Optimal metabolic states in cells

Mathematical details and proofs

Wolfram Liebermeister[1,2]

[1] INRA, UR1404, MaIAGE, Université Paris-Saclay, France
[2] Institut für Biochemie, Charité – Universitätsmedizin Berlin, Germany

Contents

S1 Metabolic states **1**

S1.1 Metabolic states . . . 1
S1.2 Metabolic state manifold and feasible metabolite-flux states . . . 2
S1.3 State manifold and the flux/metabolite states are closely related . . . 3
S1.4 Topology and dimensionality of the state manifold . . . 4

S2 Optimality problems in flux or metabolite space **4**

S2.1 Screening of metabolic states . . . 4
S2.2 Optimisation in flux/metabolite space . . . 5
S2.3 Algorithm for testing and finding locally optimal states . . . 6

S3 Proofs **7**

S3.1 Propositions 1, 2, and 3 . . . 7
S3.2 Condition for multi-objective optimality . . . 7
S3.3 Metabolite balance equation (8) . . . 8
S3.4 Balance equation for enzyme optimisation Eq. (S13) . . . 8

S1 Metabolic states

S1.1 Metabolic states

A kinetic metabolic model is defined by a list of chemical reactions, the lists of internal (balanced) and external (non-balanced) metabolites, and differentiable enzymatic rate laws $v_l = e_l\, k_l(\mathbf{c})$. For simplicity, we do not consider non-enzymatic reactions. The rate laws $k_l(\mathbf{c})$ must be reversible and thermodynamically feasible, yielding flux directions $\mathrm{sign}(k_l) = \mathrm{sign}(\theta_l) = -\mathrm{sign}(\sum_i n_{il}^{\mathrm{tot}}\, \mu_i)$, with chemical potentials given by $\mu_i(\mathbf{c}) = \mu_i^0 + RT \ln c_i$. Each of these sign conditions defines a linear inequality for the chemical potentials, and thus for the logarithmic metabolite levels. To obtain physiologically plausible states (and to construct the state manifold, the set of all these states), we further assume three kinds of constraints: (i) Physiological bounds on metabolite levels:

$\mathbf{c}^{\min} \leq \mathbf{c} \leq \mathbf{c}^{\max}$, and non-negative enzyme levels $\mathbf{e} \geq 0$. (ii) Stationary fluxes (mass balance of all internal metabolites): $\mathbf{N}_{\text{int}}\,\mathbf{v} = 0$. (iii) A predefined linear flux benefit $\mathbf{b}_{\text{v}} \cdot \mathbf{v} = b'$ (e.g. a given rate of biomass production).

To make sense of the possible metabolic states, we need to understand the possible flux patterns. A flux distribution is described by flux directions and flux magnitudes, while a flux pattern only describes which reactions are active, and what are the flux directions. The possible flux patterns in a model arise from stationarity and kinetics (or, as we see in a moment, from thermodynamics). We distinguish between a reaction rate $v_l = e_l\, k_l(\mathbf{c})$, the catalytic rate $k_l(\mathbf{c})$, and the driving force $\theta_l(\mathbf{c})$ in the same reaction. With (thermodynamically feasible) reversible rate laws, the allowed flux direction is given by the sign of the catalytic rate (or zero, if the enzyme level vanishes), which in turn is given by the sign of the driving force (or zero if the enzyme is allosterically inhibited). We can write this as

$$\text{sign}(\mathbf{v}) \sqsubseteq \text{sign}(\mathbf{k}(\mathbf{c})) \sqsubseteq \text{sign}(\boldsymbol{\theta}(\mathbf{c})), \tag{S1}$$

where $\boldsymbol{\tau} \sqsubseteq \boldsymbol{\sigma}$ denotes that $\boldsymbol{\tau}$ is conformal with $\boldsymbol{\sigma}$.

The most difficult part in constructing metabolic states is the choice of feasible flux patterns, i.e. of active reactions and flux directions (see SI S1.1). Flux directions are dictated by the thermodynamic driving forces θ_l, which depend on the metabolite profile $\mathbf{c}$. Therefore, each flux pattern requires a (physiologically plausible) metabolite profile that supports these flux directions. Moreover, reactions may be inactive either because they are in thermodynamic equilibrium ($\theta_l = 0$), because the enzymes are kinetically or allosterically inhibited ($\theta_l \neq 0$, but $k_l = 0$), or because no enzyme is present ($e_l = 0$). Each possible flux pattern corresponds to flux and metabolite polytopes, which are defined as follows. Each metabolite profile $\mathbf{c}$ defines a vector $\boldsymbol{\theta}$ of thermodynamic driving forces and a corresponding force pattern $\boldsymbol{\sigma} = \text{sign}(\boldsymbol{\theta}(\mathbf{c}))$. To be compatible with this metabolite profile, a flux mode must conform with this force pattern, i.e., it must show the same signs in all non-zero fluxes. We write this condition as $\text{sign}(\mathbf{v}) \sqsubseteq \boldsymbol{\sigma}$, and we can further split it into $\text{sign}(\mathbf{v}) \sqsubseteq \text{sign}(\mathbf{k})$ (if a reaction flux is non-zero, its sign is dictated by the catalytic rate) and $\text{sign}(\mathbf{k}) \sqsubseteq \text{sign}(\boldsymbol{\theta}$ (for all non-zero catalytic rates, the signs are dictated by the driving forces). Our patches can now be defined as follows. First, we screen the entire metabolite space (within physiological bounds) to obtain the set $S_{\boldsymbol{\sigma}}$ of feasible force patterns. Each feasible force pattern represents a non-empty M-polytope in metabolite space. The polytopes are disjoint and cover the entire metabolite space (again, within physiological bounds). Moreover, a force pattern $\boldsymbol{\sigma}$ defines a set of feasible flows in flux space (i.e., vectors $\mathbf{v}$ satisfying $\mathbf{N}_{\text{int}}\,\mathbf{v} = 0$ and $\text{sign}(\mathbf{v}) \sqsubseteq \boldsymbol{\sigma}$)), called S-polytope ("sign-restricted flux polytope"). With the extra flux benefit constraint $b(\mathbf{v}) = b'$, we can restrict our flows to a B-polytope ("benefit-restricted flux polytope"). The S-polytope (or B-polytope) of a force pattern may be empty. In this case, this force pattern (as well as the corresponding M-polytope in metabolite space) is considered infeasible and can be discarded during model construction.

S1.2 Metabolic state manifold and feasible metabolite-flux states

To enumerate the possible states of a kinetic model, we can either consider the set of thermo-physiologically feasible states $(\mathbf{v}, \mathbf{c})$ in flux-metabolite space, or the set of all kinetically feasible states $(\mathbf{v}, \mathbf{c}, \mathbf{e})$ in flux-metabolite-enzyme space, two sets that are closely related.

For the following definitions, we consider a network with n_m metabolites and n_r enzyme-catalysed reactions, and we define the thermodynamic driving forces $\theta_l(\mathbf{c}) = -\Delta\mu_l(\mathbf{c})$, where $\Delta\boldsymbol{\mu}(\mathbf{c}) = \Delta\boldsymbol{\mu}^{(0)} + RT\,\mathbf{N}_{\text{all}}^{\top} \ln \mathbf{c}$. The vector of thermodynamic driving forces can be written as $\boldsymbol{\theta} = -\Delta\boldsymbol{\mu}(\mathbf{c})/RT = -\Delta\boldsymbol{\mu}^{(0)}/RT - \mathbf{N}_{\text{all}}^{\top} \ln \mathbf{c} = \ln \mathbf{k}_{\text{eq}} - \mathbf{N}_{\text{all}}^{\top} \ln \mathbf{c}$.

Definition S1.1 Thermo-physiologically feasible states *The thermo-physiologically feasible flux-metabolite*

states form a set $\mathcal{T} = \{(\mathbf{v}, \mathbf{c}) : \mathbf{N}_{\rm int}\,\mathbf{v} = 0; \mathbf{c}_{\rm min} \le \mathbf{c} \le \mathbf{c}_{\rm max}; v_l \ne 0 \Rightarrow \mathit{sign}(v_l(\mathbf{c})) = \mathit{sign}(\theta_l(\mathbf{c})\}$. *By choosing* $c_{\rm min,l} = c_{\rm max,l}$, *the concentration of a metabolite can be predefined. The thermo-physiologically feasible states form a collection of convex polytopes in* $n_m + n_r$*-dimensional* $(\mathbf{v}, \mathbf{c})$*-space. Note that no flux bounds were used in this definition. By imposing flux bounds, we may also exclude certain flux directions, and therefore some of the convex polytopes.*

Definition S1.2 Kinetically feasible states *The kinetically feasible steady states* $(\mathbf{v}, \mathbf{c}, \mathbf{e})$ *of a kinetic model form a manifold in a* $(n_m + n_r + n_r)$*-dimensional* $(\mathbf{v}, \mathbf{c}, \mathbf{e})$*-space. The metabolic state manifold can be defined as* $\mathcal{K} = \{(\mathbf{v}, \mathbf{c}, \mathbf{e}) : \mathbf{N}_{\rm int}\,\mathbf{v} = 0; \mathbf{v} = \mathbf{e} \circ \mathbf{k}(\mathbf{c}); \mathbf{c}_{\rm min} \le \mathbf{c} \le \mathbf{c}_{\rm max}; 0 \le \mathbf{e}\}$. *We assume that the rate laws* $k_l(\mathbf{c})$ *satisfy the strict thermodynamic constraints, i.e.* $\mathit{sign}(k_l(\mathbf{c})) = \mathit{sign}(\theta_l(\mathbf{c}))$ *whenever* $k_l \ne 0$. *The strict sign constraint assumes that a reaction cannot be fully suppressed by allosteric regulation or by the lack of enzyme (it will still proceed uncatalysed, possibly with a low rate). Again, metabolite levels may be predefined by constraints.*

The metabolic state manifold can be constructed based on the possible sign patterns. We first define all possible force patterns $\boldsymbol{\sigma}$ (i.e. the possible sign patterns of thermodynamic forces)

$$S_{\boldsymbol{\sigma}} = \{\boldsymbol{\sigma} : \exists \mathbf{c} : \boldsymbol{\sigma} = \mathrm{sign}(\boldsymbol{\theta}(\mathbf{c}))\}. \tag{S2}$$

Each force pattern $\boldsymbol{\sigma}$ corresponds to a patch

$$\begin{aligned} S_{\mathbf{c}}^{(\boldsymbol{\sigma})} &= \{\mathbf{c} : \mathrm{sign}(\boldsymbol{\theta}(\mathbf{c})) = \boldsymbol{\sigma} \wedge \mathbf{c}_{\rm min} \le \mathbf{c} \le \mathbf{c}_{\rm max}\} \\ S_{\mathbf{v}}^{(\boldsymbol{\sigma})} &= \{\mathbf{v} : \mathrm{sign}(\mathbf{v}) \sqsubseteq \boldsymbol{\sigma} \wedge \mathbf{N}_{\rm int}\,\mathbf{v} = 0\} \end{aligned} \tag{S3}$$

and a sheet

$$S_{(\mathbf{v},\mathbf{c},\mathbf{e})}^{(\boldsymbol{\sigma})} = \{(\mathbf{v}, \mathbf{c}, \mathbf{e}) : \mathbf{v} \in S_{\mathbf{v}}^{(\boldsymbol{\sigma})} \wedge \mathbf{c} \in S_{\mathbf{c}}^{(\boldsymbol{\sigma})} \wedge \mathrm{Dg}(\mathbf{e})\,\mathbf{k}(\mathbf{c}) = \mathbf{v})\}. \tag{S4}$$

The patches for different flux patterns are disjoint, and the sheets for different flux patterns are disjoint, too. In Eq. (S4), the third condition can always be satisfied by a choice of non-negative enzyme levels (this is guaranteed by the previous sign conditions). The condition completely determines the enzyme levels, unless a reaction rate v_l vanishes (in this case, e_l can have any value). In Eqs (S3)-(S4), we can also use an alternative description with loose (instead of strict) flux patterns ("topes"). We define the set of feasible loose flux patterns $\boldsymbol{\tau}$,

$$S_{\boldsymbol{\tau}} = \{\boldsymbol{\tau} : \exists \boldsymbol{\sigma} \in S_{\boldsymbol{\sigma}} : \boldsymbol{\tau} \sqsubseteq \boldsymbol{\sigma}\}, \tag{S5}$$

i.e. the possible signs of catalytic rates that agree with possible thermodynamic forces. Again, each sign pattern represents a patch and a sheet

$$\begin{aligned} S_{\mathbf{c}}^{(\boldsymbol{\tau})} &= \{\mathbf{c} : \boldsymbol{\tau} \sqsubseteq \mathrm{sign}(\boldsymbol{\theta}(\mathbf{c})) \wedge \mathbf{c}_{\rm min} \le \mathbf{c} \le \mathbf{c}_{\rm max}\} \\ S_{\mathbf{v}}^{(\boldsymbol{\tau})} &= \{\mathbf{v} : \mathrm{sign}(\mathbf{v}) \sqsubseteq \boldsymbol{\tau} \wedge \mathbf{N}_{\rm int}\,\mathbf{v} = 0\} \\ S_{(\mathbf{v},\mathbf{c},\mathbf{e})}^{(\boldsymbol{\tau})} &= \{(\mathbf{v}, \mathbf{c}, \mathbf{e}) : \mathbf{v} \in S_{\mathbf{v}}^{(\boldsymbol{\tau})} \wedge \mathbf{c} \in S_{\mathbf{c}}^{(\boldsymbol{\tau})} \wedge \mathrm{Dg}(\mathbf{k}(\mathbf{c}))\,\mathbf{e} = \mathbf{v})\}. \end{aligned} \tag{S6}$$

Unlike the patches and sheets before, these patches and sheets may overlap and may be contained in each other.

S1.3 State manifold and the flux/metabolite states are closely related

The state manifold $\mathcal{K}$ and the set of thermodynamically feasible states, $\mathcal{T}$, are closely related, as stated by the following proposition.

Proposition 1 Metabolic state manifold and set of feasible flux/metabolite states *(proof in section S3.1). $\mathcal{T}$ is the projection of $\mathcal{K}$ into $(\mathbf{v}, \mathbf{c})$-space. This means: all kinetically feasible states are thermodynamically feasible, and all thermodynamically feasible flux-metabolite profiles can be kinetically realised. The projection is bijective in almost all points. The only exceptions are states in which a reaction is in thermodynamic equilibrium (i.e. $k_l = 0$ for some reactions l). In these points, the projection is bijective,* except for the enzyme levels of the equilibrium reactions*, which can assume any value. If we assume, as an extra heuristic principle, that cells switch off unnecessary enzymes ("repression of unused enzyme assumption"), the undetermined enzyme levels are set to 0, and the mapping between states in $(\mathbf{v}, \mathbf{x})$-space and $(\mathbf{v}, \mathbf{x}, \mathbf{e})$-space becomes bijective.*

The proposition implies: if we assume that all reactions are enzyme-catalysed, and that there are no bounds on enzyme levels, all thermo-physiologically feasible flows can be kinetically realised.

S1.4 Topology and dimensionality of the state manifold

Proposition 2 (Dimensionality of the state manifold) *(proof in section S3.1) We consider a model with $n_{\mathrm{m}} = n_{\mathrm{int}} + n_{\mathrm{ext}}$ metabolites (n_{int} internal, n_{ext} external) and n_{r} reactions, where all reactions are enzyme-catalysed (no spontaneous reactions) and all enzymes are specific (no promiscuous activity, no isoenzymes). Under this assumption, there are n_{r} enzymes, and the $(\mathbf{v}, \mathbf{c}, \mathbf{e})$-space has $2\,n_{\mathrm{r}} + n_{\mathrm{m}}$ dimensions. The metabolic state manifold $\mathcal{K}$ is a differentiable manifold in this space and is $n_{\mathrm{r}} + n_{\mathrm{ext}}$-dimensional in almost all points (exceptions are explained in section S3.1). If all external metabolite levels are predefined, the metabolic state manifold is an n_{r}-dimensional manifold in a $2n_{\mathrm{r}} + n_{\mathrm{int}}$-dimensional $(\mathbf{v}, \mathbf{c}, \mathbf{e})$-space.*

Proposition 3 (Topology of the state manifold) *(proof in section S3.1) The metabolic state manifold is path-connected, i.e., the cell can continuously move between any two points of the state manifold without leaving the manifold.*

Remark If "dead" states (with all fluxes being zero) are excluded or if a predefined flux benefit must be reached, the resulting constrained state manifold may not be path-connected.

S2 Optimality problems in flux or metabolite space

S2.1 Screening of metabolic states

An important lesson from the state manifold is that metabolic states can be parameterised in two ways: either by enzyme levels and conserved moiety concentrations, or by steady-state fluxes and metabolite concentrations. These two descriptions are equivalent. In the one case, if enzyme levels and conserved moiety concentrations are known, steady-state fluxes and concentrations can be computed by numerical integration. In the other case, if the steady-state fluxes and concentrations are known (which must be thermo-physiologically feasible), the enzyme levels and conserved moiety concentrations can be easily computed. These two choices of basic variables lead to two ways of screening the metabolic states. The first approach, screening the enzyme profiles and conserved moiety concentrations, is less practical because the calculation of steady-state variables is numerically expensive. The second approach can be performed step by step, by screening (i) all feasible flux patterns; (ii) all flows with a given flux pattern; (iii) and all metabolite profiles compatible with a given flux pattern; in the end, the enzyme levels can be computed easily (see Figure S1). Steps (ii) and (iii) can also be switched: we screen the space of metabolite levels, and for each choice, the possible fluxes. An advantage of the stepwise screening in flux and metabolite space is that thermodynamic correctness is ensured automatically.

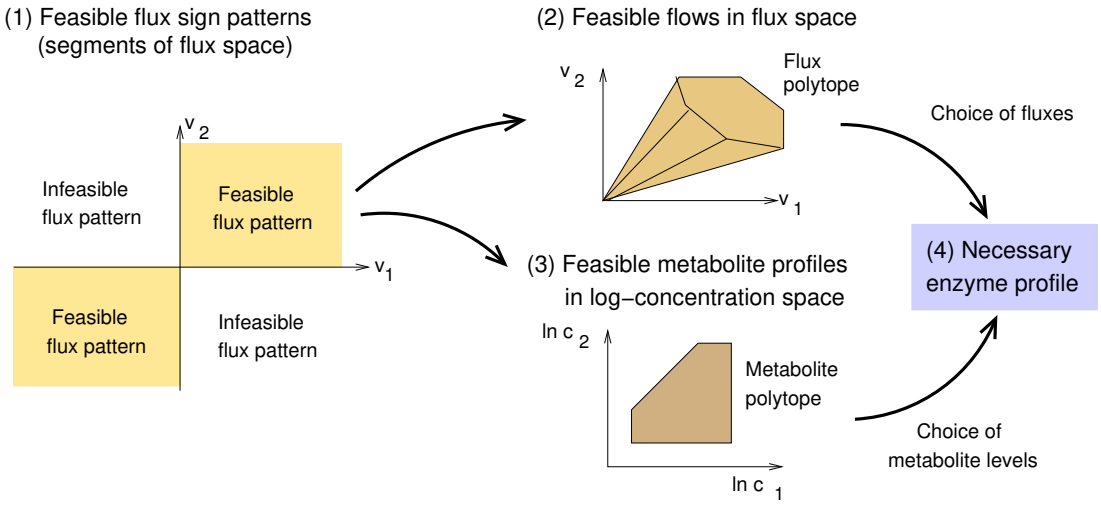

Figure S1: Screening the possible metabolic states of a kinetic model. The metabolic states (characterised by flux directions, flux magnitudes, metabolite levels, and enzyme levels) can be screened in three steps, by a layered procedure. (1) Screening of flux patterns (active reactions and their flux directions). Feasible flux patterns must allow for stationary, thermodynamically feasible flux distributions, given the external metabolite concentrations assumed in the model. (2) Screening of flux distributions ("flows") in each flux polytope. Given a flux pattern, stationarity, flux bounds, and maybe a fixed flux benefit, the possible flux distributions form a convex polytope. If there is only one flux bound, the polytope is spanned by elementary flux modes (EFMs). In theory the EFMs can be enumerated, but in practice they may be numerous. (3) Screening of metabolite profiles in the metabolite polytope. For each flow pattern, we obtain a metabolite polytope (defined by physiological concentration bounds and by thermodynamic constraints). With a given flux distribution from (2) and metabolite profile from (3), the required enzyme levels are easy to compute. If the fluxes are given, the total enzyme cost is a convex function on the metabolite polytope. In contrast, if the metabolite profile is given, the total enzyme cost is a linear function on the flux polytope.

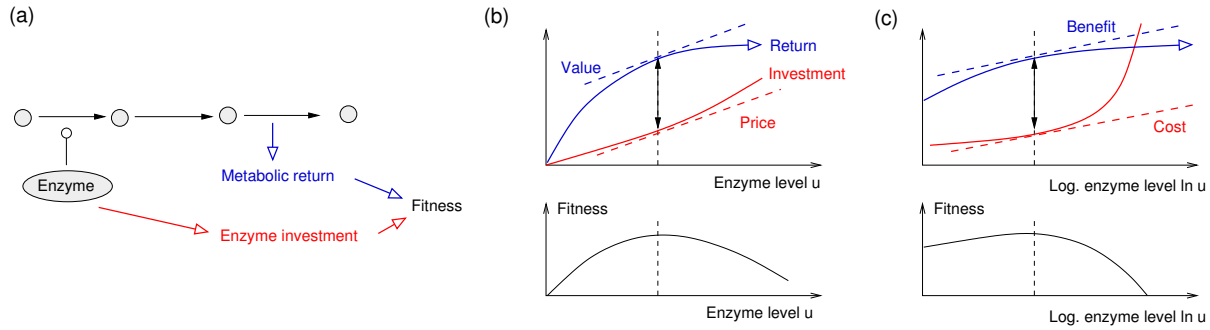

Figure S2: An optimality problem for cells' choices of enzyme concentrations. (a) Cost and benefit of an individual enzyme. By catalysing a reaction, an enzyme has an indirect impact on the metabolic fluxes and concentrations everywhere in the network. The metabolic steady state determines a metabolic objective (metabolic performance function), and the enzyme level entails an cost. The benefit is determined by scoring a metabolic flux which is indirectly supported by our enzyme. Our objective function, called fitness, is defined as the difference of metabolic performance and cost. (b) The slopes of metabolic performance and cost curves with respect to the enzyme level are called enzyme value and enzyme price, respectively. At the point of maximal fitness ($\partial g/\partial u = \partial h/\partial u$) they have equal numerical values. (c) The slopes with respect to the logarithmic enzyme level are called partial enzyme benefit and cost, respectively. Again, both slopes are balanced in the point of maximal fitness ($\partial g/\partial \ln u = \partial h/\partial \ln u$). This leads to the flux benefit equation $g^{\mathrm{v}}\, v = h_{\mathrm{u}}\, u$ (see [1]).

S2.2 Optimisation in flux/metabolite space

(see Figure 5 (a)). In the first method, we assume a fixed flow $\mathbf{v}$ and perform a loop of *enzyme cost minimisation in metabolite space*: we determine the metabolite profile $\mathbf{c}$ that minimises the cost $q^{\mathrm{kin}}(\mathbf{v}, \mathbf{c}) = q(\mathbf{c}) + h(\mathbf{e}(\mathbf{v}, \mathbf{c}))$.

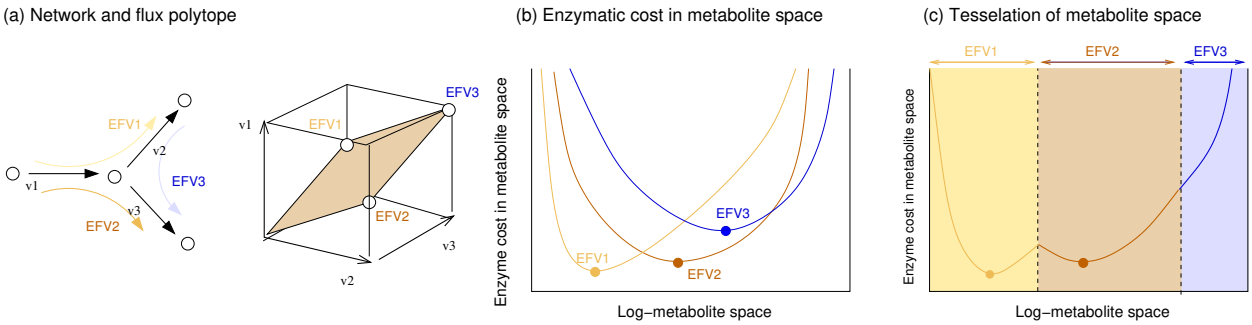

Figure S3: (a) Left: metabolic branch point model with three EFMs. Right: flux polytope spanned by three EFMs. (b) By taking the minimum value in each point, we obtain the flux-optimised enzymatic cost in metabolite space, a non-convex function. (c) Resulting tesselation of the M-polytope. Each region (marked by different background colors) belongs to one EFM and contains the metabolite profiles that favour this EFM. A minimum point from (b) can either remain a local minima (solid dots) or become irrelevant. In the example shown, the brown and yellow regions form the basis of attraction of one locally optimal state (brown dot).

The optimal value $\min_{\mathbf{c}} q^{\mathrm{kin}}(\mathbf{v}, \mathbf{c})$ is called kinetic flux cost $a^{\mathrm{kin}}(\mathbf{v})$ of our flux distribution $\mathbf{v}$. As mentioned above, it can be assigned to the flux distribution as an effective cost function, representing "hidden" enzyme and metabolite costs. Then, in an outer loop of *flux optimisation*, we consider this cost function and optimise the fluxes for a maximal fitness $f(\mathbf{v}) = b(\mathbf{v}) - a^{\mathrm{kin}}(\mathbf{v})$ (flux cost-benefit optimisation), or we minimise $a^{\mathrm{kin}}(\mathbf{v})$ at a predefined flux benefit $b(\mathbf{v}) = b'$ (*flux cost minimisation*). This outer minimisation problem resembles minimal-flux FBA, in which the sum of fluxes is used as a simple approximation of the kinetic flux cost. Around this entire optimisation, there can be another round of optimisation, over all possible flow patterns.

(see Figures S3 and 5 (b)). In our second calculation method, for cost-optimal states, we screen the M-polytope and evaluate the flux-optimal enzymatic M-cost for each point. That is, instead of choosing a flow $\mathbf{v}$ and optimising the enzymatic cost in the M-polytope, we do exactly the opposite: we fix a metabolite profile $\mathbf{m}$ and write the enzymatic cost as a function of the fluxes: $\sum h_{e_l} e_l = \sum \frac{h_{e_l}}{r_l(\mathbf{c})} v_l$, where $r_l(\mathbf{c})$ has a fixed value. This is a linear function in $\mathbf{v}$, and optimising this function in flux space is a weighted flux minimisation problem! We already know that the optimum points must be vertices of the flux polytope. Therefore, each metabolite profile $\mathbf{m}$ leads to an optimal ("favoured") flow, a polytope corner, with a corresponding optimal enzyme cost. If we now vary the metabolite profile and repeat the procedure, we obtain a flux-optimised cost function on the M-polytope. This function can also be non-convex. To see this, we write it as $q(\mathbf{m}| * \mathbf{v}) = \min_{\mathbf{v} \in \mathcal{V}} q^{\mathrm{enz}}(\mathbf{m}; \mathbf{v})$, where $\mathcal{V}$ is the set of flux polytope vertices ("candidate flows"). That is, , for each point of the M-polytope, we consider all possible candidate flows and choose the one with the minimal metabolite-constrained cost (Figure S3 (a)). The resulting cost is a piecewise convex function function on the M-polytope: to construct it, we can consider the (convex) cost functions of all candidate flows and take their minimum value. This pointwise minimum yields a piecewise convex (i.e. possibly non-convex) function (Figure S3 (b)).

S2.3 Algorithm for testing and finding locally optimal states

The criterion for locally optimal states leads to a simple algorithm for testing or constructing such states. We know that locally optimal flows must be corners of a benefit-constrained flux polytope (we shall call these corners "candidate flows").

1. **Testing a given flow** To test a candidate flow $\mathbf{v}$ for being locally optimal, we first compute its favoured metabolite profile $\mathbf{x}^{\mathrm{opt}}$. Then we evaluate the enzymatic cost of this metabolite profile for all other candidate flows $\mathbf{v}^{(k)}$. If the cost $q^{\mathrm{enz}}(\mathbf{x}^{\mathrm{opt}}, \mathbf{v})$, obtained from our original flow $\mathbf{v}$, is lower than the costs $q^{\mathrm{enz}}(\mathbf{x}^{\mathrm{opt}}, \mathbf{v}^{(k)})$

obtained from all other candidate flow, then our flow $\mathbf{v}$, together with its favoured metabolite profile and the resulting enzyme profile, represents a locally optimal state. By applying this test to all B-polytope corners $\mathbf{v}$, we can enumerate all locally optimal states. In practice, we can do this by enumerating all B-polytope corners, computing their favoured metabolite profiles, and evaluating the enzymatic cost for each possible pair. From the resulting matrix, we can easily tell which pairs are locally optimal.

2. **Constructing a locally optimal flow** To construct a locally optimal flow from scratch, we start from a given flow $\mathbf{v}$ (which need not be a polytope corner) and iteratively compute the optimal metabolic profiles and optimal flows until convergence[1]. The result is a locally optimal flow.

S3 Proofs

S3.1 Propositions 1, 2, and 3

Proposition 1, Projection: The projection of the state manifold $\mathcal{K}$ onto $(\mathbf{v}, \mathbf{c})$-space yields the set $\text{Proj}(\mathcal{K}) = \{(\mathbf{v}, \mathbf{c}) : \exists \mathbf{e} : \mathbf{N}_{\text{int}}\, \mathbf{v} = 0; \mathbf{v} = \text{Dg}(\mathbf{e})\, \mathbf{k}(\mathbf{c}); 0 \leq \mathbf{c} \leq \mathbf{c}_{\max}; 0 \leq \mathbf{e}\} = \mathcal{T}$. By setting $e_l = v_l / k_l(\mathbf{c})$ (or $e_l = 0$ whenever $v_l = 0$), we can satisfy all conditions.

Proposition 1, Bijective projection: To show that the projection is bijective (whenever $v_l \neq 0$), we need to check that two states $(\mathbf{v}_A, \mathbf{c}_A, \mathbf{e}_A)$ and $(\mathbf{v}_B, \mathbf{c}_B, \mathbf{e}_B)$ are identical if they project to the same pair $(\mathbf{v}, \mathbf{c})$, In fact, if $(\mathbf{v}_A, \mathbf{c}_A, \mathbf{e}_A)$ and $(\mathbf{v}_B, \mathbf{c}_B, \mathbf{e}_B)$ project to $(\mathbf{v}, \mathbf{c})$, we obtain $\mathbf{v}_A = \mathbf{v}_B = \mathbf{v}$ and $\mathbf{c}_A = \mathbf{c}_B = \mathbf{c}$, and thus $\mathbf{v} = \text{Dg}(\mathbf{e}_A)\, \mathbf{k}(\mathbf{c})$ and $\mathbf{v} = \text{Dg}(\mathbf{e}_B)\, \mathbf{k}(\mathbf{c})$. By taking the difference, we obtain the condition $0 = \text{Dg}(\mathbf{e}_A - \mathbf{e}_B)\, \mathbf{k}(\mathbf{c})$, which implies that $\forall l : k_l(\mathbf{c}) \neq 0 \Rightarrow e_{A,l} = e_{B,l}$.

Proposition 2: The $(\mathbf{v}, \mathbf{c}, \mathbf{e})$-space has $2n_{\text{r}} + n_{\text{m}}$ dimensions. The conditions $\mathbf{N}_{\text{int}}\, \mathbf{v} = 0$ and $\mathbf{v} = \text{Dg}(\mathbf{e})\, \mathbf{k}(\mathbf{c})$ are differentiable and define $n_{\text{int}} + n_{\text{r}}$ equality constraints. Therefore, the manifold has $n_{\text{r}} + n_{\text{ext}}$ degrees of freedom, except for points in which some of these constraints are linearly dependent (in these points, the dimensionality of the manifold could be higher). However, such points are rare: in such points, the matrix

$$\begin{pmatrix} \mathbf{N}_{\text{int}} & 0 & 0 \\ -\mathbf{I} & \text{Dg}(\mathbf{e})\, \frac{\partial \mathbf{k}}{\partial \mathbf{c}} & \text{Dg}(\mathbf{k}(\mathbf{c})) \end{pmatrix} \tag{S7}$$

must have linearly dependent rows. This requires that $\mathbf{k}(\mathbf{c}) = 0$ (thermodynamic equilibrium in all reactions) and that the elasticity matrix $\mathbf{E} = \text{Dg}(\mathbf{e})\, \frac{\partial \mathbf{k}}{\partial \mathbf{c}}$ is rank-deficient. This is unlikely to happen by chance, but it may happen in pathological cases, for example if a reaction (including its enzymatic rate law) has been deliberately duplicated in a model.

Proposition 3: Consider two points $(\mathbf{v}_A, \mathbf{c}_A, \mathbf{e}_A)$ and $(\mathbf{v}_B, \mathbf{c}_B, \mathbf{e}_B)$. Starting in the first point, we can continuously decrease all enzyme levels and fluxes by scaling them with a common prefactor until we reach the point $(0, \mathbf{c}_A, 0)$. Then we move continuously to the point $(0, \mathbf{c}_B, 0)$ (which is possible without violating any constraints). Finally, we scale up the fluxes and enzyme levels to reach the point $(\mathbf{v}_B, \mathbf{c}_B, \mathbf{e}_B)$.

S3.2 Condition for multi-objective optimality

Multi-objective ("Pareto") optimality problems, with objectives $f_1(\mathbf{x}), f_1(\mathbf{x}), ..$ lead to the same optimality conditions as simple optimality problems. To show this, we assume that the Pareto front is a manifold in $(f_1, f_2, ..)$-space, whose shape can be described by an equation $g(f_1, f_2, ..) = 0$ with a smooth function g (or that it consists

[1] If all polytope corners have different costs, the algorithm will converge to a single flow, because there is only a finite number of polytope corners, and a decrease of the overall objective in each step. If two flows have the same objective value and the same optimal metabolite profile, the iteration may end up jumping between these flows.

of a number of such manifolds). We assume that g increases with each of the objectives f_i (assuming that all metabolic objectives are meant to be maximised), so its gradient $\nabla_{\mathbf{f}} g = (\frac{\partial g}{\partial f_1}, \frac{\partial g}{\partial f_2}, ..)^\top$ has only positive elements. When moving along the Pareto front, it is impossible to move towards increasing values of g, i.e., in the direction of the gradient $\nabla_{\mathbf{f}} g$. Instead, any such movement $\mathrm{d}\mathbf{f}$ (realisable by movements $\mathrm{d}\mathbf{x}$ in $\mathbf{x}$-space) will be perpendicular to $\nabla_{\mathbf{f}}\, g$, i.e., $\nabla_{\mathbf{f}} g \cdot \mathrm{d}\mathbf{f} = 0$. More general movements $\mathrm{d}\mathbf{f}$ in $(f_1, f_2, ..)$-space can be described as $\mathrm{d}\mathbf{f} = \mathbf{F}_{\mathbf{x}} \mathrm{d}\mathbf{x}$, with the Jacobian matrix $\mathbf{F}_{\mathbf{x}} = \partial \mathbf{f} / \partial \mathbf{x}$. $\mathrm{d}\mathbf{x}$. By combining the two equations, we obtain the condition for possible movements, $0 = \sum_i \frac{\partial g}{\partial f_i} \nabla_{\mathbf{x}} f_i \, \mathrm{d}\mathbf{x}$. Since this condition must hold for *any* movements $\mathrm{d}\mathbf{x}$, we obtain the condition for Pareto-optimal points

$$0 = \sum_i \frac{\partial g}{\partial f_i} \nabla_{\mathbf{x}} f_i(\mathbf{x}). \tag{S8}$$

In other words: in each Pareto-optimal point, some positive linear combination of the gradients $\nabla_{\mathbf{x}} f_i$ must vanish. This shows that Pareto-optimal solutions are also solutions to single-objective maximisation problems, namely maximising $\sum_i \varphi_i \nabla_{\mathbf{x}} f_i(\mathbf{x})$. The prefactors are positive, $\varphi_i = \frac{\partial g}{\partial f_i}$, in multi-objective problems in which all objectives represent benefits (i.e., they are meant to be large). If objective values represent costs (i.e., are meant to be small), the corresponding prefactor $\varphi_i = \frac{\partial g}{\partial f_i}$ is negative.

S3.3 Metabolite balance equation (8)

The balance equation (8) relates the price of a metabolite to the prices of the surrounding enzymes. To derive the balance equation, we start from the optimality condition (9). After rescaling the cost functions, we obtain

$$\begin{aligned} 0 &= -\nabla_{\ln \mathbf{c}}\, q(\mathbf{c}) - \nabla_{\ln \mathbf{c}}\, h(\mathbf{e}(\mathbf{c}, \mathbf{v})) = -\mathbf{q}_{\mathrm{c}}^\top \,\mathrm{Dg}(\mathbf{c}) - \mathbf{h}_e^\top \frac{\partial \mathbf{e}}{\partial \mathbf{c}} \,\mathrm{Dg}(\mathbf{c}) \\ &= -\mathbf{q}_{\mathrm{c}}^\top \,\mathrm{Dg}(\mathbf{c}) - \mathbf{h}_e^\top \mathrm{Dg}(\mathbf{v}) \frac{\partial 1/\mathbf{v}}{\partial \mathbf{c}} \,\mathrm{Dg}(\mathbf{c}) = -\mathbf{q}_{\mathrm{c}}^\top \,\mathrm{Dg}(\mathbf{c}) - \mathbf{h}_e^\top \mathrm{Dg}(\mathbf{v}) \frac{-1}{\mathbf{v}^2} \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \,\mathrm{Dg}(\mathbf{c}) \\ &= -\mathbf{q}_{\mathrm{c}}^\top \,\mathrm{Dg}(\mathbf{c}) + \mathbf{h}_e^\top \mathrm{Dg}(\mathbf{e}) \frac{1}{\mathbf{v}} \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \,\mathrm{Dg}(\mathbf{c}) = -\mathbf{q}_{\mathrm{c}}^\top \,\mathrm{Dg}(\mathbf{c}) + \mathbf{h}_e^\top \mathrm{Dg}(\mathbf{e})\, \mathcal{E}_{\mathrm{c}}^{\mathrm{v}}. \end{aligned} \tag{S9}$$

Altogether, we obtain

$$\mathrm{Dg}(\mathbf{c})\, \mathbf{q}_{\mathrm{c}} = \mathcal{E}_{\mathrm{c}}^{\mathrm{v}\top} \mathrm{Dg}(\mathbf{e})\, \mathbf{h}_e. \tag{S10}$$

This is the metabolite balance equation in "point cost form". To obtain the balance equation in "value form", we can rewrite it as

$$\mathbf{q}_{\mathrm{c}} = \mathrm{Dg}(\mathbf{c})^{-1}\, \mathcal{E}_{\mathrm{c}}^{\mathrm{v}\top} \mathrm{Dg}(\mathbf{v})\, \mathrm{Dg}(\mathbf{v})^{-1}\, \mathrm{Dg}(\mathbf{e})\, \mathbf{h}_e = \mathbf{E}_{\mathrm{c}}^\top\, \mathbf{E}_{\mathrm{e}}^{-1}\, \mathbf{h}_e, \tag{S11}$$

where we assumed a one-to-one relationship between reactions and enzymes, and thus an enzyme elasticity matrix given by $\mathbf{E}_{\mathrm{e}} = \mathrm{Dg}(\mathbf{v})\, \mathrm{Dg}(\mathbf{e})^{-1}$.

S3.4 Balance equation for enzyme optimisation Eq. (S13)

We start from the optimality condition

$$0 = \underbrace{\nabla_{\mathbf{e}}\, b(\mathbf{v}^{\mathrm{st}}(\mathbf{e}))}_{\mathbf{b}_{\mathrm{v}}^{\star\top}\, \mathbf{C}^{\mathrm{V}}\, \mathbf{E}_{\mathrm{e}}} - \underbrace{\nabla_{\mathbf{e}}\, q(\mathbf{c}^{\mathrm{st}}(\mathbf{e}))}_{\mathbf{q}_{\mathrm{c}}{}^\top\, \mathbf{C}^{\mathrm{S}}\, \mathbf{E}_{\mathrm{e}}} - \underbrace{\nabla_{\mathbf{e}}\, h(\mathbf{e})}_{\mathbf{h}_e^\top}, \tag{S12}$$

with flux gain $\mathbf{b}_{\mathrm{v}}^{\star} = \nabla_{\mathbf{v}} b$, metabolite price $\mathbf{q}_{\mathrm{c}} = \nabla q$, control coefficient matrices $\mathbf{C}^{\mathrm{S}}$ and $\mathbf{C}^{\mathrm{V}}$, and enzyme elasticity matrix $\mathbf{E}_{\mathrm{e}}$. We can also rewrite this as

$$[\mathbf{b}_{\mathrm{v}}^{\star\top}\,\mathbf{C}^{\mathrm{V}} - \mathbf{q}_{\mathrm{c}}^{\top}\,\mathbf{C}^{\mathrm{S}}]\,\mathbf{E}_{\mathrm{e}} = \mathbf{h}_{e}^{\top}, \tag{S13}$$

our reaction balance equation (11). Knowing that $\mathbf{C}^{\mathrm{V}} = \mathbf{I} + \mathbf{E}_{\mathrm{c}}\,\mathbf{C}^{\mathrm{S}}$ and assuming one enzyme per reaction and one reaction per enzyme (i.e., $E_{e_l} = v_l/e_l$), we can write this as

$$(\mathbf{b}_{\mathrm{v}}^{\star\top} - \underbrace{[\mathbf{b}_{\mathrm{v}}^{\star\top}\,\mathbf{E}_{\mathrm{c}} - \mathbf{q}_{\mathrm{c}}{}^{\top}]}_{\mathbf{q}_{\mathrm{c}}^{\mathrm{eff}}}\,\mathbf{C}^{\mathrm{S}})\,\mathrm{Dg}(\mathbf{v}) = \mathbf{h}_{e}^{\top}\,\mathrm{Dg}(\mathbf{e}), \tag{S14}$$

and further as

$$\mathbf{b}_{\mathrm{v}}^{\star} - \mathbf{C}^{\mathrm{S}}\,\mathbf{q}_{\mathrm{c}}^{\mathrm{eff}} = \mathbf{E}_{\mathrm{e}}^{-1}\,\mathbf{h}_{e}. \tag{S15}$$

References

[1] W. Liebermeister. Metabolic economics in kinetic models. *Preprint on arXiv.org: arXiv:1404.5252*, 2014.