predicted by the analytical theory. The form of these limit cycles is also well predicted by the analytical theory. Examples of limit cycle behaviour from the simulations in the simplex geometry and three dimensions are shown in Figure 11. This confirms that also the dynamical features of the full problem are captured by the analytical model (see above). A detailed study of the dynamics in three dimensional space is beyond the scope of this article.

## Discussion and Conclusions

The framework of stochastic individual modelling of the probability dynamics of many coupled minimum sized local populations, using the analytic limit of high dimensional space, can be extended to many chemical and biological problems involving non-linear reaction kinetics and spatial effects. In this respect, we anticipate that the results of this paper will open up a range of further applications. Whether the infinite dimensional topology limit will provide a good approximation to the behaviour in three dimensional space (or lower) in general will depend on properties of the kinetic model being investigated. For the case investigated here, the extrapolation is very successful.

The probability equations described have a quadratic non-linearity in the probabilities. This is different from the standard linear equations for occupation probabilities (*e. g.* master equations) because the state space description has been condensed. Instead of considering the joint probability distribution of occupation states $\{(x_1,y_1),(x_2,y_2),\ldots,(x_M,y_M)\}$ for the two species X and Y on each of the large number M of sites

$$P((x_1,y_1),(x_2,y_2),\ldots,(x_M,y_M);t) \tag{41}$$

we have studied the single site probabilities $P(x,y;t)$. The probability distributions on different sites are not independent of one another, since they are coupled by diffusion, otherwise one could simply factorize this joint distribution in terms of the single site probabilities. What the self-consistent treatment actually entails is that the other sites have on average the same mean occupation numbers as the site under consideration. In the infinite site number ($M \to \infty$) limit, no single other site can contribute significantly to affect this statistic, and every site sees the same ($M-1 \simeq$M for M large) influx statistics. The probabilities for particle influx to any site is given exactly by the mean occupation numbers in the other sites. Certain initial conditions, which distinguish different sites probabilistically, can lead to inhomogeneous probability kinetics, but differences in probabilities should decay rapidly as a result of the common influx rates.

The importance of an individual level treatment of stochastic kinetics in biological and chemical modelling has been stressed by other authors, *e. g.* Durrett and Levin (1994a,b), as well as by one of the authors (McCaskill, 1984; Breyer *et al.,*

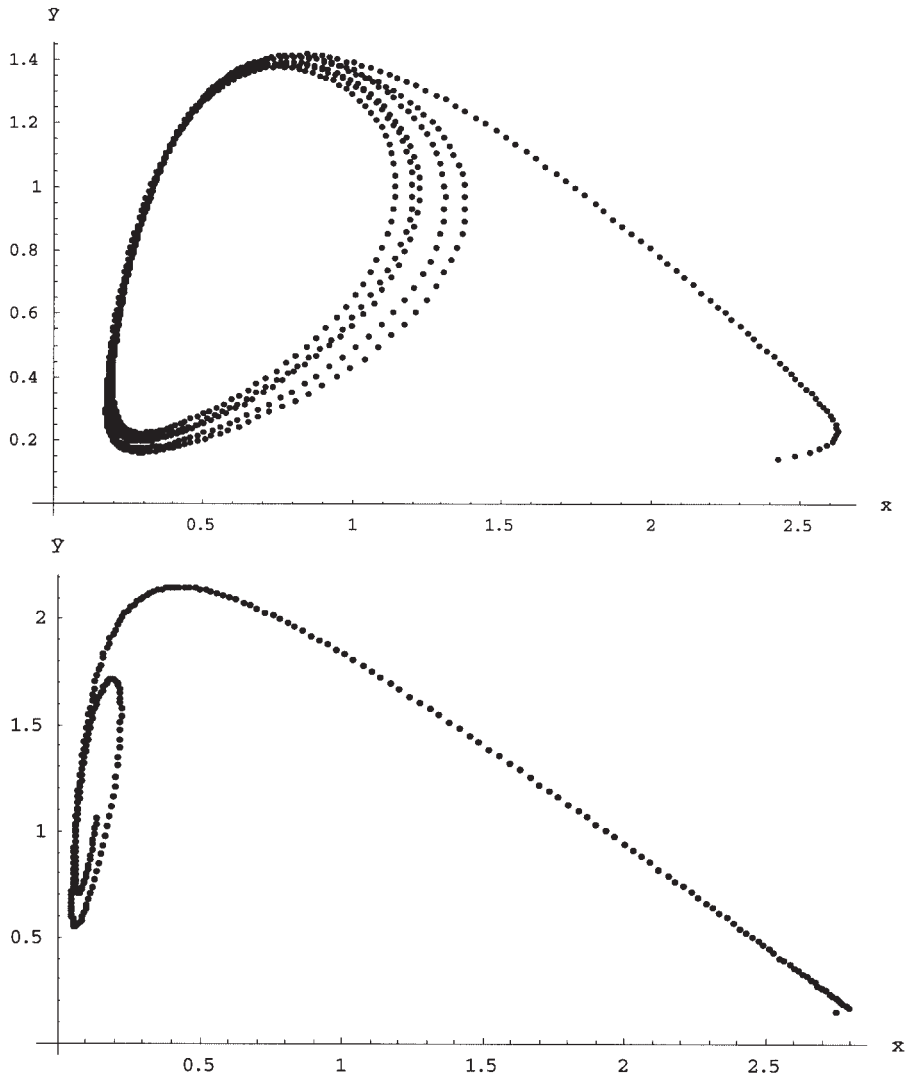


**Fig. 11** Limit Cycles in Full Simulations of the Stochastic Dynamics in the Simplex Geometry and 3D Space.
Phase portraits of the simulated stochastic dynamics in a large number of finite size sites is shown for the simplex geometry (top) and in three dimensional space (bottom). The two calculations were not dimensionality adjusted to have the same effective diffusion rate but were performed instead with the same diffusion parameter (m=0.01).

1998). In the current model, the discrete nature of local populations is necessary to allow complete local extinction of the exploiting species, prior to new catalysts reoccupying a site. Local stochastic fluctuations are important for survival of the catalytic species, as in the Stochastic Corrector model of Szathmary and Demeter (1987). The latter theory involves dividing cell-like compartments to avoid exploitation, and furthermore only approximate stochastic results were obtained. In the present theory, the stochastic dynamics and in particular the probabilistic steady-states can be described exactly, despite the non-trivial non-linear nature of the model, and no meta-dynamics of compartment division is required for survival of the catalysts.

The possibility of spatial isolation by distance playing a pivotal role in the evolution of catalytic systems was introduced in the work on spiral waves in the hypercycle model (Boerlijst and Hogeweg, 1991). The two dimensional models of discrete reaction-diffusion dynamics, using cellular automata or partial differential equations with cutoffs, demonstrated the containment of exploitation by spiral wave formation in catalytic cycles with more than four members. Subsequent work by the current author (McCaskill *et al.,* 1997) and others (Cronhjort and Blomberg, 1997) demonstrated the containment of catalytic exploitation through self-replicating spot patterns in simpler single catalyst systems. The spatial oscillatory nature of catalytic exploitation was explored in individual based evolution models in connection with an *in vitro* model of coupled macromolecular evolution (Breyer, 1998; Ackermann and Kirner, 1999). The regular limit cycle os-

cillations of the current non-linear probability dynamics were not observed, but the importance of macroscopic fluctuations in population numbers for containing parasitic exploitation was already posited in these works. Other publications dealing with early evolution have found limit cycles in different contexts, *e. g.* Chacon and Nun (1995).

The current theory of catalytic evolution goes beyond the framework of non-interacting replicators for which the quasispecies theory was developed (Eigen *et al.,* 1989). An analysis of the spatially resolved error threshold for finely spatially structured quasispecies was recently proposed by the authors (Altmeyer and McCaskill, 2001) in a simplex topology, as in the present paper, and the results compared to simulations in lower dimensional Euclidean Space. The conclusion of that work is that spatial effects lower the error threshold, decreasing the amount of information that can be assembled by limited fidelity replication, below that of the well-mixed, homogeneous limit described by the deterministic kinetic equations of the quasispecies model. In the current work, in which catalyst-template interactions are taken into account, spatial localization allows the system to deal with both neutral and hyperparasitic exploitation over a certain intermediate range of diffusion constants.

The positive effects of population structuring on the evolution of altruistic effects has been the subject of many theoretical investigations, usually within the framework of group selection or kin selection theory. In particular, Kimura (1983) studied a spatial model of the evolution of an altruistic allele within a Fokker-Planck equation (large population size or diffusion approximation) description of the probability dynamics, which like the later model of Szathmary involved two level replication. Kimura took into account migration between sites and studied the effects of an intergroup selection function on the distribution of altruists. One advantage of the model presented is that the dynamics does not refer to some phenomenological group selection function, but is based only on individual chemical rate constants. Selection is reflectd intrinsically. The basic result of that theory, describing an approximate upper limit to the migration rate of species for survival of the altruistic allele, is mirrored, despite the absence of group selection, in the upper bound for the diffusion constant m for the survival of the catalysts X in the present model.

With the introduction of more species, the present technique could also be extended to address the origin of genetic coding. Dyson (1982) discussed a model of the origin of life with isolated islands hosting molecular populations with different degrees of autocatalytic activity. Dyson's model does not capture the interplay between the dynamics on separate sites and the exchange of molecular components between sites. Though Dyson's islands are able to exchange chemical components and free energy with the environment, there is no coupling between the islands and in fact the topology plays no role in his analysis. Dyson's model also refers to a situation prior to the establishment of replication, using a characterization of catalytic activity in terms of the number of correctly placed catalysts. In previous work based on massively parallel simulation, we have shown that the explicit population dynamics modelling, as employed in the present article, is necessary to describe the qualitative features of the origin of genetic coding correctly (Füchslin and McCaskill, 2001).

What are the experimental conditions necessary for the evolutionary stabilisation of catalysts according to the proposed mechanism? In continuous aqueous solutions in three dimensional space, typical diffusion times for biopolymers such as RNA, on the length scale of 100 bases, are about $2\times10^{-10}\mathrm{m}^2/\mathrm{s}$. On the time scale of *in vitro* replication, several seconds in experiments with RNA (Wright and Joyce, 1997), a characteristic length scale $L = \sqrt{D\tau}$ of *ca.* $5\times10^{-5}$ m results. A cube of this dimension has volume *ca.* $10^{-10}$ liters, so that unrealistically small concentrations, around $10^{-13}$ M, are necessary to see the cooperative effects of spatial isolation in molecular systems in free solution. However, in gels or porous media with microscopic chambers on the spatial scale of 100 nm, more readily detectable concentrations in the nanomolar range and above may be employed. This will be the necessary environment for the autonomous evolution of heterocatalytic activity in cell-free systems, both from the point of view of the origin of life and in biotechnological applications.

In summary, we have shown analytically that spatially subdivided coupled catalytic systems can be stabilized against exploitation far beyond the condition of $2r<Q$ (which is a strong condition for well-mixed systems).Even for the case of $r\to\infty$, the catalyst can protect itself against total exploitation in the low mobility regime, making use of spatial isolation. The system exhibits four successive behaviours with increasing diffusion rate for $2r>Q$ (i) no solutions, (ii) stable fixed points, (iii) limit cycles, and (iv) no solutions. We have shown that the simplex topology is not an artificial and isolated case but exhibits the main features anticipated with spatial resolution. Furthermore, a simple scale transformation allows one to extend these results to the finite spatial dimensionality of Euclidean space.

## Appendix

The 10 by 10 matrix of transition probabilitites $\underline{\underline{W}}_{10}$ is given in terms of the basic transitions by

$$\begin{pmatrix} P[3,0]'[t] \\ P[2,0]'[t] \\ P[1,0]'[t] \\ P[0,0]'[t] \\ P[0,1]'[t] \\ P[0,2]'[t] \\ P[0,3]'[t] \\ P[1,2]'[t] \\ P[2,1]'[t] \\ P[1,1]'[t] \end{pmatrix} =$$

$$\begin{pmatrix} -a_s[3,0] & a_{xz}[2,0] & 0 & 0 & 0 & 0 & 0 & 0 & a_{xy}[2,1] & 0 \\ a_{zx}[3,0] & -a_s[2,0] & a_{xz}[1,0] & 0 & 0 & 0 & 0 & 0 & a_{zy}[2,1] & a_{xy}[1,1] \\ 0 & a_{zx}[2,0] & -a_s[1,0] & a_{xz}[0,0] & a_{xy}[0,1] & 0 & 0 & 0 & 0 & a_{zy}[1,1] \\ 0 & 0 & a_{zx}[1,0] & -a_s[0,0] & a_{zy}[0,1] & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{yx}[1,0] & a_{yz}[0,0] & -a_s[0,1] & a_{zy}[0,2] & 0 & 0 & 0 & a_{zx}[1,1] \\ 0 & 0 & 0 & 0 & a_{yz}[0,1] & -a_s[0,2] & a_{zy}[0,3] & a_{zx}[1,2] & 0 & a_{yx}[1,1] \\ 0 & 0 & 0 & 0 & 0 & a_{yz}[0,2] & -a_s[0,3] & a_{yx}[1,2] & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{xz}[0,2] & a_{xy}[0,3] & -a_s[1,2] & a_{yx}[2,1] & a_{yz}[1,1] \\ a_{yx}[3,0] & a_{yz}[2,0] & 0 & 0 & 0 & 0 & 0 & a_{xy}[1,2] & -a_s[2,1] & a_{xz}[1,1] \\ 0 & a_{yx}[2,0] & a_{yz}[1,0] & 0 & a_{xz}[0,1] & a_{xy}[0,2] & 0 & a_{zy}[1,2] & a_{zx}[2,1] & -a_s[1,1] \end{pmatrix} \cdot \begin{pmatrix} P[3,0]'[t] \\ P[2,0]'[t] \\ P[1,0]'[t] \\ P[0,0]'[t] \\ P[0,1]'[t] \\ P[0,2]'[t] \\ P[0,3]'[t] \\ P[1,2]'[t] \\ P[2,1]'[t] \\ P[1,1]'[t] \end{pmatrix} \quad (42)$$

or in terms of the parameters of the model

$$\begin{pmatrix} P[3,0]'[t] \\ P[2,0]'[t] \\ P[1,0]'[t] \\ P[0,0]'[t] \\ P[0,1]'[t] \\ P[0,2]'[t] \\ P[0,3]'[t] \\ P[1,2]'[t] \\ P[2,1]'[t] \\ P[1,1]'[t] \end{pmatrix} = \begin{pmatrix} -3(d+3m-m\bar{x}) & Q+m\bar{x} & 0 & 0 & 0 & 0 & 0 & 0 & m\bar{x} & 0 \\ 3(d+3m-m\bar{x}-m\bar{y}) & -1-2d-6m+m\bar{x}-m\bar{y} & 2m\bar{x} & 0 & 0 & 0 & 0 & 0 & d+3m-m\bar{x}-m\bar{y} & m\bar{x} \\ 0 & 2(d+3m-m\bar{x}-m\bar{y}) & -d-3m-m\bar{x}-2m\bar{y} & 3m\bar{x} & m\bar{x} & 0 & 0 & 0 & 0 & d+3m-m\bar{x}-m\bar{y} \\ 0 & 0 & d+3m-m\bar{x}-m\bar{y} & -3m(\bar{x}+\bar{y}) & d+3m-m\bar{x}-m\bar{y} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & m\bar{y} & 3m\bar{y} & -d-3m-2m\bar{x}-m\bar{y} & 2(d+3m-m\bar{x}-m\bar{y}) & 0 & 0 & 0 & d+3m-m\bar{x}-m\bar{y} \\ 0 & 0 & 0 & 0 & 2m\bar{y} & -2(d+3m)-m\bar{x}+m\bar{y} & 3(d+3m-m\bar{x}-m\bar{y}) & d+3m-m\bar{x}-m\bar{y} & 0 & m\bar{y} \\ 0 & 0 & 0 & 0 & 0 & m\bar{y} & -3(d+3m-m\bar{y}) & m\bar{y} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & m\bar{x} & 3m\bar{x} & -3(d+3m)+m\bar{x}+2m\bar{y} & 2m\bar{y} & r+m\bar{y} \\ 3m\bar{y} & 1-Q+m\bar{y} & 0 & 0 & 0 & 0 & 0 & 2m\bar{x} & -3(d+3m)+2m\bar{x}+m\bar{y} & m\bar{x} \\ 0 & 2m\bar{y} & 2m\bar{y} & 0 & 2m\bar{x} & 2m\bar{x} & 0 & 2(d+3m-m\bar{x}-m\bar{y}) & 2(d+3m-m\bar{x}-m\bar{y}) & -2d-6m-r \end{pmatrix} \cdot \begin{pmatrix} P[3,0]'[t] \\ P[2,0]'[t] \\ P[1,0]'[t] \\ P[0,0]'[t] \\ P[0,1]'[t] \\ P[0,2]'[t] \\ P[0,3]'[t] \\ P[1,2]'[t] \\ P[2,1]'[t] \\ P[1,1]'[t] \end{pmatrix} \quad (43)$$

where the sum of the terms in each column is zero, reflecting the conservation of total probability.

The matrix $\underline{\underline{W}}$ is given by:

$\underline{\underline{W}} =$

$$\begin{pmatrix} -3(d+3m-m\bar{x}) & 1-R+m\bar{x} & 0 & 0 & 0 & 0 & 0 & m\bar{x} & 0 \\ 3(d-m(-3+\bar{x}+\bar{y})) & -1-2d+m(-6+\bar{x}-\bar{y}) & 2m\bar{x} & 0 & 0 & 0 & 0 & d-m(-3+\bar{x}+\bar{y}) & m\bar{x} \\ -3m(-3+\bar{x}) & 2d+m(12-5\bar{x}-2\bar{y}) & -d-4m\bar{x}-2m\bar{y} & -2m\bar{x} & -3m\bar{x} & -3m\bar{x} & -3m(-1+\bar{x}) & -3m(-2+\bar{x}) & d-m(-6+4\bar{x}+\bar{y}) \\ -3m\bar{y} & -3m\bar{y} & -2m\bar{y} & -d-2m\bar{x}-4m\bar{y} & 2d+m(12-2\bar{x}-5\bar{y}) & -3m(-3+\bar{y}) & -3m(-2+\bar{y}) & -3m(-1+\bar{y}) & d-m(-6+\bar{x}+4\bar{y}) \\ 0 & 0 & 0 & 2m\bar{y} & -2d+m(-6-\bar{x}+\bar{y}) & 3(d-m(-3+\bar{x}+\bar{y})) & d-m(-3+\bar{x}+\bar{y}) & 0 & m\bar{y} \\ 0 & 0 & 0 & 0 & m\bar{y} & -3(d+3m-m\bar{y}) & m\bar{y} & 0 & 0 \\ 0 & 0 & 0 & 0 & m\bar{x} & 3m\bar{x} & -3d+m(-9+\bar{x}+2\bar{y}) & 2m\bar{y} & 1/2+m\bar{y} \\ 3m\bar{y} & R+m\bar{y} & 0 & 0 & 0 & 0 & 2m\bar{x} & -3d+m(-9+2\bar{x}+\bar{y}) & m\bar{x} \\ 0 & 2m\bar{y} & 2m\bar{y} & 2m\bar{x} & 2m\bar{x} & 0 & 2(d-m(-3+\bar{x}+\bar{y})) & 2(d-m(-3+\bar{x}+\bar{y})) & 1/2-2d-6m \end{pmatrix}$$

## Acknowledgements

Support from the German Research Council (DFG) and the Swiss National Science Foundation (SNF) is gratefully acknowledged. The first author particularly wishes to thank Sigrid M. for her endurance and support while this work was being completed.

## References


Ackermann, J., and Kirner, T. (1999). Parasites and pattern formation. Z. Naturforsch *54A*, 146 – 152.

Altmeyer, S., and McCaskill, J.S. (2001). Error threshold for spatially resolved evolution in the quasispecies model. Phys. Rev. Lett. *86*, 5819 – 5822.

Böddeker, B. (1995). Physikalische Modelle für Koevolution und Simulation auf einem hardwareprogrammierbaren Prozessor / Untersuchung von räumlich- und haftungsinduzierter Korrelation auf stabile Symbiose von Funktionen. Diploma thesis, University of Göttingen, Germany.

Boerlijst M., and Hogeweg P. (1991). Spiral wave structure in pre-biotic evolution – hypercycles stable against parasites. Physica D *48*, 17 – 28.

Bresch, C., Niesert, U., and Harnasch, D. (1980). Hypercycles, parasites and packages. J. Theor. Biol. *85*, 399 – 405.

Breyer, J. (1998). Anwendung eines in der Hardware konfigurierbaren Parallelrechners auf die Simulation evolutiver in vitro Amplifikationssysteme. PhD thesis, Friedrich-Schiller University, Jena, Germany.

Breyer, J., Ackermann, J., and McCaskill, J.S. (1998). Evolving reaction-diffusion ecosystems with self-assembling structures in thin films. Artif. Life 4 *1*, 25 – 40.

Buss, L. (1987). The Evolution of Individuality (Princeton, New Jersey, USA: Princeton University Press).

Chacon, P., and Nuno J.C. (1995). Spatial dynamics of a model for prebiotic evolution. Physica D *81*, 398 – 410.

Cronhjort, M.B., and Blomberg, C. (1997). Cluster compartmentalization may provide resistance to parasites for catalytic networks. Physica D *101*, 289 – 298.

Durrett, R., and Levin S. (1994a). The importance of being discrete (and spatial). Theor. Pop. Biol. *46*, 363 – 394.

Durrett, R., and Levin S. (1994b). Stochastic spatial models – a users guide to ecological applications. Phil. Trans. R. Soc. Lond. B *343*, 329 – 350.

Dyson, F. (1982). A model for the origin of life. J. Mol. Evol. *18*, 344 – 350.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften *58*, 465 – 523.

Eigen, M. (1987). Stufen zum Leben (Munich, Germany: Piper).

Eigen, M., McCaskill, J.S., and Schuster, P. (1989). Spatial dynamics of a model for prebiotic evolution. Adv. Chem. Phys. *75*, 149 – 263.

Ellinger, T., Ehricht, R. and McCaskill, J.S. (1988). *In vitro* evolution of molecular cooperation in CATCH, a cooperatively coupled amplification system. Chem. Biol. *5*, 729 – 741.

Füchslin, R.M., and McCaskill, J.S. (2001). Evolutionary self organization of cell free genetic coding. Proc. Natl. Acad. Sci. USA *98*, 9185 – 9190.

Kimura, M. (1983). Diffusion model of intergroup selection, with special reference to evolution of an altruistic character. Proc. Natl. Acad. Sci. USA *80*, 6317 – 6321.

Kirner, T., Ackermann, J., Ehricht, R. and McCaskill, J.S. (1999). Complex patterns predicted in an *in vitro* experimental model system for the evolution of molecular cooperation. Biophys. Chem. *79*, 163 – 186.

Latter, B.D.H. (1973). The island model of population differentiation: a general solution. Genetics *73*, 147 – 157.

McCaskill, J.S. (1984). A stochastic theory of macromolecular evolution. Biol. Cybern. *50*, 63 – 73.

McCaskill, J.S. (1997). Spatially resolved *in vitro* molecular ecology. Biophys. Chem. *66*, 145 – 158.

McCaskill, J.S., Tangen, U., and Ackermann, J. (1997). VLSE Very Large Scale Evolution in Hardware. Proceedings of the 4th European Conference on Artificial Life, P. Husbands and I. Harvey, eds. (Cambridge, USA: MIT Press/Bradford Books), 398 – 406.

Nowak, M., and Schuster P. (1989). Error thresholds of replication in finite populations mutation frequencies and the onset of mullers ratchet. J. Theor. Biol. *137,* 375 – 395.

Szathmary, E., and Demeter, L. (1987). Group selection of early replicators and the origin of life. J. Theor. Biol. *128*, 463 – 486.

Wright, M.C., and Joyce G.F. (1997). Continuous *in vitro* evolution of catalytic function. Science *276,* 614 – 617.

Wright, S. (1931). Evolution in Mendelian population. Genetics *16*, 97 – 159.


*Received March 21, 2001; accepted August 2, 2001*