Supplementary information for the article "How evolution draws trade-offs"

Salomé Bourg *[1], Laurent Jacob[1], Frédéric Menu[1], and Etienne Rajon[1]

[1]Univ Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR5558, F-69622 Villeurbanne, France

SI figures

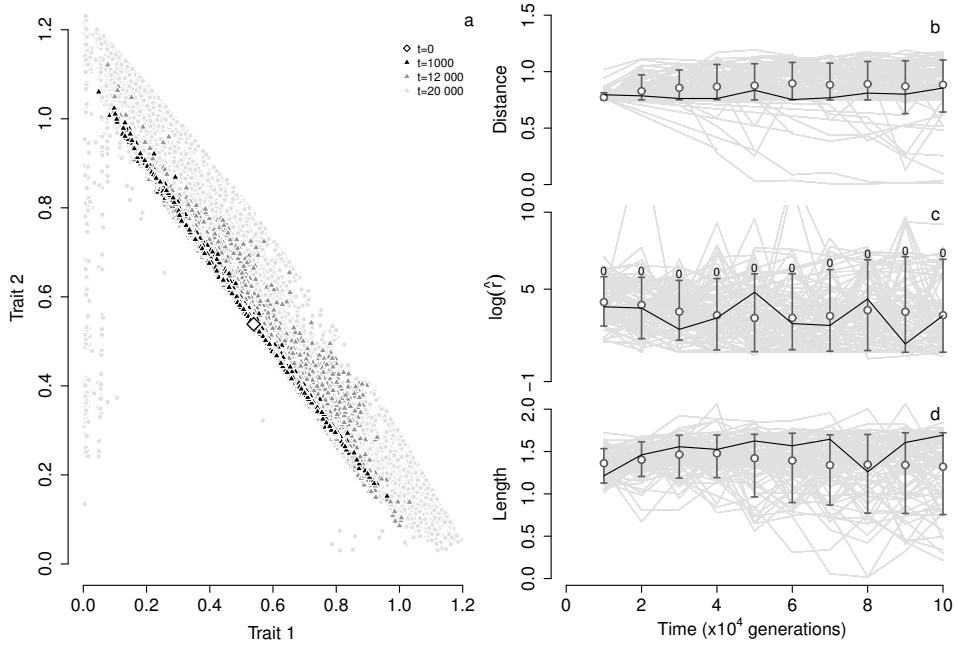

Figure S1 – Evolutionary dynamic for the neutral model. Panel b shows the temporal dynamics of the distance of the median projection on the fitted circle to the origin. The natural logarithm of the fitted circle's radius (inversely related with the trade-off curvature) is represented in panel c, and the length of the trade-off (i.e. the distance along the circle between the two most distant projections) is represented in panel d. 99 replicate simulations are represented in grey in panels b-d, and the simulation in panel a is outlined in black. Dots and bar represent respectively the mean and quantiles (0.1 and 0.9) of each parameter's distribution (across simulations) at each timestep. Standard parameter values where used (as defined in the Material and methods section), and Cstorage = 0.5. On panel b the percentage of concave trade-offs is added on the top of each error bars. In neutral context, all simulations show convex trade-offs (contrary to simulations in with selection context).

*Corresponding author: salome.bourg@univ-lyon1.fr

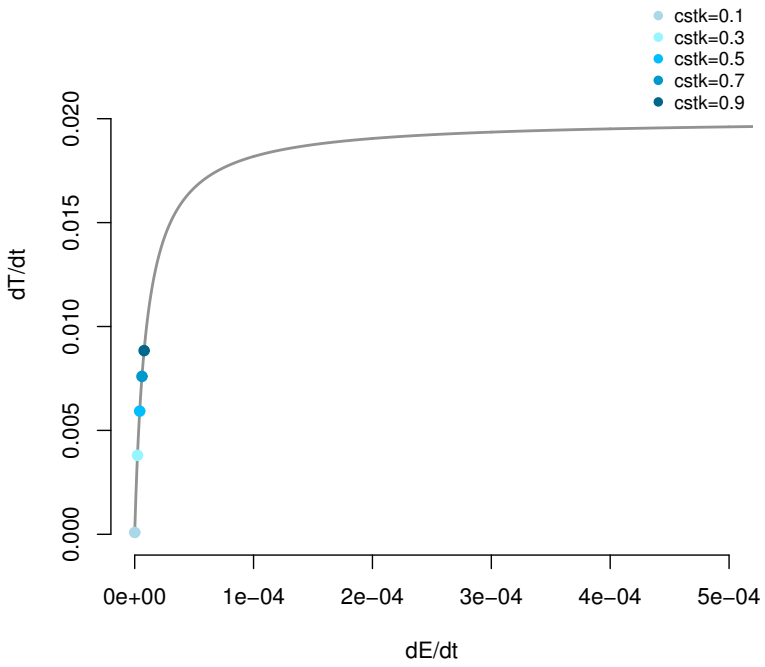

Figure S2 – The storage cost context influences directly the speed of resource consumption in order to maximize fitness. For each simulations with storage cost $\in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, we extracted the median a_j parameters (cf. Model B. Energy allocation dynamics) after evolution. On the left side of the graph, ranges from quantile at 0.1 to quantiles 0.9 are represented.

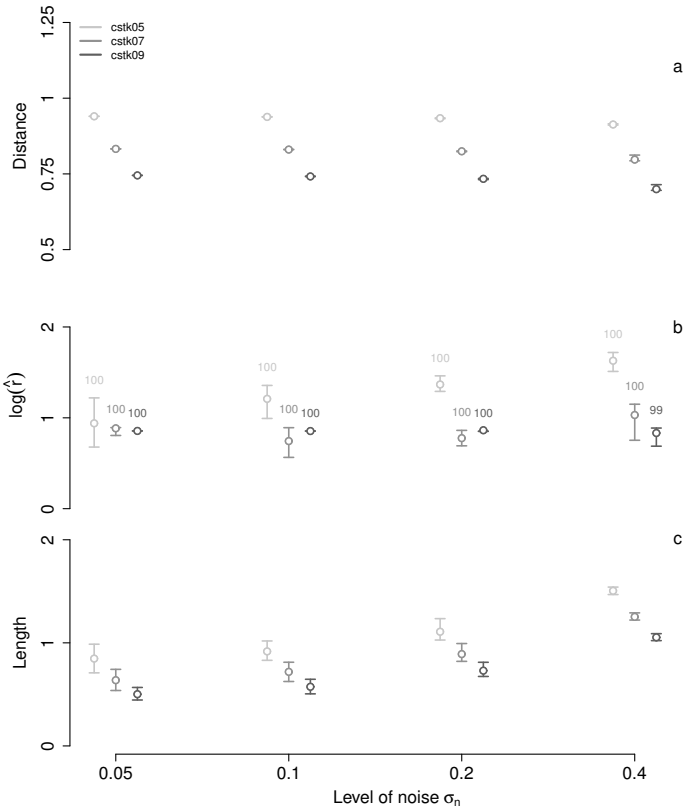

Figure S3 – The trade-off shape changes in response to the level of noise in hormones and receptors gene expressions. The noise is modeled by multiplying each genetically encoded gene expression by 10^{ϵ}, where $\epsilon = \mathcal{N}(0, \sigma_{\mathbf{n}})$. Dots and error bars represent the mean and between-quantile difference $(q(0.9) - q(0.1))$ of the distribution of shape parameters calculated in 100 replicate populations. On panel b the percentage of concave trade-offs is added on the top of each error bars. Contrary to neutral simulations, trade-offs remain concave even in high level of noise context.

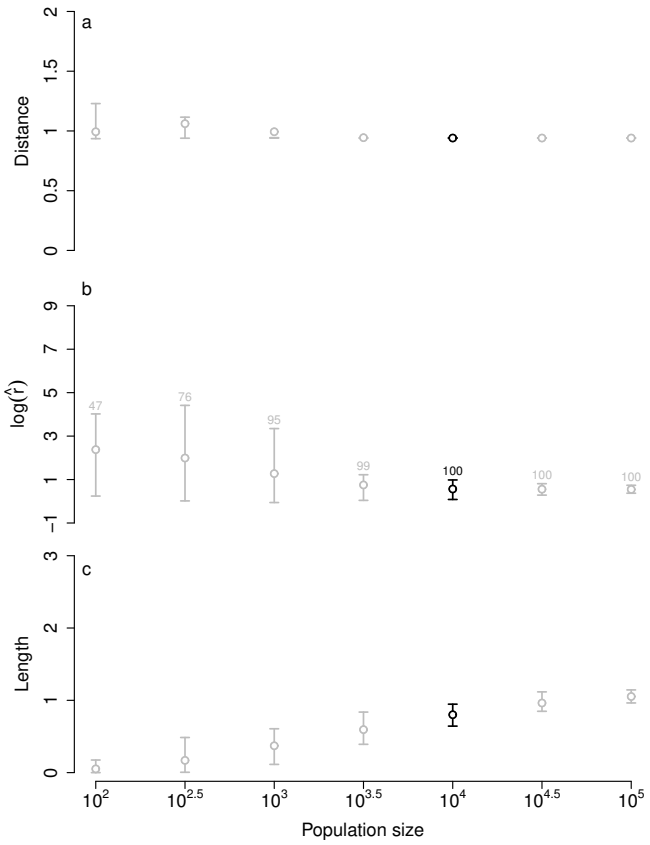

Figure S4 – Shape parameters for different population sizes. In black the population size used in our standard simulations. Dots and error bars represent the mean and between-quantile difference ($q(0.9) - q(0.1)$) of the distribution of shape parameters calculated in 100 replicate populations in a context of $C_{storage} = 0.5$. On panel b the percentage of concave trade-offs is added on the top of each error bars. As the population size increases, the trade-off length increases and the trade-off tends to curve (concave), the distance remains stable. Because of the growth of population size, more individuals can distribute on the entire trade-off and as a consequence increase its length and reveals its shape (concave).

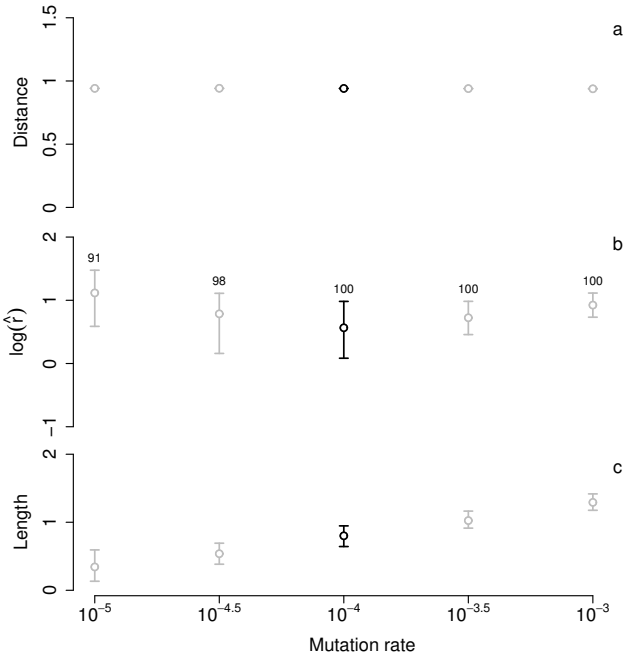

Figure S5 – Shape parameters for different mutation rate.In black the mutation rate used in our standard simulations. Dots and error bars represent the mean and between-quantile difference $(q(0.9) - q(0.1))$ of the distribution of shape parameters calculated in 100 replicate populations in a context of $C_{storage} = 0.5$. On panel b the percentage of concave trade-offs is added on the top of each error bars.

Text S1 – Detecting associations between the context and the genotype or the physiological mechanism through PCA and CCA

As shown in fig. 2, different trade-off shapes evolve in specific contexts, defined as the cost associated with the storage of a resource, C_{storage}. Here we describe our use of multidimensional statistical analyses to investigate whether specific genotypes or specific physiological mechanisms have also evolved in these different contexts. Two sets of variables were used in these analyses, which summarize different aspects of the populations at the end of our simulations :

- genotype data, namely the mean conformation of each hormone and of each receptor, the mean expression of each hormone and that of each receptor on each structure.
- physiological data, taking into account the fact that different genotypes yield strictly identical physiological mechanisms. For instance, a genotype for which both hormones have conformation 25 and both receptors conformation 35 cannot be distinguished, from a physiological point of view, from a genotype with conformations 5 and 15, respectively, all else being equal. We thus calculated the mean **distance in conformation** between all pairs of hormones and receptors, and identified the pair where this distance is minimum ('best pair' hereafter) and the other pair ('other'). To form the physiological dataset, we calculated for each pair the population mean distance in conformation, the mean expression of the hormone, the average of the sum of the expressions of the receptor on each structure, and the mean absolute difference between the receptor's expression on each structure.

We performed two different analyses on these datasets : a Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA). The PCA is aimed at identifying the combinations of variables within each dataset that explain the most overall variation. If the context has a strong influence on the evolving genotypes or physiological mechanisms, we should expect that simulations of different contexts appear distinct in this analysis. This is not the case (figure S6), possibly because the variation between populations within a single context is much higher than (and obscures) the variation between contexts. This indicates that many genotypes and many physiological mechanisms may evolve in response to a similar context, preventing any prediction of evolution at these two levels. In other words, evolution is convergent at the phenotypic level (the trade-off shape) but not at the genetic or physiological level. We note that, as mentioned above, several genotypes can yield a similar physiological mechanism, such that the observed redundancy at the genetic level is not surprising. Next we investigate if, despite the high variation observed between populations within a single context, we can identify linear combinations of the variables in our physiological dataset that covariate with the context, represented by C_{storage}. This is precisely the aim of the CCA, which we use to identify the linear combination of the variable that correlates best with C_{storage} ($\rho = 0.26$,$p = 5,09.10^{-12}$). Surprisingly, the difference in expression of the receptor of the 'other' pair between the two structures contributes negatively to this linear combination, while their summed expression contributes positively. Also surprisingly, the differences in the conformation of hormones and receptors in the two pairs contribute only little (figure S2). As expected, however, the expression of all receptors and hormones contribute positively, meaning that on average, the physiological mechanism that evolves in response to an increase in storage cost has more hormones and more receptors, which as expected increases the speed of resource consumption.

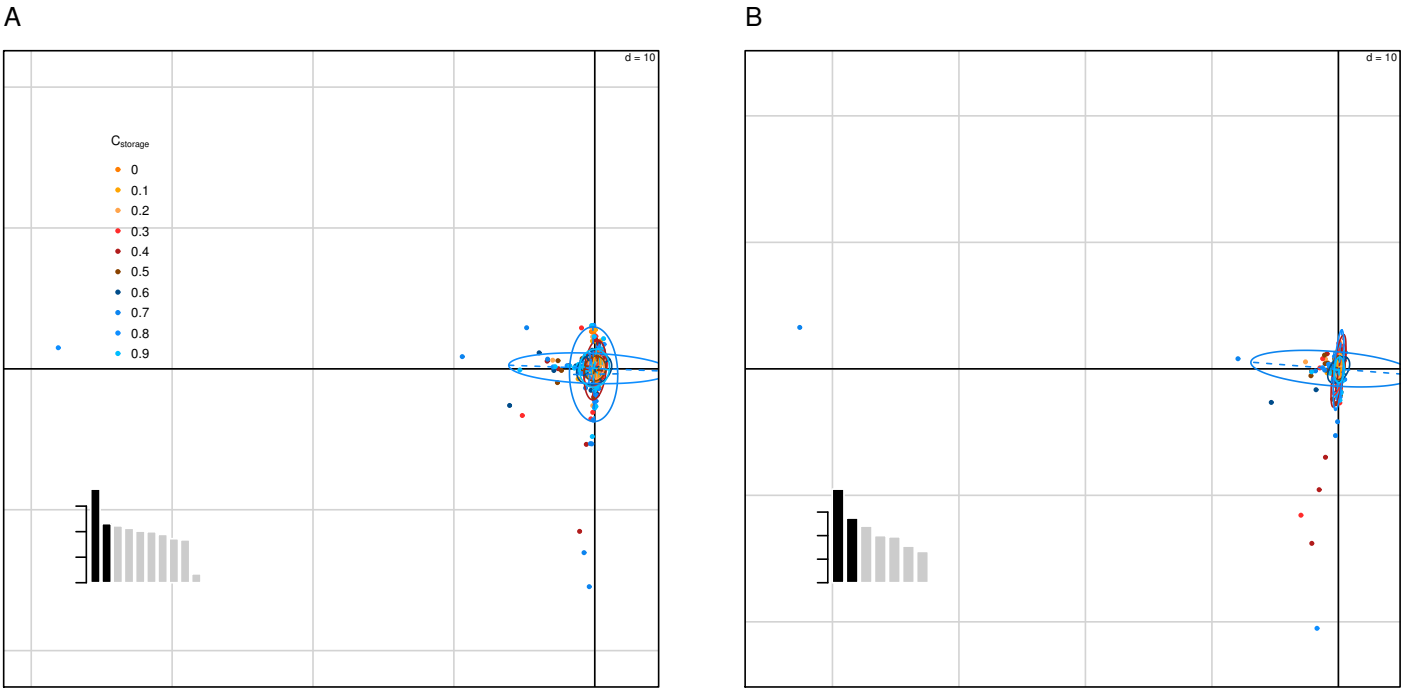

Figure S6 – The PCA on panel A (genotype data) and the PCA on panel B (physiological data) show no association between these variables and the context, here represented by the storage cost ($C_{storage}$). The two first axes of each PCA are shown, whose relative contribution to the overall variance is shown in black on barplots in the bottom left of each panel.

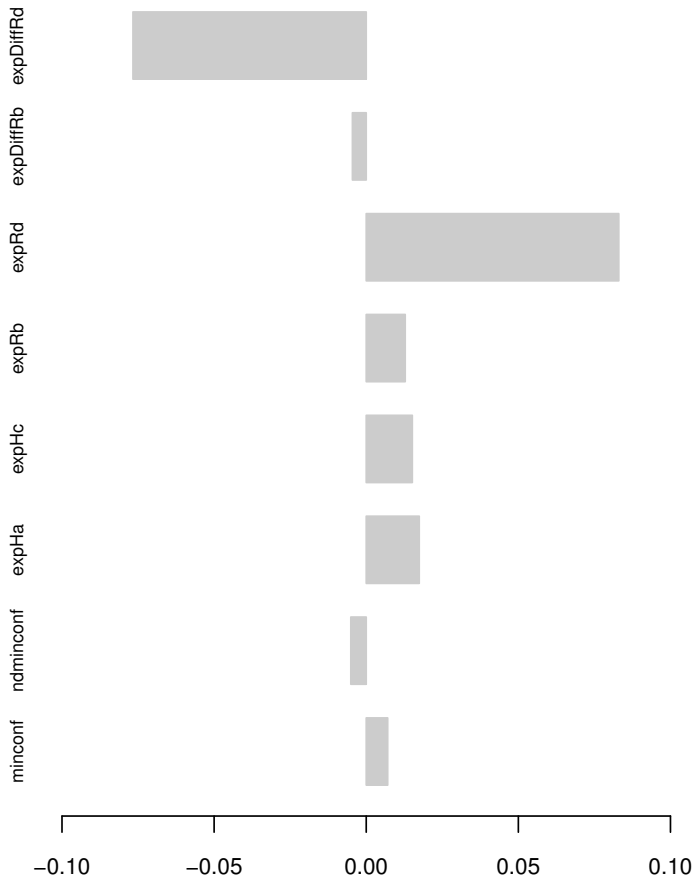

Figure S7 – CCA on physiological dataset shows a significative correlation ($p = 5,09.10^{-12}$, $\rho = 0.26$). In the legend, minconf corresponds to the mean distance in conformation of the best pair, ndminf = the mean distance in conformation of the other pair, expHa= mean hormone expression of the best paire, expHc= mean hormone expression of the other pair, expRb= mean receptor expression of the best pair, expRd= mean receptor expression of the other pair, expDiffRb= mean absolute difference between the receptor's expression on each structure for the best pair, expDiffRd= mean absolute difference between receptor's expression on each structure for the other pair.

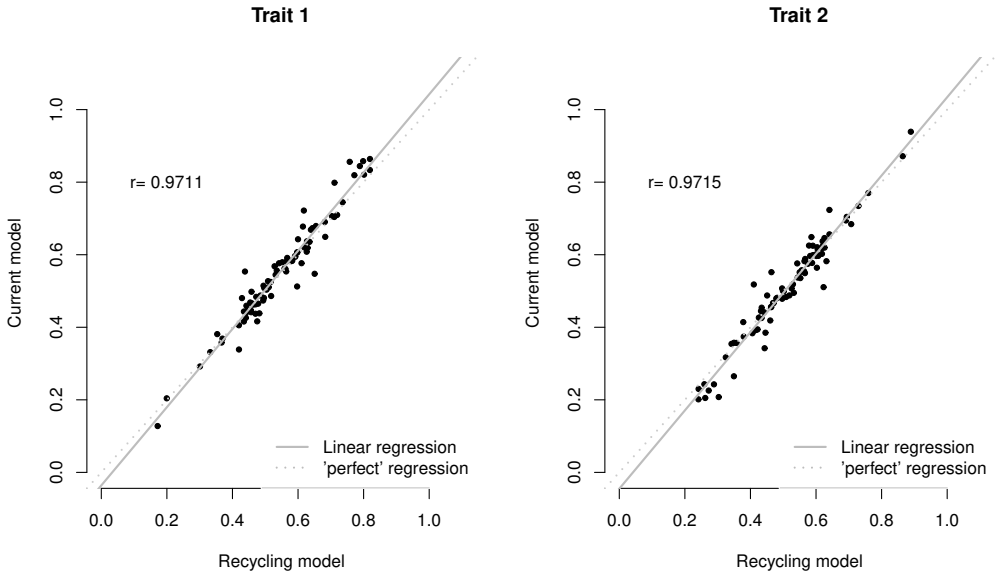

Figure S8 – Comparison of fitnesses obtained in our standard model with those obtained in a model where receptors only are recycled – instead of both the hormone and the receptor being recycled. In this latter model, eq. [1] becomes $\frac{d[H_{kp}]}{dt} = \alpha_{kp} - \sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\sum_{q=1}^{n_{ch}}\sum_{p=1}^{n_{ch}}\left(k_{\mathbf{on}_{iqkp}} \times [R_{iqj}] \times [H_{kp}]\right) - k_d \times [S_D] \times [H_{kp}]$, such that the term corresponding to the hormone released after complex dissociation is canceled. Dots correspond to different genotypes, which were randomly generated by simulating three mutations from the initial conditions defined in the main text. The line represents the situation where both models yield identical fitnesses.

SI text 2 – Solution for the energy allocation dynamics

Equations (10) and (11) in the main text yield three distinct phases for the energy allocation dynamics, as illustrated by figure 1.

Phase 1. As stated in the main text, phase 1 starts at t_0 (the meal) and goes on as long as $[E] > E_{\mathrm{om}}$, $E_{\mathrm{om}} = 0.08$ being a concentration threshold above which energy is stored. At t_0, $[E] = Eo$ and decreases until $[E] = E_{\mathrm{om}}$ (t_1). From equation (10), we obtain

$$[E](t) = \frac{b \times E_{\mathrm{om}}}{a+b} \times (1 - \mathrm{e}^{-(a+b)\times t}) + Eo \times \mathrm{e}^{-(a+b)\times t}. \tag{S1}$$

We find t_1 from the equality $[E](t_1) = E_{\mathrm{om}}$:

$$t_1 = -\ln(\frac{E_{\mathrm{om}} \times a}{Eo \times (a+b) - b \times E_{\mathrm{om}}}) \times \frac{1}{a+b} \tag{S2}$$

Finally, substituting $[E](t)$ in equation (11) we obtain :

$$[Es](t_1) = \frac{b \times (1 - C_{\mathrm{storage}})}{a+b} \times \left(-a \times E_{\mathrm{om}} \times t_1 + (Eo - \frac{b \times E_{\mathrm{om}}}{a+b}) \times (1 - e^{-(a+b)\times t_1}) \right) \tag{S3}$$

Phase 2. Here the resource is released from the storage structure until $[E_{\mathrm{s}}](t) = 0$ (t_2). At this point, we have :

In order to obtain t_2, we assume that $[E]$ is constant during this phase, so $\frac{d[E]}{dt} = 0$ and $b([E] - E_{\mathrm{om}}) = -a[E]$. Therefore,

$$\frac{d[E_{\mathrm{s}}]}{dt} = -a[E]. \tag{S4}$$

From equation (S4) we obtain

$$[E_{\mathrm{s}}](t) = [E_{\mathrm{s}}](t_1) + a \times t_1 \times \frac{bE_{\mathrm{om}}}{a+b} - a \times E_{\mathrm{om}} \times t, \tag{S5}$$

which equals 0 at t_2. We thus find t_2 :

$$t_2 = (E_{\mathrm{s}}(t_1) + a \times \frac{b \times E_{\mathrm{om}}}{(a+b)} \times t_1) \times \frac{(a+b)}{a \times b \times E_{\mathrm{om}}}. \tag{S6}$$

Finally we calculate :

$$[E](t_2) = \frac{b \times E_{\mathrm{om}}}{a+b} \times (1 - \mathrm{e}^{-(a+b)\times t_2}) + Eo \times \mathrm{e}^{-(a+b)\times t_2} \tag{S7}$$

Phase 3. Phase 3 begins when $[E_{\mathrm{s}}]$ reaches 0, such that $[E]$ decreases until it reaches the critically low value $E_{\mathrm{min}} = 0.01$ (t_3). During this phase, we find from equation (10) that

$$[E](t) = \frac{b \times E_{\mathrm{om}}}{a+b} \times \frac{e^{-a\times t}}{e^{-a\times t_2}}, \tag{S8}$$

and we find t_3 by substituting $[E](t_3)$ by E_{min} in equation (S8) :

$$t_3 = \frac{\ln(E_{\mathrm{min}} \times \frac{(a+b)}{b\times E_{\mathrm{om}}}) - a \times t_2}{-a} \tag{S9}$$