



HAL
open science

cMFA for multi-omics data integration in microbial community models

Sthyve Junior Tatho Djeanou, Simon Labarthe, Valentina Baldazzi

► **To cite this version:**

Sthyve Junior Tatho Djeanou, Simon Labarthe, Valentina Baldazzi. cMFA for multi-omics data integration in microbial community models. CMSB 2025 - 23rd International Conference on Computational Methods in Systems Biology, Sep 2025, Lyon, France. <hal-05261543>

HAL Id: hal-05261543

<https://hal.science/hal-05261543v1>

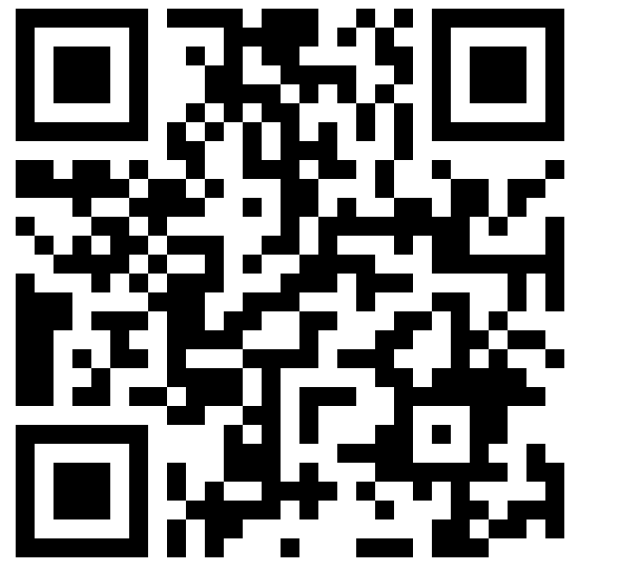
Submitted on 15 Sep 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

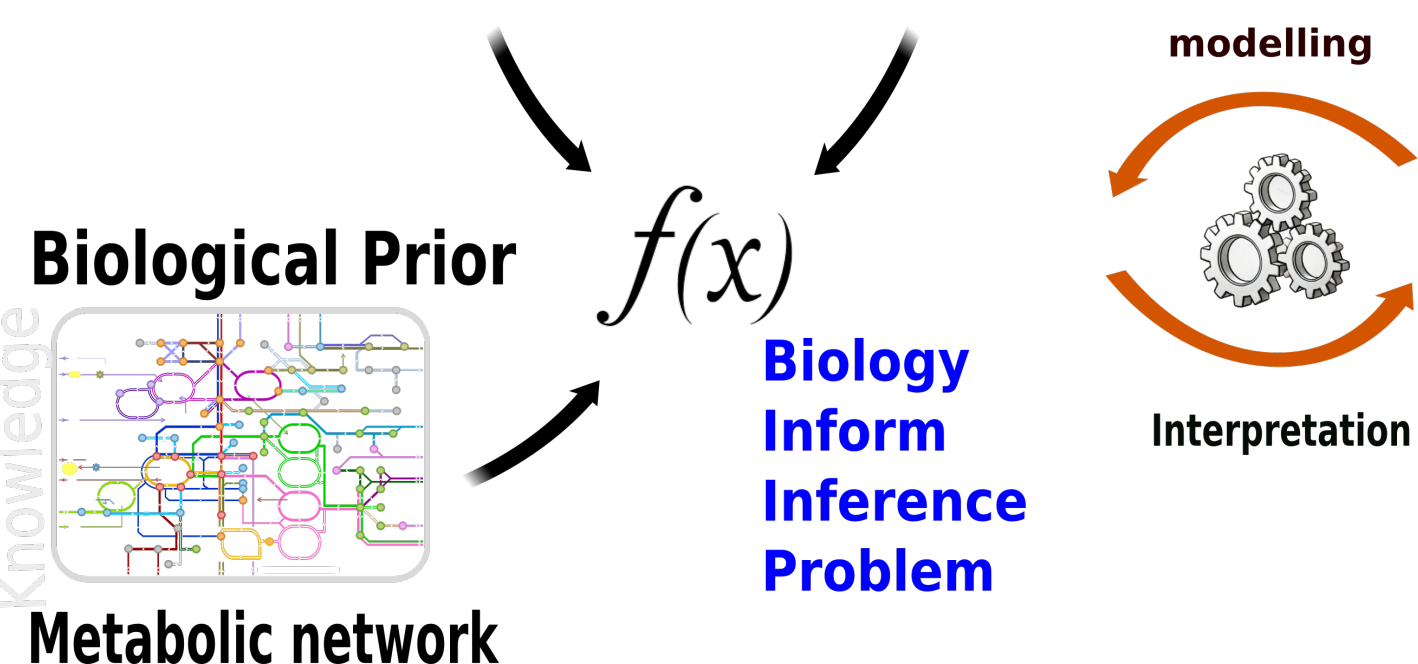
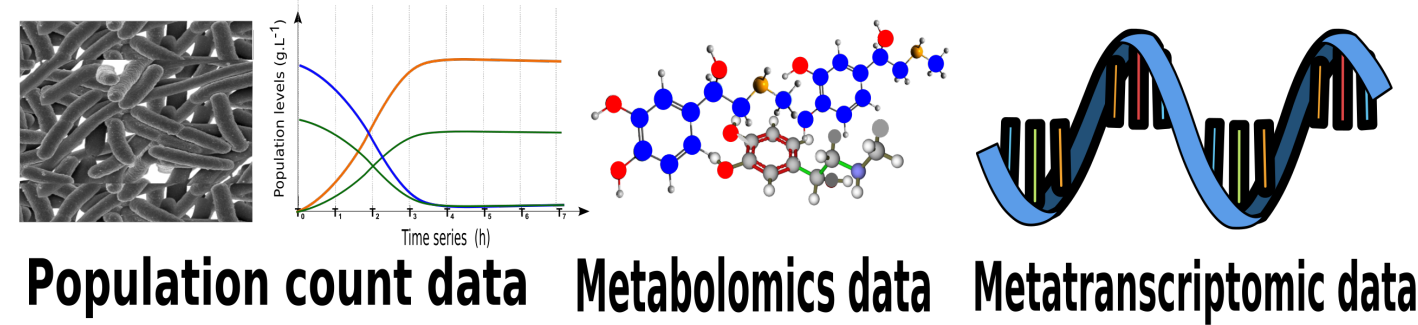


Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



INTRODUCTION

Biology-informed inference from multi-omics data



Understanding natural microbial communities presents significant challenges due to the complex interactions and metabolite exchanges between microorganisms. To gain deeper insights into these dynamics, we are focusing on:

- ➔ Developing mathematical models based on metabolic networks.
- ➔ Inferring metabolic fluxes using biology-informed high-dimensional analysis methods to integrate multi-omics data.

Scientific Question:

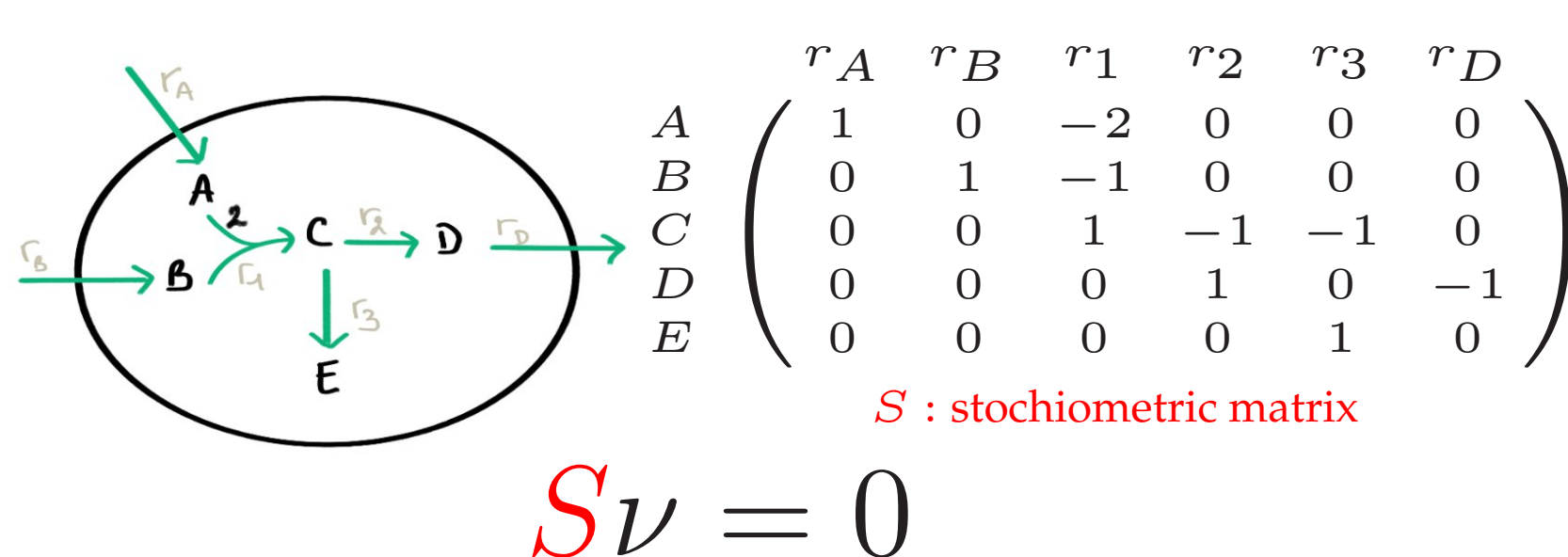
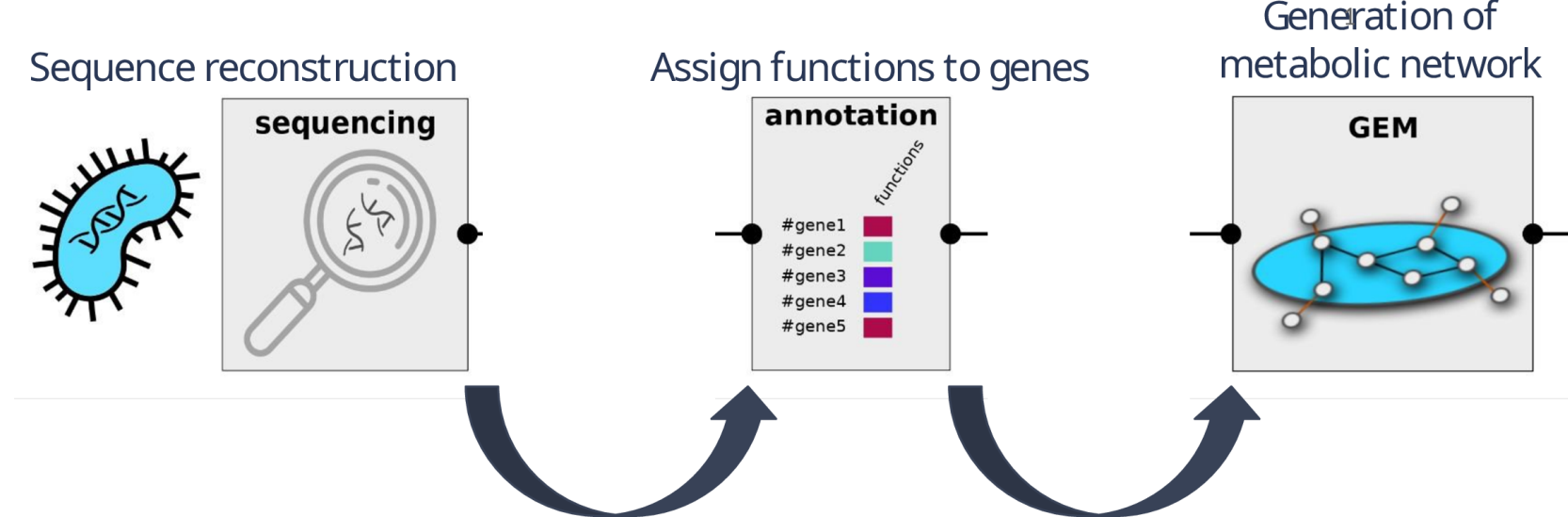
How can mathematical models help us better understand the functioning of microbial communities ?

- ➔ Internal flux reconstruction based on some external reaction rates.
- ➔ Estimating the contribution of each microorganism to the community.

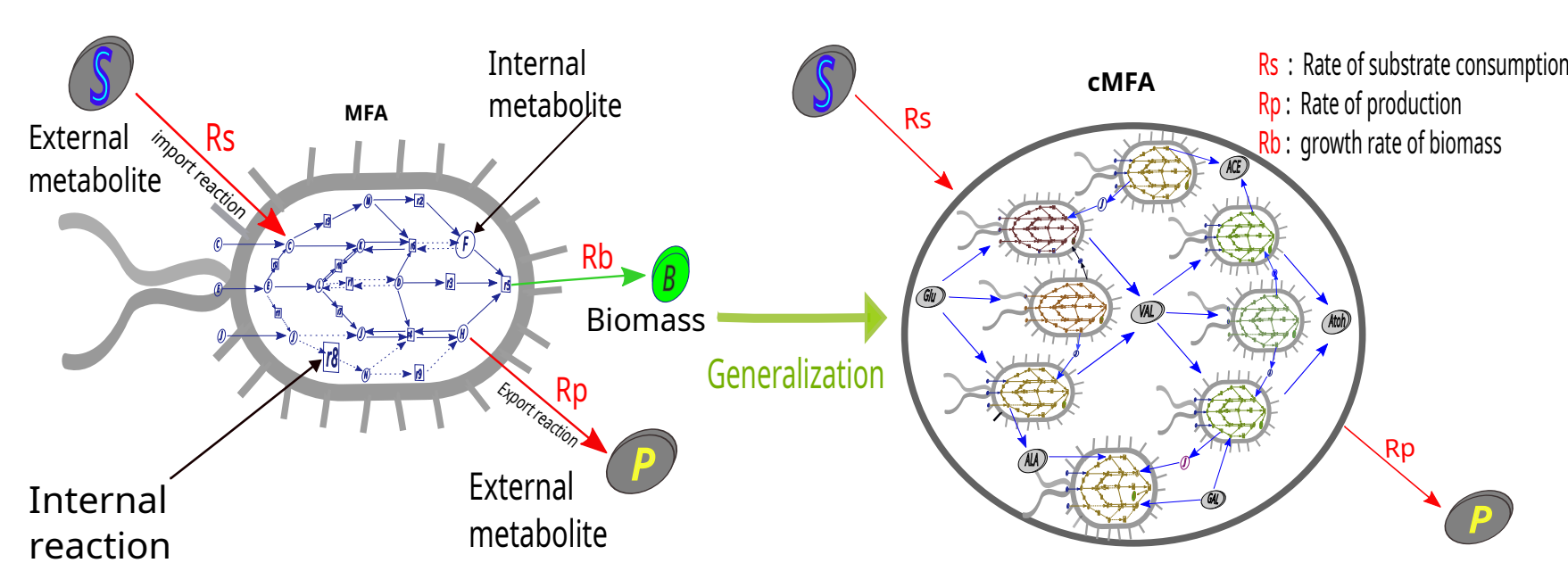
COMMUNITY METABOLIC FLUX ANALYSIS (CMFA)

METABOLIC NETWORK

A metabolic network refers to the interconnected set of biochemical reactions and pathways that occur within a cell or organism.



COMMUNITY MFA (CMFA)



MATHEMATICAL FORMULATION

Find, for $t \geq 0$, for a set \mathcal{M} of N_m metabolites and for a set \mathcal{B}_m of N_b micro-organisms the flux vector

$\hat{\nu}(t) = (\nu^{(b_1)}(t), \dots, \nu^{(b_{N_b})}(t))$ of dimension N_r such that

$$\hat{\nu}(t) := \arg \min_{\nu_t} \sum_{m \in \mathcal{M}} \left\| R_{m,t} - \sum_{b \in \mathcal{B}_m} \mathcal{X}_t^{(b)} \nu_{m,t}^{(b)} \right\|_2^2 + \lambda \mathcal{R}(\mathcal{T}\nu)$$

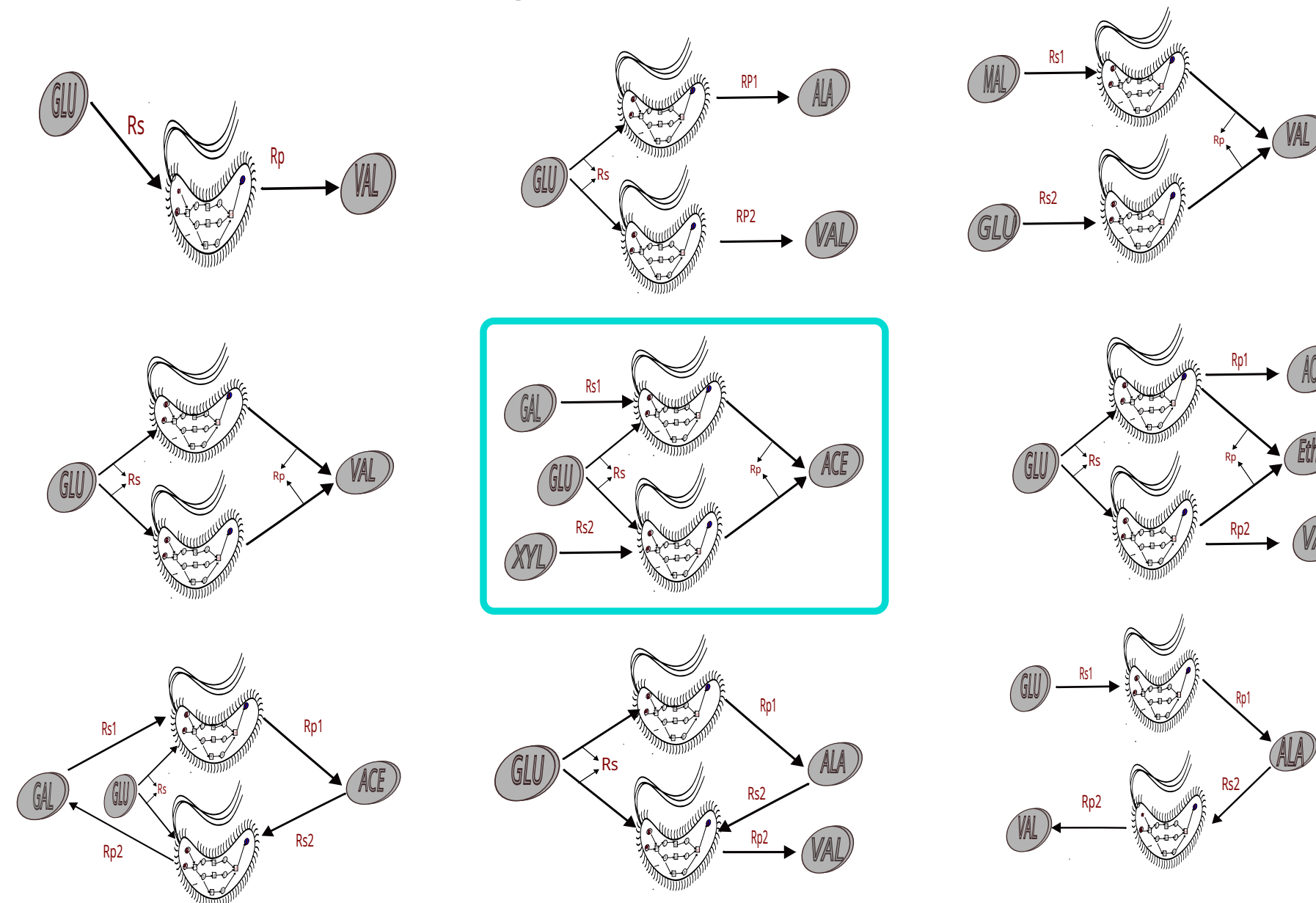
subject to: $\begin{cases} S^{(b)} \cdot \nu^{(b)} = 0 \\ C_{min}^{(b)} \leq \nu^{(b)} \leq C_{max}^{(b)}, \text{ for } b \in \mathcal{B}_m \end{cases}$

Regression problem is:

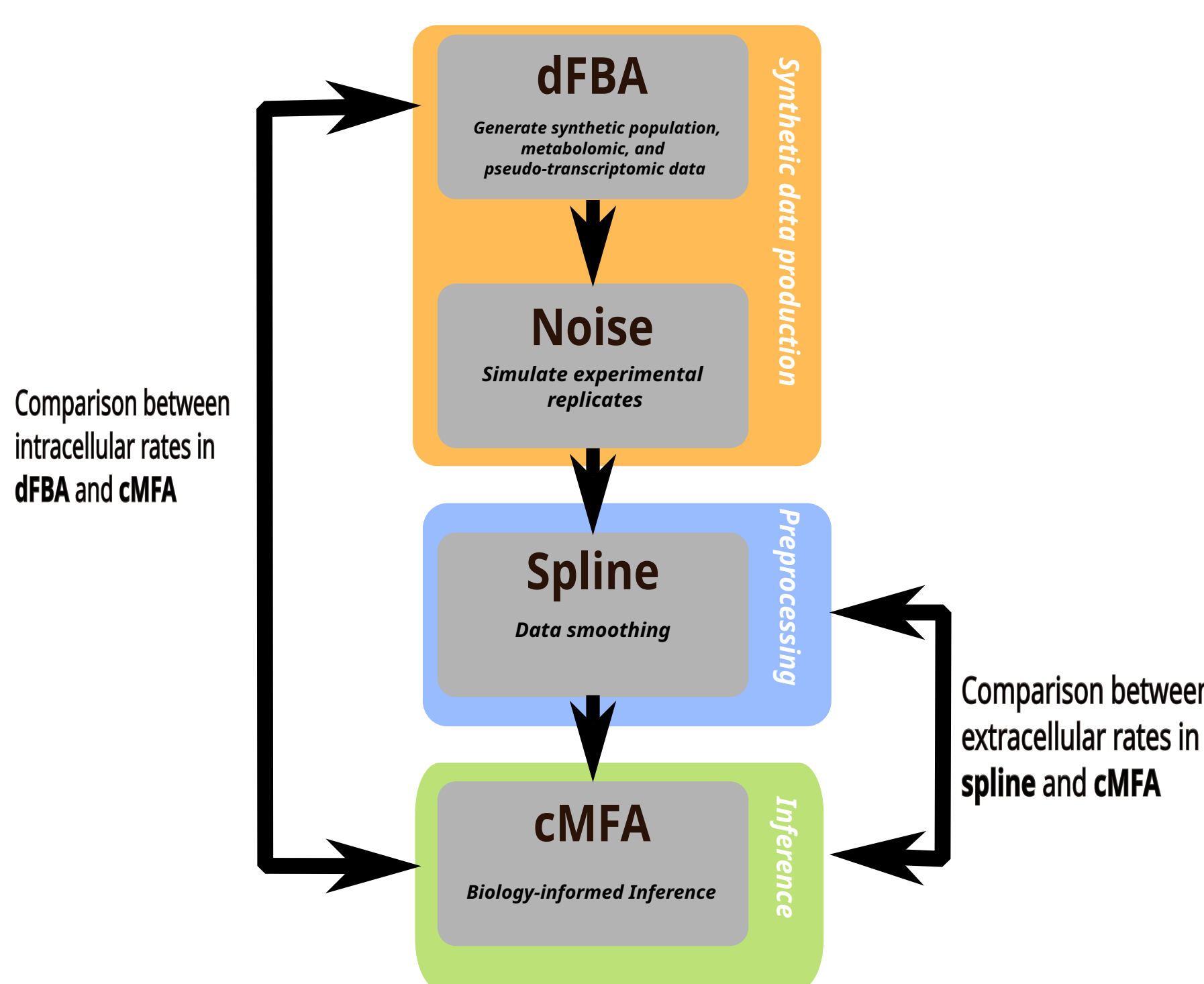
- ➔ High-dimensional ($N_r \sim 2500 \times N_b$)
- ➔ Highly-underdetermined ($N_r \gg N_m$)
- ➔ Biology-informed ($S^{(b)}$)
- ➔ Constrained by metatranscriptomics $\mathcal{R}(\mathcal{T}\nu)$: where \mathcal{T} is the reaction weight reflecting the level of gene expression.
- ➔ Regularized by a lasso ($\mathcal{R}(\nu) = \|\nu\|_1$) penalty to promote sparsity

CO-CULTURE DEFINITION

Different biological interaction of E. coli



BENCHMARK METHODOLOGY



GENERATING DATA: DFBA

Dynamic Flux Balance Analysis (DFBA) is a mathematical modeling approach to predict the metabolites fluxes in metabolic networks.

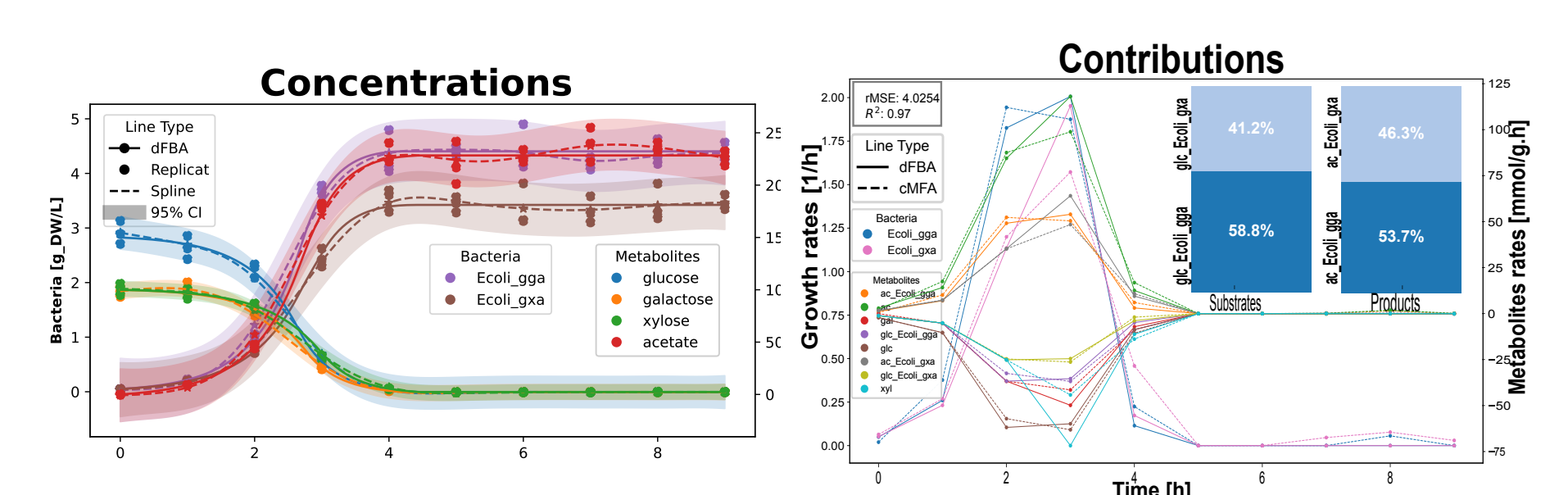
- ➔ **Model origin**
BIGG Models
Database iJO1366.
 - ➔ **model adjustment**
knocking out specific genes and reactions $S.T$
- $$\hat{\nu}_i = \arg \max B \nu_i$$
- $$\begin{cases} S_i \hat{\nu}_i = 0 \\ C_M \leq \hat{\nu}_i^s \leq 0 \\ 0 \leq \hat{\nu}_i^p \leq C_{max} \\ -C_{max} \leq \hat{\nu}_i^{in} \leq C_{max} \end{cases}$$

IMPLEMENTATION

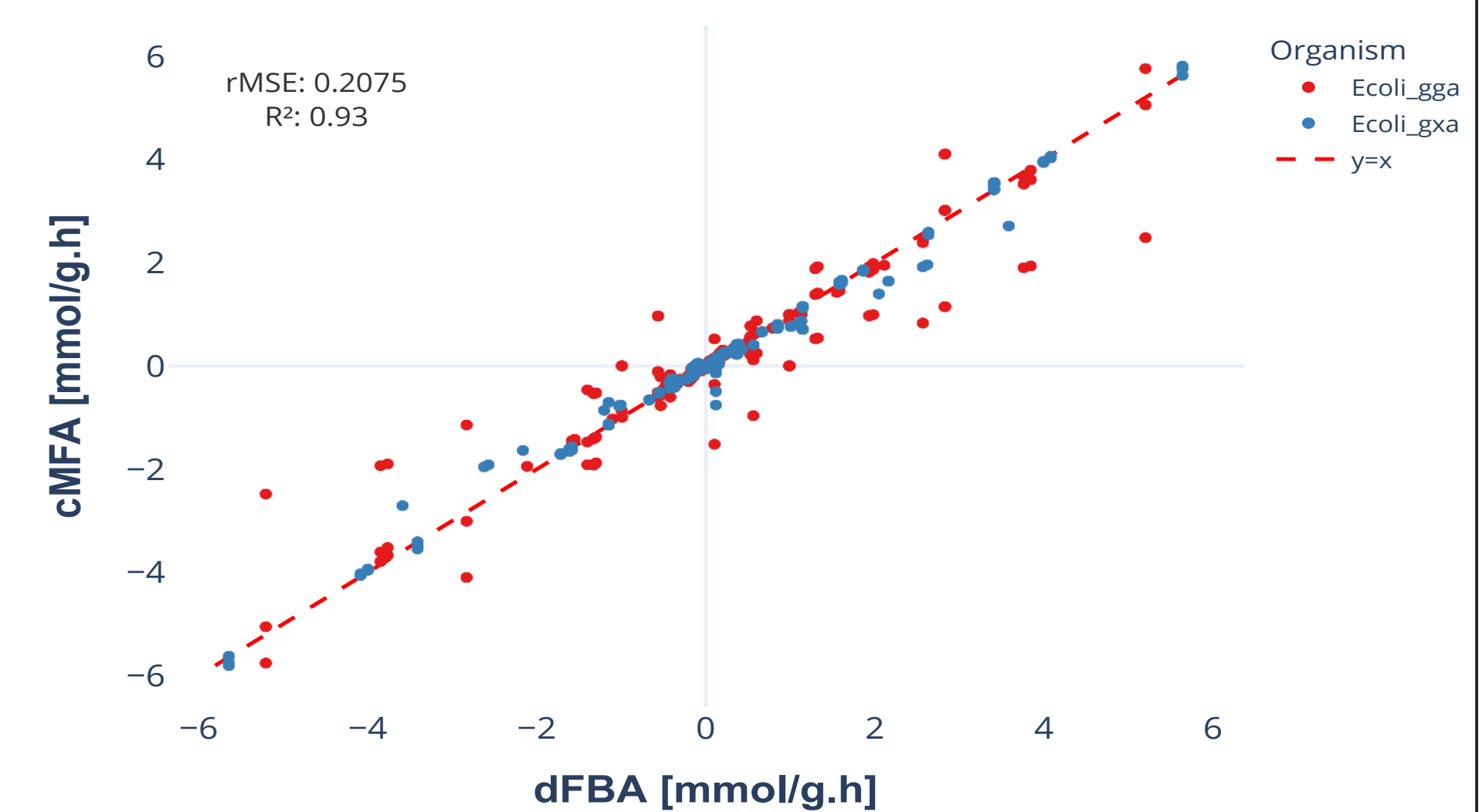
1. NumPy, Pandas for data handling and structure.
2. PyGAM for spline smoothing.
3. Matplotlib for visual representations.
4. COBRApy for constraint-based modeling.
5. OSQP for quadratic programming problems.

RESULTS

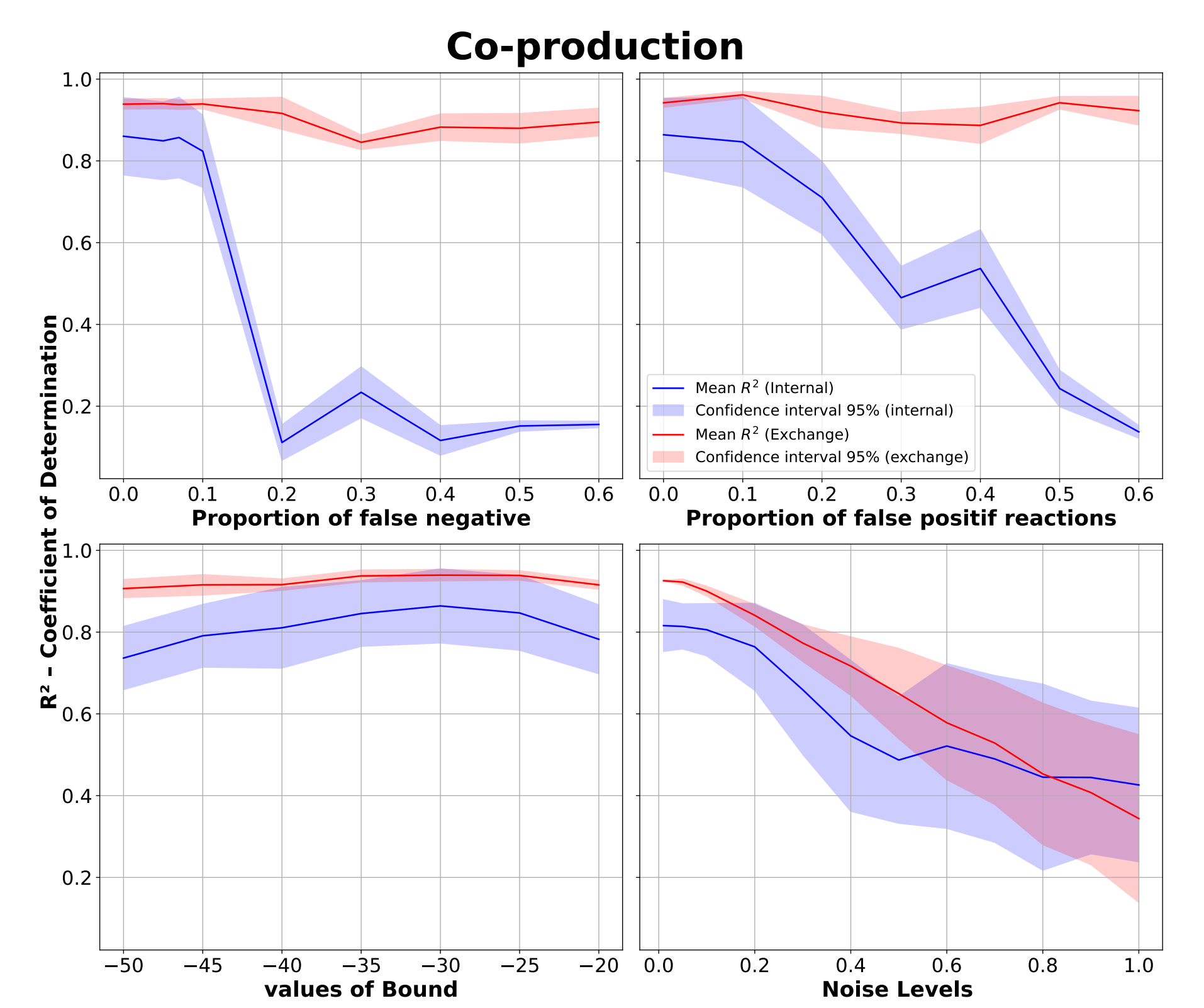
INFERRING INDIVIDUAL FLUXES: DFBA VS CMFA



Internal fluxes cMFA vs dFBA



BENCHMARK ON METHOD ROBUSTNESS



CONCLUSION & PERSPECTIVES

Conclusion : cMFA is robust to noise, metabolic import rates, and exchange rate inference with both high- and low-quality data. It scales well to large communities, but internal flux inference is less reliable with lower-quality data.

Perspectives : Large Community, missing metabolites, benchmark on real data .

REFERENCES

1. Brunner et al., 2023, *PLoS Comput. Biol.* 19, e1011661
2. Jenior, M.L. et al. (2020). *PLoS Comput. Biol.*, 16, e1007099.
3. Gerken Starepravo et al., 2022, *Digit. Chem. Eng.* 2, 100011
4. Lecomte et al., 2023, *Metab. Eng.* doi.org/10.1016/j.ymben.2024.02.014
5. Mahadevan et al., 2002, *Biophys. J.* 83, 1331–1340