

0.0.1 $\langle \Delta H_{SD} \rangle$ Derivation

The equation $H = -q(1 - q)$ defines a constant of motion for the SD part of the game, where q is the fraction of players in the 4th strategy. Using the transition probabilities of the different process we can derive an expression for the expected change in H within the simplex.

Where $i, j, k, N - i - j - k$ are the players playing R, P, S, and the 4th strategy respectively.

$$\begin{aligned}\Delta H &= H(t+1) - H(t), \\ \Delta H &= -x_{t+1}(1 - x_{t+1}) - (-x_t(1 - x_t)) \\ \Delta H &= -x_{t+1}(1 - x_{t+1}) + x_t(1 - x_t) \\ \Delta H &= x_t(1 - x_t) - x_{t+1}(1 - x_{t+1})\end{aligned}$$

$$\langle \Delta H \rangle = \sum_{i,j,k} (H_s - H_{s'}) T^{s \rightarrow s'}, \text{ } s \text{ is a particular state in the simplex.}$$

$$\begin{aligned}\langle \Delta H_{SD} \rangle &= \frac{6}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[(N-i-j-k)(1-N+i+j+k)(T^{R+} + T^{P+} + T^{S+} + T^{+R} + T^{+P} + T^{+S}) \right. \\ &\quad - (N-i-j-k+1)(-N+i+j+k)T^{R+} \\ &\quad - (N-i-j-k+1)(-N+i+j+k)T^{P+} \\ &\quad - (N-i-j-k-1)(2-N+i+j+k)T^{+R} \\ &\quad - (N-i-j-k-1)(2-N+i+j+k)T^{+P} \\ &\quad \left. - (N-i-j-k-1)(2-N+i+j+k)T^{+S} \right] \quad (1)\end{aligned}$$

The terms with transitions within the RPS simplex can be ignored as q would not change between these states, therefore the term $H_s - H_{s'} = 0$

$p = N - i - j - k$, the number of players playing the 4th strategy.

$$\begin{aligned}\langle \Delta H_{SD} \rangle &= \frac{6}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[p(1-p)(T^{R+} + T^{P+} + T^{S+} + T^{+R} + T^{+P} + T^{+S}) \right. \\ &\quad - (p+1)(-p)(T^{R+} + T^{P+} + T^{S+}) \\ &\quad \left. - (p-1)(2-p)(T^{+R} + T^{+P} + T^{+S}) \right] \\ \langle \Delta H_{SD} \rangle &= \frac{6}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[T^{R+} (p(1-p) + p(p+1)) + T^{+R} (p(1-p) - (p-1)(2-p)) \right. \\ &\quad + T^{P+} (p(1-p) + p(p+1)) + T^{+P} (p(1-p) - (p-1)(2-p)) \\ &\quad \left. + T^{S+} (p(1-p) + p(p+1)) + T^{+S} (p(1-p) - (p-1)(2-p)) \right] \\ &= \frac{6}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[2p (T^{R+} + T^{P+} + T^{S+}) + (p - p^2 - (3p - p^2 - 2))(T^{+R} + T^{+P} + T^{+S}) \right] \\ &= \frac{6}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[2p (T^{R+} + T^{P+} + T^{S+}) + (2 - 2p)(T^{+R} + T^{+P} + T^{+S}) \right]\end{aligned}$$

$$\begin{aligned}
\langle \Delta H_{SD} \rangle &= \frac{12}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[p (T^{R+} + T^{P+} + T^{S+}) + (1-p)(T^{+R} + T^{+P} + T^{+S}) \right] \\
\langle \Delta H_{SD} \rangle &= \frac{12}{N^5} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} p \left[(T^{R+} - T^{+R}) + (T^{P+} - T^{+P}) + (T^{S+} - T^{+S}) \right] + T^{+R} + T^{+P} + T^{+S} \quad (2)
\end{aligned}$$

The continuous limit, where $x = i/N$, $y = j/N$, $z = k/N$, and $q = p/N$, $p = Nq$ and $q = 1 - x - y - z$ leads to:

$$\begin{aligned}
\langle \Delta H_{SD} \rangle &= \frac{12}{N^2} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[Nq \left[(T^{R+} - T^{+R}) + (T^{P+} - T^{+P}) + (T^{S+} - T^{+S}) \right] \right. \\
&\quad \left. + (T^{+R} + T^{+P} + T^{+S}) \right]
\end{aligned}$$

Finally,

$$\begin{aligned}
\langle \Delta H_{SD} \rangle &= \frac{12}{N} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[q \left[(T^{R+} - T^{+R}) + (T^{P+} - T^{+P}) + (T^{S+} - T^{+S}) \right] \right] \\
&\quad + \frac{12}{N^2} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[T^{+R} + T^{+P} + T^{+S} \right] \quad (3)
\end{aligned}$$

This can then be solved numerically and the critical population values can be found where $\langle \Delta H_{SD} \rangle = 0$.

Critical N found matches nicely with the simulated versions. Numerical integration in python code `./augRps.py`, shows change of sign as expected. Matches nicely with the approximated values for the Moran process. The specific expression for Moran process is very long. Computed numerically and solved with `scipy.integrate` (reference `scipy`) Maybe can plot the simulated critical population sizes against the analytical on the same graph.

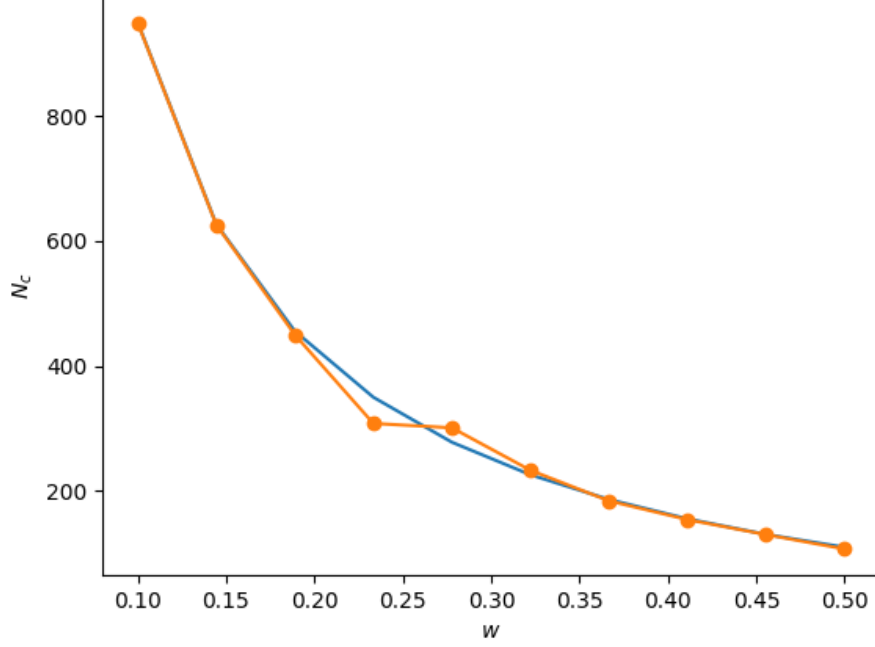


Figure 1: Comparison of simulated critical population sizes against the derived expected change in H for Moran process. Small deviations due to stochasticity in the simulations. $\gamma = 0.2, \beta = 0.1$ standard RPS.

0.0.2 $\langle \Delta H_{RPS} \rangle$ Derivation

$\langle \Delta H \rangle$ within the RPS plane $H = -xyz$.

$$\begin{aligned}
\langle \Delta H_{RPS} \rangle &= \frac{2}{N^6} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[ijk(T^{RP} + T^{RS} + T^{R+} + T^{PR} + T^{PS} + T^{P+} \right. \\
&\quad + T^{SR} + T^{SP} + T^{S+} + T^{+R} + T^{+P} + T^{+S}) \\
&\quad - k(i-1)(j+1)T^{RP} - (i-1)j(k+1)T^{RS} - jk(i-1)T^{R+} \\
&\quad - k(i+1)(j-1)T^{PR} - i(j-1)(k+1)T^{PS} - ik(j-1)T^{P+} \\
&\quad - (i+1)j(k-1)T^{SR} - i(j+1)(k-1)T^{SP} - ij(k-1)T^{S+} \\
&\quad \left. - jk(i+1)T^{+R} - ik(j+1)T^{+P} - ij(k+1)T^{+S} \right] \\
\langle \Delta H_{RPS} \rangle &= \frac{2}{N^6} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} \left[k(j-i)(T^{RP} - T^{PR}) + j(k-i)(T^{RS} - T^{SR}) + i(j-k)(T^{SP} - T^{PS}) \right. \\
&\quad + k(T^{RP} + T^{PR}) + j(T^{RS} + T^{SR}) + i(T^{SP} + T^{PS}) \\
&\quad \left. + jk(T^{R+} - T^{+R}) + ik(T^{P+} - T^{+P}) + ij(T^{S+} - T^{+S}) \right] \tag{4}
\end{aligned}$$

Compared to the derivation of simply the RPS case in[?], this has the additional terms including the difference in transition probabilities in and out of the 4th strategy.

$$\begin{aligned}
\langle \Delta H_{RPS} \rangle &= \frac{2}{N^3} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[N^2 z(y-x)(T^{RP} - T^{PR}) + N^2 y(z-x)(T^{RS} - T^{SR}) \right. \\
&\quad + N^2 x(y-z)(T^{SP} - T^{PS}) \\
&\quad + Nz(T^{RP} + T^{PR}) + Ny(T^{RS} + T^{SR}) + Nx(T^{SP} + T^{PS}) \\
&\quad \left. + N^2 xz(T^{P+} - T^{+P}) + N^2 yz(T^{R+} - T^{+R}) + N^2 xy(T^{S+} - T^{+S}) \right] \\
\langle \Delta H_{RPS} \rangle &= \frac{2}{N} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[z(y-x)(T^{RP} - T^{PR}) + y(z-x)(T^{RS} - T^{SR}) \right. \\
&\quad + x(y-z)(T^{SP} - T^{PS}) + xz(T^{P+} - T^{+P}) + yz(T^{R+} - T^{+R}) + xy(T^{S+} - T^{+S}) \left. \right] \\
&\quad + \frac{2}{N^2} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[z(T^{RP} + T^{PR}) + y(T^{RS} + T^{SR}) + x(T^{SP} + T^{PS}) \right]
\end{aligned} \tag{5}$$

Rough figures of rps and SD delta H values.

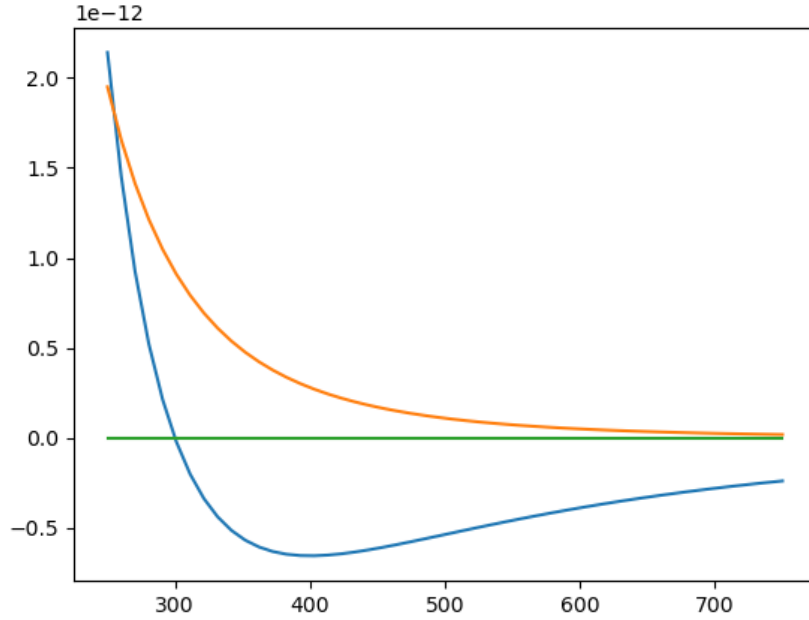


Figure 2: Blue - SD

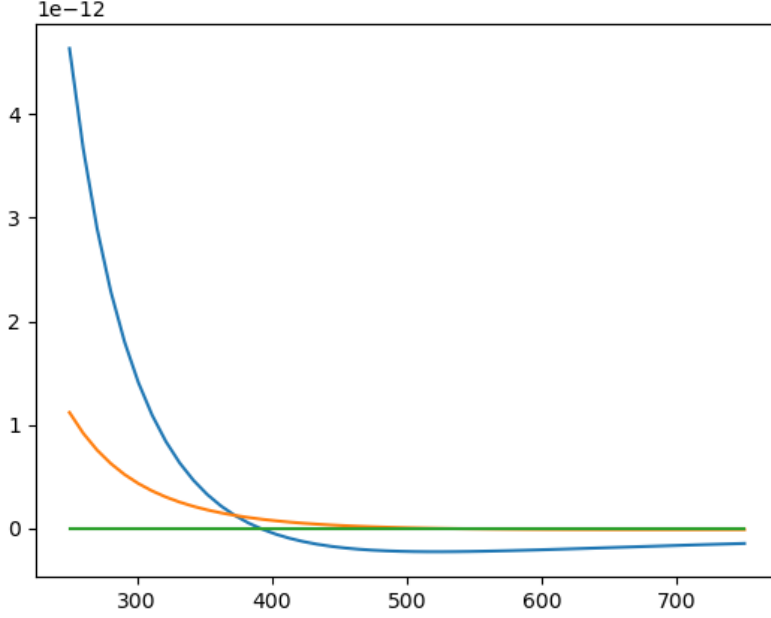


Figure 3: Double reversal case. Blue - SD

0.0.3 $\langle \Delta H_4 \rangle$ Derivation

All 4 strategies, $H = -xyz(1 - x - y - z)$, Therefore $\langle \Delta H_4 \rangle$ for the population counts can be defined as:

$$\begin{aligned}
 \langle \Delta H_4 \rangle = \frac{6}{N^7} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} & \left[ijk(N-i-j-k)(T^{RP} + T^{RS} + T^{R+} + T^{PR} + T^{PS} + T^{P+} \right. \\
 & + T^{SR} + T^{SP} + T^{S+} + T^{+R} + T^{+P} + T^{+S}) - (i-1)(j+1)k(N-i-j-k)T^{RP} \\
 & - (i-1)j(k+1)(N-i-j-k)T^{RS} - (i-1)jk(N-i-j-k+1)T^{R+} \\
 & - (i+1)(j-1)k(N-i-j-k)T^{PR} - i(j-1)(k+1)(N-i-j-k)T^{PS} \\
 & - i(j-1)k(N-i-j-k+1)T^{P+} - (i+1)j(k-1)(N-i-j-k)T^{SR} \\
 & - i(j+1)(k-1)(N-i-j-k)T^{SP} - ij(k-1)(N-i-j-k+1)T^{S+} \\
 & \left. - (i+1)jk(N-1-i-j-k)T^{+R} - i(j+1)k(N-1-i-j-k)T^{+P} - ij(k+1)(N-1-i-j-k)T^{+S} \right] \quad (6)
 \end{aligned}$$

Normalization $\frac{6}{N^7}$, as its over the whole simplex (pyramid volume 1/6) and triple summation, (N^3) , then 4 populations N^4 .

Let $p = N - i - j - k$ (4th strategy count):

$$\begin{aligned}
\langle \Delta H_4 \rangle = \frac{6}{N^7} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} & \left[ijkp(T^{RP} + T^{RS} + T^{R+} + T^{PR} + T^{PS} + T^{P+} \right. \\
& + T^{SR} + T^{SP} + T^{S+} + T^{+R} + T^{+P} + T^{+S}) - (i-1)(j+1)kpT^{RP} \\
& - (i-1)j(k+1)pT^{RS} - (i-1)jk(p+1)T^{R+} \\
& - (i+1)(j-1)kpT^{PR} - i(j-1)(k+1)pT^{PS} \\
& - i(j-1)k(p+1)T^{P+} - (i+1)j(k-1)pT^{SR} \\
& - i(j+1)(k-1)pT^{SP} - ij(k-1)(p+1)T^{S+} \\
& \left. - (i+1)jk(p-1)T^{+R} - i(j+1)k(p-1)T^{+P} - ij(k+1)(p-1)T^{+S} \right]
\end{aligned} \tag{7}$$

Looking at one pair of transitions in and out of the same state. ($R \rightarrow P, P \rightarrow R$) this can be rewritten as before for $\langle \Delta H_{RPS} \rangle$ as differences of transition probabilities in and out of a state.

$$\begin{aligned}
\langle \Delta H_4 \rangle = \frac{6}{N^7} \sum_{i=1}^N \sum_{j=1}^{N-i} \sum_{k=1}^{N-i-j} & \left[kp(j-i)(T^{RP} - T^{PR}) + jp(k-i)(T^{RS} - T^{SR}) \right. \\
& + ip(j-k)(T^{SP} - T^{PS}) + kp(T^{RP} + T^{PR}) + jp(T^{RS} + T^{SR}) \\
& + ip(T^{SP} + T^{PS}) + jk(p-i)(T^{R+} - T^{+R}) + ik(p-j)(T^{P+} - T^{+P}) \\
& \left. + ij(p-k)(T^{S+} - T^{+S}) + jk(T^{R+} + T^{+R}) + ik(T^{P+} + T^{+P}) + ij(T^{S+} + T^{+S}) \right]
\end{aligned} \tag{8}$$

As before, converting this to the integral form where $x = i/N, y = j/N, z = k/N$ and $q = p/N, p = Nq$ where $q = 1 - x - y - z$ leads to:

$$\begin{aligned}
\langle \Delta H_4 \rangle = \frac{6}{N^4} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz & \left[N^3 zq(y-x)(T^{RP} - T^{PR}) - N^3 yq(z-x)(T^{RS} - T^{SR}) \right. \\
& + N^3 xq(y-z)(T^{SP} - T^{PS}) + N^2 zq(T^{RP} + T^{PR}) + N^2 yq(T^{RS} + T^{SR}) + N^2 xq(T^{SP} + T^{PS}) \\
& + N^3 yz(q-x)(T^{R+} - T^{+R}) + N^3 xz(q-y)(T^{P+} - T^{+P}) + N^3 xy(q-z)(T^{S+} - T^{+S}) \\
& \left. + N^2 yz(T^{R+} + T^{+R}) + N^2 xz(T^{P+} + T^{+P}) + N^2 xy(T^{S+} + T^{+S}) \right]
\end{aligned} \tag{9}$$

$$\begin{aligned}
\langle \Delta H_4 \rangle = \frac{6}{N} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz & \left[zq(y-x)(T^{RP} - T^{PR}) - yq(z-x)(T^{RS} - T^{SR}) + xq(y-z)(T^{SP} - T^{PS}) \right. \\
& + yz(q-x)(T^{R+} - T^{+R}) + xz(q-y)(T^{P+} - T^{+P}) + xy(q-z)(T^{S+} - T^{+S}) \Big] \\
& + \frac{6}{N^2} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \left[zq(T^{RP} + T^{PR}) + yq(T^{RS} + T^{SR}) + xq(T^{SP} + T^{PS}) \right. \\
& \left. + yz(T^{R+} + T^{+R}) + xz(T^{P+} + T^{+P}) + xy(T^{S+} + T^{+S}) \right]
\end{aligned} \tag{10}$$

Final derivation of $\langle \Delta H_4 \rangle$. There are clear similarities to the other two expected values $\langle \Delta H_{SD} \rangle, \langle \Delta_{RPS} \rangle$ with some terms missing due to the different observation variable H used but they can all be rewritten in terms of differences of transition probabilities into and out of state (e.g. $(T^{RP} - T^{PR})$) in an integral scaled by $1/N$ plus positive terms in a second integral scaled by $1/N^2$.

0.0.4 Graphical comparison of H

Compare the 3 together on same plot for some of the interesting cases.