Construction, analysis and assessment of relevance of an algebraic model for a class of biochemical networks

Siddhartha Kundu[1, a]

[1] Department of Biochemistry,
All India Institute of Medical Sciences,
Ansari Nagar,
New Delhi 110029,
INDIA
[a] Correspondence: siddhartha_kundu@yahoo.co.in, siddhartha_kundu@aiims.edu

Abstract

The intracellular milieu presents a complex physicochemical environment where molecular redundancy prevails and infra-threshold perturbations are integrated by biochemical networks. The pathways that result from these interactions are complex and will result in a plethora of signalling cascades. The stoichiometry number matrix for a biochemical network is a suitable way to represent the interactions between various molecular species under baseline conditions and in response to stimuli. Here, we model a class of biochemical networks with a set of constrained, reaction-centric, stoichiometry equivalent and degenerate matrices. The matrices exhibit a many-to-one surjection with the null space and form a semigroup with respect to addition. The parameters for these studies are the probable dissociation constants and will be used to derive several network- and reaction-specific metrics. These will describe the class of modelled biochemical networks both, at the level of a single network and as a unit. The model is extendible and can be used to perturb a biochemical network by introducing a finite number of extraneous reactions and then comparing pairs of like-reactions. The theoretical assertions presented are complemented by detailed computational analyses of the hexose monophosphate shunt, urea cycle and folate metabolism in humans. The model provides a theoretically sound basis to interrogate the effects of molecular redundancy and perturbations in the genesis and regulation of complex biochemical function. The model is theoretically sound, mathematically rigorous, readily testable, biochemically relevant, easily parameterizable and can be used to compare biochemical networks under differing intracellular conditions, both, between cells and across taxa.

Keywords: Algebraic model for a class of biochemical networks; Constrained, reaction-centric, stoichiometry equivalent and degenerate matrices; Probable dissociation constant-based parameterization; Reaction outcome; Network-specific metrics; Computational studies.

Introduction

Biochemical networks involve enzyme-mediated transformations of internal metabolites, exchange- and diffusion-reactions, formation/disassembly of molecular complexes and are regulated by short (feedback mechanisms, post-translational modifications)- and long (*de novo* synthesis, degradation)-term mechanisms [1, 2]. A biochemical network is either studied by optimizing the biomass of a molecular species of interest and enumerating elementary vectors or is data-driven, inverse and inferential. Algorithms which optimize one or more reactants are likely to be biased in choice, type and usage of

constraints, all of which will result in a study with reduced biomedical relevance despite being numerically sound [1, 3-6]. Enumeration-based algorithms on the other hand, will involve extensive computations that are NP-hard and will become intractable for large biochemical networks [2, 6-8]. The availability of- and accessibility to- empirical- and clinical-datasets has the potential to recreate, without bias, plausible data-driven biochemical networks, infer directionality for a given biochemical network (polynomial equation) or represent an arbitrary state of a modelled network (continuous-time Lyapunov equation) [9-12]. In these approaches perturbations are modelled as stochastic fluctuations and are used to derive the covariance (weak)- or correlation (strong)-interaction matrices for the set of nodes that comprise the modelled biochemical network [13-15].. This reliance on empirical data both, *a priori* and *a posteriori*, along with the computational complexity in computing the Jacobian are limitations in the inferential study of biochemical networks [13-15].

These limitations notwithstanding, an inferred biochemical network has tremendous translational potential in precision medicine and data mining [16-20]. For example, the use of guidance-based radio- and immune-therapeutics in managing chronic inflammatory diseases, neurodegenerative pathologies and recalcitrant and recurring malignancies is being used to assess dose response and predict outcome such as the propensity, progress and resolution of an adverse drug reaction [16-20]. In particular, these investigations combine the expression/abundance profiles of system-specific panels of biomarkers (genes, proteins, small molecules) into one of several machine learning (ML)-algorithms and predict patient response and/or compliance [18-20]. However, despite these advancements the primary or raw data (time-series kinetic data, “omics”-based datasets) is dependent on the platform used, experimental design and will require several numerical assumptions and extensive- and varied-normalization protocols [9-14]. Additionally, the use of expensive equipment, reagents and manpower can potentially restrict the utility of these investigations. In contrast to the aforementioned discussion on large scale modelling and analysis, causal networks and the inferences that result thereof, examine the effects of variable stimuli on several equally plausible and well-defined outcomes of a few nodes. Although causal network modelling has had several proven successes in investigating discrete real world events, the requirement for reference data, simple architecture and simpler rules, precludes consideration of complex biochemical systems other than for mining large biological datasets for lists of statistically significant genes/proteins of well characterized networks and in response to defined perturbations [21-25].

As opposed to the data-driven approach discussed vide supra, algebraic models for biochemical networks are usually formulated as population models [26, 27]. In particular, chemical reaction networks, with regards to the partaking molecular species, are modelled as either self-sustaining/persistent once initialized or limiting (zero-occurring) [28, 29]. Although these investigations have enhanced our repertoire of modelling strategies for functional biochemical networks, novel data-analytics and parameter-free frameworks, the absence of biochemically-relevant parameterization may mitigate the potential biomedical impact and relevance of these studies [30, 31]. The dissociation (Kd)-, association (Ka)- and Michaelis-Menten (Km)-constants are important parameters in comprehending the outcome and thence the dominant direction of a biochemical reaction (non-enzyme, enzyme). There are several ways to compute the dissociation constant for a reaction. Although the vast majority of these are empirical (dual polarization interferometry, surface plasmon resonance, stopped-flow spectroscopy), data-driven inferential models (regularized least squares pKa spectrum, static binding isotherms) and theoretical assertions are not unknown [32-38]. The probable dissociation constant, for the ith-reaction is an unbiased numerical measure that is computed from a null space-generated subspace of combinatorially summed non-redundant and non-trivial vectors of the stoichiometry number matrix for a biochemical network [7, 8]. The probable dissociation constant is a strictly positive real number which can unambiguously compute the forward-, reverse- or equivalent-outcomes for a reaction [7, 8].

The studies discussed focus on specific instances of a biochemical network and will therefore, be valid for specific states of an intracellular environment. It is also not clear whether these results can be generalized to classes of biochemical networks such as those that might occur under varying physicochemical and fluctuating biochemical intracellular conditions. Here, we present an algebraic model for a class of biochemical networks where each component network is constrained by a zero-preserving change to the stoichiometry numbers for each molecular species. The model will be developed and assessed in accordance with established paradigms of biochemical kinetics. Detailed computational studies on known and well-characterized biochemical networks will be carried out and the results obtained will serve as proof-of-concept for the theoretical assertions presented. We will utilize the probable dissociation constants as a parameter for these investigations and derive metrics that will describe a class of biochemical networks both, at the level of a single network and as a unit. The manuscript is organized into an initial section where definitions and concepts in the form of

biochemically relevant axioms will be introduced. This is followed by the "Methods" section where we will present the computational tools needed to study and complement the theoretical assertions outlined. This section will also describe the derivation of network-specific metrics which will be needed to parameterize and thence compare biochemical networks. In the "Results and Discussion"-section we will develop the model for a class of biochemical networks. This will be followed by a detailed characterization and analyses of some well-characterized biochemical pathways of human metabolism. We conclude this study by summarizing the major findings of this study, discussing possible limitations and outlining future directions for this work. This is followed by detailed proofs of the definitions (D), lemmas (L), theorems (T) and corollaries (C) wherever applicable.

2 Definitions, concepts and preliminary results pertinent to modelling a class of biochemical networks

2.1 Stoichiometry number as a unit of change in the number of molecules for a species

Consider a simple unimolecular non-enzymatic reaction,

$$\mathrm{a.\,A} \underset{\mathrm{k_b}}{\overset{\mathrm{k_f}}{\rightleftarrows}} \mathbb{a}.\,\mathrm{AB} \qquad \mathrm{RXN}(1)$$

where,

$\mathrm{A} := \text{Reactant undergoing forward reaction with rate constant } \mathrm{k_f}\ (\mathrm{Rf} = \mathrm{k_f.\,A})$ (1)

$\mathrm{AB} := \text{Reactant undergoing reverse reaction with rate constant } \mathrm{k_b}\ (\mathrm{Rb} = \mathrm{k_b.\,AB})$ (2)

$\{-\mathrm{a}, \mathbb{a}\} := \text{Stoichiometry numbers for reactants and products of the forward reaction}$

$\{\mathrm{a}, -\mathbb{a}\} := \text{Stoichiometry numbers for reactants and products of the reverse reaction}$

$\mathrm{Kd} := \text{Dissociation constant for the reversible reaction RXN (1)}$

Let us subsume that this reaction is at a steady state,

$$\mathrm{Rf} = \mathrm{Rb} \qquad (3)$$

$$\Rightarrow \frac{\mathrm{AB}}{\mathrm{A}} = \frac{\mathrm{k_f}}{\mathrm{k_b}} \in \mathbb{R} \cap (0, +\infty) \text{ and } \mathrm{k_b} \neq 0 \qquad (3.1)$$

$$= \mathrm{Kd(Rf)} \qquad (3.2)$$

We can extend this annotation for the forward and reverse reactions,

$$\{\mathrm{Kd(Rf)}, \mathrm{Kd(Rb)}\} \in \mathbb{R}_+ \text{ where } \mathrm{k_f}, \mathrm{k_b} \neq 0 \quad \mathrm{Def.}\,(1)$$

Let us express the stoichiometry of a molecular species as a function of a reaction-specific dissociation constant with a linear map (f),

$$\mathrm{f}: \mathrm{Kd}(.) \in \mathbb{R} \cap (0, +\infty) \mapsto \phi(.) \in \mathbb{Z} \quad \mathrm{Def.}\,(2)$$
$$\text{where } \phi = \left\lceil \ln\left(\frac{1}{\mathrm{Kd}}\right) \right\rceil \qquad (4)$$

This map is onto and many-one since,

$$\text{if}\big(\mathrm{Kd}(.) \in \mathbb{R} \cap (0,1)\big) \text{ then } \phi(.) \in \mathbb{Z}_{-} \qquad \mathrm{Def.}\,(2a)$$
$$\text{if}(\mathrm{Kd}(.) \in \mathbb{R} \cap \{1\}) \text{ then } \phi(.) = 0 \qquad \mathrm{Def.}\,(2b)$$
$$\text{if}\big(\mathrm{Kd}(.) \in \mathbb{R} \cap (1, +\infty)\big) \text{ then } \phi(.) \in \mathbb{Z}_{+} \qquad \mathrm{Def.}\,(2c)$$

We will now describe a biochemical network with j-indexed $(\mathrm{j} = 1, \ldots, \mathrm{J})$ m-molecular species $(\mathrm{m}_1, \ldots, \mathrm{m}_\mathrm{J})$ as the vector,

$$\mathbb{m} = [\mathrm{m}_1, \ldots, \mathrm{m}_\mathrm{J}]^\mathrm{T} \quad Def.\,(3)$$

The reactions that comprise this network are i-indexed $(\mathrm{i} = 1, \ldots, \mathrm{I})$ r-reactions $(\mathrm{r}_1, \ldots, \mathrm{r}_\mathrm{I})$ each of which is represented by a dissociation constant,

$$\mathrm{r}_\mathrm{i} \equiv \mathrm{Kd}_\mathrm{i} \in \mathbb{R} \cap (0, +\infty) \quad \mathrm{Def.}\,(4)$$

We will express the stoichiometry number of a molecular species in terms of a reaction-specific dissociation constant and in accordance with the linear map (f). Since we are concerned with biochemical networks within a cell we will subsume that every participating molecular species will contribute in some measure to each and every reaction (**S1**, **Figure 1**),

$$\begin{aligned}
\mathrm{Kd}_\mathrm{i}(\mathbb{m}) &= \left[\mathrm{Kd}_\mathrm{i}(\mathrm{m}_1), \ldots, \mathrm{Kd}_\mathrm{i}(\mathrm{m}_\mathrm{j})\right]^\mathrm{T} && (5)\\
&= [\mathrm{m}_{1\mathrm{i}}, \ldots, \mathrm{m}_{\mathrm{J}\mathrm{i}}]^\mathrm{T} && (5.1)\\
&= [\mathrm{m}_1, \ldots, \mathrm{m}_\mathrm{J}]_\mathrm{i}^\mathrm{T} && (5.2)\\
&\stackrel{\mathrm{def}}{=} \mathbf{r}_\mathrm{i} \in \mathbb{Z}^\mathrm{J} && (6)
\end{aligned}$$

This means that for I-reactions of a biochemical network $(\mathscr{p})$ we have the following ordered collection of reaction vectors,

$$\mathscr{p} \stackrel{\mathrm{def}}{=} [\mathbf{r}_1, \ldots, \mathbf{r}_\mathrm{I}]^\mathrm{T} \in \mathbb{Z}^{\mathrm{J.Ix1}} \quad Def.\,(5)$$

Rewriting,

$$\mathscr{p} = \left[m_{11}, \ldots, m_{J1}, m_{12}, \ldots, m_{J2}, m_{13}, \ldots, m_{JI}, \ldots, m_{1I}, \ldots, m_{JI}\right]^{T} \quad (7)$$
$$= \mathbb{Z}^{J} \otimes \mathbb{Z}^{I} \quad (7.1)$$
$$= \begin{bmatrix} m_{11} & \cdots & m_{1I} \\ \vdots & \ddots & \vdots \\ m_{J1} & \cdots & m_{JI} \end{bmatrix}_{J\times I} \quad (7.2)$$
$$= \mathbb{Z}^{J\times I} \quad (7.3)$$

We describe a biochemical network with a sparse matrix of $J \times I$ dimensions with entries that correspond to the stoichiometry numbers of all J-molecular species that participate in I-reaction vectors (**S2**, **Figure 1**).

$$\mathscr{p} \stackrel{\text{def}}{=} \mathcal{S}_{\mathscr{p}} \in \mathbb{Z}^{J\times I} \quad (8)$$
where
$i = 1, \ldots, I$
$j = 1, \ldots, J$

2.2 Definition and relevance of a plausible reaction vector to the stoichiometry number matrix of a biochemical network

The intracellular milieu is an interplay of physicochemical- and biochemical-processes and is mediated by several molecular species. Any model for a biochemical network must necessarily include most, if not all contributors to this complexity. We define a plausible reaction vector as one which when populated will likely represent a true biochemical reaction rather than being just numerically tenable (**Def. (6)**). The subset of all such plausible reaction vectors and the revised stoichiometry number matrix that is formed is (**S1 and S2**, **Figure 1**),

$$\mathbf{R} = \left(\mathbf{r}_1, \ldots, \mathbf{r}_{\bar{\bar{I}}}\right) \subset \mathbb{A} = \left(\mathbf{r}_1, \ldots, \mathbf{r}_{\bar{\bar{I}}}, \mathbf{r}_{\bar{\bar{I}}+1}, \ldots, \mathbf{r}_I\right) \quad \text{Def. (6a)}$$
$$\mathscr{p} = \mathcal{S}_{\mathscr{p}} \in \mathbb{Z}^{J\times\bar{\bar{I}}} \quad (9)$$

where,
$I :=$ Cardinality of the set of all reaction vectors for a stoichiometry number matrix of a biochemical network — Def. (6b)
$= \#\mathbb{A}$ (10)
$\bar{\bar{I}} :=$ Cardinality of the subset of plausible reaction vectors for a stoichiometry number matrix of a biochemical network — Def. (6c)
$= \#\mathbf{R}$ (11)

Since a reaction vector can be populated combinatorially with t-combinations of J-molecular species, we can have a total of I-reaction vectors each with a different combination of J-molecular species [8],

$$I = \sum_{t=0}^{t=J} \binom{J}{t} \quad (12)$$

We will include only those reaction vectors which are likely to represent a true biochemical reaction and exclude the rest. In order to accomplish this we will present a set of axioms which when implemented will result in the subset of plausible reaction vectors (**S2**, **Figure 1**),

$$\bar{\bar{I}} < I \quad (13)$$

The occurrence of a non-reaction is highly improbable since the intracellular environment is replete with futile cycles, sequestration by membrane-bound compartments, feedback loops, partial reactions and competition for the same metabolite by opposing pathways,

Axiom 1 (A1): A plausible reaction vector is necessarily non-trivial since the converse will indicate a non-reaction,

$$\begin{aligned} &\mathbf{r}_i \neq \mathbf{0} \qquad (14)\\ &\text{where,}\\ &\mathbf{r}_i \in \mathbf{R} \subset \mathbb{A} \subset \mathbb{Z}^J\\ &j = 1, \ldots, J\\ &i = 1, \ldots, \bar{\bar{I}} \end{aligned}$$

Axiom 2 (A2): A plausible reaction vector will have at most two equally likely forms,

$$\begin{aligned} &\mathbf{r}_i \stackrel{\text{def}}{=} \begin{cases} \mathbf{r}_i \\ (-1).\mathbf{r}_i = \mathbf{r}_k \end{cases} \quad \text{Def. (7)}\\ &\text{where,}\\ &\mathbf{r}_i, \mathbf{r}_k \in \mathbf{R} \subset \mathbb{A} \subset \mathbb{Z}^J\\ &j = 1, \ldots, J\\ &i = 1, \ldots, \bar{\bar{I}} \end{aligned}$$

There are a multitude of ways by which a true biochemical reaction can be modelled by a plausible reaction vector. We can account for a large proportion of these by imposing conditions on how a plausible reaction vector should be populated.

For example, a biochemical reaction comprises molecular species which will occupy different sides of a reaction. Here, we will represent this by utilizing signed molecular species $\left((+)m_j, (-)m_j\right)$. A reactant is defined as a molecular species which will be utilized/converted by a biochemical reaction $\left((-)m_j\right)$ (**Def. (8)**). A product is defined as a molecular species which will be formed by a biochemical reaction $\left((+)m_j\right)$ (**Def. (9)**). This schema can only arise for a molecular species with a net contribution and will be expected if the generating reaction is strictly forward or reverse $(Kd_i \in \mathbb{R}_+ - \{1\})$. In

contrast, a molecular species without a net contribution $\left(m_j := 0\right)$ is one which results from an equivalent reaction $\left(Kd_i \in \mathbb{R}_+ \cap \{1\}\right)$ (**Def. (10)**).

Axiom 3 (A3): The stoichiometry contribution by any j-indexed m-molecular species $\left(m_j\right)$ to a plausible reaction vector is,

$$m_j = \begin{cases} (-)m_j \stackrel{\text{def}}{=} \text{Molecular species as a reactant} \\ (+)m_j \stackrel{\text{def}}{=} \text{Molecular species as a product} \\ 0 \stackrel{\text{def}}{=} \text{Molecular species without a net contribution} \end{cases} \quad (15)$$

where,

$m_j \in \mathbf{r}_i \in \mathbf{R} \subset \mathbb{A} \subset \mathbb{Z}^J$

$j = 1, \ldots, J$

$i = 1, \ldots, \bar{\bar{I}}$

Axiom 4 (A4): We further impose that every plausible reaction vector must have at least one pair of molecular species that are differently signed,

$$\exists\left((-)m_j, (+)m_k\right) \in \mathbf{r}_i \qquad \text{Def. (11)}$$

where

$m_j, m_k \in \mathbf{r}_i \in \mathbf{R} \subset \mathbb{A} \subset \mathbb{Z}^J$

$j, k \in [1, J] \qquad (16)$

$j \neq k \qquad (17)$

$j = 1, \ldots, J$

$i = 1, \ldots, \bar{\bar{I}}$

Axiom 4.1 (A4.1): There must be a minimum of 2 non-zero entries for any plausible reaction vector.

$$\#\mathbf{r}_i \geq 2 \qquad [\text{From (A3), (A4)}] \quad (18)$$

where

$\mathbf{r}_i \in \mathbf{R} \subset \mathbb{A} \subset \mathbb{Z}^J$

$j = 1, \ldots, J$

$i = 1, \ldots, \bar{\bar{I}}$

Although **A1-A4** are easily relatable to a true biochemical reaction, the introduction of signed stoichiometry numbers will lead to vectors that are linearly dependent. These when added to the original vector will result in one or more trivial vectors.

Axiom 5 (A5): For any set of plausible reaction vectors each constituent vector must be linearly independent,

$$\mathbf{r}_i = (-1).\mathbf{r}_k \quad (19)$$
$$\mathbf{r}_i + \mathbf{r}_k = (-1).\mathbf{r}_k + \mathbf{r}_k \quad (20)$$
$$= \mathbf{0} \quad (20.1)$$
where
$$\mathbf{r}_i, \mathbf{r}_k \in \mathbf{R} \subset \mathbb{A} \subset \mathbb{Z}^J$$
$$i \neq k \quad (21)$$
$$i, k \in [1, \bar{\bar{I}}] \quad (22)$$

Whilst **A1-A5** will assist us in populating the stoichiometry number matrix for a biochemical network, we will also require indicators of biochemical function to characterize and thence compare biochemical networks. We will assume that the reactions and the molecular species that partake in them are at equilibrium,

Axiom 6 (A6): The null space of the stoichiometry number matrix for a biochemical network $N(\mathcal{S}_p)$ is a suitable representation of a biochemical network at equilibrium,

$$\{\boldsymbol{\nu}_k \in N(\mathcal{S}_p) \subset \mathbb{R}^{\bar{\bar{I}}} | \boldsymbol{\nu}_k{}^T \times \mathcal{S}_p = \vec{\mathbf{0}}^T\} \quad \text{Def. (12)}$$
where,
$$\boldsymbol{\nu}_k \neq \vec{\mathbf{0}} \quad (23)$$
$$\boldsymbol{\nu}_k \neq \boldsymbol{\nu}_l | \boldsymbol{\nu}_k, \boldsymbol{\nu}_l \in N(\mathcal{S}_p) \quad (24)$$
$$k, l \in \mathbb{N} \quad (25)$$

Biochemical networks and thence metabolism is directional and often, stimulus dependent. This means that along with the flux and biomass of a molecular species, we will require numerical values that can characterize individual reactions of a biochemical network. We will utilize the probable dissociation constant for the ith-reaction of a biochemical network to assign a forward $(\eta_i \in \mathbb{R} \cap (1, \infty))$-, reverse $(\eta_i \in \mathbb{R} \cap (0,1))$- or equivalent $(\eta_i \in \mathbb{R} \cap \{1\})$-outcome for a reaction [7, 8]. The probable dissociation constant is an unbiased numerical measure which can be directly computed from the null space of the stoichiometry number matrix for a biochemical network [7, 8].

In order to maintain uniformity of notation we will assign the same index (i) to the probable dissociation constant for a biochemical reaction as that of a plausible reaction vector [7, 8],

$$Kd_i \equiv \eta_i \quad (26)$$
where,
$$\eta_i \in \mathbb{R} \cap (0, \infty) \quad (27)$$
$$i = 1, \ldots, \bar{\bar{I}}$$

Axiom 7 (A7): The probable dissociation constant will be used to assign an outcome to the ith-reaction,

$$\begin{aligned}\eta_i &\in \mathbb{R} \cap (0,\infty) \\ &= \mathbb{R} \cap \big((0,1) \cup \{1\} \cup (1,\infty)\big) && (28) \\ &= \big\{\big(\mathbb{R} \cap (0,1)\big) \cup (\mathbb{R} \cap \{1\}) \cup \big(\mathbb{R} \cap (1,\infty)\big)\big\} && (28.1) \\ &\stackrel{\text{def}}{=} (\text{reverse}, \text{equivalent}, \text{forward}) && (29) \\ & i = 1, \ldots, \bar{\bar{I}}\end{aligned}$$

The set of probable dissociation constants for all reactions of a biochemical network $(\forall i)$, will be used to define and derive metrics by which we will describe a class of biochemical networks both, at the level of a single network and as a unit.

Axiom 8 (A8): We can partition the outcomes for every reaction of a biochemical network into two distinct subsets of reactions,

$$\begin{aligned}\text{Equivalent (e)} &:= \text{Uncertain reaction outcome} && \text{Def. (13)} \\ &\stackrel{\text{def}}{=} \eta_i \simeq 1.0 && (30)\end{aligned}$$

$$\begin{aligned}\text{Non} - \text{equivalent (ne)} &:= \text{Defined reaction outcome} && \text{Def. (14)} \\ &= (\text{forward}, \text{reverse}) \\ &\stackrel{\text{def}}{=} \eta_i \in \mathbb{R} \cap ((0,1) \cup (1,\infty)) && (31)\end{aligned}$$

Although **A1-A8** can be utilized to model and interrogate individual instances of a biochemical network, we will require additional axioms to extend these assertions so as to model a class of biochemical networks $(\mathcal{P})$. This will assist us in exploring the role of molecular redundancy and perturbations in the genesis of complex biochemical function.

As discussed earlier, the molecular species that are purported to partake in a biochemical network are either "persistent" or modelled as "siphons" [26-28]. This is with regards to whether the molecular species is allowed to actively and persistently contribute to the overall biomass or do so intermittently. Here, we consider the "persistent' scenario wherein we examine the contribution of J-molecular species to a finite set of x-indexed $(x = 1, \ldots, X)$ and X-membered set of biochemical networks,

$$\begin{aligned}&\mathcal{P} = (\wp_1, \ldots, \wp_X) && (32) \\ &\text{where} \\ &x = 1, \ldots, X && (33)\end{aligned}$$

Axiom 9 (A9): The stoichiometry number matrices that comprise a class of biochemical networks $\mathcal{S}(\mathcal{P})$ are equivalent iff the stoichiometry numbers contributed by each molecular species to any x-indexed matrix is true for all X-members of the class.

$$\begin{aligned}
&\text{Case 1: If}\left(m_{ji}(x) = 0 \text{ where } x \in [1, X]\right) \\
&\qquad \text{then } \left(m_{ji}(x) = 0\ \forall x\right) \qquad (34) \\
&\text{Case 2: If}\left(m_{ji}(x) \neq 0 \text{ where } x \in [1, X]\right) \\
&\qquad \text{then } \left(m_{ji}(x) \neq 0\ \forall x\right) \qquad (35) \\
&\text{where,} \\
&m_{ji} \in \mathcal{S}_{\mathscr{p}_x} \in \mathcal{P} \subset \mathbb{Z}^{1\times\bar{\bar{I}}} \\
&j \in [1, \bar{J}] \\
&i \in \left[1, \bar{\bar{I}}\right] \\
&x = 1, \ldots, X
\end{aligned}$$

For a finite set of molecular species with non-zero initial values it is clear that if we assume that every x will occur, it follows that every reaction (forward, reverse, equivalent) will occur. This will establish the "persistent" nature for the class of modelled biochemical networks will be ensured at all times [26-28].

Although biochemical networks within a cell are insulated from extreme physicochemical alterations, intracellular perturbations abound and may have a significant impact on biochemical and physiological function (**Def. (15)**). These will include non-mutable sequestrations (complex formation, membrane-bound compartmentalization), post-translational modifications (acetylation, glycosylation, phosphorylation) and polymorphism-induced variants of a molecular species. This definition precludes deleterious mutations or translocations that may involve the coding strand of a protein of interest and will alter the biochemical network irrevocably. The rationale for this exclusion in the present model is that the probability that these "knock-out" events occur are extremely rare with the vast majority of perturbations being classified as single nucleotide polymorphisms [7, 38].

Axiom 10 (A10): A biochemical network $(\mathscr{p})$ is minimally perturbed $(\tilde{\mathscr{p}})$ iff on the introduction of one or more extraneous reactions we can compare and thence assess the probable dissociation constants across all pairs of reactions $(\{\eta_i(\mathscr{p}), \eta_i(\tilde{\mathscr{p}})\}\ \forall i)$,

$$\text{If}\left(\begin{array}{rl} \bar{\bar{\mathrm{I}}} := \text{Plausible reaction vectors specific for } \mathcal{S}(\mathcal{p}) \text{ then,} & \\ \tilde{\mathrm{I}} := \text{Plausible reaction vectors specific for } \mathcal{S}(\tilde{\mathcal{p}}) & \text{Def. (16)} \\ \text{where,} & \\ \tilde{\mathrm{I}} \stackrel{\text{def}}{=} \bar{\bar{\mathrm{I}}} + \rho & (36) \\ > \bar{\bar{\mathrm{I}}} & (36.1) \\ \text{and,} & \\ \rho \geq 1 & (36.2) \\ \rho \in \mathbb{N} & (36.3) \end{array}\right)$$

$$\text{then}\left(\begin{array}{c} \left\{\left(\eta_1(\mathcal{p}), \eta_1(\tilde{\mathcal{p}})\right), \eta_2(\mathcal{p}), \eta_2(\tilde{\mathcal{p}}) \ldots \left(\eta_{\bar{\bar{\mathrm{I}}}}(\mathcal{p}), \eta_{\bar{\bar{\mathrm{I}}}}(\tilde{\mathcal{p}})\right)\right\} \\ \text{where,} \\ \mathcal{p} \rightarrow \tilde{\mathcal{p}} \quad (37) \\ \mathcal{S}(\mathcal{p}) \rightarrow \mathcal{S}(\tilde{\mathcal{p}}) \quad (38) \\ \text{and,} \\ \mathrm{i} = 1, \ldots, \bar{\bar{\mathrm{I}}} \\ \mathcal{S}(\mathcal{p}) = \mathcal{S}_{\mathcal{p}} \subset \mathbb{Z}^{\mathrm{J} \times \bar{\bar{\mathrm{I}}}} \\ \mathcal{S}(\tilde{\mathcal{p}}) = \mathcal{S}_{\tilde{\mathcal{p}}} \subset \mathbb{Z}^{\mathrm{J} \times \tilde{\mathrm{I}}} \quad (39) \end{array}\right) \qquad \text{Def. (17)}$$

3 Methods

3.1 Computational tools to investigate and characterize a biochemical network

We utilize "ReDirection" to conduct detailed computational studies on known and well-characterized biochemical networks to metabolize folate, hexose monophosphate and urea in humans [7]. "ReDirection" is an R-package that deploys a mathematically rigorous algorithm to compute the probable dissociation constant and assign an outcome to every reaction of a user-defined biochemical network [7, 8]. Briefly, "ReDirection" checks the stoichiometry number matrix for a biochemical network that is entered as input and ensures that there is strict compliance with the guidelines of usage [7]. These include checking for and excluding, if necessary all linear dependent reaction vectors, verifying lower bounds (number of molecular species, number of reaction vectors) and orientation of the reaction vectors (rows, columns) [7]. "ReDirection" follows this up by computing and defining a null space-generated subspace by removing all redundant and trivial vectors and then populating a reaction-specific sequence vector with numerical values drawn from each row and across all columns of this subspace [7, 8]. These steps are done iteratively and recursively until the lower-bound of this reaction-specific sequence vector is greater than 1 or the upper-bound is less than -1 or the number of positive- and negative-terms are almost equal [7, 8]. These must hold for all reactions unequivocally, i.e., $\forall \mathrm{i}$ where $\mathrm{i} = 1,2 \ldots \bar{\bar{\mathrm{I}}}$ are the rows of a null space-generated subspace. For all iterations thereafter, "ReDirection" will compute several descriptors (arithmetic mean, standard deviation) which are then used to screen each term of the reaction-specific sequence vector and bin them to outcome-specific

subsets whose terms are summed and mapped to a positive real number [7]. These constitute the ith-reaction-specific output vector and its p1-norm is the probable dissociation constant $\left(\eta_i \in \mathbb{R} \cap (0,\infty)\right)$ [7, 8].

3.2 Assessing the baseline- and perturbed-forms for a biochemical network

The probable dissociation constants that are computed by "ReDirection" will be used to define and derive several network-specific metrics. These will be used to describe a class of biochemical networks both, at the level of an individual network and as a unit.

For a biochemical network before $(\mathcal{p})$ and after $(\tilde{\mathcal{p}})$ the introduction of one or more extraneous reactions, the corresponding stoichiometry number matrices $\left(\boldsymbol{\mathcal{S}}_{\mathcal{p}}, \boldsymbol{\mathcal{S}}_{\tilde{\mathcal{p}}}\right)$ will be used as input to "ReDirection" [7] (**Figures 1 and 2**). The probable dissociation constants will be computed, used to assign unambiguous outcomes and are compared. The number, distribution and proportion of matched (m) non-equivalent (ne)- and equivalent (e)-reactions are computed,

$$\#\mathrm{ne}_{\mathcal{p}} = \#\mathrm{f}_{\mathcal{p}} + \#\mathrm{b}_{\mathcal{p}} \tag{40}$$

$$\#\mathrm{mne}_{\mathcal{p}} = \#\mathrm{mf}_{\mathcal{p}_{\mathrm{x}}} + \#\mathrm{mb}_{\mathcal{p}} \tag{41}$$

From (40), (41)

$$\omega_{\mathcal{p}} = \left(\frac{\#\mathrm{mf}_{\mathcal{p}} + \#\mathrm{mb}_{\mathcal{p}}}{\#\mathrm{f}_{\mathcal{p}} + \#\mathrm{b}_{\mathcal{p}}}\right).100 \tag{42}$$

$$= \frac{\#\mathrm{mne}_{\mathcal{p}}}{\#\mathrm{ne}_{\mathcal{p}}}.100 \tag{42.1}$$

$$\psi_{\mathcal{p}} = \frac{\#\mathrm{me}_{\mathcal{p}}}{\#\mathrm{e}_{\mathcal{p}}}.100 \tag{43}$$

$\mathcal{p}$:= Unperturbed biochemical network
$\tilde{\mathcal{p}}$:= Perturbed biochemical network
$\mathrm{f}_{\mathcal{p}}$:= Forward reactions of the unperturbed biochemical network
$\mathrm{mf}_{\mathcal{p}}$:= Matched forward reactions after perturbation
$\mathrm{b}_{\mathcal{p}}$:= Reverse reactions of the unperturbed biochemical network
$\mathrm{mb}_{\mathcal{p}}$:= Matched reverse reactions after perturbation
$\mathrm{ne}_{\mathcal{p}}$:= Non − equivalent reactions of the unperturbed biochemical network
$\mathrm{mne}_{\mathcal{p}}$:= Matched non − equivalent reactions after perturbation
$\mathrm{une}_{\mathcal{p}}$:= Unmatched non − equivalent reactions after perturbation
$\mathrm{e}_{\mathcal{p}}$:= Equivalent reactions of the unperturbed biochemical network
$\mathrm{me}_{\mathcal{p}}$:= Matched equivalent reactions after perturbation
$\mathrm{ue}_{\mathcal{p}}$:= Unmatched equivalent reactions after perturbation

On the basis of the data obtained from these calculations we will define criteria which can be used to describe and compare a biochemical network and thence a class of biochemical networks. A globally

relevant criterion is one which is applicable to a class of biochemical networks and must therefore, be less sensitive, numerically robust and specific (**Def. (18)**). On the other hand, locally relevant criteria is expected to function at the level of a single biochemical network and is likely to influenced by the physicochemical conditions that are prevalent at the time the network is functional. This means our pre-defined criteria must necessarily be more sensitive, less robust and less specific (**Def. (19)**).

We will examine the equivalent- and non-equivalent-outcomes as possible candidates to describe a class of biochemical networks. We will compute confusion matrices for each in models of biochemical networks (**Defs. (20-27**)) (**Table 1**),

$$\begin{aligned}
\text{Case1:}\quad & \text{Positive outcome} && := \text{Equivalent reaction} && (44)\\
& \text{Confusion matrix} && \stackrel{\text{def}}{=} \begin{bmatrix} \#\text{me}_{\mathscr{p}} = \text{TP} & \#\text{une}_{\mathscr{p}} = \text{FN} \\ \#\text{ue}_{\mathscr{p}} = \text{FP} & \#\text{mne}_{\mathscr{p}} = \text{TN} \end{bmatrix} && (45)\\
& && = \mathbb{T} \subset \mathbb{Z}_{+}^{2\times 2} && (46)
\end{aligned}$$

$$\begin{aligned}
\text{Case2:}\quad & \text{Positive outcome} && := \text{Non} - \text{equivalent reaction} && (47)\\
& && = \{\text{forward}, \text{reverse}\} && (47.1)\\
& \text{Confusion matrix} && \stackrel{\text{def}}{=} \begin{bmatrix} \#\text{mne}_{\mathscr{p}} = \text{TP} & \#\text{ue}_{\mathscr{p}} = \text{FN} \\ \#\text{une}_{\mathscr{p}} = \text{FP} & \#\text{me}_{\mathscr{p}} = \text{TN} \end{bmatrix} && (48)\\
& && = \widetilde{\mathbb{T}} \subset \mathbb{Z}_{+}^{2\times 2} && (49)
\end{aligned}$$

The indices of accuracy (Ac), precision (Pr), recall (Re), false positive rate (FPr) and specificity (SPy) will be used to compare a biochemical network before and after the introduction of one- or more-extraneous plausible reaction vectors. These are defined as:

$$\begin{aligned}
\text{Pr}(\mathscr{p}) &\stackrel{\text{def}}{=} \left(\frac{\text{TP}}{\text{TP}+\text{FP}}\right).100 && (50)\\
\text{Re}(\mathscr{p}) &\stackrel{\text{def}}{=} \left(\frac{\text{TP}}{\text{TP}+\text{FN}}\right).100 && (51)\\
\text{Ac}(\mathscr{p}) &\stackrel{\text{def}}{=} \left(\frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}\right).100 && (52)\\
\text{FPr}(\mathscr{p}) &\stackrel{\text{def}}{=} \left(\frac{\text{FP}}{\text{FP}+\text{TN}}\right).100 && (53)\\
\text{SPy}(\mathscr{p}) &\stackrel{\text{def}}{=} \left(1-\text{FPr}(\mathscr{p})\right).100 && (54)\\
&= \left(1-\left(\frac{\text{FP}}{\text{FP}+\text{TN}}\right)\right).100 && (54.1)
\end{aligned}$$

3.3 Biochemical networks studied to complement theoretical assertions of the algebraic model

Biochemical networks for the hexose monophosphate shunt (HMP), urea cycle and folate metabolism in humans will be constructed and analysed with "ReDirection". These networks are well characterized,

conserved across taxa and make significant contributions to biochemical and physiological function of the cell and the individual.

The HMP-shunt or the pentose phosphate pathway (PPP) is needed for the *de novo* generation of reduced Nicotinamide adenine dinucleotide phosphate $(NADPH + H^{+})$, phosphoribosyl pyrophosphate $(PRPP)$ and several other potentially anaplerotic intermediates for Kreb's cycle (**Figure 3)** [39]. The continual presence of $NADPH$ maintains the reducing environment and a compromised, i.e., oxidative intracellular environment will lead to plasma membrane damage and an increased fragility of cells especially under physiological duress (splenic passage of the RBC). The $PRPP$ that is generated participates in de novo synthesis of purine and pyrimidine rings and thence the nucleic acids (DNA/RNA). Here, we will examine the direct conversion of β-D fructose 6-phosphate to D-ribulose 6-phosphate (r_{19}). This reaction routinely occurs in bacteria and archaea as part of the ribulose-monophosphate cycle to fix formaldehyde [39].

The human urea cycle is an important pathway in the utilization and removal of non-protein nitrogen (urea, uric acid, ammonia) and occurs, partly, in the mitochondria and the cytosol of the liver and kidneys. Although *Homo sapiens* are ureotelic, small amounts of ammonia and uric acid are routinely excreted. The inability to excrete urea over a prolonged period of time predisposes an individual to varying grades of uremia and encephalopathy. The urea cycle is characterized by several shunt pathways which includes contributions to Aspartate (Kreb's cycle, r_6) and Carbamoyl phosphate via Glutamate (Kreb's cycle, r_7; D-amino acids, r_5; Arginine-Proline, r_8). The synthesis of Citrulline (r_3) via Nitric oxide synthase (NOS, $EC1.14.13.39$) is not only a pathway that utilizes *NADPH* but also generates NO, which is an important member of reactive oxygen (RO)- and reactive nitrogen (RN)-species (ROS, RNS). This pathway has also been implicated in the observed polarization of macrophages into the M1 (pro-inflammatory)- and M2 (anti-inflammatory)-phenotypes [40]. In this study a direct route from Arginine to Carbamoyl phosphate is examined (r_{12}) (**Figure 4)**.

The biochemical network to metabolize folate connects disparate molecular species (macro, micro) such as proteins, carbohydrates, lipids, DNA/RNA and vitamins. The "Folate trap" is a clinicopathologic condition that is characterized by relative folate deficiency in the presence of vitamin B12 deficiency [41, 42]. Since folate is an important contributor to 1C-metabolism and thence DNA synthesis, its deficiency results in macrocytic anemia (large and immature RBCs). In its role as a cofactor, folate is

required by enzymes involved in the metabolism of several amino acids (Serine, Glycine, Histidine, Cysteine) [41, 42]. Here, we will introduce the direct conversion of dTMP to dUMP (r_{18}) and examine the effects of this perturbing reaction on the probable dissociation constants on a biochemical network of folate metabolism (**Figure 5)** [41, 42].

4 Results and Discussion

4.1 Populating the set of plausible reaction vectors for a class of biochemical networks

We will utilize **A1-A5** and discuss the process of populating the set of plausible reaction vectors from all possible reaction vectors (**S2; Figure 1**). Whilst the rationale to choose plausible reaction vectors from the subsets that are formed by using combinatorial indices $\mathrm{t} \geq 2$ follows directly from **A2**, the choice to include $\mathrm{t} = \mathrm{J}$ and exclude $\mathrm{t} = \mathrm{J} - 1$ will be described in greater detail.

Consider the subset of signed reaction vectors that needs to be excluded (Ω). Since this population is combinatorial we will partition this into $\Omega_{(+)}$ and $\Omega_{(-)}$ subsets,

$$\left\{\mathbf{r}_{\omega} \in \Omega_{(+)}, \tilde{\mathbf{r}}_{\tilde{\omega}} \in \Omega_{(-)}\right\} \subset \mathbf{R} \qquad \text{Def.}\,(28)$$

where,

$$\Omega = \Omega_{(+)} \cup \Omega_{(-)} \qquad (55)$$

$$\#\Omega = \#\Omega_{(+)} + \#\Omega_{(-)} \qquad (56)$$

$$\omega \neq \tilde{\omega} \qquad (57)$$

$$\omega, \tilde{\omega} = \left[1, \frac{\#\Omega}{2}\right] \qquad (58), (59)$$

Definition 28 (D28): The choice of $\mathrm{t} = \mathrm{J}$ will result in a single reaction vector with identically signed $(+)$ or $(-)$ terms,

$$\text{if}\left(\binom{J}{t = J_{(+)}} = \binom{J}{t = J_{(-)}}\right) \tag{60}$$

$$\text{then}\left(\mathbf{r}_{\omega} = \begin{bmatrix} (+)m_{1\omega} \\ (+)m_{2\omega} \\ (+)m_{3\omega} \\ \vdots \\ (+)m_{J\omega} \end{bmatrix} \neq \begin{bmatrix} (-)m_{1\tilde{\omega}} \\ (-)m_{2\tilde{\omega}} \\ (-)m_{3\tilde{\omega}} \\ \vdots \\ (-)m_{J\tilde{\omega}} \end{bmatrix} = \tilde{\mathbf{r}}_{\tilde{\omega}}\right) \tag{61}$$

where,
$j = 1,2 \ldots J$
$\omega \neq \tilde{\omega}$
$\omega, \tilde{\omega} \in \left[1, \frac{\#\Omega}{2}\right]$

$$\#\Omega_{(+)} = \binom{J}{J_{(+)}} \tag{62}$$

$$\#\Omega_{(-)} = \binom{J}{J_{(-)}} \tag{63}$$

We can immediately see from this definition that despite being numerically equal, the vectors $\mathbf{r}_{\omega}$ and $\tilde{\mathbf{r}}_{\tilde{\omega}}$ are unique and linear dependent,

$$\mathbf{r}_{\omega} \neq \tilde{\mathbf{r}}_{\tilde{\omega}} \tag{64}$$
$$\mathbf{r}_{\omega} + \tilde{\mathbf{r}}_{\tilde{\omega}} = \mathbf{0} \tag{65}$$

These vectors are therefore, in violation of **A2** and **A3** and will not represent a true biochemical reaction. If, however, we insist that there is at least one pair of differently signed reactants/products, then we will be able to assign a biochemical basis to this reaction vector and include these vectors into the set of plausible reaction vectors,

$$\text{if}\ (\mathbf{r}_{\omega} + \tilde{\mathbf{r}}_{\tilde{\omega}} \neq \mathbf{0}) \tag{66}$$
$$\text{then}\ (\mathbf{r}_{\omega}, \tilde{\mathbf{r}}_{\tilde{\omega}} \in \mathbf{R})$$
$$\text{else}\ (\mathbf{r}_{\omega}, \tilde{\mathbf{r}}_{\tilde{\omega}} \notin \mathbf{R}) \quad \text{Def. (29)}$$

Definition 30 (**D30**): The condition of $t = J - 1$ will result in reaction vectors whose terms are non-identically signed, linearly dependent and will therefore, form a trivial vector when summed,

$$\text{if}\left(\begin{pmatrix} \mathrm{J} \\ \mathrm{t} = (\mathrm{J}-1)_{(+)} \end{pmatrix} = \begin{pmatrix} \mathrm{J} \\ \mathrm{t} = (\mathrm{J}-1)_{(-)} \end{pmatrix}\right) \quad (67)$$

$$\text{then } (\exists \mathbf{r}_{\omega} | \tilde{\mathbf{r}}_{\tilde{\omega}} + \mathbf{r}_{\omega} = \mathbf{0}) \quad (68)$$

where,

$\mathrm{j} = 1,2 \ldots \mathrm{J}$

$\omega \neq \tilde{\omega}$

$\omega, \tilde{\omega} \in \left[1, \frac{\#\Omega}{2}\right]$

$$\#\Omega_{(+)} = \begin{pmatrix} \mathrm{J} \\ (\mathrm{J}-1)_{(+)} \end{pmatrix} \quad (69)$$

$$\#\Omega_{(-)} = \begin{pmatrix} \mathrm{J} \\ (\mathrm{J}-1)_{(-)} \end{pmatrix} \quad (70)$$

We will demonstrate this with a worked example,

$$\begin{pmatrix} \mathrm{J} = 3 \\ \mathrm{t} = 2(-) \end{pmatrix} = \left(\frac{3!}{(2-)!\,(1+)!}\right) \quad (71)$$

$$= 3 \quad (71.1)$$

$$\stackrel{\text{def}}{=} \#\Omega_{(-)} \quad (72)$$

$$\tilde{\mathbf{r}}_{\tilde{\omega}} | \tilde{\omega} = 1,2,3 \quad (73)$$

$$\tilde{\mathbf{r}}_1 = \begin{bmatrix} (+)\mathrm{m}_{11} \\ (-)\mathrm{m}_{21} \\ (-)\mathrm{m}_{31} \end{bmatrix}; \tilde{\mathbf{r}}_2 = \begin{bmatrix} (-)\mathrm{m}_{12} \\ (+)\mathrm{m}_{22} \\ (-)\mathrm{m}_{32} \end{bmatrix}; \tilde{\mathbf{r}}_3 = \begin{bmatrix} (-)\mathrm{m}_{13} \\ (-)\mathrm{m}_{23} \\ (+)\mathrm{m}_{33} \end{bmatrix} \quad (74-76)$$

$$\begin{pmatrix} \mathrm{J} = 3 \\ \mathrm{t} = 2(+) \end{pmatrix} = \left(\frac{3!}{(2+)!\,(1-)!}\right) \quad (77)$$

$$= 3 \quad (77.1)$$

$$\stackrel{\text{def}}{=} \#\Omega_{(+)} \quad (78)$$

$$\mathbf{r}_{\omega} | \omega = 1,2,3 \quad (79)$$

$$\mathbf{r}_1 = \begin{bmatrix} (+)\mathrm{m}_{11} \\ (+)\mathrm{m}_{21} \\ (-)\mathrm{m}_{31} \end{bmatrix}; \mathbf{r}_2 = \begin{bmatrix} (+)\mathrm{m}_{12} \\ (-)\mathrm{m}_{22} \\ (+)\mathrm{m}_{32} \end{bmatrix}; \mathbf{r}_3 = \begin{bmatrix} (-)\mathrm{m}_{13} \\ (+)\mathrm{m}_{23} \\ (+)\mathrm{m}_{33} \end{bmatrix} \quad (80-82)$$

$$\tilde{\mathbf{r}}_3 = (-1).\mathbf{r}_1 \quad (83)$$

$$\tilde{\mathbf{r}}_2 = (-1).\mathbf{r}_2 \quad (84)$$

$$\tilde{\mathbf{r}}_1 = (-1).\mathbf{r}_3 \quad (85)$$

For each $\tilde{\omega} = 1,2,3$

$$\exists \tilde{\mathbf{r}}_{\tilde{\omega}} \text{ s.t. } \tilde{\mathbf{r}}_{\tilde{\omega}} = (-1).\mathbf{r}_{\omega} \quad (86)$$

$$\Rightarrow \sum_{\omega=1}^{\omega=3} \mathbf{r}_{\omega} + \sum_{\tilde{\omega}=1}^{\tilde{\omega}=3} \tilde{\mathbf{r}}_{\tilde{\omega}} \quad (87)$$

$$= \sum_{\omega=1}^{\omega=3} (\mathbf{r}_{\omega} + (-1).\mathbf{r}_{\omega}) \quad (87.1)$$

$$= \mathbf{0} \quad (87.2)$$

We can summarize the results that will be generated by t-combinations of J-molecular species (reactants, products) where $t = 2, 3 \ldots J-2, J$. The resulting superset (**R**) of plausible reaction vectors has the cardinality $(\bar{\bar{I}})$,

$$\begin{aligned} \bar{\bar{I}} &= \#\mathbf{R} && (88) \\ &= \binom{J}{t=J} + \sum_{t=2}^{t=J-2} \binom{J}{t} && (88.1) \\ &= 1 + \sum_{t=2}^{t=J-2} \binom{J}{t} && (88.2) \end{aligned}$$

The plausible reaction vectors included are unique, non-trivial on summation and will represent one half (forward or reverse) of a biochemical reaction. Since each plausible reaction vector can be forward or reverse, the maximum number of all plausible reaction vectors is,

$$\begin{aligned} \max(\bar{\bar{I}}) &= 2.(\#\mathbf{R}) && (89) \\ &= 2.(\bar{\bar{I}}) && (89.1) \end{aligned}$$

We can revise the stoichiometry number matrix model for a biochemical network (**S3**, **Figure 1**),

$$\mathcal{p} \text{ or } \mathcal{p}_x \text{ or } \mathcal{S}(\mathcal{p}_x) \in \mathbb{Z}^{J \times \bar{\bar{I}}} \quad (90)$$

where,
$x = 1, \ldots, X$
$j = 1, \ldots, J$
$i = 1, \ldots, \bar{\bar{I}}$

4.2 Definition and description of a class of biochemical networks

It is evident from these results that to remain biochemically viable**,** a stoichiometry number matrix model for a biochemical network must possess lower bounds for the molecular species (J) and plausible reaction vectors $(\bar{\bar{I}})$ (**A9**). This will also imply the existence of a lower bound for the number of all such stoichiometry number matrices (X),

$$\begin{aligned} \min(J) \rightarrow \min(\bar{\bar{I}}) &\Rightarrow \min(X) && (91) \\ &= \min(\#\mathcal{P}) && (91.1) \end{aligned}$$

We will use these results to define and compute constraints which will be used to populate a class of biochemical networks $(\mathcal{P})$ (**S4**, **S5**, **Figure 1**). The constraints can be classified as a) architecture-preserving (minimum number of molecular species, minimum number of plausible reaction vectors) and

b) stoichiometry number equivalence for each molecular species in accordance with **A9**. The definitions for these bounds are formally presented,

$$\max\left(\bar{\bar{\mathrm{I}}}_{\mathcal{S}(\mathcal{P})}\right) \stackrel{\text{def}}{\equiv} \text{Maximum number of plausible reaction vectors for } \mathcal{S}(\mathcal{P}) \quad (92)$$
$$\equiv \text{Maximum number of plausible reaction vectors for } \mathcal{S}(\wp_{\mathrm{x}}) \quad (92.1)$$
$$= \max\left(\bar{\bar{\mathrm{I}}}_{\mathcal{S}(\wp_{\mathrm{x}})}\right) \quad (93)$$
$$= \max\left(\bar{\bar{\mathrm{I}}}_{\mathrm{x}}\right) \quad (94)$$
$$\min\left(\bar{\bar{\mathrm{I}}}_{\mathcal{S}(\mathcal{P})}\right) \stackrel{\text{def}}{\equiv} \text{Minimum number of plausible reaction vectors for } \mathcal{S}(\mathcal{P}) \quad (95)$$
$$\equiv \text{Minimum number of plausible reaction vectors for } \mathcal{S}(\wp_{\mathrm{x}}) \quad (95.1)$$
$$= \min\left(\bar{\bar{\mathrm{I}}}_{\mathcal{S}(\wp_{\mathrm{x}})}\right) \quad (96)$$
$$= \min\left(\bar{\bar{\mathrm{I}}}_{\mathrm{x}}\right) \quad (97)$$
$$\max\left(\mathrm{J}_{\mathcal{S}(\mathcal{P})}\right) \stackrel{\text{def}}{\equiv} \text{Maximum number of molculecular species for } \mathcal{S}(\mathcal{P}) \quad (98)$$
$$\equiv \text{Maximum number of molecular species for } \mathcal{S}(\wp_{\mathrm{x}}) \quad (98.1)$$
$$= \max\left(\mathrm{J}_{\mathcal{S}(\wp_{\mathrm{x}})}\right) \quad (99)$$
$$= \max(\mathrm{J}_{\mathrm{x}}) \quad (100)$$
$$\min\left(\mathrm{J}_{\mathcal{S}(\mathcal{P})}\right) \stackrel{\text{def}}{\equiv} \text{Minimum number of molculecular species for } \mathcal{S}(\mathcal{P}) \quad (101)$$
$$\equiv \text{Minimum number of molecular species for } \mathcal{S}(\wp_{\mathrm{x}}) \quad (101.1)$$
$$= \min\left(\mathrm{J}_{\mathcal{S}(\wp_{\mathrm{x}})}\right) \quad (102)$$
$$= \min(\mathrm{J}_{\mathrm{x}}) \quad (103)$$

Proposition 1 (**P1**): For a stoichiometry number matrix with J-molecular species and $\bar{\bar{\mathrm{I}}}$-plausible reaction vectors there exists a class of biochemical networks,

$$\text{if}\left(\mathcal{P} = (\wp_1, \wp_2 \ldots \wp_{\mathrm{X}})\right)$$
$$\text{then}\left(\begin{array}{c} \exists \mathcal{S}(\mathcal{P}) = \left(\mathcal{S}(\wp_1), \mathcal{S}(\wp_2) \ldots \mathcal{S}(\wp_{\mathrm{X}})\right) \\ \text{s.t.} \begin{cases} \min(\mathrm{J}_{\mathrm{x}}) \geq \mathrm{O} \\ \min\left(\bar{\bar{\mathrm{I}}}_{\mathrm{x}}\right) \geq \mathrm{Q} \\ \mathrm{m}_{\mathrm{ji}} \begin{cases} = 0 \\ \neq 0 \end{cases} \forall \mathrm{x} \end{cases} \end{array}\right) \quad (104,105)$$

Def. (30)

where,
$\mathrm{j} = 1, \ldots, \mathrm{J}$
$\mathrm{i} = 1, \ldots, \bar{\bar{\mathrm{I}}}$
$\mathrm{x} = 1, \ldots, \mathrm{X}$
$\mathrm{O}, \mathrm{Q} \in \mathbb{N}$ (106), (107)

We will now compute values for O, Q and X in accordance with (**Eq. 91**) and from the solved examples with t-combinations of J-reactants/products for each $\wp_{\mathrm{x}} \in \mathcal{P}$ where,

$$\mathrm{t} = (2,3 \ldots \mathrm{J} - 2, \mathrm{J}) \quad (108)$$
$$\mathrm{J} = (2,3,4,5) \quad (109)$$

$$\#\mathbf{R} = 1 + \sum_{t=2}^{t=0} \binom{J=2}{t=2} \quad (110)$$
$$= 1 \quad (110.1)$$
$$\Rightarrow \left(\bar{\bar{I}}_x, \max\left(\bar{\bar{I}}_x\right)\right) = (1,2) \quad (112)$$

$$\#\mathbf{R} = 1 + \sum_{t=2}^{t=1} \binom{J=3}{t=2} \quad (113)$$
$$= 1 \quad (113.1)$$
$$\Rightarrow \left(\bar{\bar{I}}_x, \max\left(\bar{\bar{I}}_x\right)\right) = (1,2) \quad (114)$$

$$\#\mathbf{R} = 1 + \sum_{t=2}^{t=2} \binom{J=4}{t=2} \quad (115)$$
$$= 1 + \binom{4}{2} \quad (115.1)$$
$$= 7 \quad (115.2)$$
$$\Rightarrow \left(\bar{\bar{I}}_x, \max\left(\bar{\bar{I}}_x\right)\right) = (7,14) \quad (116)$$

$$\#\mathbf{R} = 1 + \sum_{t=2}^{t=3} \binom{J=5}{t=2} \quad (117)$$
$$= 1 + \binom{5}{2} + \binom{5}{3} \quad (117.1)$$
$$= 21 \quad (117.2)$$
$$\Rightarrow \left(\bar{\bar{I}}_x, \max\left(\bar{\bar{I}}_x\right)\right) = (21,42) \quad (118)$$

We will subsume, for these calculations, the improbable occurrence that a single master reaction will predominate and constitute a functioning biochemical network and exclude the same,

$$\min(J_x) \geq 0 \quad [\text{From } 110 - 116] \quad (119)$$
$$\geq 4 \quad (120)$$

$$\min\left(\bar{\bar{I}}_x\right) \geq Q$$
$$\geq \bar{\bar{I}}_x - 1 \quad [\text{From } 115 - 118] \quad (121)$$
$$\geq 6 \quad (122)$$

Our final descriptor for $\mathcal{P}$ is a lower bound for the cardinality or number of constituent biochemical networks

$$X \geq \#\mathcal{P} \quad (123)$$

Since,

$$\#\mathcal{P} \stackrel{\text{def}}{=} 2^{\bar{\bar{I}}} \quad [\text{From (A2)}] \quad (124)$$

The minimum number of biochemical networks in a given class, i.e., for J-molecular species and interacting via $\min(\bar{\bar{I}}_x)$-plausible reaction vectors is,

$$\begin{aligned} X &\geq 2^{\min(\bar{\bar{I}}_x)} && [\text{From } (122), (123)] \quad (125) \\ &\geq 64 && [\text{From } (121)] \quad (126) \end{aligned}$$

From these results we formally define the set of stoichiometry number matrices for a class of biochemical networks $\mathcal{S}(\mathcal{P})$ where for each $\mathcal{S}(\mathcal{p}_x)$,

$$\begin{cases} J_x \in [4, M] \\ \bar{\bar{I}}_x \in [6, N] \\ \bar{\bar{I}}_x \geq J_x + 2 \\ X \geq 64 \end{cases} \qquad \text{Def. } (31)$$

where,
$\bar{\bar{I}}_x$:= Number of plausible reaction vectors
J_x := Number of molecular species that contribute to a plausible reaction vector
X := Lower bound for the cardinality of $\mathcal{P}$
$M, N \stackrel{\text{def}}{=}$ Finite upper bounds for molecular species and plausible reaction vectors (127), (128)

Although there is no restriction on the upper bound for the plausible reaction vectors, an analysis of the run time needed by "ReDirection" to completely and unambiguously annotate and assign an outcome to every reaction of a biochemical network has suggested that [7],

$$N \in [J_x + 2, J_x + 4] \quad (129)$$
where,
$J_x \geq 4$

4.4 The set of CReaSED-matrices is specific for a class of biochemical networks

We now formally define the set of CReaSED-matrices as a finite set of stoichiometry number matrices which are architecturally constrained and exhibit a zero-preserving stoichiometry for all the molecular species (**S5, Figure 1**),

$$\begin{aligned}
\mathcal{S}(\mathcal{P}) &\stackrel{\text{def}}{=} \mathfrak{S}(\mathcal{P}) && (130)\\
&= \mathfrak{S}\left(\mathcal{p}, \mathrm{X}, \mathrm{J}, \bar{\bar{\mathrm{I}}}\right) && (131)\\
&= \left\{\mathrm{S}_{\mathcal{p}_1}, \mathrm{S}_{\mathcal{p}_2} \ldots . \mathrm{S}_{\mathcal{p}_\mathrm{X}}\right\} && (132)\\
&\subset \mathbb{Z}^{\mathrm{J}\times\bar{\bar{\mathrm{I}}}} && \text{Def. } (32)\\
&\text{where,}\\
&\mathrm{X} \geq 2^{\min\left(\bar{\bar{\mathrm{I}}}_\mathrm{x}\right)}\\
&\bar{\bar{\mathrm{I}}}_\mathrm{x} \in [6, \mathrm{N}]\\
&\mathrm{J}_\mathrm{x} \in [4, \mathrm{M}]\\
&\bar{\bar{\mathrm{I}}}_\mathrm{x} \geq \mathrm{J} + 2\\
&\mathrm{x} = 1, \ldots, \mathrm{X}\\
&\mathrm{M}, \mathrm{N} \in \mathbb{N}
\end{aligned}$$

We will refine this definition of the CReaSED-matrices by investigating the mathematical properties and include the results, thereof.

Since the numerical contribution by each molecular species to each x-indexed stoichiometry number matrix of an X-membered class of biochemical networks will be unique in accordance with **A9**,

Theorem 1 (**T1**; **without proof**): The set of CReaSED-matrices for a class of biochemical networks are not equal,

$$\begin{aligned}
&\mathrm{S}_{\mathcal{p}_1} \neq \mathrm{S}_{\mathcal{p}_2} \ldots . \neq \mathrm{S}_{\mathcal{p}_\mathrm{X}} \quad [\text{From (A9)}] \quad (133), \text{Def. } (33)\\
&\text{where,}\\
&\mathcal{S}_{\mathcal{p}_\mathrm{x}} \in \mathfrak{S}(\mathcal{P})
\end{aligned}$$

On the basis of **T1** we can define a non-zero real valued multiplier for any $\mathcal{S}_{\mathcal{p}_\mathrm{x}}$ drawn from a set of CReaSED-matrices,

$$\begin{aligned}
&\mathcal{S}_{\mathcal{p}_\mathrm{x}} = \phi_{\mathcal{p}_\mathrm{x}}.\mathcal{S}_{\mathcal{p}_\mathrm{x}} \quad \text{Def. } (34)\\
&\text{where,}\\
&\phi_{\mathcal{p}_\mathrm{x}} \in \mathbb{R} - \{0\}\\
&\mathcal{S}_{\mathcal{p}_\mathrm{x}} \in \mathfrak{S}(\mathcal{P})
\end{aligned}$$

We will simplify the notation hereafter for a non-zero real valued multiplier for any $\mathcal{S}_{\mathcal{p}_\mathrm{x}}$ of the CReaSED-matrices,

$$\phi_{\mathcal{p}_\mathrm{x}} \stackrel{\text{def}}{=} \phi \quad (134)$$

Corollary 1 (**C1**): For any pair $\mathcal{S}_{\mathcal{p}_\mathrm{x}}, \mathcal{S}_{\mathcal{p}_\mathrm{y}}$ of distinct stoichiometry number matrices that belong to the set of CReaSED-matrices, we can define a unique non-zero real valued multiplier,

$$\mathcal{S}_{p_x} = \phi_{xy}.\mathcal{S}_{p_y} \qquad (135)$$
where,
$\phi_{xy} \in \mathbb{R} - \{0\}$
$\mathcal{S}_{p_x}, \mathcal{S}_{p_y} \in \mathfrak{S}(\mathcal{P})$
$x \neq y$

Corollary 2 (**C2**)**:** The number of real-valued multipliers for a pair of identical stoichiometry number matrices drawn from a set of CReaSED-matrices is $\bar{\bar{X}}_1$,

$$\phi_{xx} \in \Phi^T \subset (\mathbb{R} - \{0\})^{1\times\bar{\bar{X}}_1} \qquad (136)$$
where,
$x = 1, \ldots, X$
$\bar{\bar{X}}_1 = X$
$S_{p_1} \neq S_{p_2} \ldots. \neq S_{p_X}$
$S_{p_x}, S_{p_x} \in \mathfrak{C}(\mathcal{P})$

Corollary 3 (**C3**)**:** The number of real valued multipliers for a pair of distinct stoichiometry number matrices drawn from a set of CReaSED-matrices is $\bar{\bar{X}}_2$,

$$\phi_{xy} \in \Phi^T \subset (\mathbb{Q} - \{0\})^{1\times\bar{\bar{X}}_2} \qquad (137)$$
where,
$x \neq y$
$x, y = 1, \ldots, X$
$X \geq 64$
$\bar{\bar{X}}_2 = X.(X - 1)$
$S_{p_1} \neq S_{p_2} \ldots. \neq S_{p_X}$
$S_{p_x}, S_{p_y} \in \mathfrak{C}(\mathcal{P})$

We define a square matrix of all non-zero real valued multipliers with dimensions $X \times X$,

$$\mathbf{\Phi} = (\phi_{xy}) \subset (\mathbb{R} - \{0\})^{X\times X} \qquad \text{Def. (35)}$$
where,
$$\#\mathrm{Vec}(\mathbf{\Phi}) = X^2 \qquad (138)$$
$x, y = 1, \ldots, X$
$X \geq 64$

Theorem 2 (**T2**): The set of pairwise real valued multipliers for a set of CReaSED-matrices is a square matrix with $X \times X$-dimensions,

$$\phi_{xy} \in \mathbf{\Phi} \subset (\mathbb{R} - \{0\})^{X \times X} \quad (139)$$
where
$\mathbf{S}_{\mathcal{p}_1} \neq \mathbf{S}_{\mathcal{p}_2} \ldots . \neq \mathbf{S}_{\mathcal{p}_X}$
$\mathbf{S}_{\mathcal{p}_x} \in \mathfrak{S}(\mathcal{P})$
$x \neq y$
$x, y = 1, \ldots, X$
$X \geq 64$

Corollary 4 (**C4**; **without proof**): Every diagonal element of $\mathbf{\Phi}$ for a set of CReaSED-matrices corresponds to an interaction between a pair of identical stoichiometry number matrices (self-pair),

$$\phi_{xx} \in \mathbf{\Phi} \subset (\mathbb{R} - \{0\})^{X \times X} \twoheadrightarrow A = \left\{\mathbf{S}_{\mathcal{p}_x}, \mathbf{S}_{\mathcal{p}_x}\right\} \quad (140)$$
where,
$\mathbf{S}_{\mathcal{p}_x} \in \mathfrak{S}(\mathcal{P})$
$x = 1, \ldots, X$
$X \geq 64$

Corollary 5 (**C5**; **without proof**): A transposable pair of elements of $\mathbf{\Phi}$ for a set of CReaSED-matrices can be mapped to a distinct pair of stoichiometry number matrices (non-self pair),

$$\phi_{xy}, \phi_{yx} \in \mathbf{\Phi} \subset (\mathbb{R} - \{0\})^{X \times X} \twoheadrightarrow A = \left\{\mathbf{S}_{\mathcal{p}_x}, \mathbf{S}_{\mathcal{p}_y}\right\} \quad (141)$$
where,
$\mathbf{S}_{\mathcal{p}_x}, \mathbf{S}_{\mathcal{p}_y} \in \mathfrak{S}(\mathcal{P})$
$x \neq y$
$x, y = 1, \ldots, X$
$X \geq 64$

Theorem 3 (**T3**): The scalar product of all the elements of $\mathbf{\Phi}$ for a set of CReaSED-matrices is numerically equal to unity,

$$\theta = 1 \quad (142)$$
where,
$$\theta = \prod_{x=1}^{x=X} \prod_{y=1}^{y=X} \phi_{xy} \quad (143)$$
$\phi_{xy} \in \mathbf{\Phi} \subset (\mathbb{R} - \{0\})^{X \times X}$
$x, y = 1, \ldots, X$
$X \geq 64$

Corollary 6 (**C6**;): The determinant for a given $\mathbf{\Phi}$ for a set of CReaSED-matrices is identical iff for every transposable pair the elements are transposed simultaneously,

$$\text{If} \begin{pmatrix} \phi_{xy}, \phi_{yx} \in \mathbf{\Phi}_1 \\ \phi_{yx}, \phi_{xy} \in \mathbf{\Phi}_2 \end{pmatrix},$$
$$\text{then } \det(\mathbf{\Phi}_1) = \det(\mathbf{\Phi}_2) \quad (144)$$
$$\text{iff each } x \to y \text{ and each } y \to x \quad \text{Def. } (36)$$

Theorem 4 (**T4**): The stoichiometry number matrices that comprise a set of CReaSED-matrices for a class of biochemical networks are conditionally equal,

$$\mathbf{S}_{p_1} \approx \mathbf{S}_{p_2} \ldots \approx \mathbf{S}_{p_X} \quad (145)$$
where,
$\mathbf{S}_{p_x} \in \mathfrak{S}(\mathcal{P})$
$x = 1, \ldots, X$
$X \geq 64$

4.5 Elucidating the properties of the CReaSED-matrices from the null space

The null space is a numerical construct which is used to model a system of biochemical reactions at equilibrium. The null space-generated subspace for a stoichiometry number matrix has been shown to be important in the direct computation of the probable dissociation constant for every reaction of a user-defined biochemical network [7, 8]. The following results are based on elucidating the properties of a set of CReaSED-matrices from the computed null space,

Theorem 5 (**T5**): The null spaces generated by a set of CReaSED-matrices for a class of biochemical networks are equal,

$$\mathrm{N}(\mathbf{S}_{p_1}) = \mathrm{N}(\mathbf{S}_{p_2}) \ldots = \mathrm{N}(\mathbf{S}_{p_X}) \quad (146)$$

where,
$\mathbf{S}_{p_1} \neq \mathbf{S}_{p_2} \ldots \neq \mathbf{S}_{p_X}$
$\mathbf{S}_{p_x} \in \mathfrak{S}(\mathcal{P})$
$x = 1, \ldots, X$
$X \geq 64$

We will extend the definition of a set of CReaSED-matrices for a class of biochemical networks,

$$\left(\begin{array}{ll} \mathfrak{S}(p, X, J, \bar{\bar{I}}) = \{\mathbf{S}_{p_1}, \mathbf{S}_{p_2} \ldots . \mathbf{S}_{p_X}\} \subset \mathbb{Z}^{J \times \bar{\bar{I}}} & \\ \text{where,} & \text{Def. (37)} \\ \mathbf{S}_{p_1} \neq \mathbf{S}_{p_2} \ldots . \neq \mathbf{S}_{p_X} & \\ \mathrm{N}(\mathbf{S}_{p_1}) = \mathrm{N}(\mathbf{S}_{p_2}) \ldots = \mathrm{N}(\mathbf{S}_{p_X}) & \end{array}\right)$$

Theorem 6 (**T6**): The map between the set of CReaSED-matrices and the corresponding null space is many-to-one,

$$f: \mathfrak{S}(p, X, J, \bar{\bar{I}}) \twoheadrightarrow \mathrm{N}\left(\mathfrak{S}(p, X, J, \bar{\bar{I}})\right) \quad (147)$$

4.6 Algebraic properties of a set of CReaSED-matrices for a class of biochemical networks

The set of CReaSED-matrices are bonafide algebraic objects and likely to possess some interesting mathematical properties which we will investigate. These properties may be used to derive additional metrics by which biochemical networks can be characterized and compared. Unlike data-driven studies wherein biochemical fluctuations are modelled as square matrices (covariance, correlation), the CReaSED-matrices are rectangular and likely to differ algebraically.

Lemma 1 (**L1**): The set of CReaSED-matrices for a class of biochemical networks is commutative or Abelian with regards to addition,

$$\mathbf{S}_{p_{\mathrm{x}}} + \mathbf{S}_{p_{\mathrm{y}}} = \mathbf{S}_{p_{\mathrm{y}}} + \mathbf{S}_{p_{\mathrm{x}}} \quad (148)$$

where,

$\mathbf{S}_{p_{\mathrm{x}}}, \mathbf{S}_{p_{\mathrm{y}}} \in \mathfrak{S}(\mathcal{P})$

$\mathrm{x} \neq \mathrm{y}$

$\mathrm{x}, \mathrm{y} = 1, \ldots, \mathrm{X}$

$\mathrm{X} \geq 64$

Theorem 7 (**T7**): The set of CReaSED-matrices for a class of biochemical networks is a semigroup with respect to addition and is conditionally Abelian,

$$\mathrm{g}: (\mathfrak{C}(\mathcal{P}), +) \rightarrow \mathfrak{C}(\mathcal{P}) \quad (149)$$

where,

$\mathbf{S}_{p_{\mathrm{x}}} + \mathbf{S}_{p_{\mathrm{y}}} \rightarrow \mathbf{S}_{p_{\mathrm{z}}}$

$\mathbf{S}_{p_{\mathrm{x}}}, \mathbf{S}_{p_{\mathrm{y}}}, \mathbf{S}_{p_{\mathrm{z}}} \in \mathfrak{S}(\mathcal{P})$

$\mathrm{x} \neq \mathrm{y} \neq \mathrm{z}$

$\mathrm{x}, \mathrm{y}, \mathrm{z} = 1, \ldots, \mathrm{X}$

$\mathrm{X} \geq 64$

Lemma 2 (**L2**): The finite intersection of the stoichiometry number matrices that comprise a set of CReaSED-matrices for a class of biochemical networks is an empty set,

$$\bigcap_{\mathrm{x}=1}^{\mathrm{x}=\mathrm{X}} \mathbf{S}_{p_{\mathrm{x}}} = \emptyset \quad (150)$$

where,

$\mathbf{S}_{p_{\mathrm{x}}} \in \mathfrak{C}(\mathcal{P})$

$\mathrm{x} = 1, \ldots, \mathrm{X}$

$\mathrm{X} \geq 64$

Lemma 3 (**L3**): The set of CReaSED-matrices for a class of biochemical networks is the finite union of its component matrices,

$$\mathfrak{S}(\mathcal{P}) = \bigcup_{x=1}^{x=X} \boldsymbol{S}_{p_x} \quad (151)$$

where,
$\boldsymbol{S}_{p_x} \in \mathfrak{S}(\mathcal{P})$
$x = 1, \ldots, X$
$X \geq 64$

On the basis of these theoretical assertions we can refine our definition for a set of CReaSED-matrices for a class of biochemical networks (**Figure 1**) [7, 8],

$$\mathfrak{S}(\mathcal{P}) \sim \mathfrak{S}(p, X, J, \bar{\bar{I}}) = \{\boldsymbol{S}_{p_1}, \boldsymbol{S}_{p_2} \ldots . \boldsymbol{S}_{p_X}\} \subset \mathbb{Z}^{J \times \bar{\bar{I}}} \quad \text{Def. (38)}$$

where,
$\boldsymbol{S}_{p_1} \neq \boldsymbol{S}_{p_2} \ldots . \neq \boldsymbol{S}_{p_X}$
$N(\boldsymbol{S}_{p_1}) = N(\boldsymbol{S}_{p_2}) \ldots = N(\boldsymbol{S}_{p_X})$
$\boldsymbol{S}_{p_x} = \phi_{xy} . \boldsymbol{S}_{p_y}$
$\{\phi_{xy}, \phi_{yx}\} \in \mathbb{R} - \{0\}$
$x \neq y$
$x, y = 1, \ldots, X$
$X \geq 64$
$\bar{\bar{I}}_x \in [6, N]$
$J_x \in [4, M]$
$\bar{\bar{I}}_x \geq J_x + 2$
$M, N \in \mathbb{N}$

4.7 The algebraic model for a class of biochemical networks is a suitable representation of complex biochemical function

Here, we examine the effect of modelling, perturbing and analysing well-characterized biochemical networks of the HMP-shunt pathway, urea cycle and folate metabolism in humans. The CReaSED-matrices for each are constructed in accordance with the framework outlined and perturbations are introduced as single extraneous reactions. The probable dissociation constants for each pair of reactions before and after the introduction of the extraneous reaction are computed and studied (**Figures 3-5**; **Tables 2 and 3**) [7, 8]. Although a reaction with an outcome of equivalence subsumes ambiguous data, it is also probable that these reactions may act as regulatory nodes for a biochemical network. Conversely, this implies that reactions with non-equivalent (forward, reverse) outcomes will depend on the intracellular environment and thence vary. This suggests that whilst an equivalent reaction outcome may be better suited as a global parameter, non-equivalent reaction outcomes might have greater relevance as a local parameter. These hypotheses were examined, *post hoc* by the comparison of the numerical data that is generated from the confusion matrices $\mathbb{T}$ (equivalent outcome as TP) and $\tilde{\mathbb{T}}$ (non-equivalent outcome as TP) for the canonical form of a biochemical networks (**Tables 2, 3**).

The recall of a metric is a measure of how sensitive a network is to perturbation. Since these are almost identical $\left(\overline{Re}_{(HMP,Urea,Folate)} \approx 50\%\right)$ and of moderately low magnitude, the utilization of the equivalence outcome as a global $(\mathfrak{S}(\mathcal{P}))$ parameter can be easily justified (**Figures 3-5**; **Table 3**). An analysis of the average-specificity $\left(\overline{SPy}_{(HMP,Urea.Folate)} \approx 83\%\right)$ and -accuracy $\left(\overline{Ac}_{(HMP,Urea.Folate)} \approx 65\%\right)$ for the data suggests that this metric can also be utilized as an index of comparison between the baseline- and perturbed-forms of the canonical forms of a biochemical network $\left(\mathfrak{S}(\mathcal{P}), \widehat{\mathfrak{S}}(\mathcal{P})\right)$ (**Figures 3-5**; **Table 3**). The precision or positive predictive value of a metric is an indicator of robustness upon change. This, too, for the equivalence outcome is a useful parameter to comprehend and compare a biochemical network pre- and post-perturbation $\left(\overline{Pr}_{(HMP,Urea.Folate)} \approx 87\%\right)$ (**Figures 2-5**; **Table 3**). Interestingly, if the original confusion matrix was redefined with a non-equivalent reaction outcome as TP and an equivalent reaction outcome as TN $(\widetilde{\mathbb{T}})$ we get different results (**Figures 2-5**; **Table 3**). Since reactions with defined and non-equivalent outcomes are likely to be influenced by the local intracellular milieu this measure will therefore, be a poor choice as an index to compare biochemical networks $\left(\overline{Re}_{(HMP,Urea,Folate)} \approx 83\%\right)$ (**Figures 2-5**; **Table 3**). However, since forward- and reverse-reactions regulate the flux of the reactants/products through a biochemical network, an outcome of non-equivalence may be a better indicator of network directionality or pathway formation $\left(\overline{SPy}_{(HMP,Urea.Folate)} \approx 50\%\right)$ for a specific instance of the CReaSED-matrices $\left(\mathcal{S}(\wp_x) \in \mathfrak{S}(\mathcal{P})\right)$, rather than robustness $\left(\overline{Pr}_{(HMP,Urea.Folate)} \approx 54\%\right)$ (**Figures 2-5**; **Table 3**).

We conclude this discussion and suggest that an outcome of equivalence may constitute a useful metric of comparison whilst evaluating a class or even classes of biochemical networks, i.e., may function as a global parameter. On the other hand, an outcome of non-equivalence may be better suited to assess a single biochemical network, i.e., locally relevant,

$$\text{Metric: (Outcome}\sim\text{"Equivalence")} :: \left(\mathfrak{S}(\mathcal{P}), \widehat{\mathfrak{S}}(\mathcal{P})\right): \text{global} \quad (152)$$

$$\text{Metric: (Outcome}\sim\text{"Non} - \text{equivalence")} :: \mathcal{S}(\wp_{\mathrm{x}}): \text{local} \quad (153)$$

$$\begin{aligned} &\text{where,} \\ &\mathcal{S}(\wp_{\mathrm{x}}) \in \mathfrak{S}(\mathcal{P}) \\ &\mathrm{x} \in [1, \mathrm{X}] \\ &\mathrm{X} \geq 64 \end{aligned}$$

5 Concluding remarks

This study develops an algebraic model for a class of biochemical networks with a set of constrained, reaction-centric, stoichiometry equivalent and degenerate (CReaSED)-matrices. These are rigorously constructed, have distinct mathematical properties, are biochemically relevant, readily parameterizable and easily testable. The theoretical results (axioms, lemmas, theorems, corollaries) are complemented with detailed computational studies on known and well-characterized biochemical networks. The set of CReaSED-matrices provides a biochemically relevant mathematical framework to explore molecular redundancy and minimal perturbations in a class of biochemical networks. The model can be utilized to compare biochemical networks in different cells, tissues and across taxa. Future studies will endeavour to introduce and investigate alternate perturbations that may be used to simulate disease causing states and imbue existing metrics with greater universality, applicability and robustness.

6 *Detailed proofs to model, characterize and analyse a class of biochemical networks with a set of CReaSED-matrices*

Proofs of theorems to construct, formulate and define the CReaSED-matrices for a class of biochemical networks

Proof (D28):

$$\binom{J}{t = J_{(+)}} = \left(\frac{J!}{(J_{(+)})!\,(J - J_{(+)})!}\right) \quad (154)$$

$$= \left(\frac{J!}{(J_{(+)})!\,(J_{(-)})!}\right) \quad (154.1)$$

$$= \left(\frac{J!}{(J_{(+)})!\,(0)!}\right) \quad (154.2)$$

$$= 1 \quad (154.3)$$

Similarly,

$$\binom{J}{t = J_{(-)}} = \left(\frac{J!}{(J_{(-)})!\,(0)!}\right) \quad (155)$$

$$= 1 \quad (155.1)$$

From (154.3), (155.1)

■

Proof (D30):

$$\binom{J}{t=(J-1)_{(-)}} = \left(\frac{J!}{\left((J-1)_{(-)}\right)!\left(J-(J-1)_{(-)}\right)!}\right) \tag{156}$$

$$= \left(\frac{J!}{\left((J-1)_{(-)}\right)!\left(1_{(+)}\right)!}\right) \tag{156.1}$$

$$= \left(\frac{J\left((J-1)_{(-)}\right)!}{\left((J-1)_{(-)}\right)!\left(1_{(+)}\right)!}\right) \tag{156.2}$$

$$= J \tag{156.3}$$

$$= I \tag{156.4}$$

$$\tilde{\mathbf{r}}_{\tilde{\omega}} | \tilde{\omega} = 1,2 \ldots I \tag{157}$$

$$\tilde{\mathbf{r}}_1 = \begin{bmatrix} (+)m_{11} \\ (-)m_{21} \\ (-)m_{31} \\ \vdots \\ (-)m_{J1} \end{bmatrix}; \tilde{\mathbf{r}}_2 = \begin{bmatrix} (-)m_{12} \\ (+)m_{22} \\ (-)m_{32} \\ \vdots \\ (-)m_{J2} \end{bmatrix}; \ldots \tilde{\mathbf{r}}_I = \begin{bmatrix} (-)m_{1I} \\ (-)m_{2I} \\ (-)m_{3I} \\ \vdots \\ (+)m_{JI} \end{bmatrix} \quad (158.1 - 158.J)$$

$$\binom{J}{t=(J-1)_{(+)}} = \left(\frac{J!}{\left((J-1)_{(+)}\right)!\left(J-(J-1)_{(+)}\right)!}\right) \tag{159}$$

$$= \left(\frac{J!}{\left((J-1)_{(+)}\right)!\left(1_{(-)}\right)!}\right) \tag{159.1}$$

$$= \left(\frac{J\left((J-1)_{(+)}\right)!}{\left((J-1)_{(+)}\right)!\left(1_{(-)}\right)!}\right) \tag{159.2}$$

$$= J \tag{159.3}$$

$$= I \tag{159.4}$$

$$\mathbf{r}_{\omega} | \omega = 1,2 \ldots I \tag{160}$$

$$\mathbf{r}_1 = \begin{bmatrix} (-)m_{11} \\ (+)m_{21} \\ (+)m_{31} \\ \vdots \\ (+)m_{J1} \end{bmatrix}; \mathbf{r}_2 = \begin{bmatrix} (+)m_{12} \\ (-)m_{22} \\ (+)m_{32} \\ \vdots \\ (+)m_{J2} \end{bmatrix}; \ldots \mathbf{r}_I = \begin{bmatrix} (+)m_{1I} \\ (+)m_{2I} \\ (+)m_{3I} \\ \vdots \\ (-)m_{JI} \end{bmatrix} \quad (160.1 - 160.J)$$

$$\tilde{\mathbf{r}}_1 = (-1).\mathbf{r}_1 \tag{161.1}$$

$$\tilde{\mathbf{r}}_2 = (-1).\mathbf{r}_2 \tag{161.2}$$

$$\vdots$$

$$\tilde{\mathbf{r}}_J = (-1).\mathbf{r}_J \tag{161.J}$$

For each $\tilde{\omega} = 1,2 \ldots I$

$$\exists \tilde{\mathbf{r}}_{\tilde{\omega}} \text{ s.t. } \tilde{\mathbf{r}}_{\tilde{\omega}} = (-1).\mathbf{r}_{\omega} \tag{162}$$

$$\Rightarrow \sum_{\omega=1}^{\omega=I} \mathbf{r}_{\omega} + \sum_{\tilde{\omega}=1}^{\tilde{\omega}=I} \tilde{\mathbf{r}}_{\tilde{\omega}} \tag{163}$$

$$= \sum_{\omega=1}^{\omega=I} (\mathbf{r}_{\omega} + (-1).\mathbf{r}_{\omega}) \tag{163.1}$$

$$= \mathbf{0} \tag{163.2}$$

■

Proof (C1):

Since,

$\phi_x, \phi_y \in \mathbb{R} - \{0\}$ [From Def. (23)]

From (**T1**)

$$\phi_x . S_{\mathcal{p}_x} = S_{\mathcal{p}_y} \tag{164}$$

Rewriting (164),

$$S_{\mathcal{p}_x} = \left(\frac{1}{\phi_x}\right) . S_{\mathcal{p}_y} \tag{165}$$

$$= \left(\phi_x\right)^{-1} . S_{\mathcal{p}_y} \tag{165.1}$$

$$= \phi_y . \mathbf{S}_{\mathcal{p}_y} \tag{166}$$

Conversely,

$$\phi_y . S_{\mathcal{p}_y} = S_{\mathcal{p}_x} \tag{167}$$

Rewriting (167),

$$S_{\mathcal{p}_y} = \left(\frac{1}{\phi_y}\right) . S_{\mathcal{p}_x} \tag{168}$$

$$= \left(\phi_y\right)^{-1} . S_{\mathcal{p}_x} \tag{168.1}$$

$$= \phi_x . S_{\mathcal{p}_x} \tag{169}$$

From (166), (169),

$$\phi_y . S_{\mathcal{p}_y} = \phi_x . \mathbf{S}_{\mathcal{p}_x} \tag{170}$$

$$\Rightarrow \left(S_{\mathcal{p}_y} = \left(\frac{\phi_x}{\phi_y}\right) . S_{\mathcal{p}_x}\right) \vee \left(\left(\frac{\phi_y}{\phi_x}\right) . S_{\mathcal{p}_y} = S_{\mathcal{p}_x}\right) \tag{171}$$

$$\stackrel{\text{def}}{=} \left(\begin{array}{c} \left(S_{\mathcal{p}_y} = \phi_{xy} . S_{\mathcal{p}_x}\right) \vee \left(\phi_{yx} . S_{\mathcal{p}_y} = S_{\mathcal{p}_x}\right) \\ \text{where,} \\ \phi_{xy}, \phi_{yx} \in \mathbb{R} - \{0\} \end{array}\right) \tag{172}$$

$$S_{\mathcal{p}_y} = \left(\phi_{xy} \vee \phi_{yx}\right) . S_{\mathcal{p}_x} \tag{173}$$

■

Proof (C2):

$$\overline{\overline{X}}_1 \stackrel{\text{def}}{=} \text{Number of non} - \text{zero real valued multipliers for all pairs of redundant CReaSED} - \text{matrices} \tag{174}$$

$$= \binom{X}{1} \tag{174.1}$$

$$= \frac{X!}{1!\,(X-1)!} \tag{174.2}$$

$$= \frac{X(X-1)!}{(X-1)!} \tag{174.3}$$

$$= X \tag{174.4}$$

■

Proof (C3):

$$\overline{\overline{X}}_2 \stackrel{\text{def}}{=} \text{Number of non} - \text{zero real valued multipliers for all pairs of distinct CReaSED} - \text{matrices} \tag{175}$$

$$= \binom{X}{2} \tag{175.1}$$

$$= \frac{X!}{2!\,(X-2)!} \tag{175.2}$$

$$= \frac{X(X-1)(X-2)!}{2!\,(X-2)!} \tag{175.3}$$

$$= \frac{X.\,(X-1)}{2!} \cdot \left(\frac{(X-2)!}{(X-2)!}\right) \tag{175.4}$$

$$= \frac{X.\,(X-1)}{2!} \tag{175.5}$$

■

Proof (T2):

Let us populate the diagonal vector of $\mathbf{\Phi}$ with,

$$\text{If}\begin{pmatrix} \mathbf{a}_{xx} \in \text{diag}(\mathbf{\Phi}) \text{ for the self} - \text{pair } \left\{\mathbf{S}_{\mathscr{p}_x}, \mathbf{S}_{\mathscr{p}_x}\right\} \\ \text{s.t. } \mathbf{a}_{xx} = \phi_{xx} = 1 \\ \mathbf{S}_{\mathscr{p}_x} \in \mathfrak{S}\left(\mathscr{p}, X, J, \bar{\bar{I}}\right) \\ x = 1, \ldots, X \end{pmatrix}$$

then,

$$\#\text{diag}(\mathbf{\Phi}) \tag{176}$$
$$= X \qquad [\text{From } (\mathbf{C2})] \tag{176.1}$$
$$\stackrel{\text{def}}{=} \bar{\bar{X}}_1 \tag{176.2}$$

Let us populate the upper and lower triangular elements of $\mathbf{\Phi}$ with,

$$\text{If}\begin{pmatrix} \mathbf{a}_{xy}, \mathbf{a}_{yx} \in \left(\text{UT}(\mathbf{\Phi}) \cup \text{LT}(\mathbf{\Phi})\right) \text{ for } \left\{\mathbf{S}_{\mathscr{p}_x}, \mathbf{S}_{\mathscr{p}_y}\right\} \\ \text{s.t. } \mathbf{a}_{xy} = \phi_{xy}; \mathbf{a}_{yx} = \phi_{yx} = \phi_{xy}^{-1}, \\ \phi_{xy} \in \mathbb{R} - \{0\} \\ \mathbf{S}_{\mathscr{p}_x}, \mathbf{S}_{\mathscr{p}_y} \in \mathfrak{S}\left(\mathscr{p}, X, J, \bar{\bar{I}}\right); x \neq y;\ x, y = 1, \ldots, X \end{pmatrix}$$

then,

$$\#\left(\text{UT}(\mathbf{\Phi}) \cup \text{LT}(\mathbf{\Phi})\right) \tag{177}$$
$$= \#\left(\text{UT}(\mathbf{\Phi})\right) + \#\left(\text{LT}(\mathbf{\Phi})\right) \tag{177.1}$$
$$= 2.\left(\frac{X.(X-1)}{2!}\right) \qquad [\text{From } (\mathbf{C3})] \tag{177.2}$$
$$= X.(X-1) \tag{177.3}$$
$$\stackrel{\text{def}}{=} \bar{\bar{X}}_2 \tag{177.4}$$

$$\#\text{Vec}(\mathbf{\Phi})$$
$$= \#\text{diag}(\mathbf{\Phi}) + \#\left(\text{UT}(\mathbf{\Phi})\right) + \#\left(\text{LT}(\mathbf{\Phi})\right) \tag{178}$$
$$= \bar{\bar{X}}_1 + \bar{\bar{X}}_2 \tag{178.1}$$
$$= X + X.(X-1) \tag{178.2}$$
$$= X + X^2 - X \tag{178.3}$$
$$= X^2 \tag{178.4}$$

where,

$\mathbf{\Phi} \subset (\mathbb{R} - \{0\})^{X \times X}$

■

Proof (T3):

$$= \prod_{x=1}^{x=X} \prod_{y=1}^{y=X} \phi_{xy} \tag{179}$$

$$= \left(\prod_{x=2}^{x=X} \prod_{y<x}^{y=X} \phi_{xy}\right) . \left(\prod_{y=2}^{y=X} \prod_{x<y}^{y=X} \phi_{xy}\right) . \prod_{x=1}^{x=X} \phi_{xx} \tag{179.1}$$

$$= \mathrm{LT}(\boldsymbol{\Phi}) . \mathrm{UT}(\boldsymbol{\Phi}) . \mathrm{D}(\boldsymbol{\Phi}) \tag{179.2}$$

$$= \left(\left(\mathrm{LT}(\boldsymbol{\Phi}) . \mathrm{LT}^{-1}(\boldsymbol{\Phi})\right) \vee \left(\mathrm{UT}^{-1}(\boldsymbol{\Phi}) . \mathrm{UT}(\boldsymbol{\Phi})\right)\right) . \prod_{x=1}^{x=X} \phi_{xx} \tag{179.3}$$

$$= \left(\left(\mathrm{LT}(\boldsymbol{\Phi}) . \mathrm{LT}^{-1}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx}\right) \vee \left(\left(\mathrm{UT}^{-1}(\boldsymbol{\Phi}) . \mathrm{UT}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx}\right) \tag{179.4}$$

$$= 1 \tag{179.5}$$

■

Proof (C6):

Case1: $\left(\mathrm{LT}(\boldsymbol{\Phi}) . \mathrm{LT}^{-1}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx}$ Def. (39)

1a:

$$= \left(\mathrm{LT} - \mathrm{W}(\boldsymbol{\Phi}) . \mathrm{LT}^{-1} + \mathrm{W}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx} \tag{180}$$

$$= \frac{\mathrm{LT} - \mathrm{W}(\boldsymbol{\Phi})}{\mathrm{LT} + \mathrm{W}(\boldsymbol{\Phi})} . \prod_{x=1}^{x=X} \phi_{xx} \tag{180.1}$$

$$= \frac{\mathrm{LT} - \mathrm{W}(\boldsymbol{\Phi})}{\mathrm{LT} + \mathrm{W}(\boldsymbol{\Phi})} \tag{180.2}$$

$$< 1 \tag{180.3}$$

1b:

$$= \left(\mathrm{LT} - \mathrm{W}(\boldsymbol{\Phi}) . \mathrm{LT}^{-1} + \mathrm{W}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx} \tag{181}$$

$$= \frac{\mathrm{LT} - \mathrm{W}(\boldsymbol{\Phi})}{\mathrm{LT} + \mathrm{W}(\boldsymbol{\Phi})} . \prod_{x=1}^{x=X} \phi_{xx} \tag{181.1}$$

$$= \frac{\mathrm{LT} - \mathrm{W}(\boldsymbol{\Phi})}{\mathrm{LT} + \mathrm{W}(\boldsymbol{\Phi})} \tag{181.2}$$

$$> 1 \tag{181.3}$$

Case2: $\left(\mathrm{UT}^{-1}(\boldsymbol{\Phi}) . \mathrm{UT}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx}$ Def. (40)

1a:

$$= \left(\mathrm{UT}^{-1} + \mathrm{W}(\boldsymbol{\Phi}) . \mathrm{UT} - \mathrm{W}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx} \tag{182}$$

$$= \frac{\mathrm{UT} - \mathrm{W}(\boldsymbol{\Phi})}{\mathrm{UT} + \mathrm{W}(\boldsymbol{\Phi})} . \prod_{x=1}^{x=X} \phi_{xx} \tag{182.1}$$

$$< 1 \tag{182.2}$$

1b:

$$= \left(\mathrm{UT}^{-1} - \mathrm{W}(\boldsymbol{\Phi}) . \mathrm{UT} + \mathrm{W}(\boldsymbol{\Phi})\right) . \prod_{x=1}^{x=X} \phi_{xx} \tag{183}$$

$$= \frac{\mathrm{UT} + \mathrm{W}(\boldsymbol{\Phi})}{\mathrm{UT} - \mathrm{W}(\boldsymbol{\Phi})} . \prod_{x=1}^{x=X} \phi_{xx} \tag{183.1}$$

$$> 1 \tag{183.2}$$

Proof (T4):

For the matrix $\mathbf{\Phi}$ of real valued multipliers where

$$\begin{pmatrix} \mathbf{a}_{xy}, \mathbf{a}_{yx} \in \mathbf{\Phi} \subset (\mathbb{R} - \{0\})^{X \times X} \\ \text{s.t. } \mathbf{a}_{xy} = \phi_{xy} \text{ and } \mathbf{a}_{yx} = \phi_{xy}^{-1} = \phi_{yx} \\ \text{where,} \\ x, y \in [1, X] \end{pmatrix} \qquad [\text{From } (\mathbf{T2})]$$

$$\text{then } \left(\phi_{xy} = \phi_{yx} \text{ iff } \phi_{xy} = \phi_{yx} = 1\right) \tag{184}$$

$$\Rightarrow \mathcal{S}_{p_x} = \mathcal{S}_{p_y} \tag{185}$$

Extending this result $\forall x \neq y$ where $x = 1, \ldots, X$

$$\mathcal{S}_{p_1} = \mathcal{S}_{p_2} \ldots = \mathcal{S}_{p_X} \tag{186}$$

■

Proofs of theorems that investigate the properties of the CReaSED-matrices via the computed null space

Proof (T5):

$$\mathcal{S}_{p_x} = (\phi_{xy} \vee \phi_{yx}).\mathcal{S}_{p_x} \qquad \text{[From } \mathbf{T2}\text{]} \qquad (187)$$

$$N(\mathcal{S}_{p_x}) = N\left((\phi_{xy} \vee \phi_{yx}).\mathcal{S}_{p_x}\right) \qquad (188)$$

$$= N\left((\phi_{xy}.\mathcal{S}_{p_x} \vee \phi_{yx}.\mathcal{S}_{p_x})\right) \qquad (188.1)$$

$$= N\left((\phi_{xy}.\mathcal{S}_{p_x}) \vee (\phi_{yx}.\mathcal{S}_{p_x})\right) \qquad (188.2)$$

$$= N(\phi_{xy}.\mathcal{S}_{p_x}) \vee N(\phi_{yx}.\mathcal{S}_{p_x}) \qquad (188.3)$$

$$= N\left(\mathcal{S}_{p_y}\right) \qquad (188.4)$$

Conversely,

$$\mathcal{S}_{p_y} = (\phi_{xy} \vee \phi_{yx}).\mathcal{S}_{p_y} \qquad \text{[From } \mathbf{T2}\text{]} \qquad (189)$$

$$N\left(\mathcal{S}_{p_y}\right) = N\left((\phi_{xy} \vee \phi_{yx}).\mathcal{S}_{p_y}\right) \qquad (190)$$

$$= N\left(\left(\phi_{xy}.\mathcal{S}_{p_y} \vee \phi_{yx}.\mathcal{S}_{p_y}\right)\right) \qquad (190.1)$$

$$= N\left(\left(\phi_{xy}.\mathcal{S}_{p_y}\right) \vee \left(\phi_{yx}.\mathcal{S}_{p_y}\right)\right) \qquad (190.2)$$

$$= N\left(\phi_{xy}.\mathcal{S}_{p_y}\right) \vee N\left(\phi_{yx}.\mathcal{S}_{p_y}\right) \qquad (190.3)$$

$$= N(\mathcal{S}_{p_x}) \qquad (190.4)$$

From(188.4), (190.4)

$$N(\mathcal{S}_{p_x}) = N\left(\mathcal{S}_{p_y}\right) \qquad (191)$$

For any $z = 1,2 \ldots X$

$$\mathcal{S}_{p_z} \; (\phi_{xz} \vee \phi_{zx}).\mathcal{S}_{p_z} \qquad \text{[From } \mathbf{T2}\text{]} \qquad (192)$$

$$N(\mathcal{S}_{p_z}) = N\left((\phi_{xz} \vee \phi_{zx}).\mathcal{S}_{p_z}\right) \qquad (193)$$

$$= N\left((\phi_{xz}.\mathcal{S}_{p_z} \vee \phi_{zx}.\mathcal{S}_{p_z})\right) \qquad (193.1)$$

$$= N\left((\phi_{xz}.\mathcal{S}_{p_z}) \vee (\phi_{zx}.\mathcal{S}_{p_z})\right) \qquad (193.2)$$

$$= N(\phi_{xz}.\mathcal{S}_{p_z}) \vee N(\phi_{zx}.\mathcal{S}_{p_z}) \qquad (193.3)$$

$$= N(\mathcal{S}_{p_x}) \qquad (193.4)$$

From (191), (193.4)

$$N(\mathcal{S}_{p_x}) = N\left(\mathcal{S}_{p_y}\right) = N(\mathcal{S}_{p_z}) \qquad (194)$$

For any pair of redundant and non − redundant CReaSED − matrices, where $x, y = 1, \ldots, X$

$$N(\mathcal{S}_{p_x}) = N\left(\mathcal{S}_{p_y}\right) \ldots = N(\mathcal{S}_{p_X})$$

■

Proof (T6):

$$\mathbf{N}(\mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}})) \qquad \text{[From } (\mathbf{T5}), (\mathbf{L2}), (\mathbf{L3})] \tag{195}$$

$$= \mathbf{N}\left(\bigcup_{x=1}^{x=X} \mathbf{S}_{\mathcal{p}_x}\right) \tag{195.1}$$

$$= \bigcup_{x=1}^{x=X} N(\mathbf{S}_{\mathcal{p}_x}) \tag{195.2}$$

$$= N(\mathbf{S}_{\mathcal{p}_1}) \cup N(\mathbf{S}_{\mathcal{p}_2}) \ldots \cup N(\mathbf{S}_{\mathcal{p}_{x=X}}) \tag{195.3}$$

$$= N(\mathbf{S}_{\mathcal{p}}) \tag{195.4}$$

Generalising,

$$\text{If}\left(\begin{array}{c}\mathbf{S}_{\mathcal{p}_1}, \mathbf{S}_{\mathcal{p}_2} \ldots \mathbf{S}_{\mathcal{p}_{x=X}} \in \mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}}) \text{ then,} \\ \exists f\left(\mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}})\right) | f(\mathbf{S}_{\mathcal{p}_1}), f(\mathbf{S}_{\mathcal{p}_2}) \ldots, f(\mathbf{S}_{\mathcal{p}_{x=X}}) \in N\left(\mathfrak{S}(\mathcal{p}, x, X, J, \bar{\bar{I}})\right)\end{array}\right)$$

$$\Rightarrow f^{-1}\left(N\left(\mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}})\right)\right) = f^{-1} \circ f\left(\mathfrak{S}(\mathcal{p}, x, X, J, \bar{\bar{I}})\right) \in \mathfrak{S}(\mathcal{p}, x, X, J, \bar{\bar{I}}) \tag{196}$$

But,

$$f(\mathbf{S}_{\mathcal{p}_1}) = f(\mathbf{S}_{\mathcal{p}_2}) \ldots = f(\mathbf{S}_{\mathcal{p}_{x=X}}) \tag{197}$$

$$= N(\mathbf{S}_{\mathcal{p}_1}) = N(\mathbf{S}_{\mathcal{p}_2}) \ldots = N(\mathbf{S}_{\mathcal{p}_{x=X}}) \tag{197.1}$$

$$= N(\mathbf{S}_{\mathcal{p}}) \tag{197.2}$$

From (196),

$$f^{-1}\left(N(\mathbf{S}_{\mathcal{p}})\right) \in \mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}}) \tag{198}$$

$$\Rightarrow \mathbf{S}_{\mathcal{p}} \in \mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}}) \tag{199}$$

$$\Rightarrow \mathbf{S}_{\mathcal{p}} = \{\mathbf{S}_{\mathcal{p}_1}\mathbf{S}_{\mathcal{p}_2} \ldots \mathbf{S}_{\mathcal{p}_{x=X}}\} \tag{200}$$

Since,

$$\mathbf{S}_{\mathcal{p}_1} \neq \mathbf{S}_{\mathcal{p}_2} \ldots \neq \mathbf{S}_{\mathcal{p}_{x=X}} \qquad \text{[From } \mathbf{T1}]$$

$$\Rightarrow f: \mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}}) \twoheadrightarrow N\left(\mathfrak{S}(\mathcal{p}, X, J, \bar{\bar{I}})\right) \tag{201}$$

■

Proofs of theorems and lemmas to establish and investigates algebraic properties of the CReaSED-matrices for a modelled biochemical network

Proof (L1):

Case 1:

$$\mathbf{S}_{\mathcal{p}_x} + \mathbf{S}_{\mathcal{p}_y} = \mathbf{0}^{I \times J} \tag{202}$$

From (**T3**), (**T4**),

$$\phi_x . \mathbf{S}_{\mathcal{p}_x} + \phi_y . \mathbf{S}_{\mathcal{p}_y} = \mathbf{0}^{I \times J} | \phi_x, \phi_y \in \mathbb{R} - \{0\}$$

Let,

$$\mathbf{S}_{\mathcal{p}_y} = (-1) . \mathbf{S}_{\mathcal{p}_x} \tag{203}$$

Rewriting,

$$\mathbf{S}_{\mathcal{p}_x} + \mathbf{S}_{\mathcal{p}_y} = \mathbf{S}_{\mathcal{p}_x} + (-1) . \mathbf{S}_{\mathcal{p}_x} \tag{204}$$

$$= 0 . \mathbf{S}_{\mathcal{p}_x} \tag{204.1}$$

$$= \mathbf{0}^{I \times J} \tag{204.2}$$

$$= \mathbf{S}_{\mathcal{p}_x} . 0 \tag{204.3}$$

$$= (-1) . \mathbf{S}_{\mathcal{p}_x} + \mathbf{S}_{\mathcal{p}_x} \tag{204.4}$$

$$= \mathbf{S}_{\mathcal{p}_y} + \mathbf{S}_{\mathcal{p}_x} \tag{204.5}$$

Case 2:

$$\mathbf{S}_{\mathcal{p}_x} + \mathbf{S}_{\mathcal{p}_y} \neq \mathbf{0}^{I \times J} \tag{205}$$

$$\mathbf{S}_{\mathcal{p}_x} + \mathbf{S}_{\mathcal{p}_y} = \mathbf{S}_{\mathcal{p}_{xy}} \tag{206}$$

$$= \mathbf{S}_{\mathcal{p}_{yx}} \tag{206.1}$$

$$= \mathbf{S}_{\mathcal{p}_y} + \mathbf{S}_{\mathcal{p}_x} \tag{206.2}$$

From (204.5), (206.2),

■

Proof (T7):

No Additive identity:
By definition,

$$\mathbf{0}^{I\times J} \notin \mathfrak{S}\left(\mathcal{p}, X, J, \bar{\bar{I}}\right) \tag{207}$$

Conditional closure:
From (**L1**),
iff,

$$\mathbf{S}_{\mathcal{p}_x} + \mathbf{S}_{\mathcal{p}_y} = \mathbf{S}_{\mathcal{p}_{x+y}} \neq \mathbf{0}^{I\times J} | x, y = 1, \ldots, X; x \neq y$$

then,

$$\exists \mathbf{S}_{\mathcal{p}_{x+y}} \in \mathfrak{S}\left(\mathcal{p}, X, J, \bar{\bar{I}}\right) \tag{208}$$

Association:

$$\left(\mathbf{S}_{\mathcal{p}_1} + \mathbf{S}_{\mathcal{p}_2}\right) + \mathbf{S}_{\mathcal{p}_3} = \mathbf{S}_{\mathcal{p}_{1+2}} + \mathbf{S}_{\mathcal{p}_3} \tag{209}$$

$$= \mathbf{S}_{\mathcal{p}_{1+2+3}} \tag{209.1}$$

$$= \mathbf{S}_{\mathcal{p}_6} \tag{209.2}$$

$$\in \mathfrak{S}\left(\mathcal{p}, X, J, \bar{\bar{I}}\right) \tag{209.3}$$

$$\mathbf{S}_{\mathcal{p}_1} + \left(\mathbf{S}_{\mathcal{p}_2} + \mathbf{S}_{\mathcal{p}_3}\right) = \mathbf{S}_{\mathcal{p}_1} + \mathbf{S}_{\mathcal{p}_{2+3}} \tag{210}$$

$$= \mathbf{S}_{\mathcal{p}_{1+2+3}} \tag{210.1}$$

$$= \mathbf{S}_{\mathcal{p}_6} \tag{210.2}$$

$$\in \mathfrak{S}\left(\mathcal{p}, X, J, \bar{\bar{I}}\right) \tag{210.3}$$

From $(207 - 210)$,

$$\left(\mathbf{S}_{\mathcal{p}_1} + \mathbf{S}_{\mathcal{p}_2}\right) + \mathbf{S}_{\mathcal{p}_3} = \mathbf{S}_{\mathcal{p}_1} + \left(\mathbf{S}_{\mathcal{p}_2} + \mathbf{S}_{\mathcal{p}_3}\right) \tag{211}$$

Generalizing,

$$\left(\mathbf{S}_{\mathcal{p}_1} + \mathbf{S}_{\mathcal{p}_2}\right) \ldots. + \mathbf{S}_{\mathcal{p}_X} = \mathbf{S}_{\mathcal{p}_1} + \left(\mathbf{S}_{\mathcal{p}_2} \ldots. + \mathbf{S}_{\mathcal{p}_X}\right) \tag{212}$$

$$= \mathbf{S}_{\mathcal{p}_{1+2\ldots+X}} \tag{212.1}$$

$$= \mathbf{S}_{\mathcal{p}_{\left(\frac{(X).(X+1)}{2}\right)}} \tag{212.2}$$

$$\in \mathfrak{S}\left(\mathcal{p}, X, J, \bar{\bar{I}}\right) \tag{212.3}$$

■

Proof (L2):

From (**T1**)

$$\mathbf{S}_{\mathcal{p}_1} \neq \mathbf{S}_{\mathcal{p}_2} \ldots. \neq \mathbf{S}_{\mathcal{p}_{x=X}}$$

it follows,

$$\mathbf{S}_{\mathcal{p}_1} \cap \mathbf{S}_{\mathcal{p}_2} \ldots. \cap \mathbf{S}_{\mathcal{p}_{x=X}} = \emptyset \tag{213}$$

$$\bigcap_{x=1}^{x=X} \mathbf{S}_{\mathcal{p}_x} = \emptyset \tag{214}$$

■

Proof (L3):

Since,

$$\mathbf{S}_{\mathscr{p}_{\mathrm{x}}} \in \left(\bigcup_{\mathrm{x}=1}^{\mathrm{x}=\mathrm{X}} \mathbf{S}_{\mathscr{p}_{\mathrm{x}}} - \bigcap_{\mathrm{x}=1}^{\mathrm{x}=\mathrm{X}} \mathbf{S}_{\mathscr{p}_{\mathrm{x}}}\right)$$

From (**L2**),

$$\mathbf{S}_{\mathscr{p}_{\mathrm{x}}} \in \left(\bigcup_{\mathrm{x}=1}^{\mathrm{x}=\mathrm{X}} \mathbf{S}_{\mathscr{p}_{\mathrm{x}}} - \{\emptyset\}\right) \tag{215}$$

$$\in \bigcup_{\mathrm{x}=1}^{\mathrm{x}=\mathrm{X}} \mathbf{S}_{\mathscr{p}_{\mathrm{x}}} \tag{215.1}$$

$$= \mathfrak{S}\left(\mathscr{p}, \mathrm{X}, \mathrm{J}, \bar{\bar{\mathrm{I}}}\right) \tag{215.2}$$

■

Author(s) contributions

SK, designed the study; defined, formulated, developed and wrote proofs for the theorems, lemmas and corollaries; developed the assessment metrics and algorithms; conducted numerical studies and collated data; wrote all necessary code and manuscript.

Availability of data

The manuscript has data included as part of this preprint.

Conflict of interest

The author declares no conflict of interest.

Acknowledgement(s)

This study was supported in part by an extramural grant from the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India under the Mathematical Research Impact-Centric Support (MATRICS) scheme to SK (MTR/2021/000290).

References

1. Kim OD, Rocha M, Maia P. A Review of Dynamic Modeling Approaches and Their Application in Computational Strain Optimization for Metabolic Engineering. Front Microbiol. 2018;9:1690. Epub 2018/08/16. doi: 10.3389/fmicb.2018.01690. PubMed PMID: 30108559; PubMed Central PMCID: PMCPMC6079213.
2. Maarleveld TR, Khandelwal RA, Olivier BG, Teusink B, Bruggeman FJ. Basic concepts and principles of stoichiometry modeling of metabolic networks. Biotechnol J. 2013;8(9):997-1008. Epub 2013/07/31. doi: 10.1002/biot.201200291. PubMed PMID: 23893965; PubMed Central PMCID: PMCPMC4671265.
3. Klamt S, Muller S, Regensburger G, Zanghellini J. A mathematical framework for yield (vs. rate) optimization in constraint-based modeling and applications in metabolic engineering. Metab Eng. 2018;47:153-69. Epub 2018/02/11. doi: 10.1016/j.ymben.2018.02.001. PubMed PMID: 29427605; PubMed Central PMCID: PMCPMC5992331.
4. Klamt S, Regensburger G, Gerstl MP, Jungreuthmayer C, Schuster S, Mahadevan R, et al. From elementary flux modes to elementary flux vectors: Metabolic pathway analysis with arbitrary linear flux constraints. PLoS Comput Biol. 2017;13(4):e1005409. Epub 2017/04/14. doi: 10.1371/journal.pcbi.1005409. PubMed PMID: 28406903; PubMed Central PMCID: PMCPMC5390976.
5. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nat Biotechnol. 2010;28(3):245-8. Epub 2010/03/10. doi: 10.1038/nbt.1614. PubMed PMID: 20212490; PubMed Central PMCID: PMCPMC3108565.
6. Muller S, Regensburger G. Elementary Vectors and Conformal Sums in Polyhedral Geometry and their Relevance for Metabolic Pathway Analysis. Front Genet. 2016;7:90. Epub 2016/06/03. doi: 10.3389/fgene.2016.00090. PubMed PMID: 27252734; PubMed Central PMCID: PMCPMC4877377.
7. Kundu, S. ReDirection: an R-package to compute the probable dissociation constant for every reaction of a user-defined biochemical network. *Frontiers in Molecular Biosciences* **10**, doi:10.3389/fmolb.2023.1206502 (2023).
8. Kundu, S. A mathematically rigorous algorithm to define, compute and assess relevance of the probable dissociation constants in characterizing a biochemical network. Sci Rep 14, 3507, doi:10.1038/s41598-024-53231-9 (2024).
9. Srividhya J, Crampin EJ, McSharry PE, Schnell S. Reconstructing biochemical pathways from time course data. Proteomics. 2007;7(6):828-38. Epub 2007/03/21. doi: 10.1002/pmic.200600428. PubMed PMID: 17370261.
10. Oksuz M, Sadikoglu H, Cakir T. Sparsity as cellular objective to infer directed metabolic networks from steady-state metabolome data: a theoretical analysis. PLoS One. 2013;8(12):e84505. Epub 2014/01/07. doi: 10.1371/journal.pone.0084505. PubMed PMID: 24391961; PubMed Central PMCID: PMCPMC3877278.
11. Nagele T. Linking metabolomics data to underlying metabolic regulation. Front Mol Biosci. 2014;1:22. Epub 2014/01/01. doi: 10.3389/fmolb.2014.00022. PubMed PMID: 25988163; PubMed Central PMCID: PMCPMC4428386.
12. Cakir T, Hendriks MM, Westerhuis JA, Smilde AK. Metabolic network discovery through reverse engineering of metabolome data. Metabolomics. 2009;5(3):318-29. Epub 2009/09/01. doi: 10.1007/s11306-009-0156-4. PubMed PMID: 19718266; PubMed Central PMCID: PMCPMC2731157.
13. Khatibipour MJ, Kurtoglu F, Cakir T. JacLy: a Jacobian-based method for the inference of metabolic interactions from the covariance of steady-state metabolome data. PeerJ. 2018;6:e6034. Epub 2018/12/20. doi: 10.7717/peerj.6034. PubMed PMID: 30564518; PubMed Central PMCID: PMCPMC6286809.
14. Wang T, Ren Z, Ding Y, Fang Z, Sun Z, MacDonald ML, et al. FastGGM: An Efficient Algorithm for the Inference of Gaussian Graphical Model in Biological Networks. PLoS Comput Biol. 2016;12(2):e1004755. Epub 2016/02/13. doi: 10.1371/journal.pcbi.1004755. PubMed PMID: 26872036; PubMed Central PMCID: PMCPMC4752261.
15. Nagele T, Mair A, Sun X, Fragner L, Teige M, Weckwerth W. Solving the differential biochemical Jacobian from metabolomics covariance data. PLoS One. 2014;9(4):e92299. Epub 2014/04/04. doi: 10.1371/journal.pone.0092299. PubMed PMID: 24695071; PubMed Central PMCID: PMCPMC3977476.
16. Rottman BM, Hastie R. Reasoning about causal relationships: Inferences on causal networks. Psychol Bull. 2014;140(1):109-39. Epub 2013/04/03. doi: 10.1037/a0031903. PubMed PMID: 23544658; PubMed Central PMCID: PMCPMC3988659.
17. Goto T, Fernandes AFA, Tsudzuki M, Rosa GJM. Causal phenotypic networks for egg traits in an F2 chicken population. Mol Genet Genomics. 2019;294(6):1455-62. Epub 2019/06/27. doi: 10.1007/s00438-019-01588-2. PubMed PMID: 31240383.
18. Seyhan AA, Carini C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? J Transl Med. 2019;17(1):114. Epub 2019/04/07. doi: 10.1186/s12967-019-1864-9. PubMed PMID: 30953518; PubMed Central PMCID: PMCPMC6451233.

19. Biane C, Delaplace F. Causal Reasoning on Boolean Control Networks Based on Abduction: Theory and Application to Cancer Drug Discovery. IEEE/ACM Trans Comput Biol Bioinform. 2019;16(5):1574-85. Epub 2018/12/26. doi: 10.1109/TCBB.2018.2889102. PubMed PMID: 30582550.
20. Pai S, Bader GD (2018) Patient Similarity Networks for Precision Medicine. *J Mol Biol* 430: 2924-2938.
21. Sanchez P, Voisey JP, Xia T, et al. (2022) Causal machine learning for healthcare and precision medicine. *Soc Open Sci* 9: 220638.
22. Thomas JP, Modos D, Korcsmaros T, et al. (2021) Network Biology Approaches to Achieve Precision Medicine in Inflammatory Bowel Disease. *Front Genet* 12: 760501.
21. Du Q, Gong C, Wang Q, Zhou D, Yang H, Pan W, et al. Genetic architecture of growth traits in Populus revealed by integrated quantitative trait locus (QTL) analysis and association studies. New Phytol. 2016;209(3):1067-82. Epub 2015/10/27. doi: 10.1111/nph.13695. PubMed PMID: 26499329.
22. Gove RP, Chen W, Zweber NB, Erwin R, Rychtar J, Remington DL. Effects of causal networks on the structure and stability of resource allocation trait correlations. J Theor Biol. 2012;293:1-14. Epub 2011/10/19. doi: 10.1016/j.jtbi.2011.09.034. PubMed PMID: 22004994.
23. Salvador AC, Arends D, Barrington WT, Elsaadi AM, Brockmann GA, Threadgill DW. Sex-specific genetic architecture in response to American and ketogenic diets. Int J Obes (Lond). 2021;45(6):1284-97. Epub 2021/03/17. doi: 10.1038/s41366-021-00785-7. PubMed PMID: 33723359; PubMed Central PMCID: PMCPMC8159743.
24. Gopalan L, Sebastian A, Praul CA, Albert I, Ramachandran R. Metformin Affects the Transcriptomic Profile of Chicken Ovarian Cancer Cells. Genes (Basel). 2021;13(1). Epub 2022/01/22. doi: 10.3390/genes13010030. PubMed PMID: 35052372; PubMed Central PMCID: PMCPMC8774788.
25. Lu J, Dumitrascu B, McDowell IC, Jo B, Barrera A, Hong LK, et al. Causal network inference from gene transcriptional time-series response to glucocorticoids. PLoS Comput Biol. 2021;17(1):e1008223. Epub 2021/01/30. doi: 10.1371/journal.pcbi.1008223. PubMed PMID: 33513136; PubMed Central PMCID: PMCPMC7875426
26. Gross, E., Harrington, H. A., Rosen, Z. & Sturmfels, B. Algebraic Systems Biology: A Case Study for the Wnt Pathway. Bull Math Biol 78, 21-51, doi:10.1007/s11538-015-0125-1 (2016).
27. Loutchko, D. Semigroup models for biochemical reaction networks. J Math Biol 86, 78, doi:10.1007/s00285-023-01898-5 (2023).
28. Angeli, D., De Leenheer, P. & Sontag, E. D. A Petri net approach to the study of persistence in chemical reaction networks. Math Biosci 210, 598-618, doi:10.1016/j.mbs.2007.07.003 (2007).
29. Shiu, A. & Sturmfels, B. Siphons in chemical reaction networks. Bull Math Biol 72, 1448-1463, doi:10.1007/s11538-010-9502-y (2010).
30. Harrington, H. A., Ho, K. L., Thorne, T. & Stumpf, M. P. Parameter-free model discrimination criterion based on steady-state coplanarity. Proc Natl Acad Sci U S A 109, 15746-15751, doi:10.1073/pnas.1117073109 (2012).
31. Taylor, D. et al. Topological data analysis of contagion maps for examining spreading processes on networks. Nat Commun 6, 7723, doi:10.1038/ncomms8723 (2015).
32. Xue Y, Wang Y, Wang S, Yan M, Huang J, Yang X. Label-Free and Regenerable Aptasensor for Real-Time Detection of Cadmium(II) by Dual Polarization Interferometry. Anal Chem. 2020;92(14):10007-15. Epub 2020/07/04. doi: 10.1021/acs.analchem.0c01710. PubMed PMID: 32618180.
33. Da P, Li W, Lin X, Wang Y, Tang J, Zheng G. Surface plasmon resonance enhanced real-time photoelectrochemical protein sensing by gold nanoparticle-decorated TiO(2) nanowires. Anal Chem. 2014;86(13):6633-9. Epub 2014/06/11. doi: 10.1021/ac501406x. PubMed PMID: 24915128.
34. Nagai H, Ebisu S, Abe R, Goto T, Takahashi N, Hosaka T, et al. Development of a novel PPARgamma ligand screening system using pinpoint fluorescence-probed protein. Biosci Biotechnol Biochem. 2011;75(2):337-41. Epub 2011/02/11. doi: 10.1271/bbb.100810. PubMed PMID: 21307572.
35. Smith DS, Ferris FG. Specific surface chemical interactions between hydrous ferric oxide and iron-reducing bacteria determined using pK(a) spectra. J Colloid Interface Sci. 2003;266(1):60-7. Epub 2003/09/06. doi: 10.1016/s0021-9797(03)00667-2. PubMed PMID: 12957582.
36. Shindo Y, Kondo Y, Sako Y (2018) Inferring a nonlinear biochemical network model from a heterogeneous single-cell time course data. *Sci Rep* 8: 6790.
37. Koutrouli M, Karatzas E, Paez-Espino D, et al. (2020) A Guide to Conquer the Biological Network Era Using Graph Theory. Front Bioeng Biotechnol 8: 34.
38. Kundu S. Modeling ligand-macromolecular interactions as eigenvalue-based transition-state dissociation constants may offer insights into biochemical function of the resulting complexes. Mathematical Biosciences and Engineering. 2022;19:13252-13275.

39. Ferenci T, Strom T, Quayle JR. Purification and properties of 3-hexulose phosphate synthase and phospho-3-hexuloisomerase from Methylococcus capsulatus. Biochem J. 1974;144(3):477-86. Epub 1974/12/01. doi: 10.1042/bj1440477. PubMed PMID: 4219834; PubMed Central PMCID: PMCPMC1168525.
40. Yao Y, Xu XH, Jin L. Macrophage Polarization in Physiological and Pathological Pregnancy. Front Immunol. 2019;10:792. Epub 2019/05/01. doi: 10.3389/fimmu.2019.00792. PubMed PMID: 31037072; PubMed Central PMCID: PMCPMC6476302.
41. Hoffbrand AV, Weir DG (2001) The history of folic acid. Br J Haematol 113: 579-589.
42. Visentin M, Diop-Bove N, Zhao R, et al. (2014) The intestinal absorption of folates. Annu Rev Physiol 76: 251-274.

Table 1: Parameters to assess suitability of algebraic model to investigate a class of biochemical networks

Parameter	Definition	Outcome	Confusion matrix	Rmks
$\#\mathrm{me}_{\mathcal{P}}$	Number of reactions which have unchanged equivalent outcomes after a perturbation	True positive (TP)	$\mathbb{T}$	Def. (20)
$\#\mathrm{ue}_{\mathcal{P}}$	Number of equivalent reactions which have become non-equivalent, i.e., $\mathrm{equivalent} \rightarrow \{\mathrm{forward}, \mathrm{reverse}\}$ after a perturbing stimulus	False positive (FP)	$\mathbb{T}$	Def. (21)
$\#\mathrm{une}_{\mathcal{P}}$	Number of reactions with a reversed non-equivalent ($\mathrm{forward} \rightleftharpoons \mathrm{reverse}$) or changed ($\{\mathrm{forward}, \mathrm{reverse}\} \rightarrow \mathrm{equivalent}$)-outcomes after a perturbing stimulus	False negative (FN)	$\mathbb{T}$	Def. (22)
$\#\mathrm{mne}_{\mathcal{P}}$	Number of reactions which have unchanged non-equivalent outcomes after a perturbation	True negative (TN)	$\mathbb{T}$	Def. (23)
$\#\mathrm{mne}_{\mathcal{P}}$	Number of reactions which have unchanged non-equivalent outcomes after a perturbation	True positive (TP)	$\tilde{\mathbb{T}}$	Def. (24)
$\#\mathrm{une}_{\mathcal{P}}$	Number of reactions with a reversed non-equivalent (forward$\rightleftharpoons$reverse) or changed ($\{\mathrm{forward}, \mathrm{reverse}\} \rightarrow$equivalent)-outcomes after a perturbing stimulus	False positive (FP)	$\tilde{\mathbb{T}}$	Def. (25)
$\#\mathrm{ue}_{\mathcal{P}}$	Number of equivalent reactions which have become non-equivalent, i.e., equivalent$\rightarrow\{\mathrm{forward}, \mathrm{reverse}\}$ after a perturbing stimulus	False negative (FN)	$\tilde{\mathbb{T}}$	Def. (26)
$\#\mathrm{me}_{\mathcal{P}}$	Number of reactions which have unchanged equivalent outcomes after a perturbation	True negative (TN)	$\tilde{\mathbb{T}}$	Def. (27)

Table 2: Computational studies of well-characterized biochemical networks with "ReDirection" and before and after the introduction of an extraneous and perturbing reaction

	Biochemical network	**Type**	$\bar{\bar{\mathrm{I}}}_x$	J_x	$\mathcal{V}$	**#F**	**#B**	**#E**	**Time** $(\mathbf{s^{-1}})$
1.	Hexose monophosphate	N	18	16	3	6	8	4	25.68
		P	19	16	4	6	9	4	25.07
2.	Urea cycle	N	11	9	2	2	4	5	0.09
		P	12	9	3	8	1	3	1.25
3.	Folate metabolism	N	17	15	3	8	3	6	0.05
		P	18	15	3	8	3	7	0.06

Abbreviations:

$\bar{\bar{\mathrm{I}}}_x$: Number of plausible reaction vectors for a biochemical network
J_x: Number of molecular species for a biochemical network
$\mathcal{V}$: Initial nullity of null space-generated subspace of stoichiometry number matrix for a biochemical network
N: Biochemical network before the addition of extraneous reactions
P: Biochemical network after the addition of extraneous reactions
#F: Number of forward reactions for a biochemical network
#B: Number of reverse reactions for a biochemical network
#E: Number of equivalent reactions for a biochemical network

Table 3: Evaluating the confusion matrices for the modelled biochemical networks before and after the introduction of an extraneous and perturbing reaction

Confusion matrix	Network	TP	FP	FN	TN	Pr	Re	Ac	SPy
$\mathbb{T}$	HMP	4	0	5	9	100%	≈ 44%	≈ 72%	100%
	Urea	3	2	4	2	60%	≈ 43%	≈ 46%	50%
	Folate	6	0	4	7	100%	60%	≈ 77%	100%
$\widetilde{\mathbb{T}}$	HMP	9	5	0	4	≈ 64%	100%	≈ 72%	≈ 44%
	Urea	2	4	2	3	≈ 33%	50%	≈ 46%	≈ 43%
	Folate	7	4	0	6	≈ 64%	100%	≈ 77%	60%

Abbreviations:

$\mathbb{T}$: Confusion matrix with equivalence as the positive outcome
$\widetilde{\mathbb{T}}$: Confusion matrix with non-equivalence as the positive outcome
HMP: Hexose monophosphate shunt
TP: True positive
FP: False positive
FN: False negative
TN: True negative
Pr: Precision or positive predictive value
Re: Recall or sensitivity of test
Ac: Accuracy of test
SPy: Specificity of test

Figure legends

Figure 1: Algorithm to construct an algebraic model for a class of biochemical networks: A biochemical network is modelled as the stoichiometry number matrix between J-molecular species and I-reaction vectors (**S1**). In order to ensure the conformity of the model to biochemically relevant paradigms, subsets of reaction vectors are precluded on the basis of whether they represent a true biochemical reaction. These restrictions are biochemically motivated and are defined in A1-A5 (**S2**). We will annotate this reduced subset of reaction vectors as the set of plausible reaction vectors. Additionally, A6-A10 will be deployed to maintain physiological relevance such as the assumption of equilibrium, molecular redundancy, introduction of extraneous reactions as a means to perturb a biochemical network and the utilization of the probable dissociation constants (**S2**). We will compute lower bounds for the molecular species and plausible reaction vectors for this revised model of a biochemical network (**S3, S4**). In addition to these architecture-preserving constraints, we impose stoichiometry number equivalence under a zero-preserving premise for each molecular species (**S4, S5**). This set of matrices is constrained, reaction-centric, possesses stoichiometry number equivalence for the molecular species and is degenerate. We designate this collection, for a class of biochemical networks, as the set of CReaSED-matrices (**S5**). These are investigated (lemmas, theorems, corollaries), parameterized and complemented with detailed computational studies on biochemical networks which have a well-defined biochemical and physiological role (HMP shunt, Urea cycle, Folate metabolism) (**S6**). **Abbreviations: A, L, T, C** Axiom, Lemma, Theorem, Corollary; **HMP,** Hexose monophosphate shunt pathway; X**CBN**, X-membered class of biochemical networks; **ZP**, Zero-preserving;

Figure 2: The set of CReaSED-matrices is a suitable model for a class of biochemical networks: An algebraic model for a class of biochemical networks is developed with a set of CReaSED-matrices and presented. The mathematical properties of the CReaSED-matrices are studied so as to render a class of biochemical networks amenable to computational studies. In order to simulate the dynamicity of the intracellular milieu whilst ensuring the redundancy of the participating molecular species, one- or more-extraneous plausible reaction vectors are introduced. These will minimally perturb the biochemical network under study and the probable dissociation constants for each pair of reactions can be compared. The computational studies are carried out with "ReDirection", a freely available R-package, for the hexose monophosphate shunt pathway, urea cycle and folate metabolism under naïve- and perturbed-conditions (one or more extraneous reactions). This when coupled with the probable

dissociation constant-derived local- and network-specific metrics can interrogate a biochemical network under differing intracellular conditions, between cells and across taxa. **Abbreviations**: **CReaSED**, set of constrained, reaction-centric, stoichiometry equivalent and degenerate matrices for a biochemical network; $\mathcal{P}; \tilde{\mathcal{P}}$, biochemical networks before and after the addition of one- or more-extraneous reactions; **mne,** matched non-equivalent reactions; **me,** matched equivalent reactions; **f, b, e**; forward-, reverse- and equivalent-reactions;

Figure 3: A**nalysis of a constrained biochemical network of the hexose monophosphate shunt pathway in humans before and after the introduction of an extraneous reaction:** a), b) The HMP shunt can generate NADPH, PRPP and several other potential anaplerotic intermediates. An oxidative intracellular environment will occur when the plasma membrane of a cell is damaged. This may be secondary to an increased fragility of cells when subjected to physiological duress (repeated splenic passage of a red blood corpuscle), exogenous injury or under hereditary conditions (spherocytosis, elliptocytosis). The PRPP formed participates in the de novo synthesis of purine and pyrimidine rings and thence the nucleic acids (DNA/RNA). In this study the direct conversion of β-D fructose 6-phosphate to D-ribulose 6-phosphate (r_{19}) is investigated as the perturbing conditional. This conversion, although, absent in humans is a significant contributor to formaldehyde fixation in bacteria and archaea. The reactions occur via D-arabino hex-3-ulose 6-phosphate and include the enzymes (6-phospho-3-hexuloisomerase, $EC5.3.1.27$; 3-hexulose-6-phosphate synthase, $EC4.1.2.43$). These hypotheses are potentially testable and can be empirically validated. **Abbreviations**: **EC**, enzyme commission; **NADPH**, reduced Nicotinamide adenine dinucleotide phosphate; **mM**, Millimolar; **PRPP**, phosphoribosyl pyrophosphate;

Figure 4: A**nalysis of a constrained biochemical network of the human urea cycle before and after the introduction of an extraneous reaction:** a), b) The human urea cycle is an important pathway in the utilization and thence excretion of non-protein nitrogen. The cycle occurs in the cytosol and the mitochondria. In addition, the cycle is also anaplerotic for Kreb's cycle and is an important contributor to the free radical cycle vis-à-vis nitric oxide synthase $(EC1.14.13.39)$. Whilst the inability to excrete urea requires immediate intervention, impairment of the nitric oxide pathway will retard the formation of the pro-inflammatory macrophage and thence the neutralization of a noxious stimulus. In

this study, the direct conversion from Arginine to Carbamoyl phosphate is examined (r_{12}).

Abbreviations: **EC**, enzyme commission;

Figure 5: Analysis of a constrained biochemical network of folate metabolism in humans before and after the introduction of an extraneous reaction: a), b) The Folate cycle (THF, 5-methylTHF and 5,10-methyleneTHF) is responsible for the movement and exchange of several small molecules (methyl groups, vitamins, amino acids) between the larger macromolecules of DNA/RNA, proteins and lipids. The major clinical manifestation of is a state of relative deficiency and usually occurs when vitamin B12 is deficient. This "Folate"-trap results in the unavailability of Folate for other dependent biochemical processes and the consequent biomedical state, thereof. This is characterized by megaloblastic (macrocyte) anemia since the circulating RBCs are immature and unable to transport molecular oxygen. Since the methyl groups (1C) that are required to synthesize DNA are not available ("Folate"-trap), the immature RBCs have faulty DNA. **Abbreviations**: **B**, "B" group of vitamins; **1C**, 1-carbon group; **DNA/RNA**, Deoxyribonucleic acid/ Ribonucleic acid; **RBC**, Red blood cell; **THF**, Tetrahydrofolate;

Figure 1:

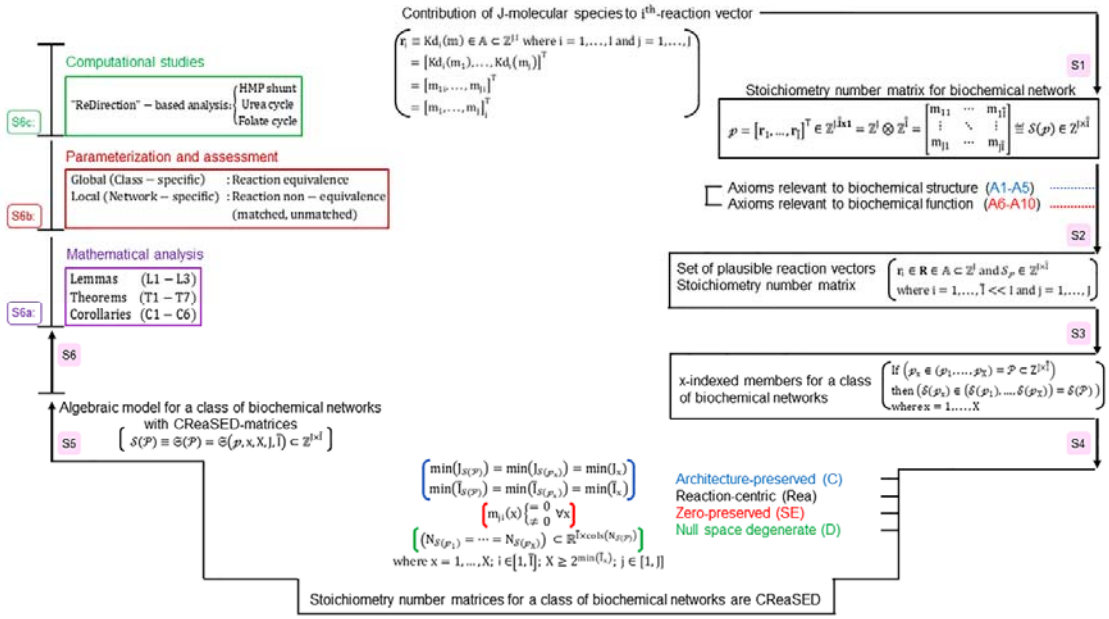

Contribution of J-molecular species to i^th-reaction vector
S1
Stoichiometry number matrix for biochemical network
Axioms relevant to biochemical structure (A1-A5)
Axioms relevant to biochemical function (A6-A10)
S2
Set of plausible reaction vectors
Stoichiometry number matrix
S3
x-indexed members for a class of biochemical networks
S4
Architecture-preserved (C)
Reaction-centric (Rea)
Zero-preserved (SE)
Null space degenerate (D)
Stoichiometry number matrices for a class of biochemical networks are CReaSED
S5
Algebraic model for a class of biochemical networks with CReaSED-matrices
S6
S6a:
Mathematical analysis
Lemmas (L1 – L3)
Theorems (T1 – T7)
Corollaries (C1 – C6)
S6b:
Parameterization and assessment
Global (Class – specific) : Reaction equivalence
Local (Network – specific) : Reaction non – equivalence (matched, unmatched)
S6c:
Computational studies
"ReDirection" – based analysis:
HMP shunt
Urea cycle
Folate cycle

Figure 2:

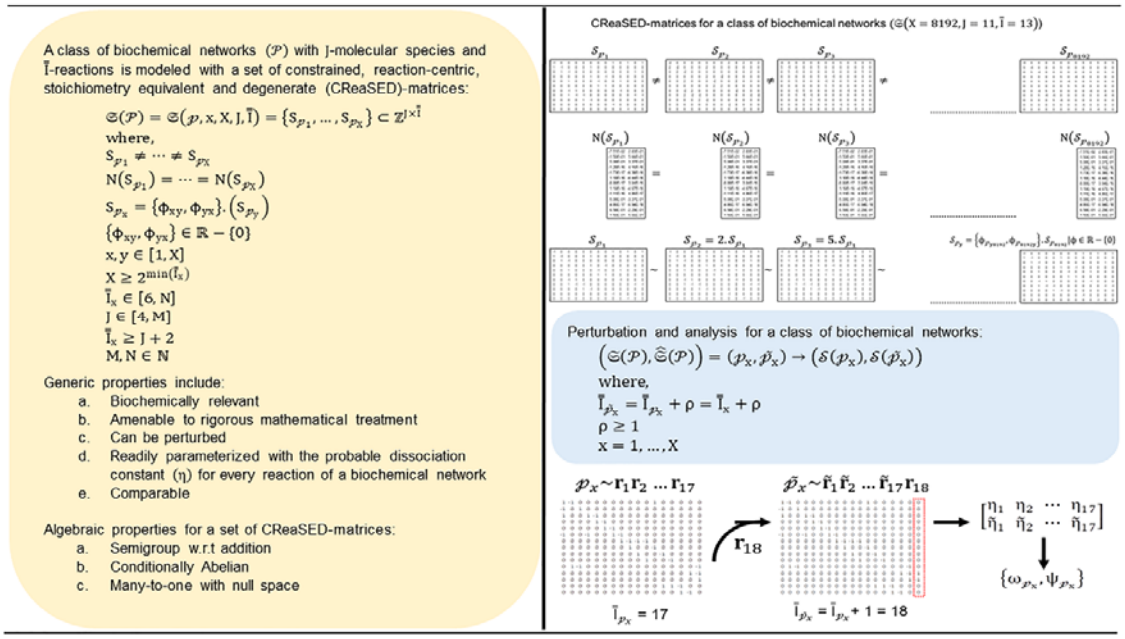

A class of biochemical networks ($\mathcal{P}$) with J-molecular species and $\bar{I}$-reactions is modeled with a set of constrained, reaction-centric, stoichiometry equivalent and degenerate (CReaSED)-matrices:
$\mathfrak{S}(\mathcal{P}) = \mathfrak{S}(\mathcal{p}, x, X, J, \bar{I}) = \{S_{\mathcal{p}_1}, \dots, S_{\mathcal{p}_X}\} \subset \mathbb{Z}^{J \times \bar{I}}$
where,
$S_{\mathcal{p}_1} \neq \cdots \neq S_{\mathcal{p}_X}$
$N(S_{\mathcal{p}_1}) = \cdots = N(S_{\mathcal{p}_X})$
$S_{\mathcal{p}_x} = \{\phi_{xy}, \phi_{yx}\}.(S_{\mathcal{p}_y})$
$\{\phi_{xy}, \phi_{yx}\} \in \mathbb{R} - \{0\}$
$x, y \in [1, X]$
$X \geq 2^{\min(\bar{I}_x)}$
$\bar{I}_x \in [6, N]$
$J \in [4, M]$
$\bar{I}_x \geq J + 2$
$M, N \in \mathbb{N}$
Generic properties include:
a. Biochemically relevant
b. Amenable to rigorous mathematical treatment
c. Can be perturbed
d. Readily parameterized with the probable dissociation constant (η) for every reaction of a biochemical network
e. Comparable
Algebraic properties for a set of CReaSED-matrices:
a. Semigroup w.r.t addition
b. Conditionally Abelian
c. Many-to-one with null space
CReaSED-matrices for a class of biochemical networks ($\mathfrak{S}(X = 8192, J = 11, \bar{I} = 13)$)
$S_{\mathcal{p}_1}$
$S_{\mathcal{p}_2}$
$S_{\mathcal{p}_3}$
$S_{\mathcal{p}_{8192}}$
$N(S_{\mathcal{p}_1})$
$N(S_{\mathcal{p}_2})$
$N(S_{\mathcal{p}_3})$
$N(S_{\mathcal{p}_{8192}})$
$S_{\mathcal{p}_2} = 2.S_{\mathcal{p}_1}$
$S_{\mathcal{p}_3} = 5.S_{\mathcal{p}_1}$
Perturbation and analysis for a class of biochemical networks:
$(\mathfrak{S}(\mathcal{P}), \tilde{\mathfrak{S}}(\mathcal{P})) = (\mathcal{p}_x, \tilde{\mathcal{p}}_x) \rightarrow (\mathcal{S}(\mathcal{p}_x), \mathcal{S}(\tilde{\mathcal{p}}_x))$
where,
$\bar{I}_{\mathcal{p}_x} = \bar{I}_{\tilde{\mathcal{p}}_x} + \rho = \bar{I}_x + \rho$
$\rho \geq 1$
$x = 1, \dots, X$
$\mathcal{p}_x \sim r_1 r_2 \dots r_{17}$
$\tilde{\mathcal{p}}_x \sim \tilde{r}_1 \tilde{r}_2 \dots \tilde{r}_{17} r_{18}$
r_{18}
$\begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_{17} \\ \tilde{\eta}_1 & \tilde{\eta}_2 & \cdots & \tilde{\eta}_{17} \end{bmatrix}$
$\{\omega_{\mathcal{p}_x}, \psi_{\mathcal{p}_x}\}$
$\bar{I}_{\mathcal{p}_x} = 17$
$\bar{I}_{\tilde{\mathcal{p}}_x} = \bar{I}_{\mathcal{p}_x} + 1 = 18$

Figure 3:

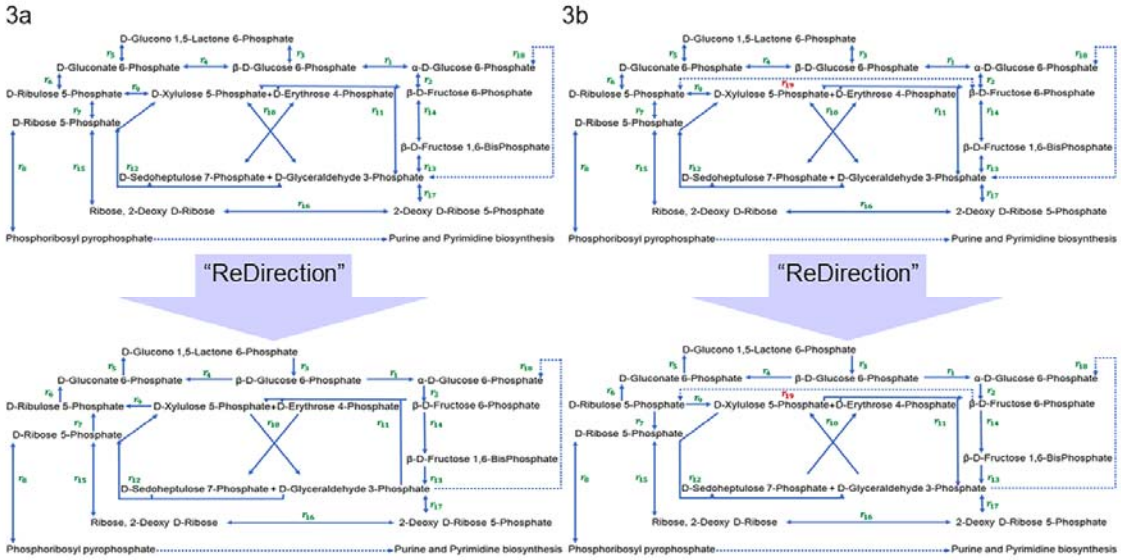

3a
3b
D-Glucono 1,5-Lactone 6-Phosphate
D-Gluconate 6-Phosphate
β-D-Glucose 6-Phosphate
α-D-Glucose 6-Phosphate
D-Ribulose 5-Phosphate
D-Xylulose 5-Phosphate+D-Erythrose 4-Phosphate
β-D-Fructose 6-Phosphate
D-Ribose 5-Phosphate
β-D-Fructose 1,6-BisPhosphate
D-Sedoheptulose 7-Phosphate + D-Glyceraldehyde 3-Phosphate
Ribose, 2-Deoxy D-Ribose
2-Deoxy D-Ribose 5-Phosphate
Phosphoribosyl pyrophosphate
Purine and Pyrimidine biosynthesis
"ReDirection"
"ReDirection"

Figure 4:

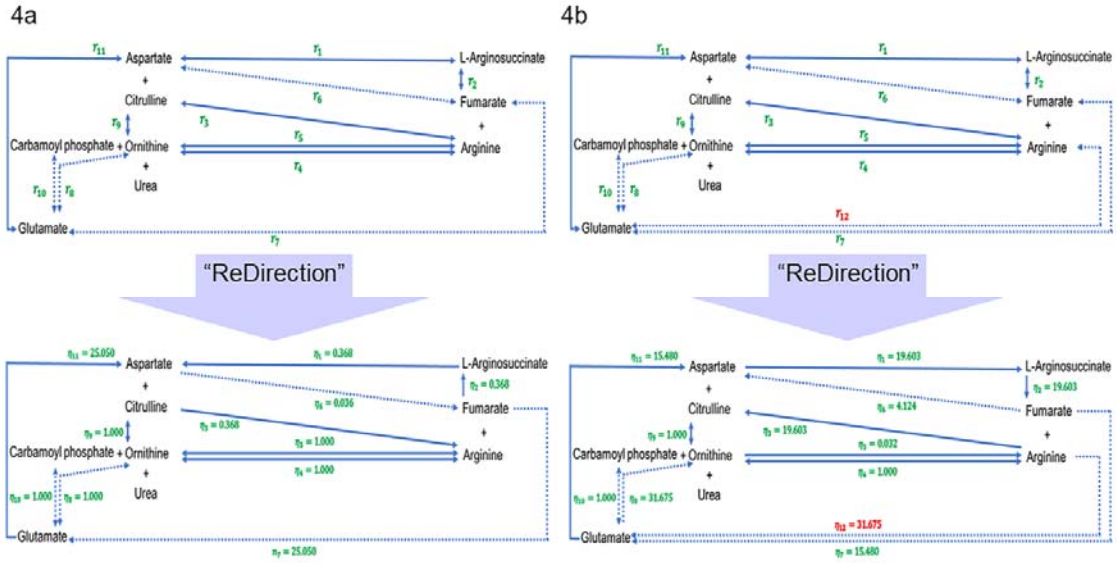
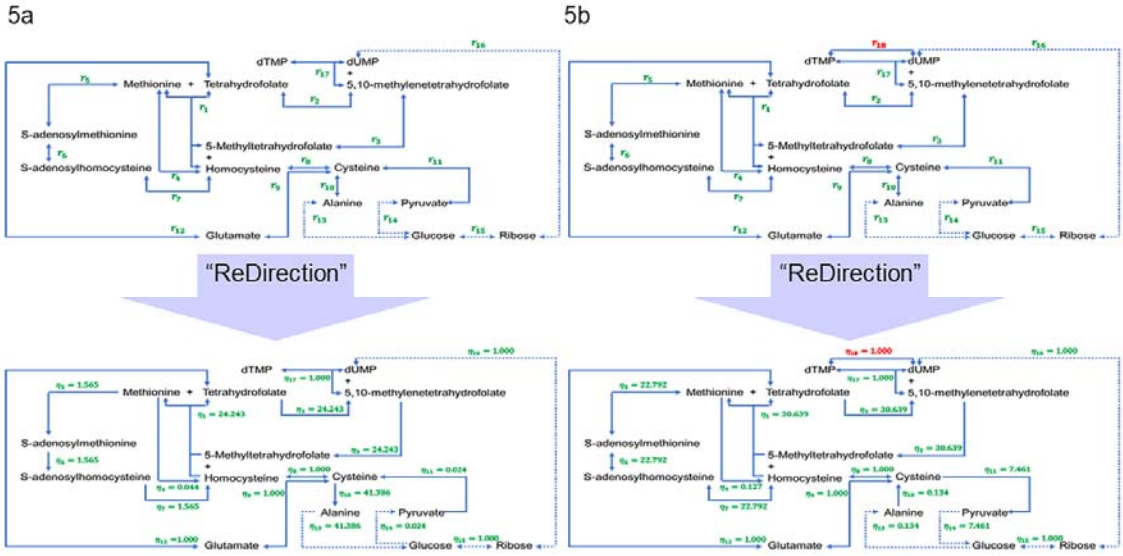

4a
4b
Aspartate
L-Arginosuccinate
Citrulline
Fumarate
Carbamoyl phosphate
Ornithine
Arginine
Urea
Glutamate
"ReDirection"
"ReDirection"

Figure 5:


5a
5b
dTMP
dUMP
5,10-methylenetetrahydrofolate
Methionine
Tetrahydrofolate
S-adenosylmethionine
5-Methyltetrahydrofolate
S-adenosylhomocysteine
Homocysteine
Cysteine
Alanine
Pyruvate
Glutamate
Glucose
Ribose
"ReDirection"
"ReDirection"