Optimization problem for estimation of kinetic parameters from noisy experimental data. The noisy data is generated through the use of additive noise in flux equations that use a known kinetic model to predict the responses of a metabolic network to different perturbations.

$$v^* = g(x,p) + \mathcal{N}(0,1) \tag{1}$$

where $\mathcal{N}(0,1)$ is Gaussian noise with zero mean and a standard deviation of 1.

$$\frac{d}{dt}pep = v_1 - v_2 - v_4 \tag{2}$$

$$\frac{d}{dt}fdp = v_2 - v_3 \tag{3}$$

$$\frac{d}{dt}E = v_{e,max}\left(\frac{1}{1+\left(\frac{fdp}{K_e^{fdp}}\right)^{n_e}}\right) - dE \tag{4}$$

The kinetic expressions for fluxes v_1 through v_4 are given below. The consumption of acetate through v_1 and conversion of *pep* through v_2 are expressed in Equations (5) and (6) respectively using Michaelis-Menten kinetics. The acetate flux through v_1 is also governed by the quantity of available enzyme E.

$$v_1 = k_1^{cat}E\frac{acetate}{acetate + K_1^{acetate}} \tag{5}$$

$$v_2 = V_2^{max}\frac{pep}{pep + K_2^{pep}} \tag{6}$$

$$v_3 = V_3^{max}\frac{\tilde{fdp}\left(1+\tilde{fdp}\right)^3}{\left(1+\tilde{fdp}\right)^4 + L_3\left(1+\frac{pep}{K_3^{pep}}\right)^{-4}} \tag{7}$$

The allosterically regulated flux v_3 for the consumption of *fdp* is expressed in Equation (7) using the Monod-Wyman-Changeux (MWC) model for allosterically regulated enzymes, where $\tilde{fdp}$ refers to the ratio of *fdp* with respect to its allosteric binding constant K_3^{fdp}. The added flux v_4 for the export of *pep* is expressed as a linear equation dependent on *pep* in Equation (8).

$$v_4 = k_4^{cat}.pep \tag{8}$$

We denote the known noisy flux and concentration information with the superscript *. Accordingly, fluxes v^* and concentrations x^* are the noisy information that will be used for estimation. Fluxes (v) and concentrations (x) that predicted by the model (without noise) are denoted without the superscript.

The optimization formulation to estimate fluxes given below is based on the minimization of the tolerance for the L2-norm of the difference between the model predicted flux and the noisy flux.

$$\min_{x,p,e} \; e \tag{9a}$$

$$\text{st } |v - v^*| \leq \varepsilon \tag{9b}$$

$$|x - x^*| \leq \varepsilon \tag{9c}$$

$$Sv = 0 \tag{9d}$$

$$v = f(x, p) + e \tag{9e}$$

$$x_{min} \leq x \leq x_{max} \tag{9f}$$

$$p_{min} \leq p \leq p_{max} \tag{9g}$$

$$e_{min} \leq e \leq e_{max} \tag{9h}$$

In Equation (9) above, ε is the absolute tolerance between the predicted and the measured quantities. The relationship between the predicted and the measured quantities are given by the nonlinear constraints in Equations (9b) and (9c). The stoichiometric constraints for steady state are given by Equation (9d), and Equation (9e) specifies the kinetic rate laws for the predicted fluxes as a function of the concentrations (x) and the parameters (p). Equations (9f) and (9g) are the bounds for the concentrations and parameters that are to be estimated by the optimization problem. The above optimization problem is to be solved for fixed values of $\varepsilon \in \{.01, .05, .1\}$.

Changes in problem formulation due to feasibility issues: Since the above problem (Equation 9) does not result in feasible global optimal solutions (using SCIP), the constraints are relaxed and the feasible problem is presented below.

$$\min_{x,p_i,e_i} \; e_i \tag{10a}$$

$$\text{st } |v_i - v^*| \leq 0.2 \tag{10b}$$

$$|x - x^*| \leq 0.3 \tag{10c}$$

$$|Sv| \leq 1\text{x}10^{-8} \tag{10d}$$

$$v_i = f(x, p_i) + e \tag{10e}$$

$$v_{j \neq i} = f(x, p_{j \neq i}) \tag{10f}$$

$$x_{min} \leq x \leq x_{max} \tag{10g}$$

$$p_{i,min} \leq p_i \leq p_{i,max} \tag{10h}$$

$$e_{i,min} \leq e_i \leq e_{i,max} \tag{10i}$$

In Equation (10) above, the index i represents the flux whose parameters are being optimized and index j represents all other fluxes whose parameters are held constant.

The steady state constraint has been relaxed (from a equality constraint to an inequality constraint) along with permitting the concentrations and the estimated fluxes to deviate significantly from their experimentally observed values (20 - 30%).

The bounds for the flux noise were set at (0,20) in the above problem.

Suggested modifications to address infeasibility (July 4) After discussion with Prof. M, the following changes were made to the problem formulation:

- Experimental fluxes may/may not be at steady state. Accordingly, using the steady state constraint in the above formulation can render the optimization problem infeasible. Suggestion is to remove this constraint(s).
- Include the norm minimization constraint as an objective.
- Add the noise in fluxes and concentrations as part of the bound constraint.

$$\min_{x,p_j} \; \|v_i - v^*\| \qquad i \in \mathcal{R}^n, \; j \in \mathcal{R}^l \tag{11a}$$

$$\text{st } v_i = f(x, p_j) \tag{11b}$$

$$x_{min}(1-\epsilon) \leq x \leq x_{max}(1+\epsilon) \tag{11c}$$

$$p_{j,min} \leq p_j \leq p_{j,max} \tag{11d}$$

- The noise ϵ is fixed at 5% or 10% to signify experimental noise.

More changes to address nonlinearity due to division(July 13) Prof. M suggested changing the nonlinear formulation of the flux equations into just a bilinear form by removing the denominator and formulating it as nonlinear algebraic equation.

$$\min_{x,p_j,v_i} \; \|v_i - v^*\| \qquad i \in \mathcal{R}^n, \; j \in \mathcal{R}^l \tag{12a}$$

$$\text{st } N(x, p_j) - v_i D(x, p_j) = 0 \tag{12b}$$

$$x_{min}(1-\epsilon) \leq x \leq x_{max}(1+\epsilon) \tag{12c}$$

$$p_{j,min} \leq p_j \leq p_{j,max} \tag{12d}$$

$$v_{i,min}(1-\epsilon) \leq v_i \leq v_{i,max}(1+\epsilon) \tag{12e}$$

- Also, the initial noise added to the steady state solutions was reduced from 50% to 5% or 10%.
- ϵ is fixed at 50%.
- Solution is obtained using multisolve with IPOPT instead of aiming to obtain a global solution with SCIP.

- Using approximately 14 million points for the second phase of multistart in OPTI