Supplementary Material for
"A rate-distortion scenario for the emergence and evolution of noisy molecular codes"

Tsvi Tlusty
Department of Physics of Complex Systems, Weizmann Institute of Science

Contents

A. The Bayes theorem

In principle, to find the optimal code one needs to optimize the fitness function H with respect to all its $2 \times m \times s$ degrees of freedom, the $m \times s$ entries of the encoder $e_{\alpha i}$ and the $m \times s$ entries of the decoder $d_{j\omega}$, taking into account the m conservation relations for the encoder $\sum_i e_{\alpha i} = 1$ and the s conservation relations for the decoder $\sum_\omega d_{j\omega} = 1$. However, although the degrees of freedom of the encoder and the decoder are independent, most encoder-decoder combinations would yield high distortion and therefore an improbable code. Instead, it proves convenient to sum only over the most probable codes by considering for each encoder $e_{\alpha i}$ only the most probable decoder $d_{j\omega}$ (This is somewhat similar to a saddle-point approximation). An optimal decoder must take into account the knowledge about the encoder, that is if the encoder $e_{\alpha i}$ tends to encode the meaning ω as the symbol j then it is likely that the decoder $d_{j\omega}$ will interpret j as ω. The mathematical manifestation of this intuitive observation is through the Bayes theorem, $P(j \to \omega) = P(j)P(j|\omega) = P(\omega)P(j|\omega) = P(\omega \to j)$. For the present channel, Bayes theorem can be expressed as

$$p_j d_{j\omega} = f_\omega P(j|\omega), \tag{1}$$

where $P(j|\omega) = \sum_i e_{\omega i} r_{ij}$ sums all the possible paths to encode the meaning ω and to read it as the symbol j, and $p_j = \sum_{\alpha,i} f_\alpha e_{\alpha i} r_{ij} = \sum_i u_i r_{ij}$ sums all the probabilities to read the symbol j, starting from any meaning. Eq. 1 gives the optimal decoder as a function of the reader and the optimal encoder [1, 2],

$$d_{j\omega} = \frac{f_\omega \sum_i e_{\omega i} r_{ij}}{\sum_{\alpha,i} f_\alpha e_{\alpha i} r_{ij}}. \tag{2}$$

B. The optimal code

To calculate the optimal encoder (Eq. 4) we use the following Lagrangian,

$$H_T = H + \sum_\alpha \mu_\alpha \sum_i e_{\alpha i} = -D - \kappa^{-1} I + \sum_\alpha \mu_\alpha \sum_i e_{\alpha i}, \tag{3}$$

which adds to the fitness H Lagrange multipliers for each of the m conservations $\sum_i e_{\alpha i} = 1$. The distortion D and the cost I in Eq. 3 are given by Eqs. 1-2,

$$D = \sum_{\alpha,i,j,\omega} f_\alpha e_{\alpha i} r_{ij} d_{j\omega} c_{\alpha\omega}, \tag{4}$$

$$I = \sum_{\alpha,i} f_\alpha e_{\alpha i} \ln \frac{e_{\alpha i}}{u_i},$$

where $u_i = \sum_\alpha f_\alpha e_{\alpha i}$. The extremum of H that obeys the constraints occurs at $\partial H_T / de_{\alpha i} = 0$. To calculate this extremum we need following derivatives,

$$\frac{\partial I}{\partial e_{\alpha i}} = f_\alpha \ln \frac{e_{\alpha i}}{u_i}, \tag{5}$$

$$\frac{\partial D}{\partial e_{\alpha i}} = f_\alpha \sum_{m,\gamma} r_{im} d_{m\gamma} c_{\alpha\gamma} + \sum_{\beta,k,m,\gamma} f_\beta e_{\beta k} r_{km} \frac{\partial d_{m\gamma}}{\partial e_{\alpha i}} c_{\beta\gamma},$$

where the derivative of the decoder (given by Eq. 2) is

$$\frac{\partial d_{j\omega}}{\partial e_{\alpha i}} = \frac{r_{ij}}{p_j} \left(f_\omega \delta_{\alpha\omega} - f_\alpha d_{j\omega} \right). \tag{6}$$

The extremum is located at

$$\frac{\partial H_T}{\partial e_{\alpha i}} = -\frac{\partial D}{\partial e_{\alpha i}} - \kappa^{-1} \frac{\partial I}{\partial e_{\alpha i}} + \mu_\alpha = 0. \tag{7}$$

Then, by substitution of Eqs. 4-6 into Eq. 7, while taking into account the Bayes relation (Eq. 2), we find that the optimal encoder is given by the Boltzmann-like distribution

$$e_{\alpha i} = u_i e^{\kappa(\mu_\alpha / f_\alpha - \Omega_{\alpha i})}, \tag{8}$$

with the effective distortion energies

$$\Omega_{\alpha i} = \sum_{j,\omega} r_{ij} d_{j\omega} \left(2c_{\alpha\omega} - \sum_\gamma d_{j\gamma} c_{\gamma\omega} \right). \tag{9}$$

The normalization constants $e^{\kappa\mu_\alpha / f_\alpha}$ are found by demanding $\sum_i e_{\alpha i} = 1$ for all α and, finally, we obtain Eq. 4,

$$e_{\alpha i} = \frac{u_i e^{-\kappa\Omega_{\alpha i}}}{\sum_j u_j e^{-\kappa\Omega_{\alpha j}}}. \tag{10}$$

Boltzmann partitions such as Eq. 10 (Eq. 4 in the text) are common in rate-distortion theory [2–9], where functionals similar to H are minimized to find an optimal channel. Expressions somewhat similar to Eqs. 9-10 have been derived in the context of deterministic annealing and clustering [2, 7, 9] by considering a canonical ensemble of binary encoders, $e_{\alpha i} \in \{0, 1\}$, and calculating the mean-field approximation for the maximum entropy [10–12] of such ensemble with a distortion that plays the role of a Hamiltonian.

The numerical solution of Eq. 4 is obtained by numerical iteration of Eqs. 2, 9 and 10, in the spirit of the Blahut-Arimoto algorithm [13, 14] (Matlab files are available by request). An example for such solution is shown in Fig. 2.

C. The critical coding transition

In this section we locate the critical coding/no-coding transition (Eq. 5) by examining the Hessian of the fitness H. At low values of the gain κ, the system is the non-coding state with $e^0_{\alpha i} = u_i$, where u_i can be any arbitrary probability distribution, $u_i \in [0, 1]$ and $\sum_i u_i = 1$. The non-coding decoder is obtained from the Bayes theorem (Eq. 2), $d^0_{j\omega} = f_\omega$. Since the encoder u_i is independent of the meanings α, it conveys no information and the cost therefore vanishes, $I_{nc} = 0$ (Fig. 2B). The only contribution to the fitness at the non-coding state comes from the distortion. This contribution is invariant with respect to changes in u_i and its invariant value is the average distance

$$H_{nc} = -D_{nc} = -\sum_{\alpha,\beta} f_\alpha f_\beta c_{\alpha\beta}. \tag{11}$$

To examine the stability of the fitness H, we expand it around non-coding state up to second order in the variation $\delta e_{\alpha i} = e_{\alpha i} - e^0_{\alpha i} = e_{\alpha i} - u_i$,

$$H \simeq H_{nc} - \frac{1}{2} \sum\nolimits_{\alpha,i,\omega,j} Q_{\alpha i \omega j} \delta e_{\alpha i} \delta e_{\omega j}, \tag{12}$$

where the Hessian tensor is $Q_{\alpha i\omega j} = -(\partial^2 H/\partial e_{\alpha i} \delta e_{\omega j})_{nc}$. To calculate the Hessian we need the following second derivatives

$$\begin{aligned} \frac{\partial^2 I}{\partial e_{\alpha i} \partial e_{\omega j}} &= f_\alpha \delta_{ij} \left(\frac{\delta_{\alpha\omega}}{e_{\alpha i}} - \frac{f_\omega}{u_i} \right), \\ \frac{\partial^2 D}{\partial e_{\alpha i} \partial e_{\omega j}} &= f_\alpha \sum_{m,\gamma} r_{im} \frac{\partial d_{m\gamma}}{\partial e_{\omega j}} c_{\alpha\gamma} + f_\omega \sum_{m,\gamma} r_{jm} \frac{\partial d_{m\gamma}}{\partial e_{\alpha i}} c_{\omega\gamma} \\ &\quad + \sum_{\beta,k,m,\gamma} f_\beta e_{\beta k} r_{km} \frac{\partial^2 d_{m\gamma}}{\partial e_{\alpha i} \partial e_{\omega j}} c_{\beta\gamma}. \end{aligned} \tag{13}$$

At the non-coding state, the first and the second derivatives of the decoder with respect to the encoder are

$$\begin{aligned} \left(\frac{\partial d_{k\gamma}}{\partial e_{\alpha i}} \right)_{nc} &= \frac{f_\gamma r_{ik}}{p_k} \left(\delta_{\alpha\gamma} - f_\alpha \right), \\ \left(\frac{\partial d_{k\gamma}}{\partial e_{\alpha i} \partial e_{\omega j}} \right)_{nc} &= \frac{r_{ik} r_{jk}}{p_k} \times \\ &\quad \left(2 f_\alpha f_\omega d_{k\gamma} - f_\alpha f_\gamma \delta_{\gamma\omega} - f_\omega f_\gamma \delta_{\gamma\alpha} \right), \end{aligned} \tag{14}$$

where $p_k = \sum_{\eta,m} f_\eta e_{\eta m} r_{mk}$. By substitution of Eq. 14 in Eq. 13 we find that the Hessian at the non-coding state is

$$\begin{aligned} Q_{\alpha i\omega j} &= -\left(\frac{\partial^2 H}{\partial e_{\alpha i} \partial e_{\omega j}} \right)_{nc} = \\ &= \kappa^{-1} \frac{\partial^2 I}{\partial e_{\alpha i} \partial e_{\omega j}} + \frac{\partial^2 D}{\partial e_{\alpha i} \partial e_{\omega j}} = \\ &\kappa^{-1} \frac{\delta_{ij}}{u_i} \times f_\alpha \left(\delta_{\alpha\omega} - f_\omega \right) + 2 \sum_k \frac{r_{ik} r_{jk}}{p_k} \times \\ &f_\alpha f_\omega \Big(\sum_\beta f_\beta c_{\beta\omega} + \sum_\gamma f_\gamma c_{\alpha\gamma} - \sum_{\beta,\gamma} f_\beta f_\gamma c_{\beta\gamma} - c_{\alpha\omega} \Big). \end{aligned} \tag{15}$$

For convenience, we scale the variation as $\delta \tilde{e}_{\alpha i} = \sqrt{f_\alpha / u_i} \delta e_{\alpha i}$. The scaled Hessian becomes

$$\begin{aligned} \tilde{Q}_{\alpha i\omega j} &= -\left(\frac{\partial^2 H}{\partial \tilde{e}_{\alpha i} \delta \tilde{e}_{\omega j}} \right)_{nc} = \\ &= \kappa^{-1} \delta_{ij} \left(\delta_{\alpha\omega} - \sqrt{f_\alpha f_\omega} \right) - 2 R_{ij} C_{\alpha\omega}, \end{aligned} \tag{16}$$

where $C_{\alpha\omega}$ is the normalized distance

$$\begin{aligned} C_{\alpha\omega} &= \sqrt{f_\alpha f_\omega} \times \\ &\left(\sum_\beta f_\beta c_{\beta\omega} + \sum_\gamma f_\gamma c_{\alpha\gamma} - \sum_{\beta,\gamma} f_\beta f_\gamma c_{\beta\gamma} - c_{\alpha\omega} \right), \end{aligned} \tag{17}$$

and R_{ij} is

$$R_{ij} = \sqrt{u_i u_j} \sum_k \left(\frac{r_{ik} r_{kj}}{\sum_t u_t r_{tk}} \right). \tag{18}$$

To test whether the Hessian is positive-definite, it proves convenient to express the vectors as a sum of the normalized eigenvectors of $\tilde{Q}$, $\delta\tilde{e} = \sum_n a_n \delta\tilde{e}^{(n)}$, where $\delta\tilde{e}^{(n)} \cdot \delta\tilde{e}^{(t)} = \delta_{nt}$. Then, the quadratic form is simply $\delta\tilde{e}^\dagger \tilde{Q}\, \delta\tilde{e} = \sum_n \lambda_n a_n^2$. The eigenvalues λ_n depend on the gain κ. It follows that the critical point occurs exactly at the lowest value of gain κ_c where one of the eigenvalues (which may be a degenerate one) becomes non-positive. It is evident from Eqs. 16-18 that the Hessian is a sum of two tensor products

$$\tilde{Q} = \kappa^{-1} I_s \times \left(I_m - \tilde{f} \right) - 2R \times C, \tag{19}$$

where I_s and I_m are the identity matrices in the spaces of symbols and meanings, respectively, and $\tilde{f}_{\alpha\omega} = \sqrt{f_\alpha f_\omega}$. An immediate result is that the eigenvectors of $\tilde{Q}$ are also tensor products of eigenvectors in the spaces of symbols and meanings,

$$\delta\tilde{e}_{\alpha i} = \delta\tilde{e}_{\alpha} \times \delta\tilde{e}_{i} = \sqrt{f_\alpha}\delta e_\alpha \times \sqrt{u_i}\delta e_i. \tag{20}$$

Since the variations must obey the conservation relations, $\sum_i \delta e_{\alpha i} = 0$, the symbol component of the eigenvector obeys

$$\sum_i \delta e_i = \sum_i \sqrt{u_i}\delta\tilde{e}_i = 0. \tag{21}$$

The vector $\delta\tilde{e}_i = \sqrt{u_i}$ is an eigenvector of R with the maximal eigenvalue $\lambda_R^{\max} = 1$, which corresponds to the uniform non-coding state. From Eq. 21 it follows that all the relevant eigenvectors are orthogonal to $\sqrt{u_i}$ in the symbol-space. Similarly, the vector $\delta\tilde{e}_\alpha = \sqrt{f_\alpha}$ is an eigenvector of the normalized distance C with the eigenvalue $\lambda_C = 0$. It is easy to see that $\sqrt{f_\alpha}$ is also an eigenvector of $\left(I_m - \tilde{f}\right)$ with the same $\lambda = 0$. Hence, all the relevant eigenvectors are orthogonal to $\sqrt{f_\alpha}$ in the meaning-space,

$$\sum_i \sqrt{f_\alpha}\delta\tilde{e}_\alpha = 0. \tag{22}$$

An immediate consequence of Eq. 22 is that $(I_m - \tilde{f})\delta\tilde{e}_\alpha = I_m\delta\tilde{e}_\alpha = \delta\tilde{e}_\alpha$. From all this we conclude that the relevant eigenvalues of $\tilde{Q}$ are

$$\lambda_{\tilde{Q}} = \kappa^{-1} - 2\lambda_R\lambda_C. \tag{23}$$

Now, it is evident from Eq. 23 that the first eigenvalue to turn negative is $\lambda^*_{\tilde{Q}} = \kappa^{-1} - 2\lambda^*_R\lambda^*_C$, where λ^*_C is the maximal eigenvalue of the normalized distance C, and λ^*_R is the *second-largest* eigenvalue of R. This occurs exactly at the critical gain (Eq. 5)

$$\frac{1}{\kappa_c} = 2\lambda^*_R\lambda^*_C. \tag{24}$$

Both λ^*_R and λ^*_C are non-negative by the symmetry of C and R, which ensures that κ_c is non-negative as well.

The reader is related to the graph-Laplacian Δ_{ij}, which describes random walk on the graph via misreading, $r_{ij} = \delta_{ij} - \Delta_{ij}$. When the non-coding state is uniform, $e_{\alpha i} = 1/s$, the critical gain is $\kappa_c = (2(1-\lambda^*_\Delta)^2\lambda^*_C)^{-1}$, where λ^*_Δ is the second-smallest eigenvalue of the graph-Laplacian.

D. The quasi-species model

In this section we derive the solution to the quasi-species equation (Eqs. 6-7). The space of all possible codes is an $m \times s-$dimensional unit cube cube $e_{\alpha i} \in [0,1]$, where each axis corresponds to an entry of the encoder matrix $e_{\alpha i}$. Each point in this cube is an $m \times s-$dimensional radius-vector $e_{\alpha i}$ that represents a possible code. The codes are constrained to reside within the intersection of this unit cube with m hyperplanes of dimension $(m-1) \times s$, which are defined by the m conservation relations $\sum_i e_{\alpha i} = 1$. The population is described in terms of a probability density $\Psi(e_{\alpha i})$ that an organism would have the code $e_{\alpha i}$. The quasi-species equation describes the dynamics of this probability density,

$$\frac{\partial\Psi}{\partial t} = \left[H(e_{\alpha i}) - \bar{H}\right]\Psi + \mu\sum_{\alpha,i}\frac{\partial^2\Psi}{\partial e^2_{\alpha i}}. \tag{25}$$

Eq. 25 is actually a linear reaction-diffusion equation The reaction term $\left[H(e_{\alpha i}) - \bar{H}\right]\Psi$ accounts for exponential growth at a rate which is equal to the fitness of the code $H(e_{\alpha i})$ (Here H is the Malthusian fitness). This growth rate is H is normalized by the average fitness $\bar{H} = \int\Psi(e_{\alpha i})H(e_{\alpha i})de_{\alpha i}$ to conserve the probability distribution $\int\Psi(e_{\alpha i})de_{\alpha i} = 1$. The diffusion term $\mu\sum_{\alpha,i}\partial^2\Psi/\partial e^2_{\alpha i}$ accounts for mutations that change the code. This continuous approximation to the random walk via mutations assumes that changes in the code are relatively small $\delta e_{\alpha i} \ll 1$. This approximation is plausible for molecular codes where the typical binding sites include a dozen or so of DNA bases or amino-acids. In such code, a point mutation in the binding site will change the binding energies $\varepsilon_{\alpha i}$ and the binding probabilities only by little.

Typically, the probability distribution Ψ approaches asymptotically a steady-state $\partial\Psi/\partial t = 0$, which represents a population of constant composition $\Psi(e_{\alpha i})$, or a "quasi-species", that grows as a whole at a rate $\bar{H}$ [15]. The growth rate is the maximal eigenvalue of the time-independent solution of Eq. 25. To find this steady-state we approximate the fitness in the vicinity of an optimum H_* by a quadratic expansion in the variation $\delta e_{\alpha i}$,

$$H \simeq H_* - \frac{1}{2}\sum\nolimits_{\alpha,i,\omega,j} Q_{\alpha i\omega j}\delta e_{\alpha i}\delta e_{\omega j}, \tag{26}$$

where the Hessian tensor is $Q_{\alpha i\omega j} = -(\partial^2 H/\delta e_{\alpha i}\delta e_{\omega j})_*$. Next, we assume the following Gaussian ansatz for Ψ,

$$\Psi = A\prod_\alpha\delta(\sum_i\delta e_{\alpha i})\exp(-\frac{1}{2}\sum_{\alpha,i,\omega,j} b_{\alpha i\omega j}\delta e_{\alpha i}\delta e_{\omega j}), \tag{27}$$

where the $\delta-$functions take care of the conservation relations and A is a normalization constant. To use standard Gaussian methods, it is convenient to write the $\delta-$functions as $\delta(x) = \lim_{\gamma\to\infty}(\gamma/2\pi)^{1/2}e^{-\gamma x^2/2}$. Then Eq. 27 can be rewritten as

$$\Psi = \lim_{\gamma\to\infty}(2\pi)^{-m\cdot s/2}\sqrt{\det\tilde{b}}\times \tag{28}$$

$$\exp(-\frac{1}{2}\sum_{\alpha,i,\omega,j}\tilde{b}_{\alpha i\omega j}\delta e_{\alpha i}\delta e_{\omega j}),$$

where $\tilde{b}_{\alpha i\omega j} = b_{\alpha i\omega j} + \gamma\delta_{\alpha\omega}$. Substitution of Eq. 28 in the time-independent quasi-species equation (Eq. 25) yields

$$\frac{1}{2}\sum_{\alpha,i,\omega,j} Q_{\alpha i\omega j}\left(\langle\delta e_{\alpha i}\delta e_{\omega j}\rangle - \delta e_{\alpha i}\delta e_{\omega j}\right) + \tag{29}$$

$$+\mu \lim_{\gamma\to\infty}\left(\sum_{\alpha,i}\tilde{b}_{\alpha i\alpha i} - (\sum_{\omega,j}\tilde{b}_{\alpha i\omega j}\delta e_{\omega j})^2\right) = 0,$$

where $\langle\delta e_{\alpha i}\delta e_{\omega j}\rangle$ is the population average $\langle\delta e_{\alpha i}\delta e_{\omega j}\rangle = \int \delta e_{\alpha i}\delta e_{\omega j}\Psi de_{\alpha i}$. This average is calculated by

$$\langle\delta e_{\alpha i}\delta e_{\omega j}\rangle = -2\lim_{\gamma\to\infty}(2\pi)^{-m\cdot s/2}\sqrt{\det\tilde{b}}\times \tag{30}$$

$$\frac{\partial}{\partial\tilde{b}_{\alpha i\omega j}}\int\exp(-\frac{1}{2}\sum_{\alpha,i,\omega,j}\tilde{b}_{\alpha i\omega j}\delta e_{\alpha i}\delta e_{\omega j})de_{\alpha i} =$$

$$= \lim_{\gamma\to\infty}\frac{\partial\log(\det\tilde{b})}{\partial\tilde{b}_{\alpha i\omega j}} = \lim_{\gamma\to\infty}\tilde{b}^{-1}_{\alpha i\omega j}.$$

By substitution of Eq. 30 into Eq. 29 we find that

$$b = \frac{\sqrt{Q}}{\sqrt{2\mu}}, \tag{31}$$

$\sqrt{Q}$ is the square root of the Hessian $Q_{\alpha i\omega j}$. The normalization constant is found by equating the prefactors of the Gaussians in Eqs. 27-28,

$$A = (2\pi)^{-(m-1)\cdot s/2}\lim_{\gamma\to\infty}\gamma^{-m/2}\sqrt{\det\tilde{b}} = \tag{32}$$

$$= (2\pi)^{-(m-1)\cdot s/2}\left(\prod\lambda_Q\right)^{1/4},$$

where the product is over $m\cdot(s-1)$ of the eigenvalues of Q excluding the m eigenvalues that correspond to uniform eigenvector in the symbol-space. To conclude, from Eqs. 31-32 it follows that the steady-state distribution (Eq. 7) is

$$\Psi = A\prod_{\alpha}\delta(\sum_{i}\delta e_{\alpha i})\times \tag{33}$$

$$\exp\left(-(8\mu)^{-1/2}\sum_{\alpha,i,\omega,j}\sqrt{Q}_{\alpha i\omega j}\delta e_{\alpha i}\delta e_{\omega j}\right).$$

Next, we calculate the reduction in the fitness (i.e. growth rate) of the quasi-species due to the diffusive widening of the Gaussian peak. We find that the leakage by mutations from the optimal code to lesser codes reduces the average fitness by an amount proportional to $\mu^{1/2}$,

$$\bar{H} = H_* - \frac{1}{2}\sum\nolimits_{\alpha,i,\omega,j}Q_{\alpha i\omega j}\langle\delta e_{\alpha i}\delta e_{\omega j}\rangle = \tag{34}$$

$$= H_* - (\mu/2)^{1/2}\sum_{\alpha,i}\sqrt{Q}_{\alpha i\alpha i}$$

$$= H_* - (\mu/2)^{1/2}\mathrm{Tr}\sqrt{Q}$$

$$= H_* - (\mu/2)^{1/2}\sum_{\alpha,i}\sqrt{\lambda_Q},$$

where the sum is over the square roots of all the eigenvalues of Q. At the coding transition (Eq. 24), one or more of the eigenvalues λ_Q vanish, and the Gaussian distribution Ψ becomes infinitely wide in the direction of the emergent coding eigenvector $\delta e^*_{\alpha i}$.

E. Effects of genetic drift

To include the stochastic effects of genetic drift we consider a family of models, in which the typical dynamics exhibits long periods of time when the population resides in the neighborhood of fitness maxima separated by relatively fast transitions by genetic drift between the maxima. Neglecting the fine details of the diffusive transitions, we can regard the dynamics as instantaneous random transitions between fitness maxima, which play the role of energy levels. This type of dynamics has been analyzed in detail in [16] and in the context of TRNs in [17]. The common mathematical property to this family of models is that the ratio of the transition rates forward ω_{12} and backward ω_{21} between two maxima, H_1 and H_2, depends only on the ratio of a fitness "potential" $U(H)$ at these two points, $\omega_{12}/\omega_{21} = U(H_1)/U(H_2)$. One such example is Kimura's process [18], where the transition rate is

$$\omega_{12} = n\mu\frac{1-e^{-2(H_2-H_1)}}{1-e^{-2n(H_2-H_1)}}, \tag{35}$$

with the population size n and the mutation rate μ between states 1 and 2. For this model, $\omega_{12}/\omega_{21} = e^{2(n-1)(H_1-H_2)}$, with a Boltzmann-like potential $U(H) = e^{2(n-1)H}$. Then, if the system is ergodic, i.e. there is a mutation path between any two states, it reaches an asymptotic equilibrium state, in which the population is partitioned according to a Boltzmann exponential $\Psi \sim e^{2(n-1)H}$. In this distribution the fitness H plays the role of (minus) the Hamiltonian and population size n is the "inverse temperature" $\beta = 1/T$ (up to a prefactor that depends on the specific model) [16]. One can then formalize an equivalent of a free-energy,

$$F = \langle -H\rangle - n^{-1}S = -\int\Psi(e_{\alpha i})H(e_{\alpha i})de_{\alpha i} \tag{36}$$

$$+ n^{-1}\int\Psi(e_{\alpha i})\ln\Psi(e_{\alpha i})de_{\alpha i},$$

where the entropy $S = -\int\Psi\ln\Psi de_{\alpha i}$ measures the randomness due to genetic drift.

A mean-field estimate gives an expression that is practically equivalent to the free energy of a Potts model (e.g. [19]). The "energy" term is (minus) the fitness at the average code, $\bar{e}_{\alpha i} = \int e_{\alpha i}\Psi(e_{\alpha i})de_{\alpha i}$. The entropy S is due to the usual $\sum_i\phi_i\ln\phi_i$ mixing term, where the fractions are $\phi_i = f_\alpha\bar{e}_{\alpha i}$ and there are m such sums, one for each meaning α (Eq. 8),

$$F = -H(\bar{e}_{\alpha i}) + n^{-1}\sum_{\alpha,i}f_\alpha\bar{e}_{\alpha i}\ln f_\alpha\bar{e}_{\alpha i}, \tag{37}$$

We can now calculate the effect of the genetic drift entropy S on the coding transition. As in section C we examine the stability of the free energy F by examining its Hessian,

$$K_{\alpha i\omega j}{=}\left(\frac{\partial^2F}{\partial e_{\alpha i}\delta e_{\omega j}}\right)_{nc} = -\left(\frac{\partial^2H}{\partial e_{\alpha i}\delta e_{\omega j}}\right)_{nc} \tag{38}$$

$$+n^{-1}\left(\frac{\partial^2S}{\partial e_{\alpha i}\delta e_{\omega j}}\right)_{nc} = Q_{\alpha i\omega j} + n^{-1}\frac{f_\alpha}{u_i}\delta_{\alpha\omega}\delta_{ij},$$

where $Q_{\alpha i\omega j}$ is given by Eq. 15. Again, we transform the Hessian into scaled coordinates $\delta\tilde{e}_{\alpha i} = \sqrt{f_\alpha/u_i}\delta e_{\alpha i}$ and obtain

$$\tilde{K}_{\alpha i\omega j} = \left(\frac{\partial^2 F}{\partial\tilde{e}_{\alpha i}\delta\tilde{e}_{\omega j}}\right)_{nc} = \tilde{Q}_{\alpha i\omega j} + n^{-1}\delta_{\alpha\omega}\delta_{ij}, \quad (39)$$

where $\tilde{Q}_{\alpha i\omega j}$ is given by Eq. 16. Since the additional term $n^{-1}\delta_{\alpha\omega}\delta_{ij}$ is diagonal, we can use the same procedure that follows Eq. 19, and easily find the criticality condition,

$$\frac{1}{\kappa_c} + \frac{1}{n_c} = 2\lambda_R^*\lambda_C^*. \quad (40)$$

This equation indicates four possible pathways towards the coding transition: (i) increasing the gain k (ii) increasing the population n (iii) increasing the accuracy of the reader (increasing λ_R^*) and (iv) increasing the distance (which amounts to increasing λ_C^*).

[1] S. P. Luttrell, Neural Comput. **6**, 767 (1994).
[2] T. Graepel, M. Burger, and K. Obermayer, Neurocomputing **21**, 173 (1998).
[3] T. Berger, *Rate distortion theory* (Prentice-Hall, NJ, 1971).
[4] T. M. Cover and J. A. Thomas, *Elements of information theory* (Wiley-Interscience, Hoboken, NJ, 2006), 2nd ed.
[5] K. Rose, E. Gurewitz, and G. C. Fox, Phys. Rev. Lett. **65**, 945 (1990).
[6] K. Rose, Proc IEEE **86**, 2210 (1998).
[7] T. Graepel, M. Burger, and K. Obermayer, Physical Review E **56**, 3876 (1997).
[8] N. Tishby, F. Pereira, and W. Bialek, in *The 37-th Allerton Conf. on Communication, Control and Computing* (1999), pp. 368–377.
[9] T. Hofmann and J. M. Buhmann, IEEE Trans. Patt. Anal. Machine Intell. **19**, 1 (1997).
[10] E. T. Jaynes, Physical Review **106**, 620 (1957).
[11] E. T. Jaynes, Physical Review **108**, 171 (1957).
[12] E. T. Jaynes, Proc IEEE **70**, 939 (1982).
[13] R. E. Blahut, Ieee T Inform Theory **It18**, 460 (1972).
[14] S. Arimoto, Ieee T Inform Theory **It18**, 14 (1972).
[15] M. Eigen, J. McCaskill, and P. Schuster, Adv. Chem. Phys. **75**, 149 (1989).
[16] G. Sella and A. E. Hirsh, Proc. Nat. Acad. Sci. U.S.A. **102**, 9541 (2005).
[17] J. Berg, S. Willmann, and M. Lassig, BMC Evol Bio **4**, (2004).
[18] J. F. Crow and M. Kimura, *An introduction to population genetics theory* (Harper and Row, New York, 1970).
[19] F. Y. Wu, Rev Mod Phys **54**, 235 (1982).