Scrambling in the Spin-Glass Evolution Simulation

August 6, 2024

Observe the "scrambling" effect in the simulation. Say we define some initial time t_0 and time variable t that starts from t_0. Remember $\vec{\alpha}$ is the binary genome vector, $\hat{J}$ is the epistatic matrix.

$$F(\vec{\alpha}) = \sum_{i=1}^{L} h_i \alpha_i + \sum_{i,j=1}^{L} \alpha_i J_{ij} \alpha_j$$

And the fitness effect of flipping $\alpha_i \to -\alpha_i$:

$$\Delta_i = -2\alpha_i \left(h_i + \sum_{j=1}^{L} J_{ij} \alpha_j \right)$$

Define the local field f_i:

$$f_i := \sum_{j=1}^{L} J_{ij} \alpha_j$$

$$F(\vec{\alpha}) = \sum_{i=1}^{L} (h_i + f_i) \alpha_i$$

$$\Delta_i = -2\alpha_i \left(h_i + f_i \right) \approx -2\alpha_i f_i$$

Where we disregard the h_i for this effect, as we assume it happens for $\beta \approx 1$. Define "Forward DFE" or "Forward propagation" the act of taking the beneficial DFE (BDFE) at time t_0, and observing the genes responsible for this distribution (these would be the genes that have $\alpha_i f_i < 0$). Propagate the system to time t, and plot the distribution of fitness effects of these same genes. Are they still all beneficial, or perhaps the distribution changes? The

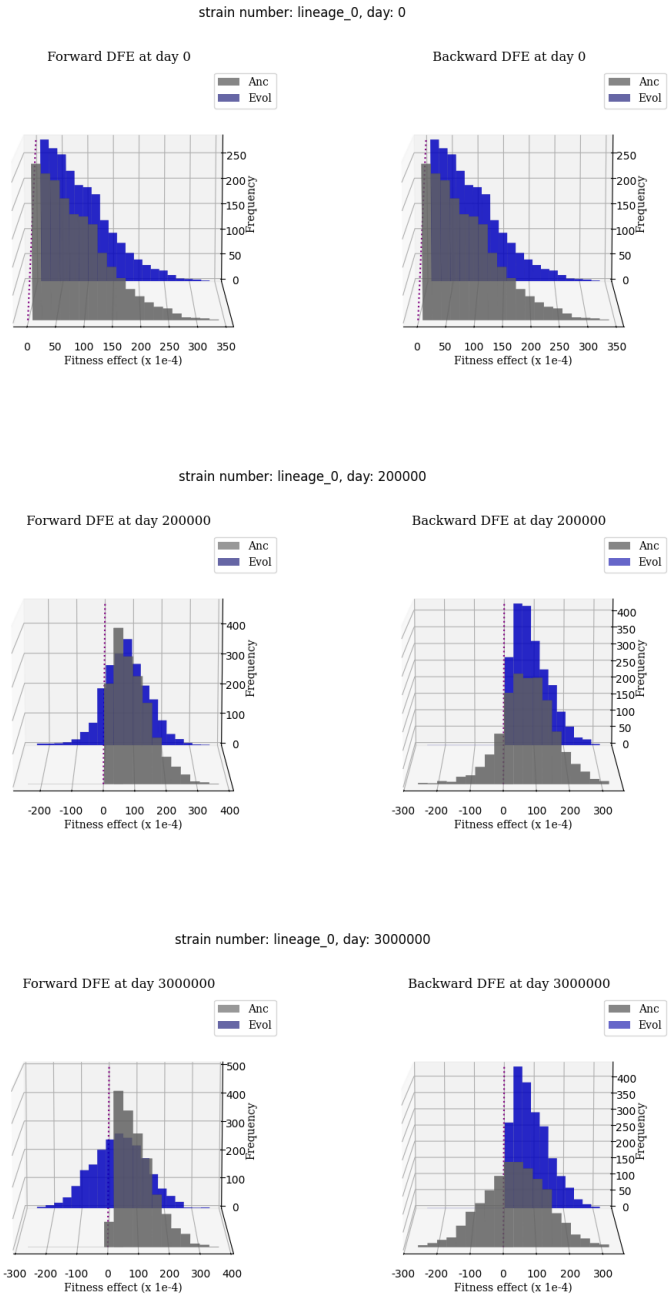

Figure 1: Forward and backward propagations for $t_0 = 0$, $t = [0, 2 \cdot 10^5, 3 \cdot 10^6]$. We can observe the slow evolution of the propagated DFEs into what seems to be a gaussian.

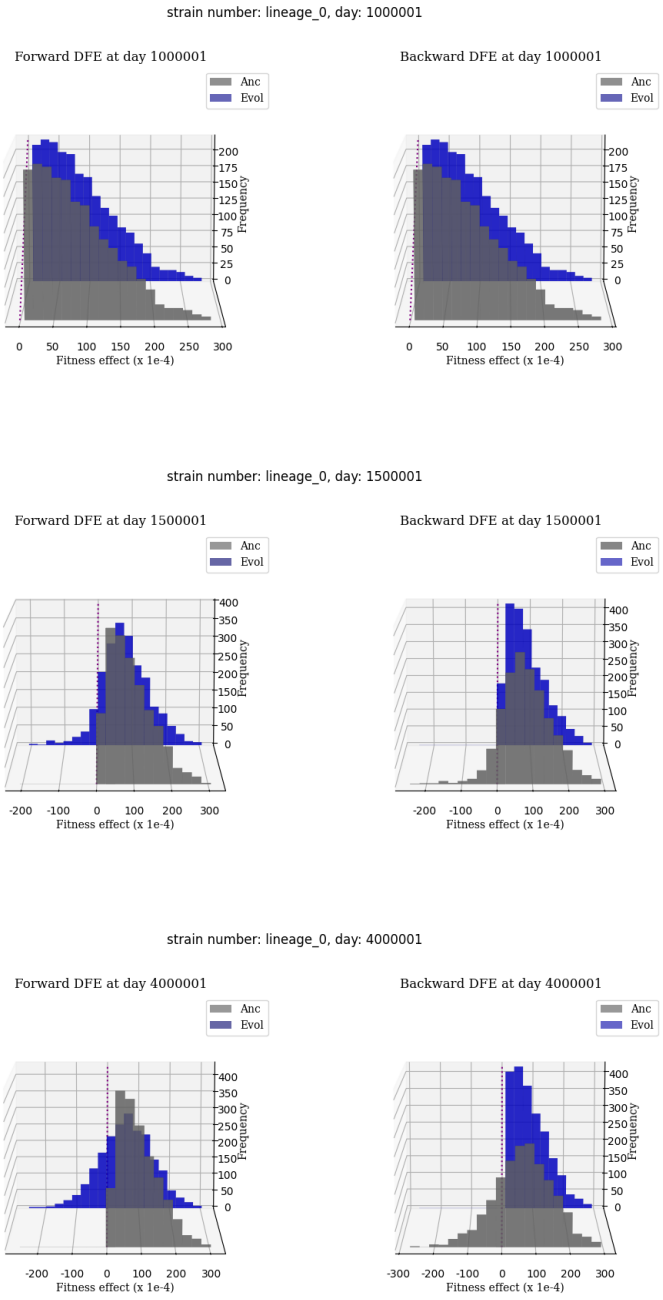

Figure 2: Forward and backward propagations for $t_0 = 1+10^6$, $t = [1+10^6, 1+15{\cdot}10^5, 1+4{\cdot}10^6]$. Timescale is larger because of diminishing magnitude of fixated fitness effects, leading to larger fixation times.

same can be defined as "Backward DFE" or "Backward propagation". Take the BDFE at time t and propagate the system back to time t_0, and observe the distribution of fitness effects of thees genes at t_0. We plot the propagations for $t_0 = 0$ (1) and for $t_0 = 1 + 10^6$ (2).

So, it seems the value of t_0 doesn't change this effect. We shall immediately attempt to analyze this effect, but beforehand we must state the obvious - It must be that this Gaussian is a transitionary effect, and such is the one observed in Bayms paper as well. I say this because it is obvious that the BDFE loses its meaning/nature when we approach a local fitness peak and there are almost no beneficial distribution effects left. Thus, we must think of this creature in the right regimes of existence. Now, to analyze:

$$\Delta_i = -2\alpha_i \left(h_i + f_i\right) \approx -2\alpha_i f_i$$

Define $N_+(t)$ as the number of $\alpha_i = +1$ at time t and $N_-(t) = L - N_+(t)$ the number of $\alpha_i = -1$. Δ_i are the fitness effects. If we look at the BDFE, we are **conditioning on** i **s.t** $\Delta_i > 0$, i.e. $\alpha_i f_i < 0$. Observe that in general and regardless of time there exists a symmetry of J_{ij}, h_i in the system around 0, thus we assume $N_+(t) \approx N_-(t)$, regardless of t. The only thing that changes is that α_i flip to match their corresponding f_i , and $+ \to -$ as much as $- \to +$. But, when we condition on $\Delta_i > 0$ we choose the i s.t their f_i are the ones that this symmetry is broken in a way that $f_i \neq 0$, and in particular $sign(f_i) \neq sign(\alpha_i)$. For example, if $\alpha_i < 0$ and the fitness effect of gene i is beneficial, this means $f_i = \sum_{j=1}^{L} J_{ij}\alpha_j > 0$, and in particular, considering $N_+(t) \approx N_-(t)$, it must be that specifically for row i we have and/or:

1. We have an untypical amount of J_{ij} that have the same sign as $\alpha_i\alpha_j$. Name this "quantity skew". We will use the name later.
2. We have that $\alpha_i\alpha_j$ are the same sign as J_{ij} for the largest values of J_{ij}. Name this "quality skew".

Now we assume that evolving in time randomly flips α_j , at least as far as f_i are concerned. For $f_i > 0$ we had for the above example more terms of $J_{ij}\alpha_j > 0$ than $J_{ij}\alpha_j < 0$. In this case, flipping random α_j more likely "evens out" the sum, such that there roughly the same number of $J_{ij}\alpha_j > 0$ as $J_{ij}\alpha_j < 0$. Observe that now, for large $L\rho$ and an unbiased sum:

$$f_i := \sum_{j=1}^{L} J_{ij}\alpha_j \approx \sum_{j=1}^{L} J_{ij} \to X_i \sim N(0, \sigma_J\sqrt{L\rho})$$

Where we assumed that for the now "unbiased" sum, randomly flipping the signs of J_{ij} by multiplying with α_j does not change the result, because J_{ij} are drawn from a normal distribution. So, we see that via the CLT, f_i becomes a normally distributed RV. Now, because we have:

$$\Delta_i \approx -2\alpha_i f_i$$

The other terms can, at most, flip the sign of f_i. So, we see that the DFE is the distribution of f_i who are perhaps flipped, which does not change the distribution as they are normally distributed RVs. This can perhaps explain the "scrambling" effect we observe in the simulation, and that which is observed in Bayms paper.

To Emphasize this point I did the following simulation:

1. Save the set of genes that have beneficial fitness effects at t_0, call them the set G_0
2. Define $N_+^i$the number of positive elements in the sum $\sum_{j=1}^{L} -2\alpha_i J_{ij}\alpha_j = -2\alpha_i f_i = \Delta_i$, which basically quantifies the "quantity skew" as defined above.
3. For a general i we we expect $\frac{N_+^i}{\rho L} \approx 0.5$, but for $i \in G_0$ we expect more!
4. Now plot the averaged $\frac{N_+^i}{\rho L} - \frac{1}{2}$ for $i \in G_0$ (averaged over G_0) to quantify this "quantity skewness" over time.
5. Do it for different t_0.

We plot exactly this for 2 different clones, the first from $t_0 = 0$ to $t = 1 \cdot 10^6$ (3) and the second from $t_0 = 1 \cdot 10^6$ to $t = 4 \cdot 10^6$ (4) (same times as plotted above). We can observe what seems to be the same kind of (exponential?) decay for the number of untypically aligned members regardless of t_0!

So, to summarize:

Observing the BDFE, we condition on the fact that Δ_i, the fitness effects of mutation in these genes, will be > 0, and this is true regardless of t_0 because this is the definition of the BDFE. Simplified, this means that for given gene i, we have an untypical amount(defined as N_+) of $J_{ij}\alpha_j$ such that $sign(\alpha_i) = sign(J_{ij}\alpha_j)$. This amount is typically just $1/2$, but not for $i \in G_0$. During evolution, α_j flip randomly(at least randomly from the POV of row i of J_{ij}), and $N_+/\rho L \to 1/2$. In other words, mutations "randomly" become beneficial, but their autocorrelation function decays in time.

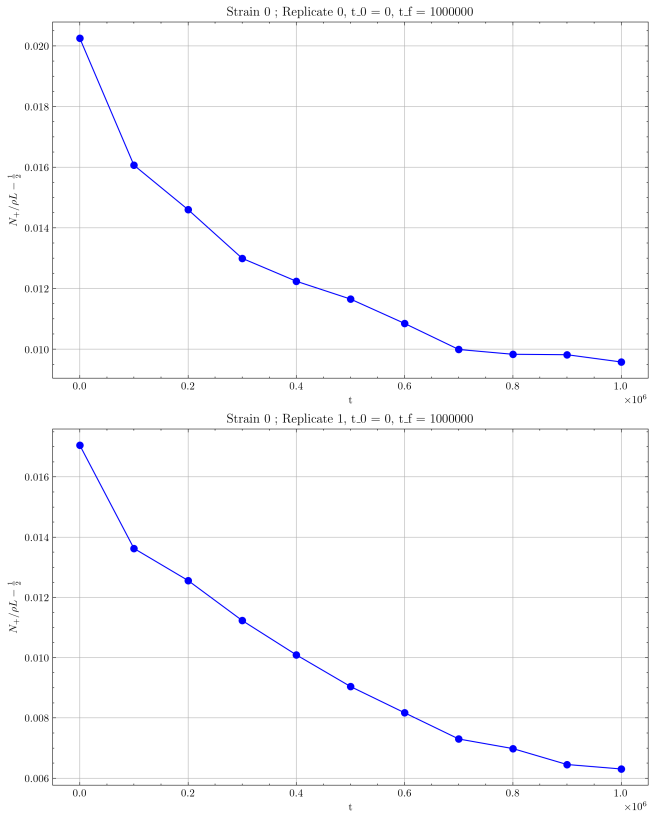

Figure 3: We plot $N_+/\rho L - 1/2$ for 2 different clone evolutions, from $t_0 = 0$ to $t = 10^6$. Decay can be observed.

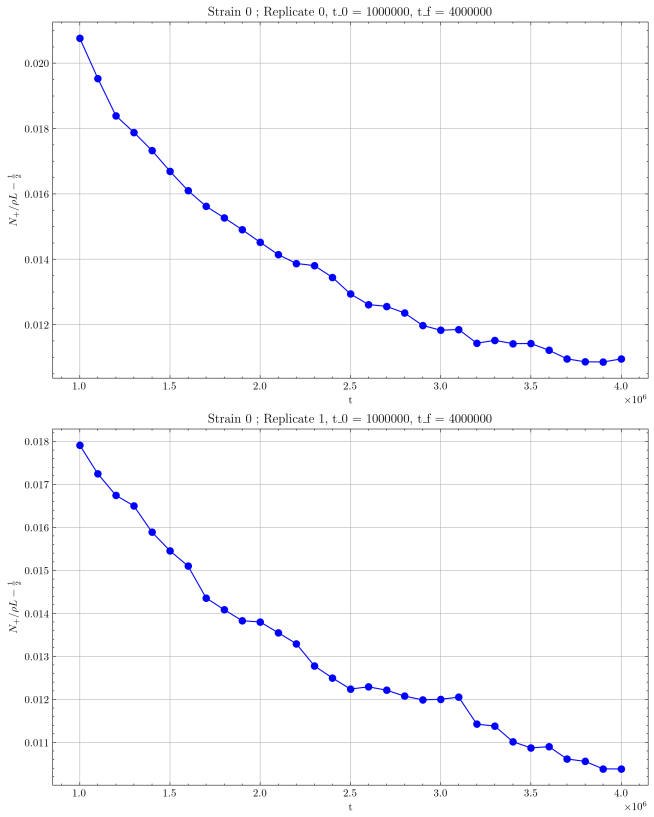

Figure 4: We plot $N_+/\rho L - 1/2$ for 2 different clone evolutions, from $t_0 = 1 \cdot 10^6$ to $t = 4 \cdot 10^6$. Decay can be observed.