2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Natural host feedback simplifies the design of metabolic switches

Alexander P.S. Darlington[1], Ahmad A. Mannan[2] and Declan G. Bates[1]

***Abstract*— The performance of microbial chemical production can be improved by incorporating inducible synthetic gene circuits which 'switch' the microbial cell factories from growth to production. Here, we consider the design of an inducible switch, implemented as a small-scale gene regulatory network, in the context of host processes. We show that by accounting for the non-regulatory interactions which arise between host and circuit processes due to natural metabolic and ribosomal constraints the design of the gene circuit can be simplified with little cost to performance. We show that whilst optimal performance is achieved by engineering the full three-gene controller, a partial system composed of fewer regulated genes can still achieve near optimal performance. This may allow for engineering controllers in living cells with fewer time consuming and expensive experimental steps.**

I. INTRODUCTION

Engineering metabolic pathways in microorganisms provides a route to the manufacture of commodity and high value chemicals from sustainable feedstocks. However, engineered processes impose a "burden" on the cell: engineered microbes show reduced growth rate due to both (i) co-option of the cell's ribosomes for new enzyme production and (ii) diversion of metabolic fluxes from key growth processes to the new metabolic product. These processes are interlinked as gene expression utilises metabolic resources (e.g. ATP, amino acids) and metabolic precursors are generated by enzymes, which themselves are produced by gene expression resources (Fig. 1). These interactions limit the production efficiency of microbial cell factories by reducing volumetric productivity (the total product synthesised per unit total batch culture time, limited by poor growth) and yield (the substrate to product conversation ratio, limited due to host cell production). Temporally separating growth and production processes enables improved performance of microbial cell factories (reviewed in [1]). This can be achieved using synthetic genetic control systems (depicted as blue blocks in Fig. 1) which activate the engineered process (the green blocks of E') and concurrently inhibit host processes (e.g. deactivating E) at some optimal batch culture time [2]. To date, a number of simple systems have been designed and successfully engineered into living cells (reviewed in [3]). However, how to construct the circuits to maximise the production performance in light of the complex constraints imposed by host metabolic and gene expression interactions remains an open question.

Gene expression in *E. coli* is known to be limited by the availability of free ribosomes (e.g. [4]) which results in the emergence of non-regulatory interactions between genes. For example, as expression of one gene increases, the expression of others must fall [5]. These interactions can complicate synthetic circuit design [6], [7]. Recent results also demonstrate a complex interplay between the cell's metabolism and expression of synthetic genes (e.g. [8]). These resource limitations create an additional layer of feedback, which we term natural feedback, but is often neglected during controller design. In this paper, we analyse the performance of a simple growth-production switch in the presence of natural feedback mediated by the host's processes. In Section II we develop multi-scale models of the controller in the absence and presence of these natural feedbacks. In Section III we demonstrate that natural feedback can enhance the production of an engineered pathway. In Section IV, we use multi-objective optimisation to design controllers which maximise key industrial performance metrics (yield and volumetric productivity). We then show that natural host feedback can simplify the design of the proposed control system by enabling near optimal performance to be achieved by regulating fewer genes.

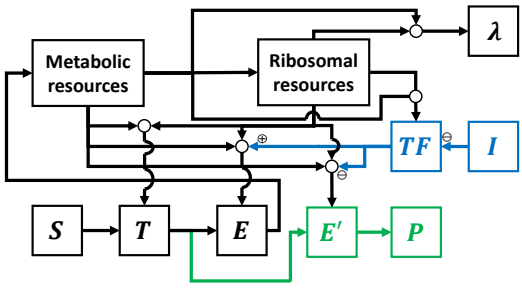

Fig. 1. Block diagram of host-pathway interactions. Host processes are shown in black, the synthetic pathway is highlighted in green and control system depicted in blue. The extracellular substrate S is imported by transporters T and converted, by enzyme E, to metabolic resources. The engineered process produces P via enzyme E'. The controller input (inducer) I acts via the transcription factor TF on E and E'.

II. CONTEXT-BASED MODEL OF THE GENETIC SWITCH

A. Model of the pathway and regulatory switch

We consider a simple dynamic model of metabolism where an extracellular substrate S is imported into the cell at rate v_{uptake} to form a pool of internalised substrate s_{in}. This substrate is then utilised for growth processes at rate v_{host} (e.g. energy production) or the engineered process at rate

Funding to AD from Royal Academy of Engineering under the Research Fellowship scheme. Funding to AM, DB and high performance computing facilities provided by Warwick Integrative Synthetic Biology Centre from BBSRC grant BB/M017982/1. For the purpose of open access, the authors has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising. [1]School of Engineering, University of Warwick, Coventry, UK. [2]Department of Bioengineering, Imperial College London, London, UK (correspondence to a.darlington.1@warwick.ac.uk)

v_{prod}. Growth at rate λ leads to dilution of all species. Applying the law of mass action, the substrate s_{in} and product p_{in} dynamics within the cell are:

$$\dot{s}_{in} = v_{uptake}(S, p_T) - v_{host}(s_n, p_E) - v_{prod}(s_{in}, p_{E'}) - \lambda \cdot s_{in} \quad (1)$$

$$\dot{p}_{in} = v_{prod}(s_{in},\ p_{E'}) - \lambda \cdot p_{in} \quad (2)$$

We model the reaction rates with Michaelis-Menten kinetics, where the rate of reaction catalysed by p_Y is:

$$v_{rxn}(x, p_Y) = \frac{v_Y \cdot x \cdot p_Y}{\kappa_Y + x} \quad (3)$$

where x is the substrate concentration and p_Y is the concentration of the catalysing enzyme with a turnover number of v_Y and Michaelis constant of κ_Y. We denote the enzyme of the engineered production pathway as E'.

The gene expression model of protein Y is composed of an mRNA (m_Y), which is born spontaneously at rate $T_X(e,\ p_{TF})$ (where e is the cell's internal energy supply and p_{TF} is the regulating transcription factor (see Eq. 10). The mRNA (reversibly) binds to free ribosomes R to produce translation complexes c_X. These undergo translation (protein birth) at rate $T_L(c_Y,\ e)$ to produce proteins (p_Y). All species dilute due to cell growth at rate λ and mRNAs are also subject to decay at rate δ_m. Applying the law of mass action, we model the dynamics of mRNA, translation complex and proteins as:

$$\dot{m}_Y = T_X(e, p_{TF}) - b_Y \cdot R \cdot m_Y + u_Y \cdot c_Y - (\lambda + \delta_m) \cdot m_Y, \quad (4)$$

$$\dot{c}_Y = b_Y \cdot R \cdot m_Y - u_Y \cdot c_Y - T_L(c_Y,\ e) - \lambda \cdot c_Y, \quad (5)$$

$$\dot{p}_Y = T_L(c_Y,\ e) - \lambda \cdot p_Y. \quad (6)$$

Note that for ease of comparision to the host model described in Section II-D, we do not reduce the dimension of this model by assuming the quasi-steady states as is typically done.

We consider a genetic switch system based on a single transcription factor which regulates both host enzymes, E, and pathway enzymes, E', as shown in Fig. 1. This transcription factor exists in two states: p_{TF} (the (functional) DNA-binding form) and p_{TFc} (the non-functional inert complex bound by two molecules of the inducer I_{in}). The inducer is introduced as I which is then internalised at rate v_I to form I_{in}. The switch dynamics are:

$$\dot{p}_{TF} = T_L(c_{TF}, e) - \lambda \cdot p_{TF} - k_{I,f} \cdot I_{in}^2 \cdot p_{TF} + k_{I,r} \cdot p_{TFc} \quad (7)$$

$$\dot{p}_{TFc} = k_{I,f} \cdot I_i^2 \cdot p_{TF} - k_{I,r} \cdot p_{TFc} - \lambda \cdot p_{TFc} \quad (8)$$

$$\dot{I}_{in} = v_I(I) - 2 \cdot k_{I,f} \cdot I_{in}^2 \cdot p_{TF} + 2 \cdot k_{I,r} \cdot p_{TFc} - \lambda \cdot I_{in} \quad (9)$$

where $k_{I,f}$ and $k_{I,r}$ are the rates of I_{in} binding and unbinding from the transcription factor species p_{TF} and p_{TFc}, respectively. $v_I(I)$ captures the internalisation of the inducer and is described in Section II-B.

The core controller motif is based on a dual regulation scheme, as depicted in Fig. 1, where the transcription factor activates expression of the host p_E and inhibits expression of the pathway $p_{E'}$. Upon addition of I, the TF is sequestered causing p_E transcription to fall (due to loss of activation) and $p_{E'}$ transcription to rise (due to loss of repression). The regulation is captured in the $T_X(\cdot)$ transcription rate function.

The transcription rate $T_X(e,\ p_{TF})$ of a regulated gene is a nonlinear function of the cell's internal energy (e) and regulation by p_{TF}, where p_{TF} can be either an activator or inhibitor:

$$T_X(e,\ p_{TF}) \quad (10)$$
$$= \begin{cases} \theta_Y \cdot \left(\dfrac{\omega_Y \cdot e}{\pi_Y + e}\right) \cdot \left(\omega_0 + \dfrac{K_Y \cdot p_{TF}}{K_Y \cdot p_{TF} + 1}\right) & \text{[Activation]} \\ \theta_Y \cdot \left(\dfrac{\omega_Y \cdot e}{\pi_Y + e}\right) & \text{[Constitutive (no regulation)]} \\ \theta_Y \cdot \left(\dfrac{\omega_Y \cdot e}{\pi_Y + e}\right) \cdot \left(\omega_0 + \dfrac{1}{K_Y \cdot p_{TF} + 1}\right) & \text{[Repression]} \end{cases}$$

where ω_Y is the maximal expression rate, ω_0 is the constant 'leaky' expression, K_Y is the effective strength of p_{TF}. Here we set synthetic gene expression $\omega_{E'}$ to 20 mRNAs per min and set the host values ω_Y where $Y = \{T,\ E,\ H,\ R, r\}$ to the values reported in [10]. Our nominal value for synthetic gene expression represents approximately five times the mRNA synthesis rate for the host enzymes. We introduce θ_Y as an adjunct scaling factor which enables us to tune these nominal values in our simulations and optimisations (see Section IV). The $e/(\pi_Y + e)$ expression scales the promoter activity by the cell's internal energy which is either constant (see II-C) or dynamically calculated (see II-D). The parameter θ_Y is an adjunct scaling factor required for our optimization approach (see IV).

Proteins are produced by translation at rate $T_L(c_Y,\ e)$:

$$T_L(c_Y,\ e) = \frac{1}{n_Y} \cdot \left(\frac{\gamma_{max} \cdot e}{\kappa_\gamma + e}\right) \cdot c_Y \cdot \quad (11)$$

where n_Y, κ_γ and γ_{max} represent the protein peptide length, translational energy threshold and maximum translation rate. For the model without host processes $e = e_0$ (a constant) while for the host-aware model e is calculated dynamically.

B. Batch fermentation model

The four state biotechnological process is composed of: the cell population (N), the external substrate (S), which is taken up by cells, the extracellular product (P) which is exported from the cell and the extracellular inducer (I). The dynamics of this system are:

$$\dot{N} = \lambda \cdot N \quad (12)$$

$$\dot{S} = -v_{uptake}(S, p_T) \cdot N \quad (13)$$

$$\dot{P} = v_{prod}(s_{in}, p_{E'}) \cdot N \quad (14)$$

$$\dot{I} = -v_I(I) \cdot N \quad (15)$$

The inducer uptake rate v_I is calculated based on the difference between the internal and external concentration:

$$v_I = k_{I,D} \cdot (I \cdot (V_{cell}/V_{cult} - I_{in})\lambda \cdot N \quad (16)$$

C. Model in the absence of host feedback

To model the process and switch controller in Section II-A in the absence of host-circuit interactions, we set the host energy e and free ribosome concentration R to constants e_0

and R_0, equivalent to the steady state values of the host model energy species e and free ribosomes R as calculated numerically (see Appendix). We also set the rate $v_{uptake}(S, p_T)$ to its steady state value. To assess the impact of diverting s_{in} from growth to product synthesis, we estimate the growth rate based on the s_{in} and p_E concentration

$$\lambda = \left(\frac{\Phi_E}{\tau_E}\right) \cdot \left(\frac{s_{in}}{\kappa_E + s_{in}}\right) \tag{17}$$

where the mass fraction of the host protein E (Φ_E) is given by $(n_E \cdot p_E)/M_0$ and the time taken to generate sufficient energy to produce a protein is given by $\tau_E = n_E/(\phi_e \cdot v_E)$. The parameters n_E, M_0, ϕ_e and v_E correspond to enzyme E length in amino acids, total cell mass in amino acids, nutrient quality and enzyme E turnover number, respectively. See [9] for the derivation. We hereafter refer to this model as the "original" model.

D. Model augmented with host feedback

We embed the model from Section II-A in a previously described coarse grained model of *E. coli* [10]. This nonlinear 'self-replicator' model, hereafter referred to as the "host-aware model", is composed of 16 ordinary differential equations that capture the interaction between the cell's metabolism and gene expression by capturing the time evolution of a simple metabolism (consisting of substrate s_{in} and universal energy carrier e), gene expression of a coarse grained proteome (consisting of transporters p_T, metabolic enzymes p_E, host proteins p_H and ribosomal proteins p_R, each with a corresponding mRNA m_i and translation complex c_i modelled as in Eq. 4-6), translational resource biogenesis (consisting of the transcription of rRNAs, r, and the formation of ribosomes from r and p_R) and cell growth. The metabolic model in Section II-A is augmented by the introduction of a dynamic e species, produced from s_{in} and consumed by translation:

$$s_{in} \xrightarrow{v_{host}(s_{in},\ p_E)} e \xrightarrow{\sum(T_L(c_Y,\ e))} \emptyset \tag{18}$$

The dynamic nature of e causes the transcription and translation rates of the model's genes to dynamically respond to the impact of the engineered pathway flux (through Eq. 10 and 11). The model captures gene expression resource limitations and competition through the free ribosome dynamics:

$$\dot{R} = b_\rho \cdot p_R \cdot r - u_\rho \cdot R - \lambda \cdot R ... \tag{19}$$
$$... + \sum_X \left(T_L(c_X,\ e) - b_X \cdot m_X \cdot R + u_X \cdot c_X\right)$$

where p_R and r are ribosomal (r)-proteins and rRNAs, respectively, which form the functional free ribosome and construct all other proteins: $X = \{T,\ E,\ H,\ R,\ E', TF\}$. All genes within the model are connected to this single pool by ribosome-mRNA association reactions through the $\sum_X(\cdot)$ term. As one gene Y increases the $R \rightarrow c_Y$ rate increases which perturbs the other genes in the set X through the $\sum_X(\cdot)$ term. We define the growth rate as a function of the total cellular translation rate:

$$\lambda = \frac{1}{M_0} \cdot \left(\frac{\gamma_{max} \cdot e}{\kappa_\gamma + e}\right) \cdot \sum_X \left(\ c_X\ \right) \tag{20}$$

See [10] for a full description.

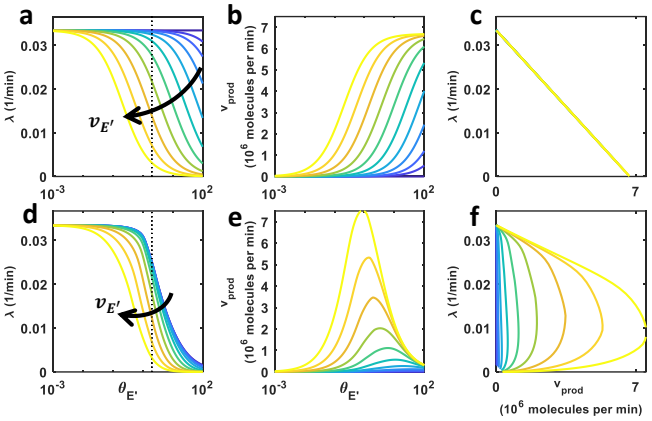

Fig. 2. Comparing the impact of varying $\theta_{E'}$ in both the original and host-aware model. $\theta_{E'}$ is the scaling factor for the transcription rate $p_{E'}$, see Eq. 10. $v_{E'}$ is varied on a log-scale from 0 (blue) to $5,800\text{min}^{-1}$ (yellow). (a) The impact of varying enzyme induction (production) rate $\theta_{E'}$ and enzyme turnover number (speed of action) $v_{E'}$ on λ in the original model. (b) The impact of varying $\theta_{E'}$ and $v_{E'}$ on v_{prod} in the original model. (c) The relationship between λ and v_{prod} in the original model. (d) The impact of varying $\theta_{E'}$ and $v_{E'}$ on λ in the host-aware model. (e) The impact of varying $\theta_{E'}$ and $v_{E'}$ on v_{prod} in the host-aware model. (f) The relationship between λ and v_{prod} in the host-aware model.

III. Natural Feedback Enhances Pathway Performance

To understand how the synthetic enzyme expression (and subsequent flux redirection) impacts host growth we simulated the original and host-aware models with varying transcription rates and turnover number of E'. We find that the original model (Section II-C) shows an inverse linear relationship between growth rate and enzyme turnover number $v_{E'}$ (Fig. 2a,c) with production flux (v_{prod}) increasing monotonically with enzyme transcription ($\theta_{E'}$) across all $v_{E'}$ (Fig. 2b). However, the host-aware model (Section II-D) shows a more complex relationship due to $p_{E'}$ production dynamics and host feedback. Whilst the impact of $\theta_{E'}$ is qualitatively the same, λ falls significantly at high $\theta_{E'}$ due to ribosomal burden (Fig. 2d). Therefore the v_{prod} flux is non-monotonic with $\theta_{E'}$ with a peak in flux at intermediate values of $\theta_{E'}$. Increasing $v_{E'}$ reduces the growth rate for a given $\theta_{E'}$ and results in the peak of production flux occuring at a lower transcription rate. In the presence of host feedback the v_{prod}-λ relationship is no longer linear (Fig. 2f). The point at which the system is 'overloaded' (i.e. when more $p_{E'}$ does not increase v_{prod} but instead causes λ to fall), increases with increasing $v_{E'}$.

Assessing the dynamics of the internal processes of the host-aware model shows that host-mediated feedback enhances pathway performance (Fig. 3) [9]. The impact of the engineered pathway ($s_{in} \rightarrow p_{in}$) catalysed by the pathway

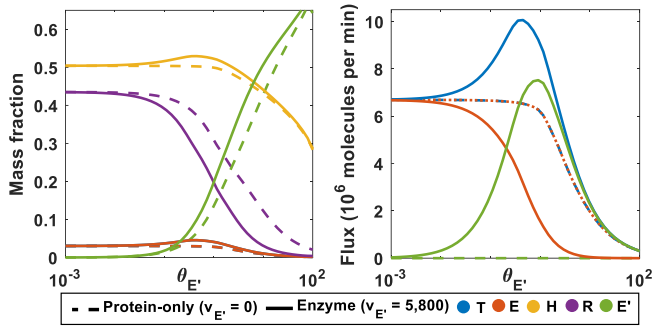

Fig. 3. Natural feedback enhances pathway performance. The model is simulated with $v_{E'} = 0$ min^{-1} to assess the impact of protein burden and $v_{E'} = 5{,}800$ min^{-1} to assess the combined impact of protein burden and $s_{in} \rightarrow e$ reduction. The colours in the right hand panel correspond to the flux through the respective enzyme. T represents the host transport proteins; E, the host metabolic enzyme; H, the host q-proteins which are nonfunctional in the model but generate biomass (see [10]); R, the ribosomes (both free and translating); E', the synthetic pathway enzymes.

enzyme $p_{E'}$ has an additive impact to protein production despite the fact that it increases burden (and reduces growth). Graphically this can be seen as a shift in the host proteome to the left in its response to $\theta_{E'}$ (see Fig. 3). For example, $p_{E'}$ reaches a mass fraction of 0.2 at a lower $\theta_{E'}$ value when $v_{E'} > 0$ than when $v_{E'} = 0$. As π_X (where $X \neq r,\ R$) is less than π_r and π_R, the non-ribosomal proteins are less sensitive to these changes in e. As the ribosomal mass fraction (R in Fig. 3) falls, the mass fraction of other proteins (including host enzymes) rises. Therefore, the impact of these non-regulatory interactions (i.e. host feedback) counter-intuitively enhances pathway performance: the addition of a foreign enzyme causes the cell to become an enzyme-dominated proteome which can further enhance the foreign enzyme production. This results in an increase in flux through the transporter, T. Even though the host enzyme p_E rises, the flux through v_{host} decreases as the s_{in} decreases due to the increased flux through v_{prod}.

IV. Design of Switching Systems in Original and Host-Aware Frameworks

The purpose of our control strategy is to optimise two key biotechnology metrics: yield (which governs process/production efficiency) and volumetric productivity (which governs production times and is a key driver of process costs). These objectives are often opposed, therefore we employ a multiobjective optimisation routine (see Appendix) to evaluate a number of controllers across any potential trade-off. To characterise a controller design, the system of equations are simulated until all substrate is depleted ($S(t) = 0$), at which point we define time as t_{end} and calculate the yield (J_y) and volumetric productivity (J_{vp}) as:

$$J_y = P(t_{end})/S(0) \quad \text{and} \quad J_{vp} = P(t_{end})/t_{end}. \tag{21}$$

For each controller we optimise the parameters in the set k representing the expression and control of the pathway enzymes, including their transcriptional scaling factor θ_Z, $Z\ (= \{E,\ E',\ TF\})$, and regulation strength, if any, K_Z. The induction time t_{ind} (when $I = 0$ to $I = I_{max}$) is also optimised.

Given these objectives vary over orders of magnitude we took a multi-step optimisation approach to the controller parametric design. Firstly, we utilised a genetic algorithm to solve two single-dimensional optimisation problems which maximise yield (J_y) and volumetric productivity alone (J_{vp}):

$$\begin{aligned} &\underset{k=\{\theta,\ K,\},\ t_{ind}}{\text{maximise}} \quad (J_y) \text{ or } (J_{vp}) \\ &\text{subject to} \\ &\qquad lb \leq k \leq ub,\ 0 \leq t_{ind} \leq t_{max}. \end{aligned} \tag{22}$$

We denote these maximal yield and volumetric productivity as $J_{y,max}$ and $J_{vp,max}$, respectively and their designs as k_y and k_{vp}. Next to identify the Pareto optimal designs and the performance trade-off, if any, between these two extremes, we used a multi-objective genetic algorithm to solve the following optimisation problem:

$$\begin{aligned} &\underset{k=\{\theta,\ K\},\ t_{ind}}{\text{maximise}} \quad \left(J_y/J_{y,max},\ J_{vp}/J_{vp,max}\right) \\ &\text{subject to} \\ &\qquad lb \leq \theta \leq ub,\ 0 \leq t_{ind} \leq t_{max}. \end{aligned} \tag{23}$$

To increase the efficiency of the optimisation we initialised the multi-objective optimisation with a population of initial points draw between the optimal designs k_y and k_{vp}.

The normalisation of each objective (e.g. J_y) by its previously identified optimal value (for J_y this is $J_{y,max}$) scales each objective to vary between 0 and 1 ensuring they are on the same scale and increasing the performance of the multi-objective genetic algorithm.

We optimised the control system in the two proposed modelling frameworks; the original model without host-circuit interactions (Section II-C) and the host-aware model (Section II-D). We optimised the strength of the growth branch (via θ_E), the strength of the production branch (via $\theta_{E'}$) and the regulation parameters (K_E, $K_{E'}$, θ_{TF}), and the induction time t_{ind}. We set the lower and upper bound of θ values to 10^-3 and 2, corresponding to biologically realistic values of nearly abolishing expression and modest two fold up-regulation. The bounds of association constant K bounds were 10^{-6} and 1, corresponding to dissociation constants in the nM range (e.g. LacI, tetR) and no binding. We set the induction time bounds to 1 min (i.e. instantly) and 24 h.

Both model frameworks show a trade-off between volumetric productivity and yield (Fig. 4a,b) with similar trends in their design rules across the Pareto fronts (Fig. 4c,d). Note, however, that the host-aware model suggests higher θ_E values are needed than in the original model to offset the effects of competition between p_E and other genes. The strength of the production branch (via $\theta_{E'}$) varies significantly. In the original model, without host context, $\theta_{E'}$ should be maximised while the host-aware model implies it should be minimised (although still non-zero). This suggests that incorporation of host processes changes the nature of the control problem.

To assess the necessity of each regulatory linkage, we set the regulation of each arm of the controller to 0 (i.e. we remove the regulation by rendering the node constitutive).

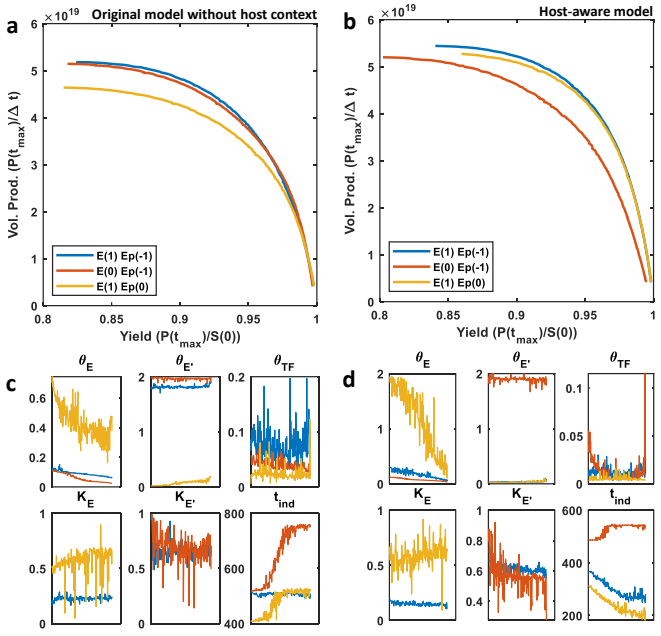

Fig. 4. Trade-off between the yield and volumetric productivity. The Pareto front identified in the optimisation routine described in Section IV. The lines are denoted $G(C)$ where G is the gene and N is the regulation with $C=+1$ is activation by p_{TF}, $C=-1$ is inhibition by p_{TF} and $C=0$ is unregulated (constitutive) expression. The x-axis of both (c) and (d) is the order of the points along the yield axis. (a) The three Pareto fronts for the original model without host constraints outlined in Section II-C. The three Pareto fronts for the host-aware model which accounts for host constraints as outlined in Section II-D. (c) The optimal designs corresponding to the fronts shown in (a). (d) The optimal designs corresponding to the fronts shown in (b). (Note that $\theta_{E'}$ has identified the upper bound. The optimisation was repeated with an upper bound of 10 and still identified the upper bound: these high θ values yield non-biologically feasible solutions as the synthetic protein mass fraction exceeds those tolerated [4].)

We name these controllers based on their topology where each gene is noted $G(C)$ with G being the gene and C being the regulation ($+1$, activation, -1 inhibition, and 0, unregulated (constitutive) expression). In the absence of host constraints, it appears that regulation of only the E' branch is required for optimal performance with $E(1)E'(-1)$ showing the same performance as $E(0)E'(-1)$ (Fig. 4a). However, this relationship is inverted in the presence of host constraints: where the regulation of the E is crucial. The $E(+1)E'(0)$ topology shows near equivalent performance to $E(+1)E'(-1)$. Both frameworks show simplified motifs can be optimal, however, they suggest *opposite* motifs. This demonstrates again that the incorporation of host-mediated feedback significantly changes the nature of the control problem, which in this case means host-feedback can allow construction of simpler circuit designs without much cost to performance.

V. Analysis of Simplified Motifs in Presence of Host Constraints

Although the single branch and dual branch regulation topologies ($E(+1)E'(0)$ and $E(+1)E'(-1)$, respectively) have the same performance, they have significantly different design rules (Fig. 4d). The switch induction time t_{ind} governs the trade-off between yield and productivity in both systems with shorter t_{ind} increase yield at a cost of volumetric productivity. The dual regulation system has low expression (θ_E and $\theta_{E'}$) while the $E(+1)E'(0)$ topology requires much higher θ_E values to overcome the impact of the constant $s_{in} \rightarrow p_{in}$ flux (and therefore lower e) due to constant E' activity and the competition for translational resources between E and E'.

To establish why the two topologies produce similar performance, we simulated the controller dynamics for a similar volumetric productivity (Fig. 5b). This shows that in addition to the $E(+1)E'(-1)$ controller acting to decrease p_E and increase $p_{E'}$, it also increases p_T due to (i) reducing e increasing $T_X(e)$ of the T species (Fig. 5a) and (ii) reducing translational competition enabling higher T translation (Fig. 5b). This results in a concurrent increase in flux through T (Fig. 5d). The $E(+1)E'(0)$ topology shows similar qualitative dynamics (Fig. 5). As E falls after t_{ind}, non-regualtory interactions enable T and E' to rise. Whilst both T and E' rise by 58%, the higher mRNA birth rate of T over E' means the impact of the former is more significant.

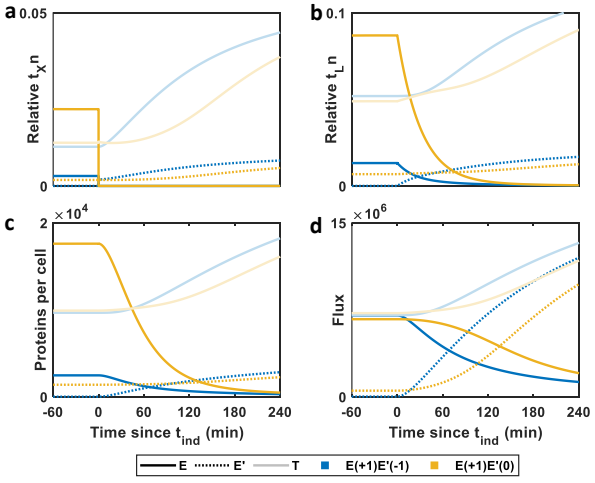

Fig. 5. Internal dynamics of the cell for selected $E(+1)E'(0)$ and $E(+1)E'(-1)$ controllers. Selected dynamics of controllers from Fig. 4b are shown for systems with the same volumetric productivity ($\approx 5\times10^{19}$ molecules per min). (a) Relative transcription rates, scaled by the sum of the transcription rates. (b) Relative translation rates, scaled by the sum of the translation rates. (c) Protein dynamics for E, E' and the transporter T. (d) Flux catalysed by the protein shown. E: $s_{in} \rightarrow e$, E': $s_{in} \rightarrow p_{in}$, T: $S \rightarrow s_{in}$.

VI. Generalising to Other Host Parameterisations

To assess the robustness of our observations to other parameterisations of the host model, we varied a selection of parameters in the host and repeated our analysis. Results are shown in Fig. 6 for the following representative parameters from [10]: v_T (the turnover rate - the maximum efficiency, of the import protein p_T which governs substrate import), k_H (the regulation threshold of the host q-proteins, a smaller value increases the proteome space occupied by host enzymes and ribosomes), γ_{max} (the maximal translation rate with increasing rates increasing the speed of protein production), ω_R (the maximal r-protein production rate, reducing this value reduces the number of functional ribosomes R in

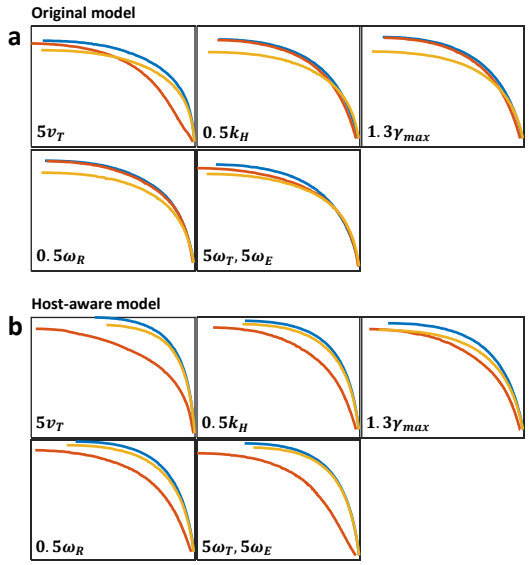

Fig. 6. The results are generalisable to different parameterisations of the host model. Pareto fronts were determined as outlined in Figure 4. The x-axis is the yield between 0.8 and 1. The y-axis is the volumetric productivity between 0 and 6.1×10^{19}. The y-axis of the panels representing v_T and ω_T/ω_E is between 0 and 1.1×20. The different colored fronts represent different controllers using the same legend in Figure 4. v_T, k_H, γ_{max}, ω_R, ω_T, and ω_E are parameters in the host model [10] representing biological rates depicted in the main text. (a) The original model in the absence of host constraints was simulated parameters optimised as detailed in the main text. For each new host parameterisation the parameters e_0, R_0 and v_{uptake} were calculated as in Appendix. (b) The Pareto fronts of the controllers implemented different parameterised host models.

the model), and ω_T and ω_E (the two rates which determine transporter and host metabolic enzyme mRNA transcription rate). In all settings, except where v_T is varied, the original model predicts that the E(0)E'(-1) topology performs nearly as well as the E(+1)E'(-1) topology. The model with host-constraints predicts that the E(+1)E'(0) topology gives near optimal performance in all tested settings.

VII. CONCLUSIONS

In the paper we developed a dynamic model of an engineered metabolic pathway subject to substrate uptake, metabolic competition and ribosome limitations. We showed that the perturbation engineered pathways cause to their host cell can lead the cell to adopt a favourable proteome allocation: expression of pathway enzymes enhances their own expression as pathway burden causes the host to adopt an 'enzyme dominated' growth strategy. We proposed the topology of genetic growth-production switch and show that the non-regulatory interactions which emerge from the presence of metabolic and ribosomal limitations can enable redesign of the switch with fewer regulatory linkage making the systems easier to engineer experimentally. These results enable the engineering of simple growth-production switches enabling the advantages of such systems to be achieved with fewer costly and time consuming experimental steps.

APPENDIX

The model was implemented in MATLAB 2020a with dynamics simulated with the in-built stiff solver *ode15s* with tolerances of 10^{-6} and *NonNegative* such that no states could be negative. To achieve specific designs (i.e. design parametrisations), we utilised the *ga* and *gamulti* functions from MATLAB's Global Optimisation Toolbox (version 4.4). The initial conditions for the multiscale simulation were determined by simulating the cell and process model (i.e. $dN/dt = dS/dt = dP/dt = dI/dt = 0$) to steady state from $N(0) = 0$, $S(0) = 10^4$, $s_{in}(0) = 10^6$, $e = 10^6$, $p_T(0) = p_E(0) = R(0) = 100$ and $p_{TF}(0) = 1$. Multiscale model simulations were then initialised with $N(0) = 10^6$, $S(0) = 4.18 \times 10^{22}$ (10 g glucose in a culture vessel of 1.25 L), $P(0) = I(0) = 0$. At $t = t_{ind}$, I is raised to 10^{23} as a step input. Models were simulated to a t_{max} equivalent to 7 days. The host model parameters are $\phi_e = 0.5$, $v_T = 728$ min^{-1}, $v_E = 5800$ min^{-1}, $\kappa_E = \kappa_T = 1000$ molecules, $\omega_T = \omega_E = 4.14$ mRNAs $\cdot$ min^{-1}, $\omega_H = 948.93$ mRNAs $\cdot$ min^{-1}, $\omega_r = 3170$ mRNAs $\cdot$ min^{-1}, $\omega_R = 930$ mRNAs $\cdot$ min^{-1}, $\pi_T = \pi_E = 4.38$ molecules, $\pi_R = 426.87$ molecules, $n_T = n_E = n_H = 300$ amino acids, $n_R = 7459$ amino acids, $b_T = b_E = b_H = b_R = 1$, $u_T = u_E = u_H = u_R = 1$, $b_\rho = 1$, $u_\rho = 1$, $\delta_m = 0.1$ min^{-1}, $\kappa_H = 121775.2$ molecules, $h_H = 8$, $\gamma_{max} = 1260$ amino acids per min, $\gamma_\kappa = 7$ molecules, $M_0 = 10^8$ amino acids. The circuit and pathway parameters (where $X = \{E', TF\}$) are $w_0 = 10^{-4}$, $\omega_X = 20$ mRNAs $\cdot$ min^{-1}, $\pi_X = 4.38$ molecules, $n_X = 300$ amino acids, $b_X = 1$, $u_X = 1$, $v_{E'} = 5800$ min^{-1}, $\kappa_{E'} = 1000$ molecules. $k_{I,D} = 3600/60$ min^{-1}, $V_{cell} = 10^{-15}$ L, $V_{cult} = 1.25$ L, $k_{I,f} = 1057.7/60$ molecules$^{-2} \cdot$ min^{-1}, $k_{I,r} = 1292.1/60$ min^{-1}.

REFERENCES

[1] C. J. Hartline, A. C. Schmitz, Y. Han, and F. Zhang, "Dynamic control in metabolic engineering: Theories, tools, and applications," *Metabolic Engineering*, vol. 63, pp. 126–140, 2021.

[2] K. G. Gadkar, F. J. Doyle III, J. S. Edwards, and R. Mahadevan, "Estimating optimal profiles of genetic alterations using constraint-based models," *Biotechnology and Bioengineering*, vol. 89, no.2 pp. 243–251, 2005.

[3] D. Liu, A. Mannan, Y. Han, D. A. Oyarzún and F. Zhang, "Dynamic metabolic control : towards precision engineering of metabolism," *Journal of Industrial Microbiology & Biotechnology*, vol. 45, no. 7, pp. 535-543 2018.

[4] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, "Interdependence of cell growth and gene expression: origins and consequences." *Science*, vol. 330, no. 6007, pp. 1099–1102, 2010.

[5] A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio, "Isocost Lines Describe the Cellular Economy of Genetic Circuits," *Biophysical Journal*, vol. 109, no. 3, pp. 639–646, 2015.

[6] A. P. S. Darlington and D. Bates, "Host-aware modelling of a synthetic genetic oscillator," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2016-Octob, 2016.

[7] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource Competition Shapes the Response of Genetic Circuits," *ACS Synthetic Biology*, vol. 6, no. 7, pp. 1263–1272, jul 2017. [Online].

[8] J. Kim, A. P. S. Darlington, D. G. Bates, and J. I. Jimenez, "The interplay between growth rate and nutrient quality defines gene expression capacity," *bioRxiv* preprint, doi. 10.1101/2021.04.02.438188, 2021.

[9] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain, "Mechanistic links between cellular trade-offs, gene expression, and growth," *Proceedings of the National Academy of Sciences*, vol. 112, no. 9, pp. E1038–E1047, 2015.

[10] A. P. S. Darlington, J. Kim, J. Jiménez, and D. Bates, "Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes," *Nature Communications*, vol. 9, no. 1, 2018.