Hybrid Cybernetic Modelling for studying the behaviour of metabolic systems towards the variation of its environment

Lalit S. Khot[1]
Prof. Arvind M. Lali[1]

Dr. Hyun-Seob Song[2]
Prof. Doraiswami Ramkrishna[2]

DBT-ICT-Centre for Energy Biosciences, Institute of Chemical Technology, Mumbai, India[1].
School of Chemical Engineering, Purdue University, West Lafayette, Indiana, USA[2].

Satellite Conference of ICM 2010

on

APPLICATION OF CONTROL THEORY AND OPTIMIZATION TECHNIQUES IN BIOCHEMICAL PATHWAYS

HICC, Hyderabad, India
16-18 August, 2010

Outline of the presentation

- Cybernetic approach
- Cybernetic control laws
- Incorporation of cybernetic control laws
- Transitions in Cybernetic Models
- Pathway Modeling
- Hybrid Cybernetic Model
  - Model formulation
  - Case study of *Clostridium acetobutylicum*
- Conclusion

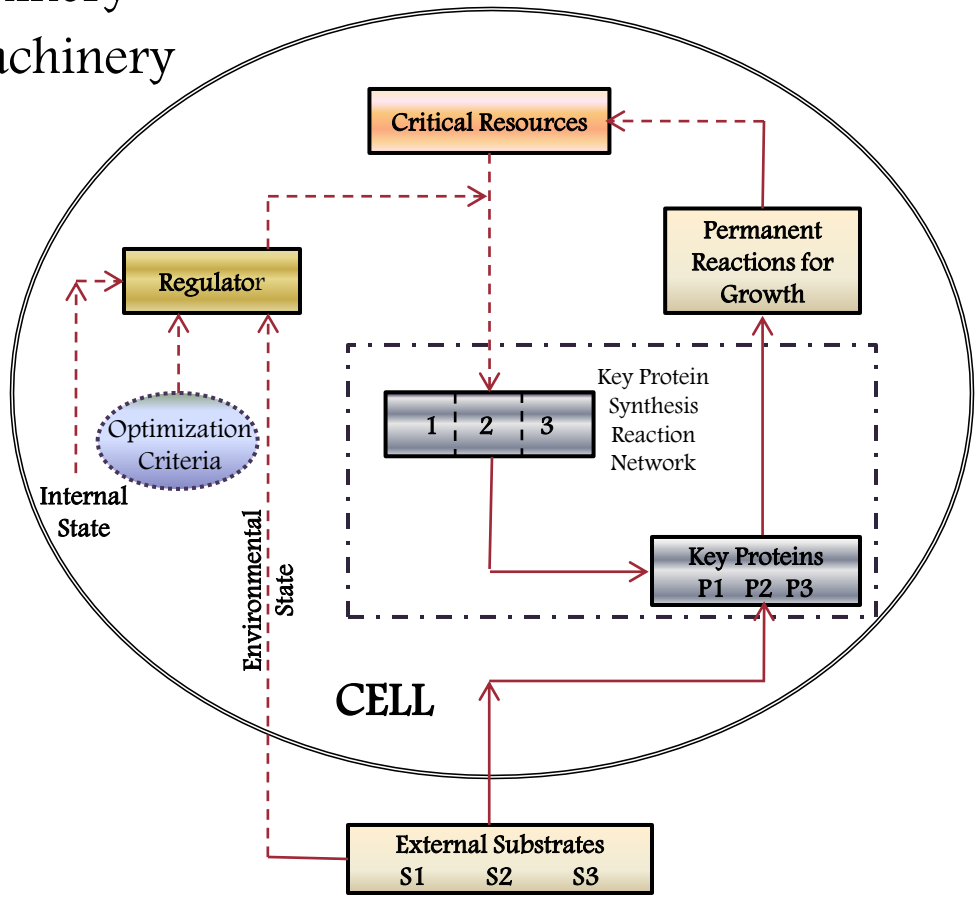
Introduction to Cybernetic Approach

- This approach differs from kinetic modelling efforts by incorporation of the optimal nature of microbial regulatory processes
- "Cybernetic" is derived from the Greek ***"χυβερτησ"*** which means steersman approach
- Microorganisms are optimal control strategists, use their internal regulatory machinery to "steer" themselves toward maximization of performance index (goal) while interacting with the environment
- Microbes have acquired the capability to control their regulatory processes to optimize their growth pattern
- The complex regulatory processes are reflected in terms of the cell's accomplishment of its optimal control objectives
- Microbial response in multiple substrate environments is a judicious investment of cellular resources in synthesizing different key proteins according to an optimal regulatory strategy

Cybernetic view of Cell

Cells can be viewed as a combination of machineries

- Adaptive machinery
- Permanent machinery
- Regulators



Kinetic Model

For Multiple Substrate Growth

Microbial growth on substrate (S_i) : $B + S_i \xrightarrow{E_i} (1+Y_i)B + ...$

The synthesis of enzyme (E_i) : $B \xrightarrow{S_i} E_i + B$

- The rate equation for biomass formation is : $r_i = \frac{\mu_i e_i s_i B}{K_i + s_i}$
- The rate equation for enzyme synthesis : $r_{Ei} = \frac{\alpha_i s_i B}{K_i + s_i}$

B - biomass; **S** - substrate; **Y** - yield coefficient;
E - enzyme; **r** - rate of synthesis (hr^{-1}); **μ** - specific growth rate (hr^{-1});
K - Michaelis constant (g/L); **α** - enzyme synthesis rate constant;

Kompala et. al., (1986)

Regulatory variables

When multiple substrates are present, the cellular regulatory processes of **repression/induction** and **inhibition/activation** affects the growth

Hence, the actual rate of synthesis of enzyme : $r_{E_i} u_i$; $(0 \le u_i \le 1) \& (\sum u_i = 1)$

u_i is fractional allocation of critical resources for the **synthesis** of E_i

- It incorporates regulatory action of repression and induction

& the total growth rate : $\sum_i r_i v_i$; $(0 \le v_i \le 1)$

v_i is fractional allocation of critical resources for the **activity** of E_i

- It incorporates regulatory action of inhibition and activation

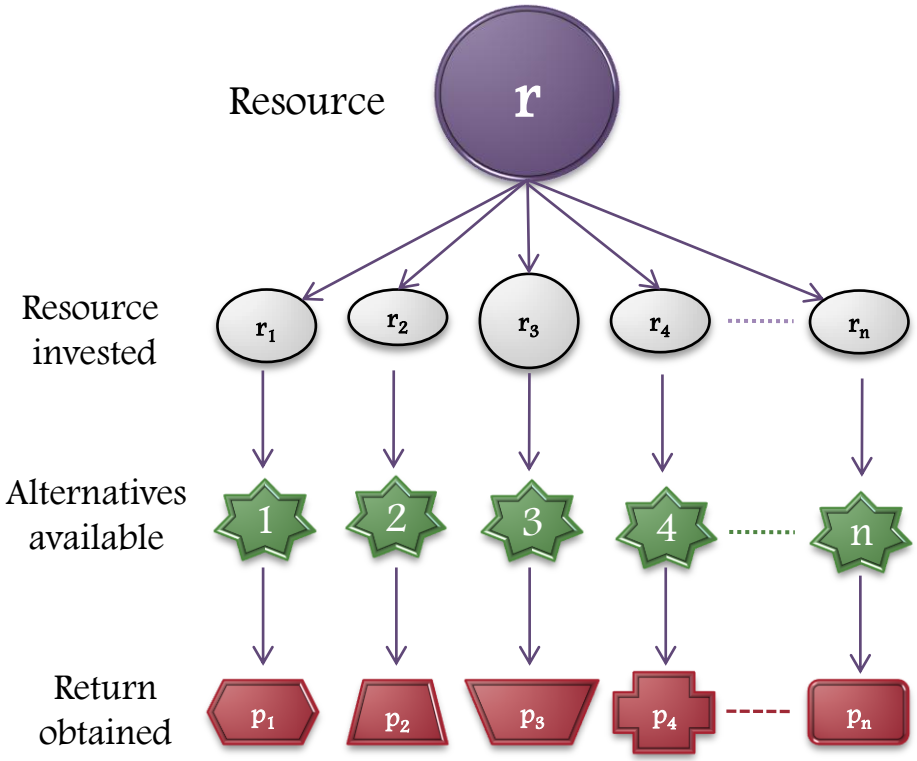
Kompala et. al., (1986)

$u_i \Longrightarrow$ Matching Law

$v_i \Longrightarrow$ Proportional Law

Matching Law

The regulatory action of repression/induction is incorporated by variable u



According to the law of diminishing marginal utility :

- Maximum for Total Returns $= \sum_i p_i(r_i)$

S.T. : Total Resources $\sum_i r_i = r$

can be obtained when,

$$\frac{dp_1}{dr_1} = \frac{dp_2}{dr_2} = \ldots\ldots\ldots = \frac{dp_n}{dr_n}$$

Hence, Fractional Allocation (u_i) :

$$u_i = \frac{dr_i}{\sum_j dr_j} = \frac{dp_i}{\sum_j dp_j}$$

$$(0 \le u_i \le 1) \,\&\, (\sum u_i = 1)$$

Fractional allocation must match fractional return

Proportional Law

Multiple substrate environment causes inhibition/activation of enzymes

The control action of inhibition/activation governed by v_i is proportional to maximum specific growth rate r_i

$$v_i = \lambda\, r_i$$

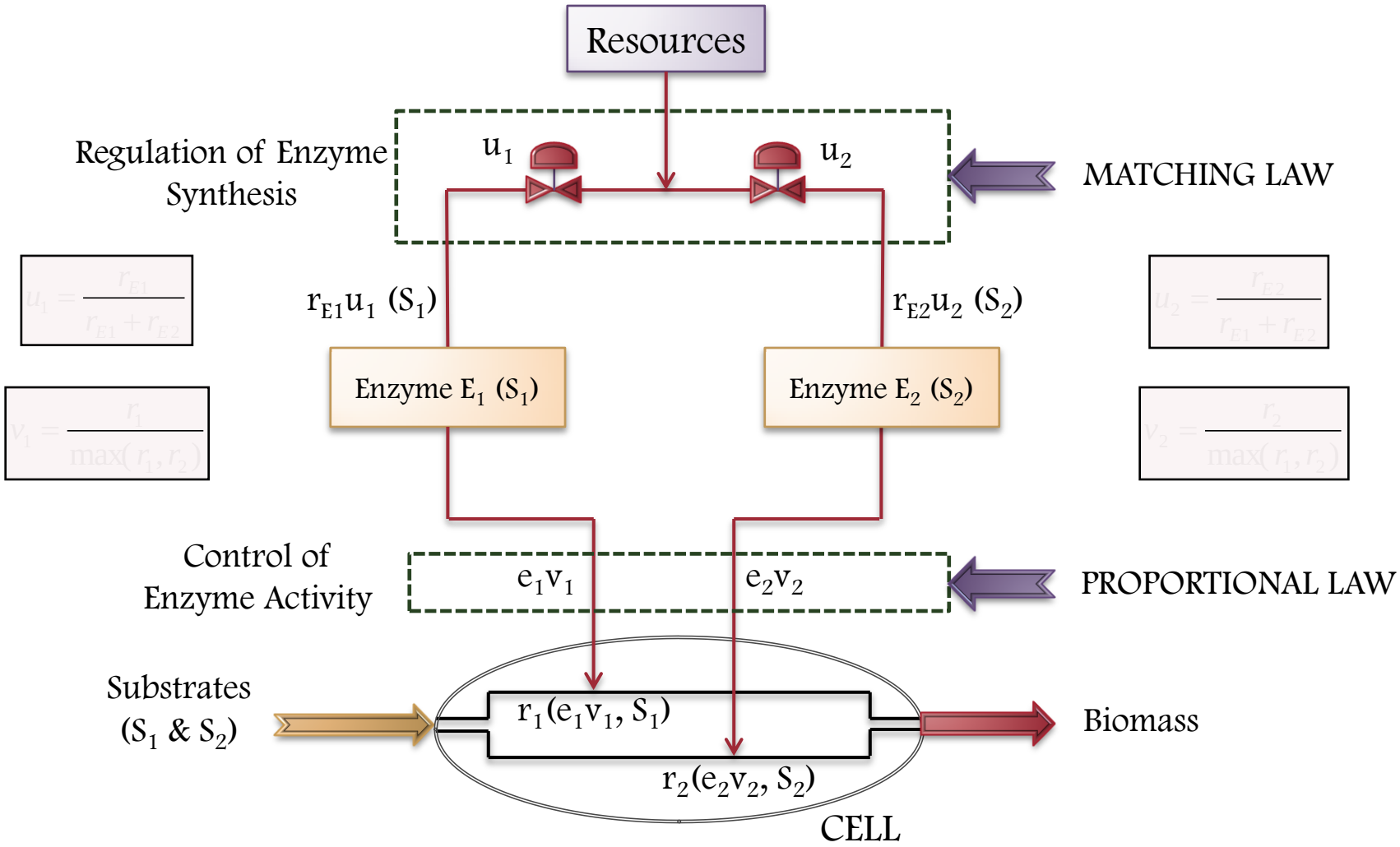
This proportionality combined with constraints determine the bound on λ

$$0 \le v_i \le 1 \;\Rightarrow\; 0 \le \lambda \le \frac{1}{r_i} \;\; or \;\; \lambda \le \frac{1}{\max_j (r_j)}$$

- The actual growth rate $\sum_i r_i v_i = \lambda \sum_i r_i^2 \le \frac{1}{\max_j (r_j)} \sum_j r_j^2$

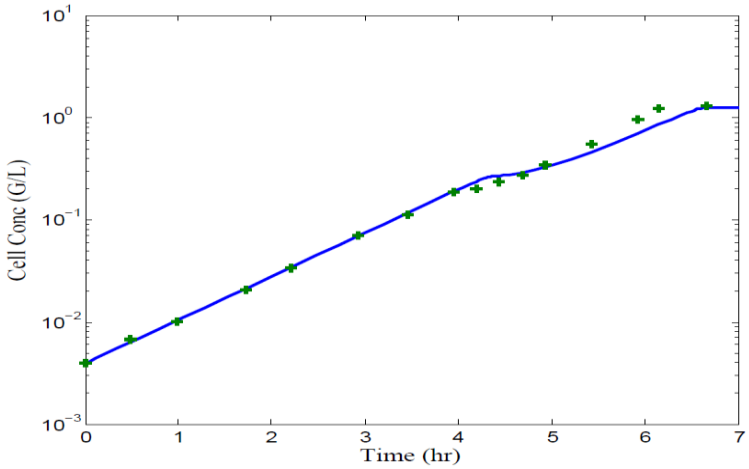
- Therefore, for maximum of growth rate $\lambda = \frac{1}{\max_j (r_j)}$ and $v_i = \frac{r_i}{\max_j (r_j)}$

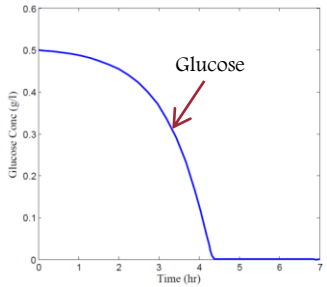
Incorporation of cybernetic variables

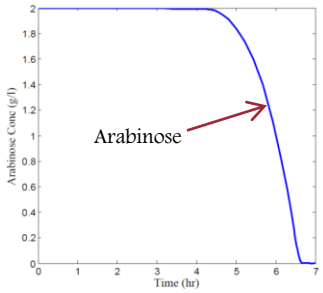


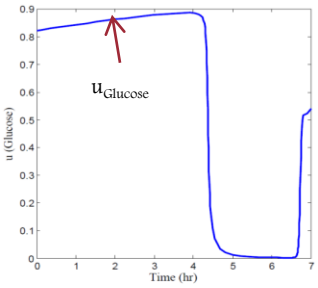
Diauxic Growth

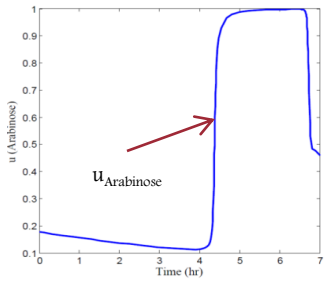
- Diauxic growth of *Klebsiella oxytoca* on mixed carbon source of Glucose and Arabinose

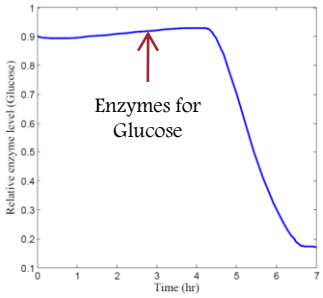


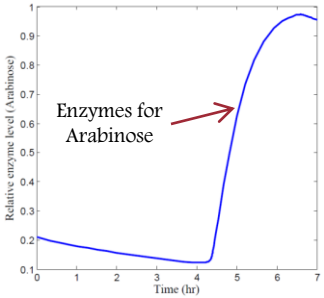


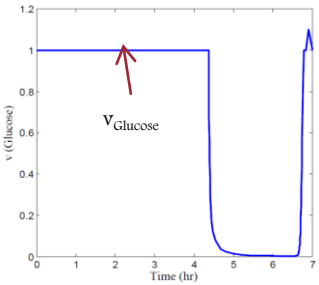


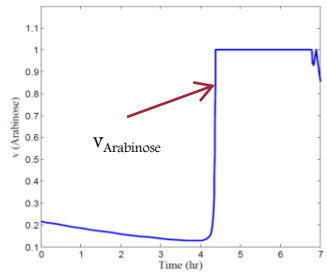








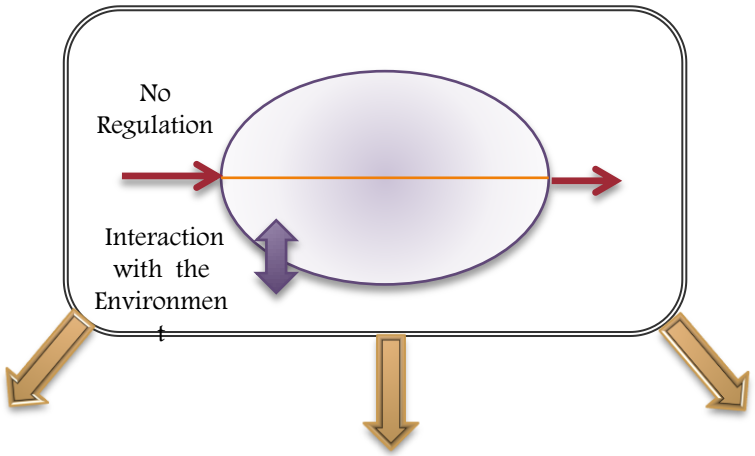




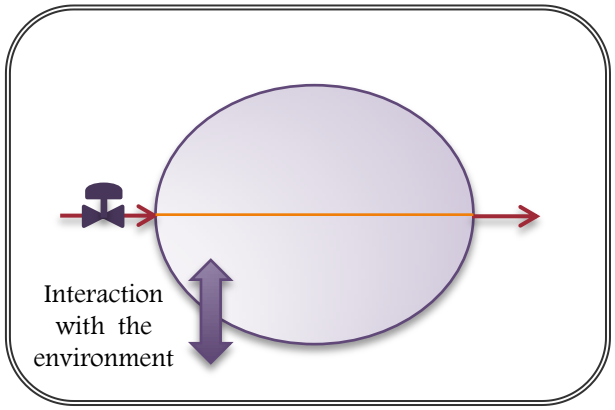
Kompala et. al, 1986

Cybernetic Models

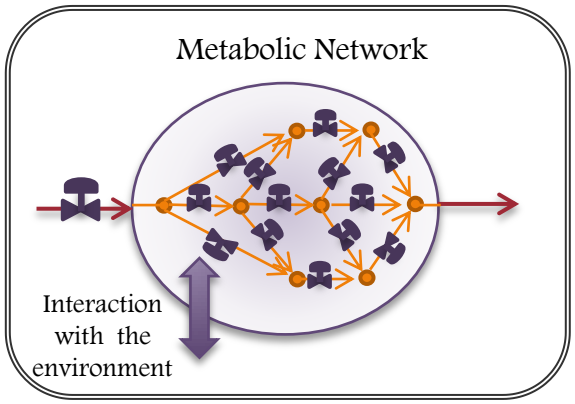
Lumped Kinetic Model



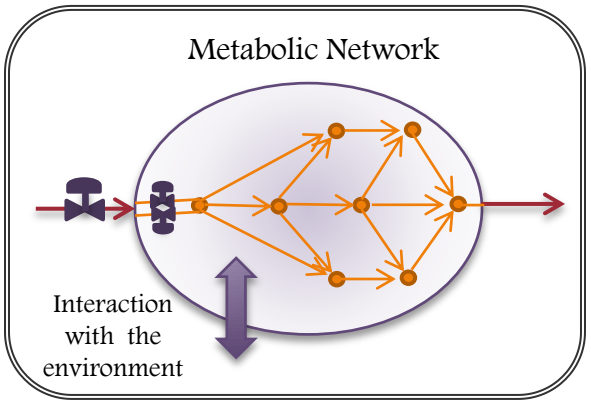
Lumped Cybernetic Model



Young's Model

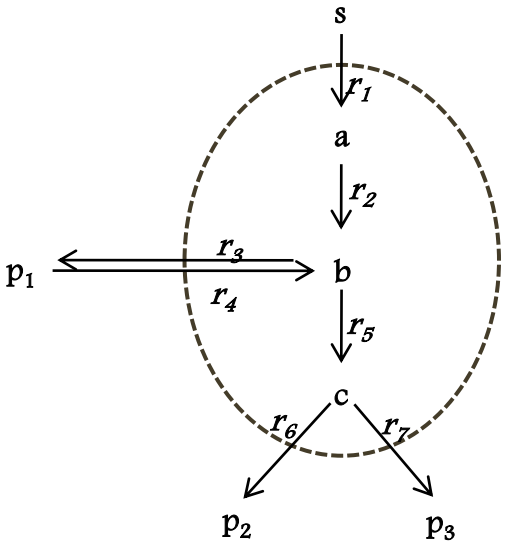


Hybrid Cybernetic Model



Modeling of Metabolic Systems

Representation of Metabolic Systems



Metabolism can be represented in the algebraic form :

$$\frac{d[w]}{dt} = S\,r$$

S – stoichiometric matrix
r – flux vector of reactants
LHS – flux exchange vector

- s – substrate
- a, b & c – intermediate metabolites
- p_1, p_2 & p_3 – extracellular products

[w]	d[w]/dt
s	$-r_1$
a	$r_1 - r_2$
b	$r_2 + r_4 - r_3 - r_5$
c	$r_5 - r_6 - r_7$
p_1	$r_3 - r_4$
p_2	r_6
p_3	r_7

$$\frac{d}{dt}\begin{bmatrix} s \\ a \\ b \\ c \\ p1 \\ p2 \\ p3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \end{bmatrix}$$

Metabolites

X

Reaction rates

Hybrid Cybernetic Model (HCM)

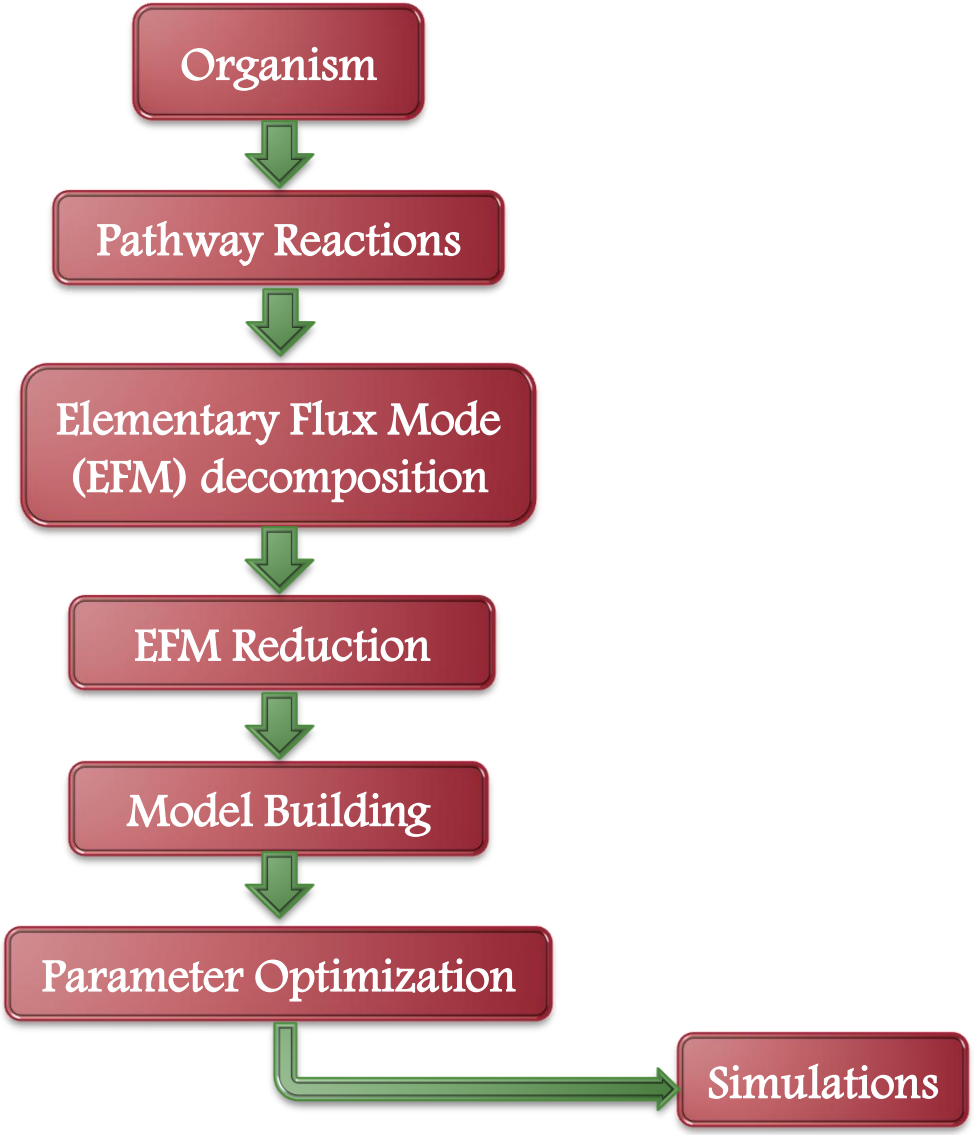
- It is the combination of FBA and classical lumped cybernetic model
- HCM considers the metabolic map of organism for the model building
- Provides the dynamic framework for modeling metabolic systems
- It requires only the measurement of extracellular component fluxes to estimate the coupled intracellular fluxes
- HCM considers the decomposition of network into several elementary modes
- HCM considers the pathway as convex combination of elementary modes, which can obtain by METATOOL 5.1

Hybrid Cybernetic Model (HCM)

Assumptions:

- Organism adapts itself to extracellular environment which continuously changes with time
- Quasy-steady state for intracellular metabolites
- Extracellular metabolites are considered as dynamic
- Slow dynamic intracellular metabolite are considered as extracellular
- Response of an organism is summation of response obtained through different elementary modes

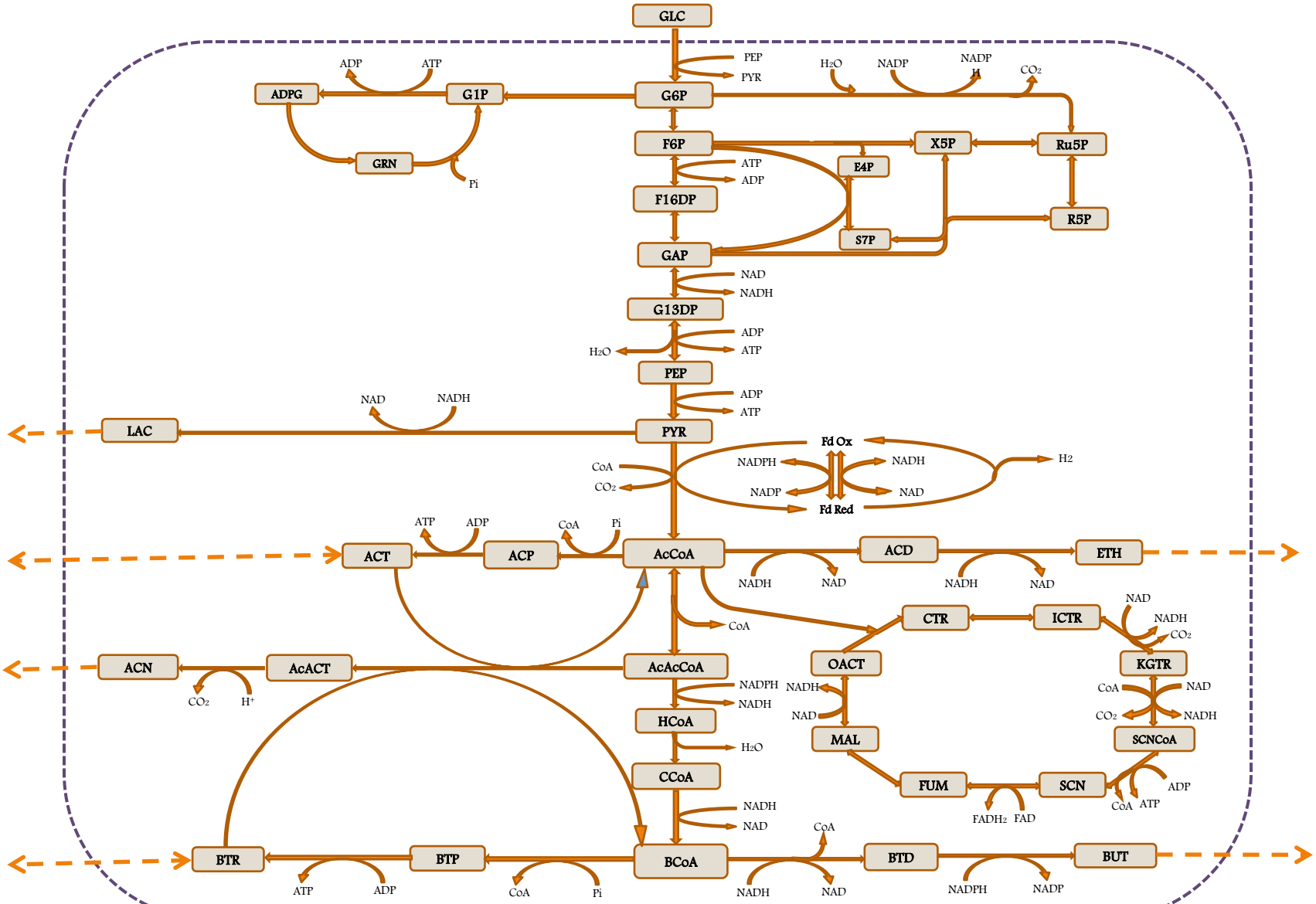
Model Formulation



HCM of *Clostridium acetobutylicum*

- *Clostridium acetobutylicum* is used in ABE fermentation
- Main products are acetic acid, butyric acid, acetone, butanol, and ethanol
- Its growth is combination of two phases; namely acidogenic and solventogenic
  - Acidogenic – acetic acid and butyric acid
  - Solventogenic – acetone, butanol and ethanol
- Dynamic data of these 5 products, glucose uptake and biomass formed are considered for model building and parameter estimation

1. *Clostridium acetobutylicum* pathway



2. Clostridium acetobutylicum pathway reactions

From KEGG (KYOTO ENCYCLOPEDIA OF GENES AND GENOMES) Pathway Database

Sr.	Glycolysis
R1	GLC + PEP = G6P + PYR
R2f	G6P = F6P
R2b	F6P = G6P
R3	F6P + ATP = F16DP
R4f	F16DP = 2 GAP
R4b	2 GAP = F16DP
R5f	GAP = G13DP + NADH
R5b	G13DP + NADH = GAP
R6f	G13DP = PEP + ATP
R6b	PEP + ATP = G13DP
R7	PEP = PYR + ATP

Sr.	Pentose Phosphate Pathway
R8	G6P = Ru5P + CO2 + 2 NADPH
R9f	Ru5P = X5P
R9b	X5P = Ru5P
R10f	Ru5P = R5P
R10b	R5P = Ru5P
R11f	X5P + R5P = GAP + S7P
R11b	GAP + S7P = X5P + R5P
R12f	S7P + GAP = F6P + E4P
R12b	F6P + E4P = S7P + GAP
R13f	X5P + E4P = F6P + GAP
R13b	F6P + GAP = X5P + E4P

Sr	Pyruvate Metabolism
R14	PYR + 2 FDO = AcCoA + 2 FDR + CO2
R15	CO2 = CO2x
R16f	2 FDO + NADPH = 2 FDR
R16b	2 FDR = 2 FDO + NADPH
R17f	2 FDO + NADH = 2 FDR
R17b	2 FDR = 2 FDO + NADH
R18	2 FDR = 2 FDO + H2
R19	H2 = H2x
R20f	PYR + NADH = LAC
R20b	LAC = PYR + NADH
R21	LAC = LACx

Sr.	Citric Acid Cycle
R22	OACT + AcCoA = CTR
R23f	CTR = ICTR
R23b	ICTR = CTR
R24	ICTR = AKG + NADH + CO2
R25	AKG = SCNCoA + NADH + CO2
R26f	SCNCoA = SCN + ATP
R26b	SCN + ATP = SCNCoA
R27f	FUM + FADH2 = SCN
R27b	SCN = FUM + FADH2
R28f	FUM = MAL
R28b	MAL = FUM
R29f	MAL = OACT + NADH
R29b	OACT + NADH = MAL

Sr.	Butanoate Matabolism
R30f	2 AcCoA = AcAcCoA
R30b	AcAcCoA = 2 AcCoA
R31f	AcAcCoA + NADPH = HCoA
R31b	HCoA = AcAcCoA + NADPH
R32f	HCoA = CCoA
R32b	CCoA = HCoA
R33f	CCoA + NADH = BCoA
R33b	BCoA = CCoA + NADH

Sr.	Butanoate Matabolism
R30f	2 AcCoA = AcAcCoA
R30b	AcAcCoA = 2 AcCoA
R31f	AcAcCoA + NADPH = HCoA
R31b	HCoA = AcAcCoA + NADPH
R32f	HCoA = CCoA
R32b	CCoA = HCoA
R33f	CCoA + NADH = BCoA
R33b	BCoA = CCoA + NADH

Sr.	Acid Phase Reactions
R34f	AcCoA = ACP
R34b	ACP = AcCoA
R35f	ACP = ACT + ATP
R35b	ACT + ATP = ACP
R36f	ACT = ACTx
R36b	ACTx = ACT
R37f	BCoA = BTP
R37b	BTP = BCoA
R38f	BTP = BTR + ATP
R38b	BTR + ATP = BTP
R39f	BTR = BTRx
R39b	BTRx = BTR

Sr.	Solvent Phase Reactions
R40f	ACD = AcCoA + NADH
R40b	AcCoA + NADH = ACD
R41f	ACD + NADH = ETH
R41b	ETH = ACD + NADH
R42	ETH = ETHx
R43f	AcAcCoA = AcACT
R43b	AcACT = AcAcCoA
R44	AcACT = ACN + ATP
R45	ACN = ACNx
R46	BCoA + NADH = BTD
R47f	BTD + NADPH = BUT
R47b	BUT = BTD + NADPH
R48	BUT = BUTx
R49	ACT + AcAcCoA = AcACT + AcCoA
R50	BTR + AcAcCoA = BCoA + AcACT

Sr.	Granulose Accumulation
R51f	G6P = G1P
R51b	G1P = G6P
R52	G1P + ATP = ADPG
R53	ADPG = GRN
R54	GRN = G1P

Sr.	Anapleurotic Reactions
R61	PEP + CO2 = OACT
R62	PYR + ATP = OACT

Sr	Maintenance/Transhydrogenation/Oxidative Phosphorilation
R56	ATP = MAINT
R57f	NADPH = NADH
R57b	NADH = NADPH
R58	NADH = FADH2
R59	NADH = 2 ATP
R60	FADH2 = ATP

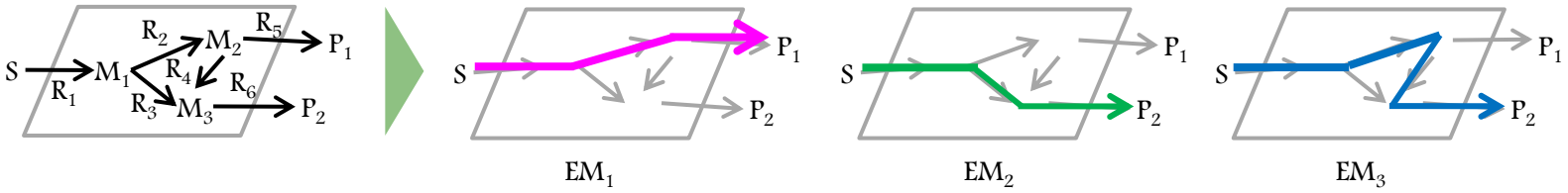
Sr.	
R63	NH3x = NH3

Biomass Formation
0.20 G6P + 0.81 R5P + 0.356 E4P + 2.29 PEP + 2.95 PYR + 2.24 AcCoA + 1.12 AKG + 1.83 OACT + 40.06 ATP + 12.69 NADH + 10.09 NH3 = BIOM + 0.30 CO2

3. Decomposition of network into EFMs

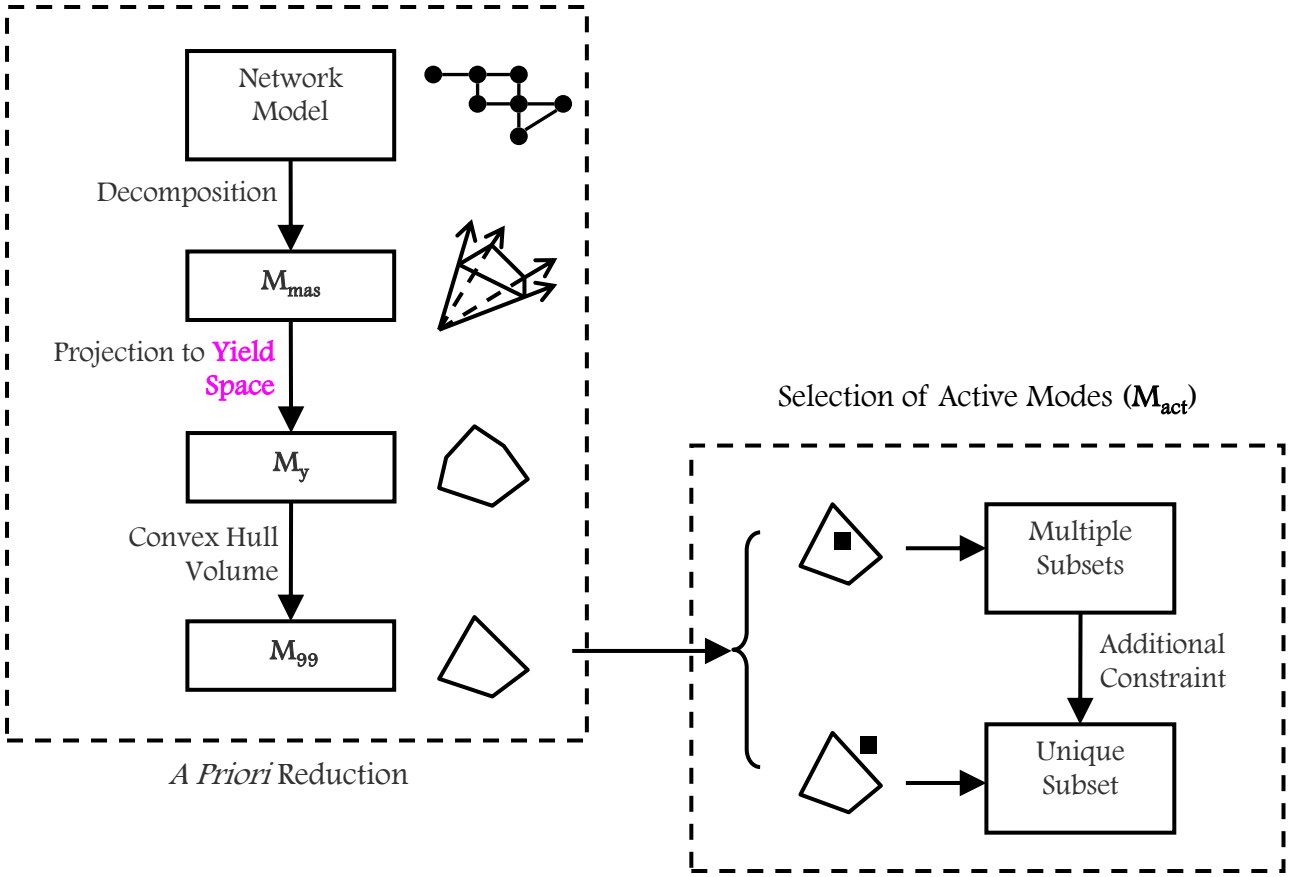
- Elementary Flux Modes are a set of nondecomposable pathways consisting of a minimal set of reactions that function in steady state

- Using METATOOL 5.1
  - Kamp A. and Schuster S.
    (Department of Bioinformatics, Friedrich-Schiller-University, Jena, Germany)



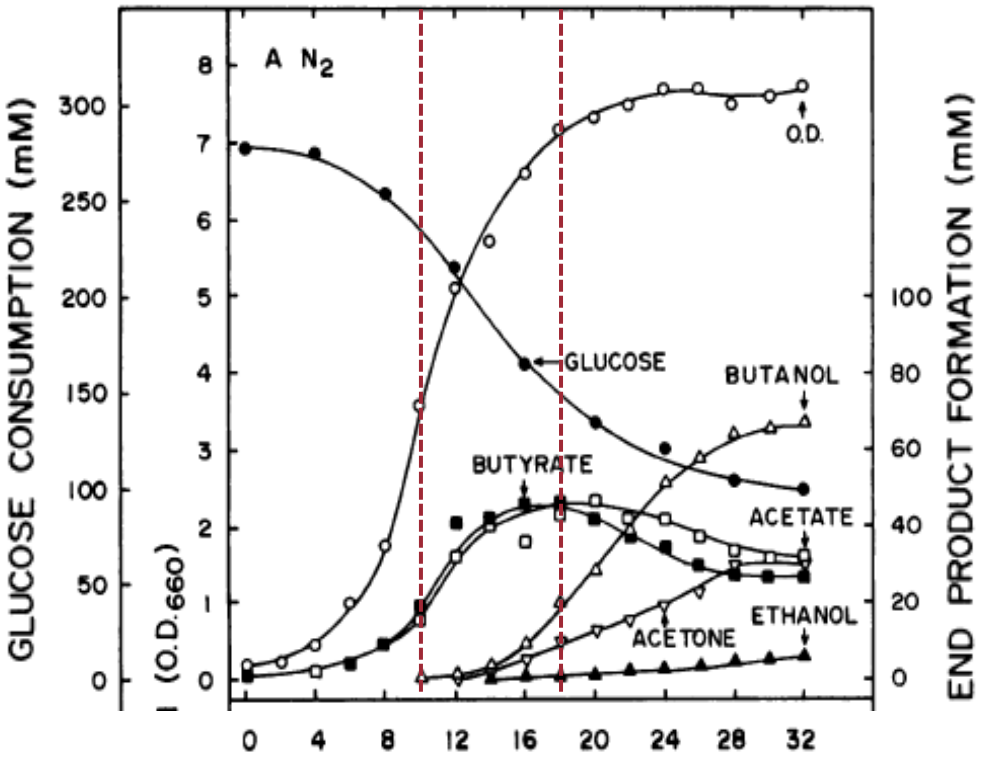
- Number of total EFMs = 30091
- Number of Glucose consuming EFMs = 19715

4. Reduction of set of EFMs



Song et. al., (2009)

Fermentation profile of *Clostridium acetobutylicum* ATCC 4259



Kim et al. 1984

Selection of M_{act} based on experimental yield data

- Phase I

Extracellular products	Biomass	Acetic acid	Butyric acid
Y_{Model}	0.0250	0.3580	0.3900
Y_{Expt}	0.0250	0.3580	0.3900

- Phase II

Extracellular products	Biomass	Acetic acid	Butyric acid	Ethanol	Acetone	Butanol
Y_{Model}	0.0142	0.3655	0.3199	0.0729	0.0080	0.1169
Y_{Expt}	0.0150	0.3940	0.3940	0.0750	0.0080	0.1260

- Phase III

Extracellular products	Biomass	Acetic acid	Butyric acid	Ethanol	Acetone	Butanol
Y_{Model}	0.0050	-0.1843	-0.3028	0.3398	0.0798	0.8405
Y_{Expt}	0.0050	-0.1800	-0.2800	0.3800	0.0800	0.8940

Reduced EFMs

Phase	M_{mas}	M_y	M_{99}	M_{act}	M_{act}
I	2127	6	6	4	15
II	6181	47	28	6	
III	18529	113	47	5	

Group	EFM	Net Reactions
GLC	2	GLC = BTR
	3	GLC = 2 ACT
	4	GLC = 0.1080 BIOM + 0.5120 ACT
GLC + ACT	5	GLC + 35 ACT = 18 ACN
	6	GLC + 24 ACT = 10 BTR
	10	GLC + 15 ACT = 7 BTR
GLC + BTR	7	GLC = 10 BTR + 9BUT
	8	GLC + 4 BTR = 6 ETH + 2 ACN
	9	GLC + 1.8664 BTR = 0.1920 BIOM + 3.0876 ACT
	11	GLC + 2 BTR = 2 ACN + BUT
	12	GLC + 2 BTR = 2 ETH + 2 CAN
	13	GLC + BTR = 2 ACT + BUT
GLC + ACT + BTR	14	GLC + 15.4229 ACT + 0.3245 BTR = 0.1920 BIOM +7.7134 ACN

$Z =$

MODES	1	2	3	4	5	6	7	8	9	10	11	12	13	14
GLU	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
BIOM	0	0	0	0.1080	0	0	0	0	0.1920	0	0	0	0	0.1920
ACT	0	0	2	0.5120	-35	-24	0	0	3.0876	-15	0	0	2	-15.4229
BTR	0	1	0	0	0	10	-10	-4	-1.8664	7	-2	-2	-1	-0.3245
ETH	0	0	0	0	0	0	0	6	0	0	0	2	0	0
ACN	0	0	0	0	18	0	0	2	0	0	2	2	0	7.7134
BUT	0	0	0	0	0	0	9	0	0	0	1	0	1	0

5. Model Building

- Metabolism in the algebraic form : $Sr = w$ (1)
- The vector of extracellular variables : $x = \begin{bmatrix} s \\ p \\ c \end{bmatrix}$ (2)
- The intracellular variables are represented by vector $\mathbf{m}$
- The dynamic model then can be represented by,

$$Sr = \begin{bmatrix} \frac{1}{c}\frac{dx}{dt} \\ \frac{dm}{dt} \end{bmatrix} \quad (3)$$

- Applying pseudo-steady state hypothesis on internal metabolites,

$$\frac{dm}{dt} = 0 \quad (4)$$

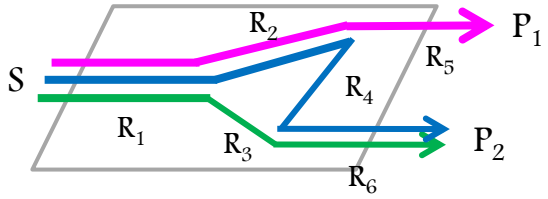
Kim et.al.,(2008)

- So, the dynamic model of interest reduces to :

$$S_x r = \frac{1}{c}\frac{dx}{dt} \qquad (5)$$

where, $\mathbf{S_x}$ is the stoichiometric matrix of extracellular fluxes

- The reaction rate vector is expressed in terms of elementary mode decomposition as :



$$\frac{1}{c}\frac{dx}{dt} = S_x r \qquad (6)$$

$$r = Zr_M$$

$$\frac{1}{c}\frac{dx}{dt} = S_x Z r_M$$

$$r = Zr_M = \begin{bmatrix} z_1 & z_2 & \ldots z_n \end{bmatrix} \begin{bmatrix} r_{M_1} \\ r_{M_2} \\ \vdots \\ r_{M_r} \end{bmatrix}$$

- The differential equation for enzymes,

$$\frac{de_{M_i}}{dt} = \alpha + r_{E_i} u_i - (\beta + r_G) e_{M_i} \qquad (7)$$

- Growth rates through all modes,

$$r_G = \sum_{i=1}^{n_z} Z_{i,n_f} v_i r_{M_i} \qquad (8)$$

- where specific uptake rate (r_M) of each mode and the enzyme synthesis rate (r_E)is given by,

$$r_{M_i} = k_i^{\max} e_i v_i \frac{s_1}{K_i + s_1} \qquad r_{E_i} = k_E u_i \frac{s_1}{K_i + s_1} \qquad i = 1,2,\ldots..,n_z - 1 \qquad (9 \& 10)$$

k^{max} - maximum uptake rate

n_z - number of elementary modes

Model Equations

- Fluxes of extracellular species:

$$\frac{dx}{dt} = S_x Z r_M c$$

- Uptake rate:

Modes	Substrate uptake rates	
1, 2, 3 & 4	$r_{M,i} = v_i e_i k_i^{max} \frac{x_{GLC}}{K_G + x_{GLC}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right]$	➢ GLC
5, 6 & 10	$r_{M,i} = v_i e_i k_i^{max} \frac{x_{GLC}}{K_G + x_{GLC}} \frac{x_{ACT}}{K_A + x_{ACT}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right]$	➢ GLC + ACT
7, 8, 9, 11, 12 & 13	$r_{M,i} = v_i e_i k_i^{max} \frac{x_{GLC}}{K_G + x_{GLC}} \frac{x_{BTR}}{K_B + x_{BTR}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right]$	➢ GLC + BTR
14	$r_{M,i} = v_i e_i k_i^{max} \frac{x_{GLC}}{K_G + x_{GLC}} \frac{x_{ACT}}{K_A + x_{ACT}} \frac{x_{BTR}}{K_B + x_{BTR}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right]$	➢ GLC + ACT + BTR

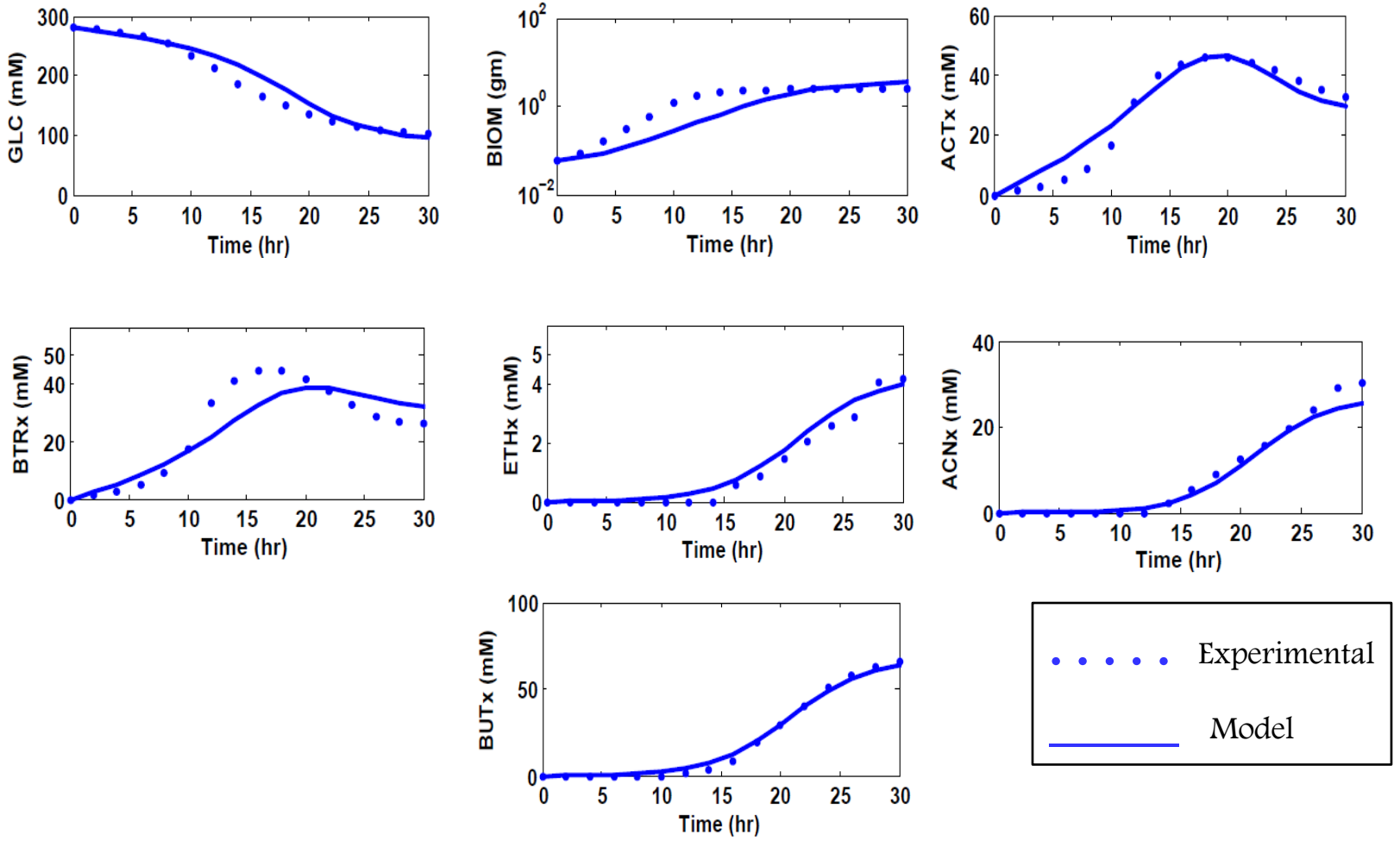
Model Equations

- Enzyme synthesis rates:

Modes	Enzyme synthesis rates
1, 2, 3 & 4	$\frac{de_i}{dt} = \alpha_i + u_i k_{E,i} \frac{x_{GLC}}{K_G + x_{GLC}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right] - (\beta_i + \mu) e_i$
5, 6 & 10	$\frac{de_i}{dt} = \alpha_i + u_i k_{E,i} \frac{x_{GLC}}{K_G + x_{GLC}} \frac{x_{ACT}}{K_A + x_{ACT}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right] - (\beta_i + \mu) e_i$
7, 8, 9, 11, 12 & 13	$\frac{de_i}{dt} = \alpha_i + u_i k_{E,i} \frac{x_{GLC}}{K_G + x_{GLC}} \frac{x_{BTR}}{K_B + x_{BTR}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right] - (\beta_i + \mu) e_i$
14	$\frac{de_i}{dt} = \alpha_i + u_i k_{E,i} \frac{x_{GLC}}{K_G + x_{GLC}} \frac{x_{ACT}}{K_A + x_{ACT}} \frac{x_{BTR}}{K_B + x_{BTR}} \left[1 - \left(\frac{x_{BUT,i}}{x_{BUT,m}}\right)^n\right] - (\beta_i + \mu) e_i$

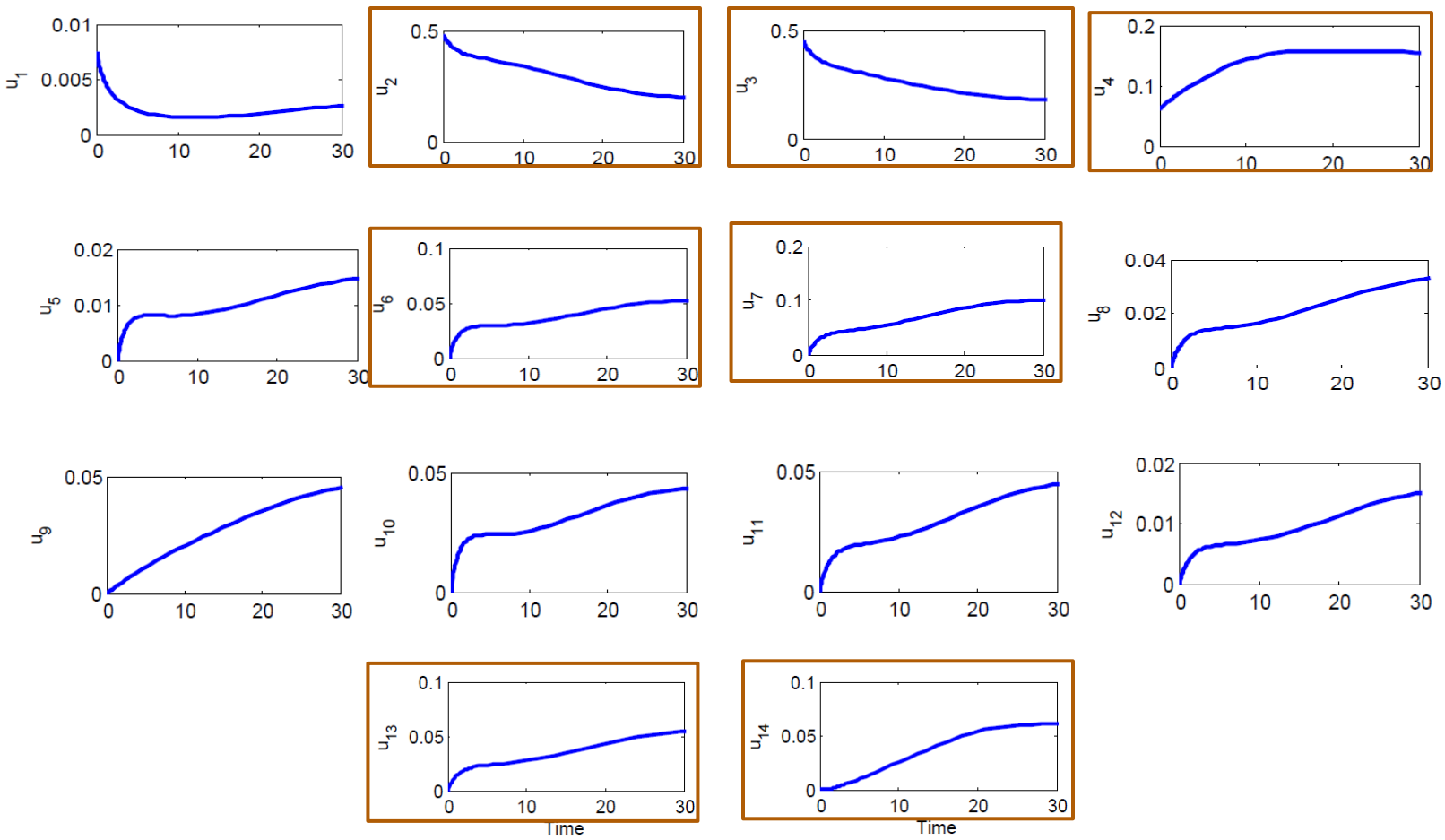
Metabolites profile

Metabolites profile of *Clostridium acetobutylicum* ATCC 4259



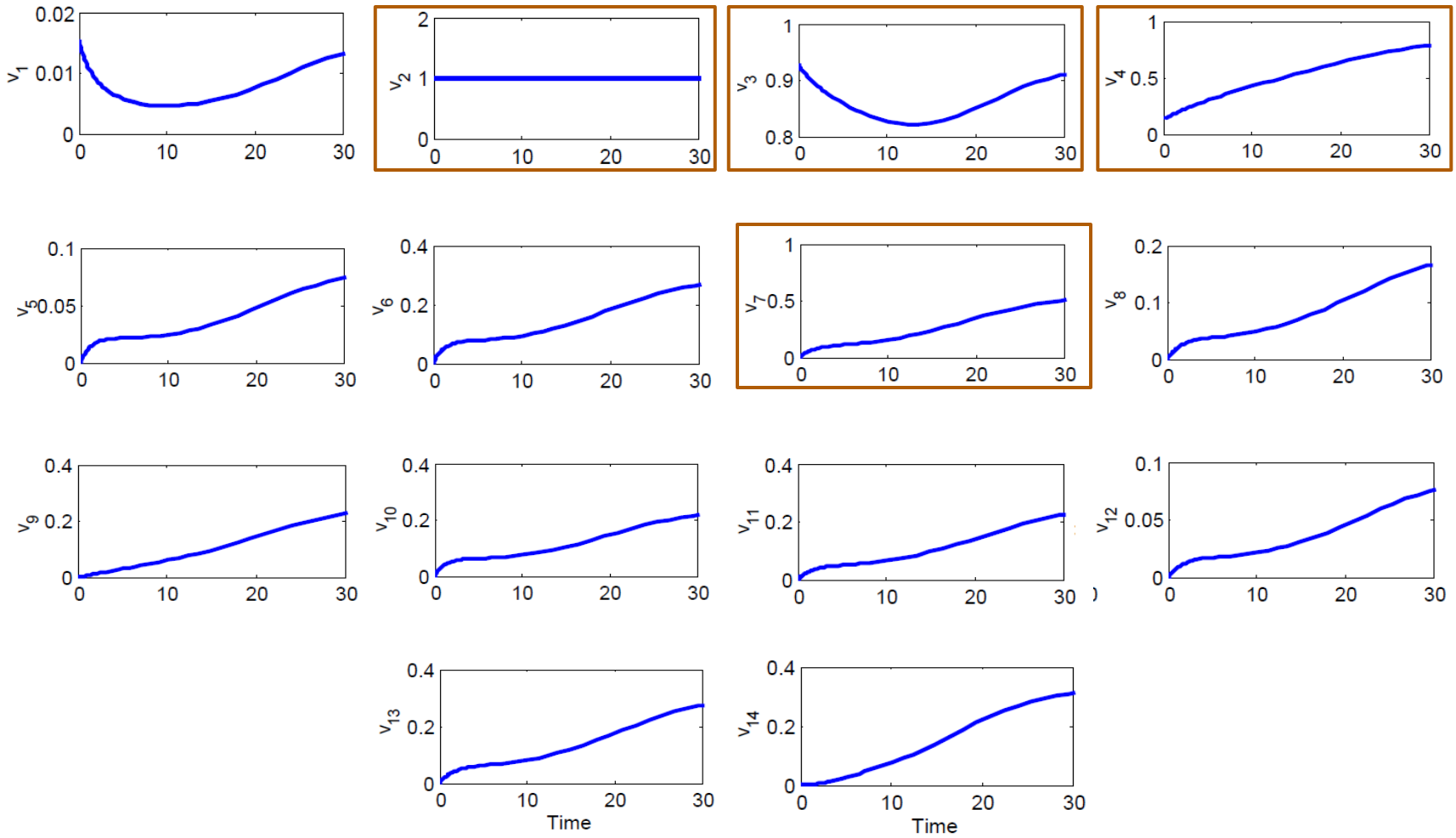
Cybernetic variable (u)

Cybernetic variable (u) of each elementary flux mode



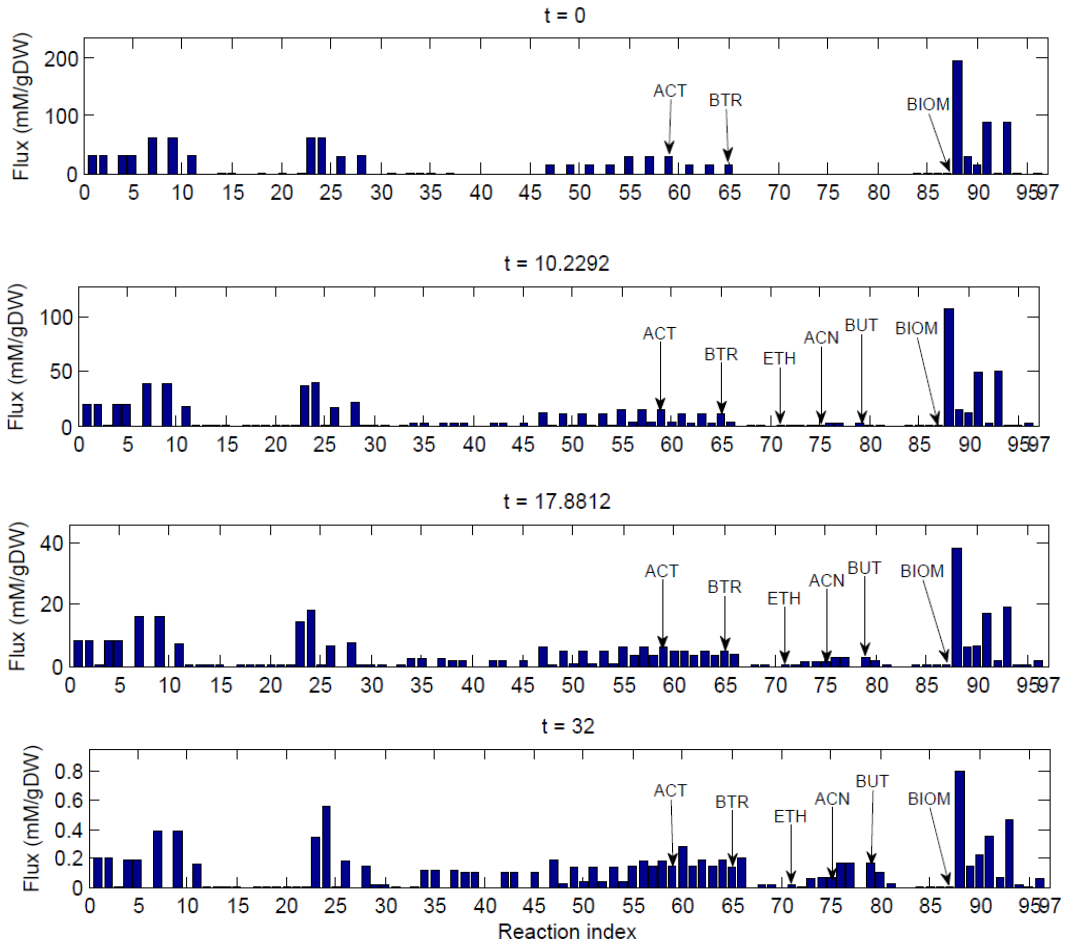
Cybernetic variable v

Cybernetic variable (v) of each elementary flux mode



Fluxes through metabolic pathway

Fluxes of metabolic pathway at different time intervals



Conclusion

- Cybernetic modeling framework best describes the control action of regulatory processes
- The dynamic framework resulting from cybernetic models have been shown to describe dynamic data on concentrations of biomass, substrate and extracellular variables
- Metabolic pathway of *Clostridium acetobutylicum* is modeled from limited set of data
- Hybrid Cybernetic Modeling reduces the burden of parameter estimation by reducing the number of elementary modes, while describing the metabolic network