POSTER

Balancing Metabolic Homeostasis and Enzyme Cost with Multi-Objective Evolutionary Algorithms

ARTHUR LEQUERTIER, National Research Institute for Agriculture, Food and Environment, Paris, Ile-de-France, France

ALBERTO PAOLO TONDA, National Research Institute for Agriculture, Food and Environment, Paris, Ile-de-France, France

WOLFRAM LIEBERMEISTER, National Research Institute for Agriculture, Food and Environment, Paris, Ile-de-France, France

Open Access Support provided by:

National Research Institute for Agriculture, Food and Environment

Published: 14 July 2025

Citation in BibTeX format

GECCO '25 Companion: Genetic and Evolutionary Computation Conference Companion
July 14 - 18, 2025
Malaga, Spain

Conference Sponsors:
SIGEVO

https://doi.org/10.1145/3712255.3726597
ISBN: 9798400714641

Balancing Metabolic Homeostasis and Enzyme Cost with Multi-Objective Evolutionary Algorithms

Arthur Lequertier
arthur.lequertier@inrae.fr
Université Paris-Saclay, INRAE, MaIAGE
Jouy-en-Josas, France

Alberto Tonda
alberto.tonda@inrae.fr
UMR 518 MIA-PS, INRAE, Université Paris-Saclay
Palaiseau, France
UAR 3106 Institut des Systèmes Complexes
Paris, France

Wolfram Liebermeister
wolfram.liebermeister@inrae.fr
Université Paris-Saclay, INRAE, MaIAGE
Jouy-en-Josas, France

Abstract

The metabolic fluxes in cells are regulated by metabolites that bind to enzymes and modulate their activities. Metabolites can stabilize their own concentrations by exerting a negative feedback on their production pathways. Although metabolic networks have been studied extensively, many regulation arrows remain unknown. To model the costs and benefits of direct enzyme regulation, we apply multi-objective optimization with one loss function describing how regulation stabilizes the metabolic state against random perturbations and another loss function scoring the extra enzyme amounts required by regulation arrows. Using the number of arrows as a third objective for ease of interpretation, we study how activating and inhibiting arrows should be arranged in the system. Using an evolutionary multi-objective approach, simulating an evolution under biological trade-offs, we explore optimal arrow configurations. Our framework can also be used with other biological objectives, including optimal adaptation or information transmission in metabolic networks, to study their trade-offs with objectives such as energy or enzyme investment in cells.

CCS Concepts

• **Applied computing** → **Computational biology**; • **Theory of computation** → **Evolutionary algorithms**.

Keywords

Metabolic model, Enzyme regulation, Robustness, Enzyme cost, Multi-objective optimization, NSGA-II

ACM Reference Format:
Arthur Lequertier, Alberto Tonda, and Wolfram Liebermeister. 2025. Balancing Metabolic Homeostasis and Enzyme Cost with Multi-Objective Evolutionary Algorithms. In *Genetic and Evolutionary Computation Conference (GECCO '25 Companion), July 14–18, 2025, Malaga, Spain.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3712255.3726597

GECCO '25 Companion, July 14–18, 2025, Malaga, Spain

ACM ISBN 979-8-4007-1464-1/2025/07
https://doi.org/10.1145/3712255.3726597

1 Introduction

The metabolic fluxes in cells are constantly adapted to environmental changes and internal demands. An important regulation mechanism behind this is the activation or inhibition of enzymes by small effector molecules. Such regulation arrows in metabolic networks can stabilize metabolic states in changing environments. For example, some metabolites stabilize their concentrations by negatively regulating enzymes that catalyze their own producing reactions. Although the metabolic network structures of many microbial species have been uncovered, the regulation arrows in these networks are often unknown.

Different species tend to share similar metabolic network structures, but appear to implement regulation in different ways. However, there are typical arrangements of regulation arrows, reminiscent of network motifs in gene regulation [9]. To explain these regulation motifs, we may claim that they are evolutionarily selected for specific benefits and costs, for example, providing homeostasis, but also requiring costly investments in enzymes. Assuming such compromises, we may ask whether evolution favors "wiring economy", that is, whether regulation should be implemented by few, strong arrows or rather by larger numbers of weaker arrows [1].

The stabilizing and destabilizing effects of enzyme regulation have been explored in [13, 14] using computational models. Structural Kinetic Modeling [7, 13] is a method to study the dynamic effects of network structures, for example of the presence and signs of regulation arrows, regardless of quantitative parameters. Here we employ this method to optimize the arrangement of regulation arrows in cells that live in stochastic environments. In our model, we consider a cell with a given metabolic reference state. A stochastic environment perturbing the system leads to a random ensemble of metabolic states. We consider possible arrangements of activating or inhibiting regulatory arrows, each with a specific strength, and apply multiobjective optimization to explore arrangements providing optimal compromises between favorable control properties and low enzyme cost.

2 Background

2.1 Metabolic network models

Our framework for metabolic models is based on Structural Kinetic Modeling and follows the description in [7]. A metabolic model describes the biochemical compounds and reactions within a cell. Network nodes i represent chemical species called metabolites,

while hyperedges l represent chemical reactions. Assuming a given reference state of the system, we describe how small changes in metabolite concentrations would influence the reaction rates. The scaled elasticities of a reaction, defined as $\mathcal{E}_{li} = \partial \ln v_l / \partial \ln c_i$ and parameterized by enzyme saturation coefficients φ_{li}, quantify how sensitive the reaction rate v_l is to small changes in the concentrations c_i of metabolites.

Reaction elasticities. To study the effects of regulation arrows in a network, we write the elasticity matrix $\boldsymbol{\mathcal{E}}$ as the sum of two sparse matrices:

$$\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}}^{\text{kin}} + \boldsymbol{\mathcal{E}}^{\text{reg}}. \tag{1}$$

The matrix of kinetic elasticities $\boldsymbol{\mathcal{E}}^{\text{kin}}$, describing the effects of the enzymes' own reactants, is assumed to be known. The matrix $\boldsymbol{\mathcal{E}}^{\text{reg}}$ represents additional regulatory effects that metabolites can have on enzyme catalytic activities. If an enzyme catalyzes a reaction, the reaction rate v_l depends on the activity of its enzyme, and metabolites regulating the enzyme will lead to non-zero elasticities $\mathcal{E}^{\text{reg}}_{li}$.

Metabolic steady-state responses to external perturbations. To model variability in metabolic states, we describe all concentrations and reaction rates with random variables [8]. From given random distributions of the external variables (external metabolite concentrations and enzyme levels), we obtain the distributions of all state variables (internal metabolite concentrations and reaction rates). All model variables are described on a logarithmic scale. Assuming small variations, we replace the system dynamics by a linear approximation using concepts from Metabolic Control Analysis (MCA) [6, 10]: the linearized steady-state responses $\delta\mathbf{z} = \boldsymbol{\mathcal{R}}\ \delta\mathbf{x}$ of logarithmic state variable vectors to logarithmic external perturbation vectors are described by the scaled response matrix $\boldsymbol{\mathcal{R}}$ [6], which can be computed from the elasticity matrix $\boldsymbol{\mathcal{E}}$. To compute the variances and covariances, we assume that all variables (on logarithmic scale) follow normal distributions centered around the reference state. For logarithmic external variables (in a random vector $\mathbf{x}$), we predefine a given diagonal covariance matrix $\text{Cov}(\mathbf{x})$. In our linear approximation, the logarithmic state variables (in a random vector $\mathbf{z}$) will then follow a multivariate normal distribution with covariance matrix [8]

$$\text{Cov}(\mathbf{z}) = \boldsymbol{\mathcal{R}}\ \text{Cov}(\mathbf{x})\ \boldsymbol{\mathcal{R}}^{\top}. \tag{2}$$

2.2 Multi-objective evolutionary optimization

In traditional optimization methods, conflicting objectives can be addressed by constructing a composite objective function, typically as a weighted sum of individual objectives. In contrast, multi-objective optimization [2] does not require predefined weights, and instead explores the possible trade-offs among conflicting objectives. The result is not just single solution, but a set of solutions that are not Pareto-dominated by others. Evolutionary algorithms (EAs) are well suited for multi-objective optimization and are widely seen as the state of the art in this domain [12]. For this study, we chose the well-established Non-Dominated Sorting Genetic Algorithm II (NSGA-II) [3]. Despite not being the most recent method, NSGA-II remains highly competitive for optimization problems involving up to three objectives and is available across various programming languages, including C++ and Python.

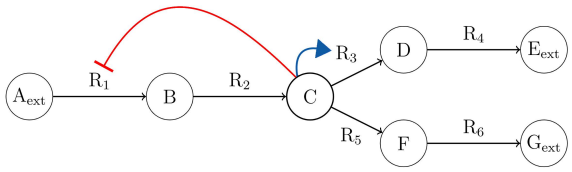

Figure 1: Metabolic network model considered in this study: a branched metabolic pathway. As an example, the figure shows two regulation arrows (red: inhibition; blue: activation). Through product inhibition and substrate activation, the arrows tend to stabilize the concentration of metabolite C.

3 Proposed approach

To model metabolic dynamics, we start from a metabolic network with a given reference state, add some regulation arrows, and score the system by objectives describing the total enzyme demand, the resulting (non)robustness of the metabolic steady state, and the number of regulation arrows. Using multi-objective optimization, we determine arrangements of regulation arrows that represent optimal compromises between the objectives.

3.1 Candidate solutions

In our optimality problem, each candidate solution represents an arrangement of arrows defining a regulatory elasticity matrix $\boldsymbol{\mathcal{E}}^{\text{reg}}$. For convenience, a candidate solution is described by a bit vector $\mathbf{b}$ (denoting the presence or absence of each possible arrow) and a real-valued vector $\mathbf{a}$ encoding the arrow strengths and signs (activating or inhibiting arrow)

$$\begin{aligned} \mathbf{b} &\in \{0,1\}^d \\ \mathbf{a} &\in [0,1]^d, \end{aligned} \tag{3}$$

where d is the number of possible arcs. The binary vector $\mathbf{b}$ encodes the existence of arcs between different metabolites and reactions in the network; the real-valued numbers in $\mathbf{a}$ encode the arrow signs as $q = \text{sign}(a - 0.5) \in \{0, 1\}$ and the strengths as $\varphi = b\ |2\,a - 1| \in [0, 1]$, ranging between 0 (no effect) and 1 (maximal effect), and thus determine the elasticities $\mathcal{E}^{\text{reg}}_{li} = q_{li}\ \varphi_{li}$.

3.2 Objective functions

If a new regulation arrow is added to a cell's metabolic network, this will have two effects. First, it decreases the average activity of the regulated enzyme, and to restore the activity (and keep the reference state unchanged) the enzyme level must be increased by a factor that depends on the strength of the arrow. Hence, each arrangement of arrows and arrow strengths leads to an increase in the cell's enzyme demand. Second, a new regulation arrow also changes the way state variables respond to external perturbations. In a random environment, each additional arrow will reshape the

probability distribution of all state variables. Assuming that variability of state variables may incur fitness losses, we can characterize each arrangement of arrows by an average fitness loss. The three loss functions are computed as follows.

Loss function f_1: Enzyme cost. The "enzyme cost" loss of an arrow arrangement (described by vectors **b** and **a**) describes the extra amount of enzyme that a cell needs to invest to maintain all steady-state fluxes and metabolite concentrations at their reference values. It is given by

$$f_1 = \sum_l \left[\left[\prod_i \frac{1}{1-\varphi_{li}} \right] - 1 \right] e_l \tag{4}$$

where e_l denotes the reference enzyme levels and φ_{li} is the regulation strength of enzyme l with respect to effector metabolite i, determined by the respective element in **a**. Without regulation arrows, the enzyme cost would be 0. Using arguments based on whole cell models, an increase in enzyme amounts could be translated into a decrease in cell growth rate, a proxy for Darwinian fitness [11, 15].

Loss function f_2: Non-robustness. The "non-robustness" loss function penalizes variability of metabolic state variables, described by the variances from Eq. (2). It is given by

$$f_2 = \sum_k m_k \, \mathrm{Var}(z_k) \tag{5}$$

with positive weights m_k. In a network without arrows, there will be a positive loss. As arrows are added, this loss may increase or decrease.

Loss function f_3: Number of regulation arrows. To study wiring economy – that is, a possible parsimonious usage of arrows – we treat the number of regulation arrows as a third loss function f_3.

3.3 Evolutionary operators

Given the structure of a candidate solution, we designed structure-aware operators that generate valid offspring and preserve semantic information. The operators are applied with probability p_m for structure-aware mutations and p_c for structure-aware crossovers.

Structure-aware mutation. With uniform probability, the mutation operator can change an individual by (i) flipping a bit in the vector **b** or (ii) performing a Gaussian mutation on a single random element of **a**. For a Gaussian mutation, we add a Gaussian random number to one of the values a_j (choosing an element j for which $b_j = 1$) and bound the result within the allowed domain $[0, 1] \subset \mathbb{R}$. This mutation has two hyperparameters: a mean μ_m and a standard deviation σ_m.

Structure-aware crossover. The crossover operator used is a one-point crossover, choosing randomly with uniform probability a cut point for both **b** and **a**. This preserves the semantic information linking the existence of regulation arrows (encoded by **b**) with the regulation signs and strengths (encoded by **a**).

3.4 A test case: metabolic branch-point model

For our experiments we used the simple pathway model shown in Figure 1. Among the 7 metabolites, A_{ext}, E_{ext}, and G_{ext} are external metabolites with concentrations treated as sources of variability. The concentrations of the internal metabolites B, C, D, and F and the reaction rates v_l (for reaction indices $l = 1..6$) are state variables. As a simple objective to describe homeostasis, we use a loss function $f_2 = \mathrm{Var}(c_C)$ that penalizes variability in the concentration of metabolite C. In cells, stabilizing the concentration of intermediates can be important in reducing toxicity, changes in osmotic pressure, or losses caused by membrane leakage. To focus our simulations on the most relevant part of the Pareto front, we capped the enzyme cost function f_1 at an upper value of 50 (arbitrary choice).

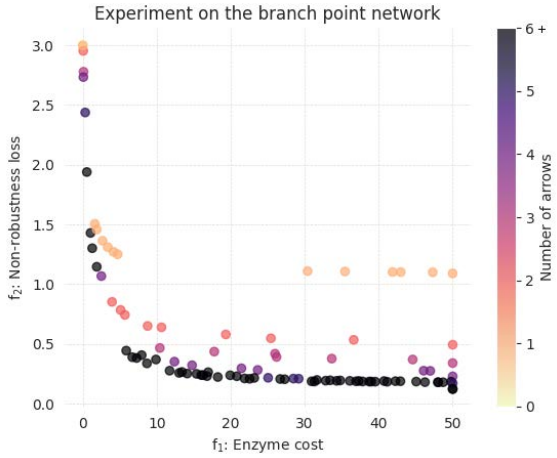

Figure 2: Pareto front for the model in Figure 1, obtained after 230 generations of evolutionary optimization with 3 objectives: enzyme cost f_1 (x-axis, capped at 50), non-robustness loss f_2 (y-axis), and arrow count number f_3 (shown in color, where "6+" represents numbers ≥ 6). The plot represents a projection of the 3d Pareto front on the f_1/f_2 plane.

4 Experimental evaluation

4.1 Set-up for multi-objective optimization

As our multi-objective evolutionary algorithm for experimental evaluation, we chose the Non-Sorting Genetic Algorithm II (NSGA-II) [4], an established choice for problems with two or three objectives. For all experiments, the hyperparameters of the algorithm were set as follows: population size $\mu = 100$, offspring size $\lambda = 100$, maximum number of generations $G = 1000$, $p_\mathrm{c} = 0.8$, $p_\mathrm{m} = 0.5$, $\mu_\mathrm{m} = 0.0$, $\sigma_\mathrm{m} = 0.1$. In our case study, each individual (a possible arrow arrangement) was encoded by vectors **b** and **a** of length $d = 6 \cdot 7 = 42$. With this configuration, a single run took around 4 hours on an end-user laptop. The scripts for the experiments, coded in Python, use the `inspyred` [5] library[1] for the implementation of NSGA-II. All the code and data necessary to reproduce the experiments are available in a public GitHub repository[2].

[1] Inspyred, https://github.com/aarongarrett/inspyred
[2] https://github.com/albertotonda/evolutionary-optimization-cell-models

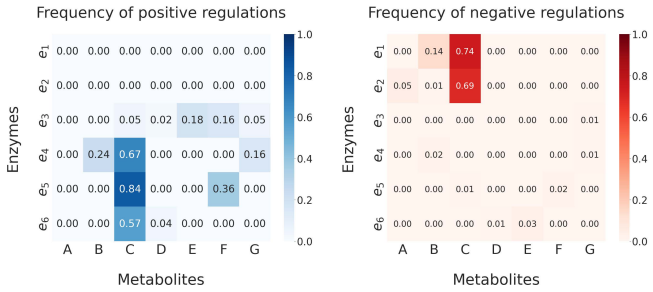

Figure 3: Relative frequencies of occurrence of regulation arrows in the solutions. Each number refers to an arrow between an effector metabolite and an enzyme (see Figure 1).

4.2 Results

Our optimization results are shown in Figures 2 and 3. Figure 2 shows the Pareto front of our 3-objective problem projected onto the plane of f_1 (enzyme cost) and f_2 (non-robustness of the central metabolite C). The third objective f_3, representing the number of regulation arrows, is shown in color.

In the 2d projection in Figure 2, the 3d Pareto front consists of a set of separate fronts corresponding to different total numbers of arrows: for each count number f_3 we observe a simple curved front with a trade-off between f_1 and f_2. If we see f_1 and f_2 as fitness relevant (and f_3 as a mere feature of the solutions), the optimization does not seem to favor parsimonious solutions. At a given enzyme investment (x-axis), better homeostasis (that is, lower non-robustness) requires larger numbers of arrows. However, for count numbers larger than 6, the resulting extra benefits are negligible.

Figure 3 shows how often each of the possible arrows appears in the solutions. The predicted favorable arrows are as expected: the most common arrows to stabilize the metabolite C are feedback inhibitions from C to its producing reactions 1 and 2, or forward activations of the consumption pathways, like the arrows displayed in Figure 1.

5 Conclusions

We developed a framework for assessing the benefit and cost of enzyme regulation in cells. Regulatory interactions were scored by loss functions that reflect biological fitness objectives: regulation incurs an enzyme cost, as higher enzyme levels are needed to compensate for reduced enzyme activities, but it can also improve the robustness of metabolic variables. The evolutionary algorithm used here should not just be seen as a computational tool, but resembles a model of biological evolution, potentially explaining diversity in metabolic regulation across microbial species. It can simulate, for example, how cell populations might evolve new regulation arrows under changing environmental selection pressures.

Although some biological regulation systems show wiring economy [1] – a preference for sparse connections – our model predicts a preference for multiple, weaker arrows, at least up to a certain number. This result depends on our cost function, which penalizes large regulation strengths very strongly and makes distributed regulation more favorable – although above a certain number of arrows the extra benefits become very small. Even if our cost function is biologically realistic, this prediction is not fully conclusive: considering other costly effects in our model, giving rise to cost functions similar to a L^1 norm term, may have led to a different result. Moreover, our results reflect many other model assumptions made: for example, enzyme levels were treated as simple sources of noise, while in reality they are controlled by transcriptional feedback regulation, which comes at its own benefits and costs. In real cells, only a limited number of enzyme regulations are known, although many more may exist. If wiring economy is not a general principle, then weak, numerous, undetected regulatory interactions might play a larger role in metabolism than previously assumed.

Our framework, applied here to a simple pathway, also applies to larger, realistic metabolic networks. Having gained confidence in our approach, we would like to explore how other network structures and other assumptions about noise variables will shape regulation arrows. By optimizing not only regulation, but also other details of enzyme kinetics (such as enzyme saturation with reaction substrates and products), our framework may be used to tackle a wide range of biological questions.

Acknowledgments

This work was supported by the French National Research Agency (grant ANR-21-CE45-0021).

References

[1] Eda Cakir, Annick Lesne, and Marc-Thorsten Hütt. 2021. The economy of chromosomal distances in bacterial gene regulation. *npj Systems Biology and Applications* 7, 1 (2021). https://doi.org/10.1038/s41540-021-00209-2
[2] K. Deb. 2001. *Multi-Objective Optimization Using Evolutionary Algorithms*. Vol. 16. John Wiley & Sons.
[3] K. Deb and A. Pratap. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6 (2002), 182–197. Issue 2.
[4] K. Deb and A. Pratap. 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 2 (2002), 182–197.
[5] A. Garrett. 2012–2024. inspyred (Version 1.0.1) [software]. Inspired Intelligence. https://github.com/aarongarrett/inspyred.
[6] J. Hofmeyr. 2001. Metabolic control analysis in a nutshell. In *ICSB 2001 Online Proceedings, http://www.icsb2001.org/toc.html.*
[7] W. Liebermeister. 2022. Structural Thermokinetic Modelling. *Metabolites* 12, 5 (May 2022), 434. https://doi.org/10.3390/metabo12050434
[8] W. Liebermeister and E. Klipp. 2005. Biochemical networks with uncertain parameters. *IEE Proc. Sys. Biol.* 152, 3 (2005), 97–107.
[9] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. 2002. Network motifs: simple building blocks of complex networks. *Science* 298 (2002), 824–827.
[10] C. Reder. 1988. Metabolic control theory: A structural approach. *Journal of Theoretical Biology* 135, 2 (Nov. 1988), 175–201. https://doi.org/10.1016/s0022-5193(88)80073-0
[11] M. Scott, C.W. Gunderson, E.M. Mateescu, Z. Zhang, and T. Hwa. 2010. Interdependence of cell growth and gene expression: origins and Consequences. *Science* 330 (2010), 1099.
[12] S. Sharma and V. Kumar. 2022. A Comprehensive Review on Multi-objective Optimization Techniques: Past, Present and Future. *Archives of Computational Methods in Engineering* 29, 7 (July 2022), 5605–5633. https://doi.org/10.1007/s11831-022-09778-9
[13] R. Steuer, T. Gross, J. Selbig, and B. Blasius. 2006. Structural kinetic modeling of metabolic networks. *Proceedings of the National Academy of Sciences* 103, 32 (2006), 11868–11873.
[14] J. Uhlendorf. 2009. *Prediction of optimal enzymatic regulation architectures.* Master's thesis. Department of Mathematics and Computer Science Bioinformatics Program.
[15] M.T. Wortel, E. Noor, M. Ferris, F.J. Bruggeman, and W. Liebermeister. 2018. Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield. *PLoS Computational Biology* 14, 2 (2018), e1006010.