

SSI: A Statistical Sensitivity Index for Chemical Reaction Networks in cancer

GIORGIA BIDDAU*

MIDA, Dipartimento di Matematica, Università di Genova, Genova, Italy

GIACOMO CAVIGLIA

MIDA, Dipartimento di Matematica, Università di Genova, Genova, Italy

MICHELE PIANA

MIDA, Dipartimento di Matematica, Università di Genova, Genova, Italy

IRCCS Ospedale Policlinico San Martino, LISCOMP, Genova, Italy

SARA SOMMARIVA

MIDA, Dipartimento di Matematica, Università di Genova, Genova, Italy

SUMMARY

At the cellular level, cancer is triggered by mutations of the proteins involved in signalling networks made of hundreds of reacting species. The corresponding mathematical model consists of a large system of non-linear Ordinary Differential Equations for the unknown proteins concentrations depending on a consistently large number of kinetic parameters and initial concentrations. For this model, the present paper considers the problem of assessing the impact of each parameter and initial concentration on the system's output. More specifically, we introduced a statistical

*To whom correspondence should be addressed.

sensitivity index whose values can be easily computed by means of principal component analysis, and which leads to the partition of the parameters' and initial concentrations' sets into sensible and non-sensible families. This approach allows the identification of those kinetic parameters and initial concentrations that mostly impact the mutation-driven modification of the proteomic profile at equilibrium, and of those pathways in the network that are mostly affected by the presence of mutations in the cancer cell.

Key words: Chemical Reaction Networks; sensitivity analysis; principal component analysis; colon-rectal cancer; numerical simulations

1. INTRODUCTION

The mathematical modeling of chemical reactions networks (CRNs) may involve a large number of differential equations and parameters (Zi, 2011). By instance, a model recently proposed for the colon-rectal cancer (CRC) incorporates the available biological knowledge into a system of 419 proteins interacting through 850 chemical reactions with as many rate constants (Tortolina *and others*, 2015; Santra, 2018; Sommariva *and others*, 2021*a*). In this model, under the assumption that the chemical reactions rely on mass action kinetics, the network accordingly corresponds to a non linear system of 419 ordinary differential equations (ODEs), involving 419 protein concentrations as unknowns, and 850 rate constants as equations' coefficients (Sommariva *and others*, 2021*a*). Following an assignment of 419 initial conditions, the solution of the system of ODEs provides the simulated time course of the concentrations, and the (possibly allowed) asymptotic stationary state, to be identified with an equilibrium state. Numerical simulations may be used to predict changes in the model response resulting from varied external signals or internal conditions (Zi, 2011), to put into evidence feedback effects (Santra, 2018; Sommariva *and others*, 2021*b*), and to determine the impact of the introduction of targeted drugs (Santra, 2018; Krishnan *and*

others, 2020; Sommariva *and others*, 2021*b*).

It is well known that the results obtained from simulations depend on the values assigned to the parameters and to the initial conditions, which, in the case of the CRC CRN in the G1-S phase (Bartek and Lukas, 2001), were fixed on the basis of literature data (Tortolina *and others*, 2015). This fact opens a crucial issue about how to quantitatively assess the impact of the parameters' uncertainty on the CRC CRN equilibrium state. The present paper is devoted to a general analysis of this problem and illustrates a novel statistical tool for its solution arising from sensitivity analysis. In fact, our approach is specifically tailored to static conditions, which provides notable formal and computational advantages in contrast with the complexity of global dynamical approaches (Saltelli *and others*, 2000, 2005; Quaiser and Mönnigmann, 2009; Saltelli *and others*, 2010; Zi, 2011). Specifically, this analysis relies on two steps: an application of the implicit function theorem, whereby equilibrium states are obtained as functions of the parameters; and a principal component analysis (PCA), which leads to the distinction between sensible and non sensible rate constants and initial conditions via the definition of a statistical sensitivity index (SSI) weighing the relevance of either a specific constant or of the corresponding reaction. We point out that this second PCA-based step is inspired by an approach utilized by Liu *and others* (2005). However, differently than our case, that sensitivity index was introduced under dynamic conditions and, further, it suffered an intrinsic eigenvector sign ambiguity. The application of our analysis to the CRC CRN shows that SSI provides a reliable measure of the mutations' impact on the protein expression, and allows the identification of the reactions that are mostly triggered by those mutations.

The plan of the paper is as follows. Section 2 setups the formalism of the mathematical description of a CRN and introduces SSI. Section 3 shows some results concerned with the CRC CRN. Our conclusions are offered in Section 4.

2. METHODS

2.1 Equilibrium point for CRNs

We consider a CRN comprising r reactions involving n well-mixed chemical species and assume that the law of mass action kinetics holds. In the following we shall denote with $x_i(t)$, $i = 1, \dots, n$, the molar concentration (nM) at time t of the i -th species A_i , and with k_j , $j = 1, \dots, r$, the rate constant corresponding to the j -th reaction. The dynamic of the species concentration can be described by a system of n ordinary differential equations (ODEs) of the form

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}(\mathbf{x}, \mathbf{k}) \quad (2.1)$$

where $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}_+^n$, $\mathbf{k} = (k_1, \dots, k_r)^T \in \mathbb{R}_+^r$ is the vector of rate constants, $\mathbf{S} \in \mathbb{Z}^{n \times r}$ is the stoichiometric matrix, $\mathbf{v}(\mathbf{x}, \mathbf{k}) \in \mathbb{R}_+^r$ is the vector of reaction fluxes defined through the law of mass action, and the superimposed dot denotes the time derivative (Feinberg, 1987; Chellaboina *and others*, 2009; Yu and Craciun, 2018).

We further assume the CRN to be weakly elemented (Sommariva *and others*, 2021a), that is $p = n - \text{rank}(\mathbf{S})$ conservation laws exist described by the linear system

$$\mathbf{N}\mathbf{x} = \mathbf{c}, \quad (2.2)$$

where $\mathbf{c} \in \mathbb{R}_+^p$ is the vector of conservations laws' constants, and the matrix $\mathbf{N} \in \mathbb{N}^{p \times n}$ is determined by computing the kernel of \mathbf{S}^T and is assumed to contain a minor equal to the identity matrix of size p . Therefore, up to a change in the order of the components of \mathbf{x} , we can consider the decompositions

$$\mathbf{N} = \begin{bmatrix} \mathbf{I}_p & \mathbf{N}_2^{p \times (n-p)} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} \mathbf{S}_1^{p \times r} \\ \mathbf{S}_2^{(n-p) \times r} \end{bmatrix}, \quad (2.3)$$

where \mathbf{x}_1 is a p -vector of the so called elemental species (Shinar *and others*, 2009; Sommariva *and others*, 2021a) and $\mathbf{S}_2 \in \mathbb{R}^{(n-p) \times r}$ has rank $(n-p)$ and is the submatrix of \mathbf{S} corresponding to the non-elemental species.

It follows from the definition of conservation laws that $\mathbf{N}\mathbf{S} = 0$, whence $\mathbf{S}_1 = -\mathbf{N}_2\mathbf{S}_2$. It can be shown (Chellaboina *and others*, 2009; Berra *and others*, 2022) that the equilibrium (stationary) solution of the Cauchy problem is determined by the algebraic system

$$\begin{cases} \mathbf{S}_2\mathbf{v}(\mathbf{x}, \mathbf{k}) = 0 \\ \mathbf{N}\mathbf{x} - \mathbf{c} = 0 \end{cases}. \quad (2.4)$$

We observe that seeking the solution of this algebraic system is equivalent to the root-finding problem

$$\mathbf{f}(\mathbf{x}, \mathbf{k}, \mathbf{c}) = 0, \quad (2.5)$$

where $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ is a continuously differentiable function defined as

$$\mathbf{f}(\mathbf{x}, \mathbf{k}, \mathbf{c}) = \begin{bmatrix} \mathbf{S}_2\mathbf{v}(\mathbf{x}, \mathbf{k}) \\ \mathbf{N}\mathbf{x} - \mathbf{c} \end{bmatrix}. \quad (2.6)$$

In the case of the CRC CRN, to any fixed \mathbf{k} and \mathbf{c} there corresponds a unique equilibrium/stationary state \mathbf{x}^e , which may be determined by either a dynamic simulation (Sommariva *and others*, 2021a) or the solution of (2.4) (Berra *and others*, 2022). Moreover, since each reaction involves up to two reactants, from the law of mass action it follows that the components of \mathbf{f} are polynomials of degree two or lower.

2.2 Analytic computation of the local sensitivity matrix

Local sensitivity analysis consists in determining how small variations of a parameter input influence the output of a certain model (Goulet, 2016; Varma and Morbidelli, 1999). In this particular case, we are interested in studying the local sensitivity of an equilibrium/stationary state \mathbf{x}^e with respect to the parameters \mathbf{k} and \mathbf{c} by constructing the local sensitivity matrix. Given a weakly elemented CRN, we denote with \mathbf{k}^0 and \mathbf{c}^0 the vectors whose components are the values of the network rate constants and conservation laws' constants, respectively, when the network's equilibrium is a given point \mathbf{x}^e . Further, we define the vector $\mathbf{h}^0 = (\mathbf{k}^0, \mathbf{c}^0) \in \mathbb{R}^{r+p}$,

and we assume that

$$\det[(\nabla_{\mathbf{x}}\mathbf{f})(\mathbf{x}^e, \mathbf{h}^0)] \neq 0 \quad , \quad (2.7)$$

where henceforth, given a multivariate, vector-valued, continuously differentiable function \mathbf{g} , $\nabla_{\mathbf{z}}\mathbf{g}$ will denote the Jacobian matrix of \mathbf{g} with respect to the variable(s) \mathbf{z} . On account of the implicit function theorem (Krantz and Parks, 2002), equations (2.4) and (2.7) imply that there exists one and only one $\mathbf{x} = \mathbf{x}(\mathbf{h})$ such that

$$\mathbf{f}(\mathbf{x}(\mathbf{h}), \mathbf{h}) = 0 \quad (2.8)$$

for $\mathbf{h} = (\mathbf{k}, \mathbf{c})$ in a neighborhood of \mathbf{h}^0 , and thus $\mathbf{x}(\mathbf{h})$ is the equilibrium point of the CRN defined by the set of parameters \mathbf{h} . Hence, the local sensitivity matrix is defined as

$$\nabla_{\mathbf{h}}\mathbf{x}(\mathbf{h}^0) = \begin{bmatrix} \frac{\partial x_1}{\partial k_1}(\mathbf{h}^0) & \dots & \frac{\partial x_1}{\partial k_r}(\mathbf{h}^0) & \frac{\partial x_1}{\partial c_1}(\mathbf{h}^0) & \dots & \frac{\partial x_1}{\partial c_p}(\mathbf{h}^0) \\ \vdots & & \vdots & & & \\ \frac{\partial x_n}{\partial k_1}(\mathbf{h}^0) & \dots & \frac{\partial x_n}{\partial k_r}(\mathbf{h}^0) & \frac{\partial x_n}{\partial c_1}(\mathbf{h}^0) & \dots & \frac{\partial x_n}{\partial c_p}(\mathbf{h}^0) \end{bmatrix} \quad (2.9)$$

and can be computed by using

$$\nabla_{\mathbf{h}}\mathbf{x}(\mathbf{h}^0) = -[\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x}^e, \mathbf{h}^0)]^{-1} \nabla_{\mathbf{h}}\mathbf{f}(\mathbf{x}^e, \mathbf{h}^0). \quad (2.10)$$

Further, the definition (2.6) of \mathbf{f} implies that

$$\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x}^e, \mathbf{h}^0) = \begin{bmatrix} \mathbf{S}_2 \nabla_{\mathbf{x}}\mathbf{v}(\mathbf{x}^e, \mathbf{k}^0) \\ \mathbf{N} \end{bmatrix} \quad (2.11)$$

and

$$\nabla_{\mathbf{h}}\mathbf{f}(\mathbf{x}^e, \mathbf{h}^0) = \begin{bmatrix} \mathbf{S}_2 \nabla_{\mathbf{k}}\mathbf{v}(\mathbf{x}^e, \mathbf{k}^0) & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_p \end{bmatrix}. \quad (2.12)$$

The submatrix of (2.9) formed by the first r columns is the sensitivity matrix with respect to the rate constants, while the submatrix formed by the last p columns is the sensitivity matrix with respect to the conservation laws' constants.

2.3 Statistical Sensitivity Index

The first order Taylor expansion for the implicit function $\mathbf{x} = \mathbf{x}(\mathbf{h})$ leads to

$$\Delta \mathbf{x} = \nabla_{\mathbf{h}} \mathbf{x}(\mathbf{h}^0) \Delta \mathbf{h} + o(|\Delta \mathbf{h}|) \quad , \quad (2.13)$$

where

$$\Delta \mathbf{x} := \mathbf{x}(\mathbf{h}) - \mathbf{x}(\mathbf{h}^0) = \mathbf{x}(\mathbf{h}) - \mathbf{x}^e \quad (2.14)$$

and

$$\Delta \mathbf{h} := \mathbf{h} - \mathbf{h}^0 \quad , \quad (2.15)$$

while the $n \times (r+p)$ matrix $\nabla_{\mathbf{h}} \mathbf{x}(\mathbf{h}^0)$ is computed as in equation (2.10) and describes the variations of equilibrium points in terms of the variations of the set \mathbf{h} of parameters.

By using (2.13) at first order and by applying some straightforward algebraic computations we obtained

$$\hat{\Delta} \mathbf{x} = \mathbf{S} \hat{\Delta} \mathbf{h} \quad , \quad (2.16)$$

with

$$\hat{\Delta} \mathbf{x} = \left[\frac{\Delta x_1}{x_1^e}, \dots, \frac{\Delta x_n}{x_n^e} \right]^T, \quad \hat{\Delta} \mathbf{h} = \left[\frac{\Delta h_1}{h_1^0}, \dots, \frac{\Delta h_r}{h_{r+p}^0} \right]^T, \quad (2.17)$$

$$\mathbf{S} = (\mathbf{D}_{\mathbf{x}}^e)^{-1} \nabla_{\mathbf{h}} \mathbf{x}(\mathbf{h}^0) \mathbf{D}_{\mathbf{h}}^0 \quad , \quad (2.18)$$

and

$$\mathbf{D}_{\mathbf{x}}^e = \text{diag}[x_1^e, \dots, x_n^e], \quad \mathbf{D}_{\mathbf{h}}^0 = \text{diag}[h_1^0, \dots, h_{r+p}^0] \quad . \quad (2.19)$$

Following Liu *and others* (2005), we found that the function

$$Q(\hat{\Delta} \mathbf{h}) = \hat{\Delta} \mathbf{x}^T \hat{\Delta} \mathbf{x} \quad (2.20)$$

represents a measure of the relative deviation of concentrations resulting from a relative change in \mathbf{h} at equilibrium. Clearly, Q is a positive semi-definite quadratic form, so that

$$Q(\hat{\Delta} \mathbf{h}) = \hat{\Delta} \mathbf{h}^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \hat{\Delta} \mathbf{h} \quad , \quad (2.21)$$

where the matrices $\mathbf{\Lambda} = \text{diag}[\lambda_1, \dots, \lambda_{r+p}]$, with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{r+p} \geq 0$, and $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{r+p}]$ are formed by the eigenvalues and the corresponding normalized eigenvectors of $\mathbf{S}^T \mathbf{S}$, respectively. Equation (2.21) allows weighing the relevance of the j -th parameter, corresponding to the fractional change $\hat{\Delta}h_j$. Indeed, when $\hat{\Delta}\mathbf{h}$ has $\hat{\Delta}h_j$ as unique non-vanishing component, (2.21) leads to

$$Q(\hat{\Delta}h_j) = \sum_{a=1}^{r+p} \hat{\Delta}h_j u_{ja} \lambda_a u_{ja} \hat{\Delta}h_j \quad . \quad (2.22)$$

This equation inspires the definition of the j -th statistical sensitivity index (SSI)

$$e_j^{\mathbf{h}} = \frac{\sum_{a=1}^{r+p} \lambda_a u_{ja}^2}{\sum_{a=1}^{r+p} \lambda_a} \quad , \quad (2.23)$$

which is the average value of the squared components of the eigenvectors of $\mathbf{S}^T \mathbf{S}$ weighed by the corresponding eigenvalues, and whose values range in $[0, 1]$. Getting back to the original rate constant vector \mathbf{k} and conservation laws' vector \mathbf{c} , (2.23) naturally implies that $e_j^{\mathbf{k}}$ provides a synthetic and compact estimate of the weight of reaction j in the whole CRN, while $e_j^{\mathbf{c}}$ refers to the overall sensitivity of the network to the j -th component of vector \mathbf{c} . For illustrative purpose, the entire workflow for computing the SSI values is summarized in Figure 1.

2.3.1 Sensitivity analysis for a subnetwork As schematically illustrated in Figure 1, SSI can be used also for the sensitivity analysis of subnetworks of a CRN. Indeed, let $\mathcal{A} = \{A_{i_1}, \dots, A_{i_t}\}$, $t < n$, be the set of chemical species involved in a specific subnetwork. Then, the submatrix $\tilde{\mathbf{S}}$ of \mathbf{S} , formed by the t rows referred to \mathcal{A} , is the searched sensitivity matrix. From this, the diagonalization of $\tilde{\mathbf{S}}^T \tilde{\mathbf{S}}$ provides the SSI values referred to \mathcal{A} . Specifically, such SSI values allow discriminating the parameters of the network whose variation mainly affects the values of the concentrations at equilibrium for the subset \mathcal{A} of chemical species.

Remark: Let us consider the case where $t=1$, i.e. $\mathcal{A} = \{A_i\}$, and let \mathbf{s} be the i -th row of \mathbf{S} .

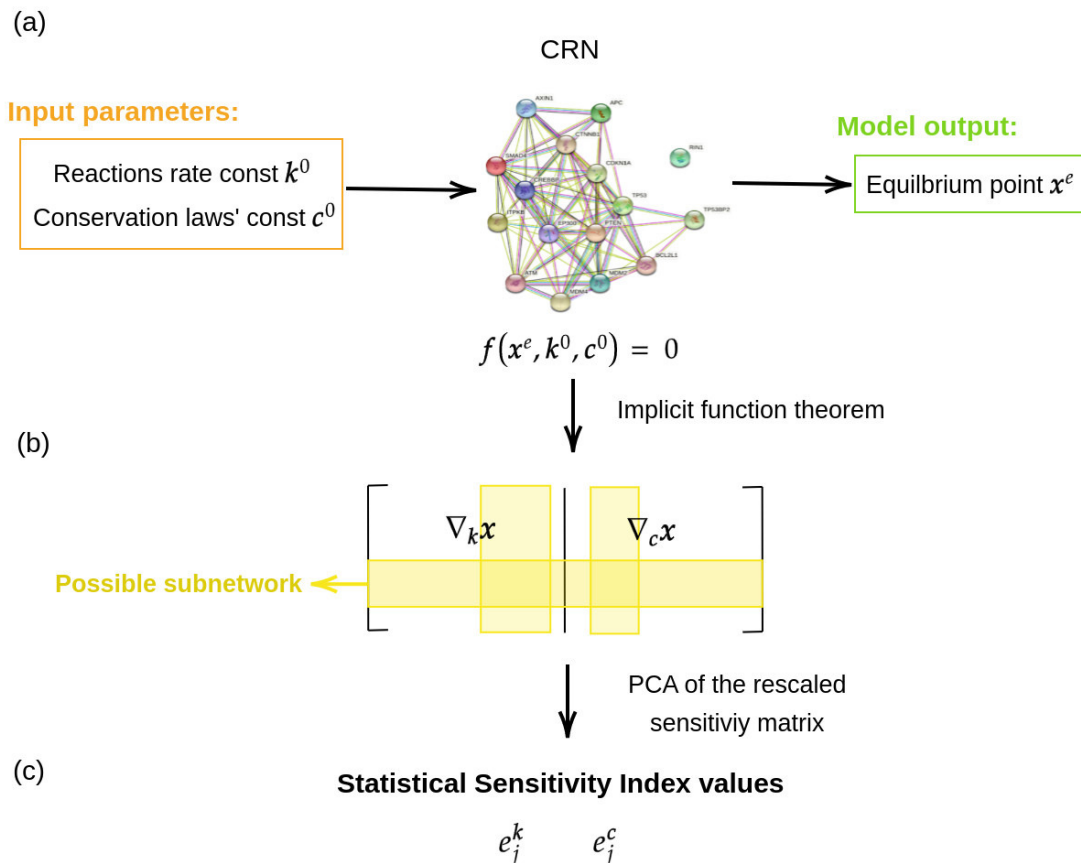


Fig. 1. Schematic representation of the workflow for computing the SSI: (a) sensitivity analysis allows the quantification of how small changes in the input parameters, namely the rate constants k^0 and the conservation laws' constants c^0 , impact the protein concentrations x^e at equilibrium; (b) by applying the implicit function theorem, the sensitivity matrix $\nabla_{\mathbf{h}} \mathbf{x} = [\nabla_{\mathbf{k}} \mathbf{x}, \nabla_{\mathbf{c}} \mathbf{x}]$ can be analytically computed; (c) SSI values are then determined by performing a PCA of the whole rescaled sensitivity matrix or of a proper submatrix (yellow bands) representing a given subnetwork.

Given the diagonalization $\mathbf{s}^T \mathbf{s} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$, then

$$\mathbf{s}_j^2 = (\mathbf{s}^T \mathbf{s})_{jj} = (\mathbf{U} \mathbf{\Lambda} \mathbf{U}^T)_{jj} = \sum_{a=1}^{r+p} \lambda_a u_{ja}^2. \quad (2.24)$$

Hence, in this case, the statistical sensitivity indices can be computed as

$$e_j^h = \frac{\mathbf{s}_j^2}{\sum_{a=1}^{r+p} \lambda_a}. \quad (2.25)$$

We notice that, since the denominator of equation (2.25) does not depend on j , the parameter

rankings determined by SSI and by the square of the components of \mathbf{s} coincide. This is coherent with the fact that the components of \mathbf{s} are the elements of the rescaled sensitivity matrix \mathcal{S} describing the relative sensitivity of the equilibrium concentrations of the species \mathcal{A} with respect to the parameters of the network.

2.3.2 Sensitivity analysis of a mutated CRN The mathematical description of a pathological CRN can be obtained by implementing the presence of mutations in the Cauchy problem associated to the healthy CRN. When the Loss of Function (LoF) mutation of an elemental species occurs (Sommariva *and others*, 2021b), the corresponding conservation law is null. It follows that all the species involved in the conservation law are null during time. Let us call \mathcal{A} this set of species and \mathcal{B} the set of reactions where these species are involved and that consequently do not occur in the network. The species in \mathcal{A} and the reactions in \mathcal{B} should not be considered in the matrices and vectors describing the network. Therefore, the rows corresponding to \mathcal{A} and the columns corresponding to \mathcal{B} of the stoichiometric matrix \mathcal{S} shall be eliminated, such as the entries of the flux vector \mathbf{v} referred to \mathcal{B} . Further, the row referred to the elemental specie must be eliminated both from the moiety vector \mathbf{c} and the conservation laws matrix, whose columns associated to \mathcal{A} shall be removed as well.

The Gain of Function (GoF) mutation of a specific species implies that a set of reactions involving it do not occur in the network any longer (Sommariva *and others*, 2021b). Let us define with \mathcal{B} the set of these reactions. All the chemical species that are involved only in the reactions of \mathcal{B} , as either reactants or products, shall be inserted in a set \mathcal{A} of species to be eliminated from the network. In this case, only the physiological stoichiometric matrix and the fluxes vector are modified from the physiological system. The columns of \mathcal{S} and the entries of \mathbf{v} referred to the reactions in \mathcal{B} are eliminated, such as the rows of \mathcal{S} corresponding to the species in \mathcal{A} .

3. RESULTS

SSI has been defined in a general setting and therefore can be utilized for sensitivity analysis for all possible CRNs. However, here we focused on applications concerned with the colon-rectal CRN in both healthy and mutated conditions.

3.1 *Sensitivity analysis for the physiological colon-rectal CRN*

We first computed the SSI values in the case of the healthy colon-rectal CRN. As said, in the G1-S phase this network comprises $n = 419$ proteins involved in $r=850$ chemical reactions and $p=81$ independent conservation laws. We first pointed out the components of \mathbf{k} and \mathbf{c} characterized by the maximum, median, and minimal SSI values, respectively. The results of this computation are presented in Figure 2 and clearly show that these components are, respectively, the ones with highest, median, and lowest global sensitivity, this latter being assessed as the Euclidean norm of $\hat{\Delta}\mathbf{x}$.

As said in subsection 2.3.1, one of the nice aspects of SSI is that it allows a reliable estimate of the sensitivity of specific subnetworks of the CRN. For example, Figure 3 shows the results of the same analysis as in Figure 2, but this time focused on the sensitivity of the subnetwork involving gene TP53 (Iacopetta, 2003), which is made by a single equation. Also in this case, the SSI values are nice indicators of the sensitivity level associated to each component of the conservation laws' vector \mathbf{c} . Even more interestingly, in this case the conservation law associated with the maximum SSI value is the one associated to MDM2, which is in fact directly involved in the degradation of TP53 (Sommariva *and others*, 2021*b*). Coherently, a further analysis not shown here demonstrated that TP53 concentration is reduced by increasing the component of \mathbf{c} corresponding to the MDM2 conservation law.

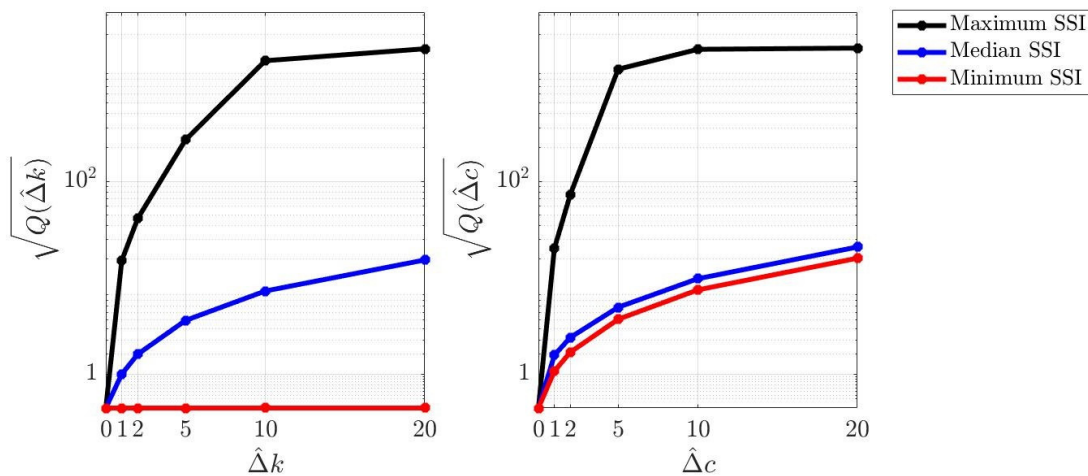


Fig. 2. Global sensitivity and SSI values in the physiological colon-rectal CRN. Left panel: $\sqrt{Q(\hat{\Delta}k)}$ as a function of the relative variation of the rate constants with minimum, median, and maximum SSI values. Right panel: $\sqrt{Q(\hat{\Delta}c)}$ as a function of the relative variation of the conservation laws' constants with minimum, median, and maximum SSI values.

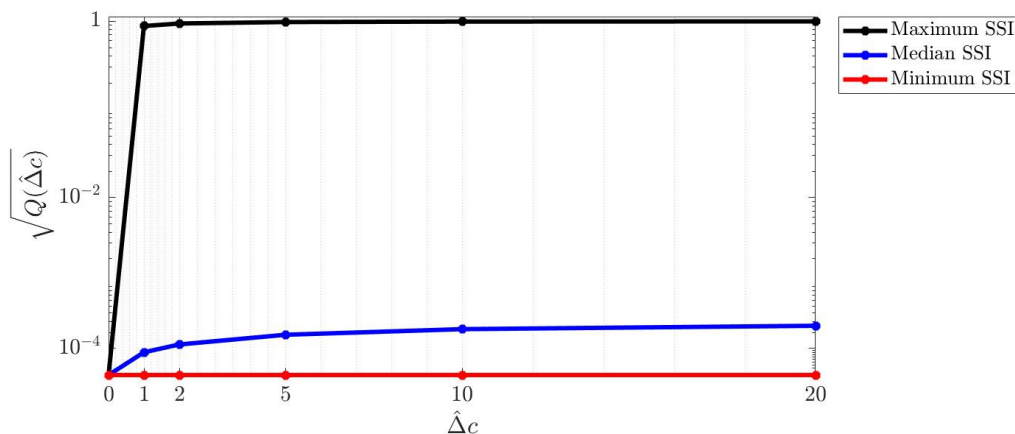


Fig. 3. Sensitivity of the equilibrium concentration of TP53: $\sqrt{Q(\hat{\Delta}c)}$ for the TP53 subnetwork is represented versus the relative variation of the conservation laws' constants with minimum, median, and maximum SSI values.

3.2 *Sensitivity analysis for the CRC CRN*

A greater significance of our approach to sensitivity analysis is concerned with applications to mutated CRN. Indeed, in this context, SSI plays two roles.

As the protein concentration at equilibrium, SSI is an indicator of the impact that each mutation has on the expression of that specific protein. This role is clearly shown by Figure 4 and Figure 5, which represent the SSI values for all components of the rate constant (Figure 4) and conservation laws' (Figure 5) vectors associated to the colon-rectal CRN when four CRC mutations are included in the network. These figures indicate that the KRAS GoF mutation strongly modifies the SSI values of most indexes and this is coherent with the fact that the concentration values at equilibrium are significantly altered by the presence of such GoF mutation in the network (Sommariva *and others*, 2021b).

More than this, SSI also allows the identification of the specific reactions in the network that were mostly affected by the mutation. This second role of SSI is pointed out by Table 1 that contains a list of reaction ranked according to the values of the relative difference between the SSI values in the physiological and mutated cases (the first column of the table contains the reference numbers attributed to the reactions in the supplementary material of Tortolina *and others* (2015)). All these reactions are related to the signalling pathways of the Ras protein, which is involved in the most important signalling cascades regulating cell growth, proliferation, and survival for many cancer types (Orton *and others*, 2005; Levine *and others*, 2019). Via R41, free Ras is freezed in the inactive chemical compound *Ras_GDT*, while R52 leads to a reduction of the active form *Ras_GTP*, which is expected to be available in large amounts because of the mutation; the same compound of R52 is generated in reaction R49. Further, among the three reactions R294, R297, and R302, only the last one involves Ras explicitly, via the compound *Ras_GDP*. The other two reactions deal with the same compound *ERBP_ShP_G_S* but, more importantly, they are deeply involved in the process of signal transmission from the activation

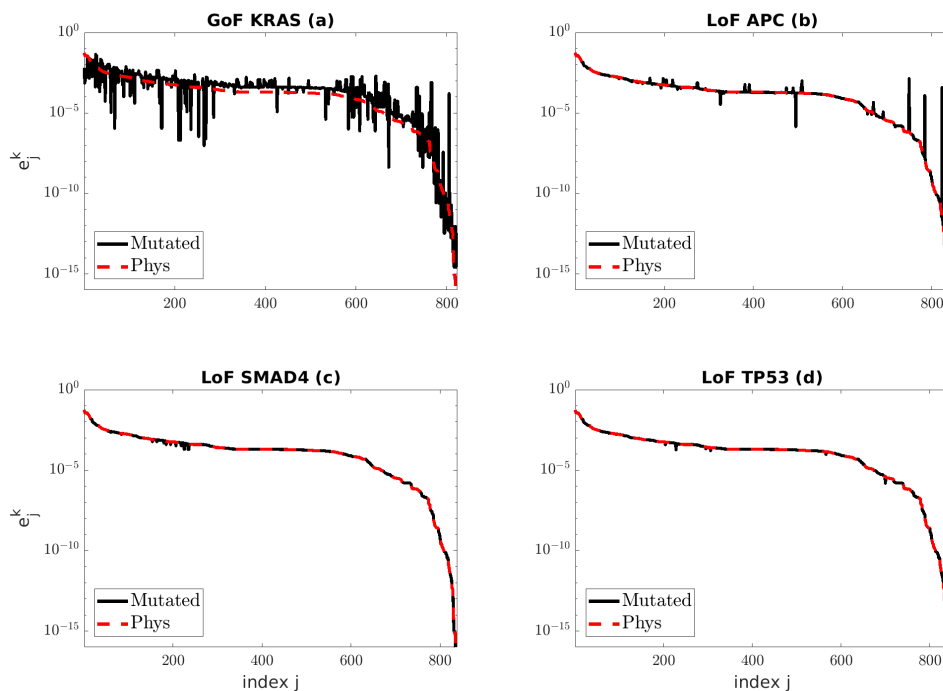


Fig. 4. SSI values for all rate constants in the case of four mutated networks (black solid line) with respect to the SSI values of the physiological CRN (red dashed line).

of receptors on the cell membrane to the molecules of Ras. Similar remarks hold true for the remaining reactions R412, R415, R420 (see the Supplementary Material in Tortolina *and others* (2015)).

4. DISCUSSION AND CONCLUSIONS

CRNs are typically characterized by a huge number of variables, and although static formulations reduce the number of parameters in the game, yet calibration of a CRN is often a challenging issue. Sensitivity analysis of both physiological and mutated CRNs is an effective way to reduce the complexity of the network, since just sensitive rate and conservation laws' constants should be regarded as parameters to finely tune, while insensitive variables can be coarsely approximated.

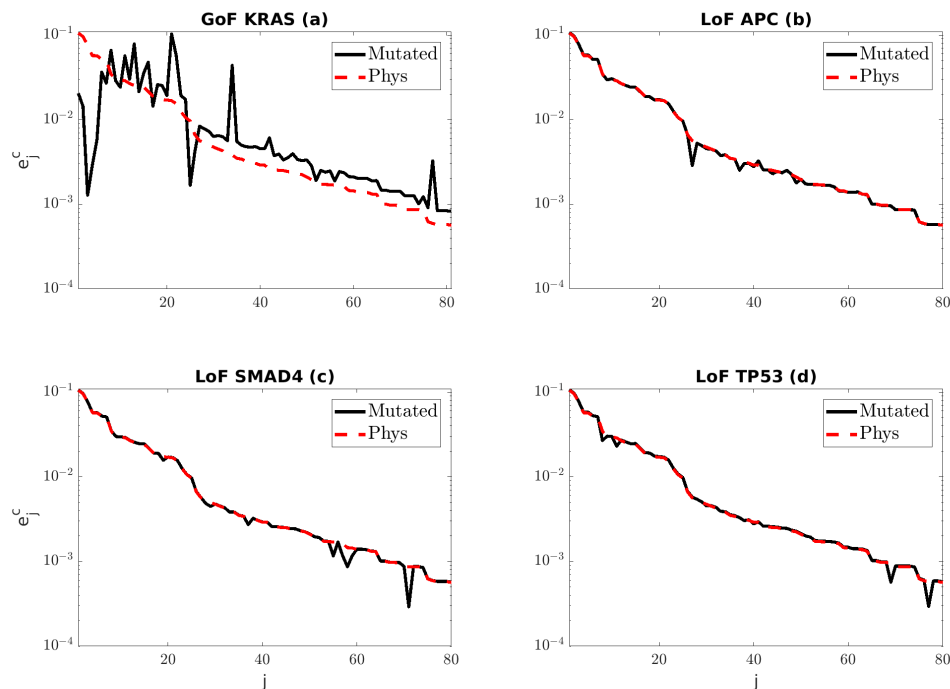


Fig. 5. SSI values for all conservation laws' constants in the case of four mutated networks (black solid line) with respect to the SSI values of the physiological CRN (red dashed line).

Table 1. Reactions whose sensitivity indexes are mostly effected by the GoF mutation of KRAS. The third column contains the relative differences of e_j^k between the mutated and the physiological CRC CRNs.

Reference #	Reaction	δe_j^k
R41	Ras + GDP -> Ras_GDP	1.91e + 07
R294	ERBP_ShP_G_S -> ERBP_ShP_G + SOS	2.91e + 04
R49	RP_ShP_G_S_Ras + GTP -> RP_ShP_G_S_Ras_GTP	2.88e + 04
R412	ERB3P_Sh -> ERB3P + Shc	2.86e + 04
R415	ShP + ERB3P -> ERB3P_ShP	1.23e + 04
R297	ERBP_ShP + G_S -> ERBP_ShP_G_S	8.64e + 03
R302	ERBP_ShP_G_S_Ras + GDP -> ERBP_ShP_G_S_Ras_GDP	5.78e + 02
R57	RP_G_S_Ras_GDP -> RP_G_S_Ras + GDP	5.78e + 02
R420	ERB3P_ShP_G + SOS -> ERB3P_ShP_G_S	5.78e + 02
R52	RP_ShP_G_S + Ras_GTP -> RP_ShP_G_S_Ras_GTP	4.95e + 02

In our view, the sensitivity index introduced in this paper has several advantages. It is specifically tailored to the static case and it is based on PCA, which makes its computation rather straightforward. From a technical viewpoint, it does not suffer any intrinsic sign ambiguity, as it occurs in the case of the index introduced by Liu *and others* (2005). Further, it can be applied for the sensitivity analysis of a specific single species, of a specific pathway, and of the whole network; and, also, it can be used to nicely rank the reactions with respect to the impact that a specific mutation has on the signalling pathways. Finally, although we have validated SSI in the case of the CRC CRN, this index has a more general value, and can be utilized for assessing the equilibrium condition of any mutated signalling network.

As far as future work is concerned, the next step will be to use SSI to select the sensitive pathways within the global CRC CRN, and then to calibrate the simplified network by applying computational algorithms based on inverse problems theory (Bertero and Piana, 2006) to experimental measurements of the proteins' concentrations at equilibrium.

ACKNOWLEDGEMENTS

The support of the Ecosistema 'RAISE – Robotics and AI for Socio-economic Empowerment' (PNRR - Missione 4 Componente 2 Investimento 1.5) is kindly acknowledged. S.S. and M.P. have been partially supported by Gruppo Nazionale per il Calcolo Scientifico (GNCS-INdAM)

SOFTWARE

Software in the form of Matlab code is available at https://github.com/theMIDAgrouP/CRN_sensitivity

REFERENCES

17

REFERENCES

- BARTEK, JIRI AND LUKAS, JIRI. (2001). Pathways governing g1/s transition and their response to dna damage. *FEBS letters* **490**(3), 117–122.
- BERRA, SILVIA, LA TORRACA, ALESSANDRO, BENVENUTO, FEDERICO AND SOMMARIVA, SARA. (2022). A fast and convergent combined newton and gradient descent method for computing steady states of chemical reaction networks. Preprint at <https://arxiv.org/abs/2212.14252>.
- BERTERO, M AND PIANA, M. (2006). Inverse problems in biomedical imaging: modeling and methods of solution. In: *Complex systems in biomedicine*. Springer, pp. 1–33.
- CHELLABOINA, VIJAYSEKHAR, BHAT, SANJAY P, HADDAD, WASSIM M AND BERNSTEIN, DENNIS S. (2009). Modeling and analysis of mass-action kinetics. *IEEE Control Systems Mag* **29**(4), 60–78.
- FEINBERG, MARTIN. (1987). Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chem Engin Sci* **42**(10), 2229–2268.
- GOULET, DAVE. (2016). Modeling, simulating, and parameter fitting of biochemical kinetic experiments. *siam REVIEW* **58**(2), 331–353.
- IACOPETTA, BARRY. (2003). Tp53 mutation in colorectal cancer. *Human mutation* **21**(3), 271–276.
- KRANTZ, STEVEN GEORGE AND PARKS, HAROLD R. (2002). *The implicit function theorem: history, theory, and applications*. Springer Science & Business Media.
- KRISHNAN, JEYASHREE, TORABI, REZA, SCHUPPERT, ANDREAS AND NAPOLI, EDOARDO DI. (2020). A modified ising model of barabási–albert network with gene-type spins. *Journal of mathematical biology* **81**(3), 769–798.

- LEVINE, ARNOLD J, JENKINS, NANCY A AND COPELAND, NEAL G. (2019). The roles of initiating truncal mutations in human cancers: the order of mutations and tumor cell type matters. *Cancer cell* **35**(1), 10–15.
- LIU, GANG, SWIHART, MARK T AND NEELAMEGHAM, SRIRAM. (2005). Sensitivity, principal component and flux analysis applied to signal transduction: the case of epidermal growth factor mediated signaling. *Bioinformatics* **21**(7), 1194–1202.
- ORTON, RICHARD J, STURM, OLIVER E, VYSHEMIRSKY, VLADISLAV, CALDER, MUFFY, GILBERT, DAVID R AND KOLCH, WALTER. (2005). Computational modelling of the receptor-tyrosine-kinase-activated mapk pathway. *Biochemical Journal* **392**(2), 249–261.
- QUAISER, TOM AND MÖNNIGMANN, MARTIN. (2009). Systematic identifiability testing for unambiguous mechanistic modeling—application to jak-stat, map kinase, and nf- κ b signaling pathway models. *BMC systems biology* **3**(1), 1–21.
- SALTELLI, ANDREA, ANNONI, PAOLA, AZZINI, IVANO, CAMPOLONGO, FRANCESCA, RATTO, MARCO AND TARANTOLA, STEFANO. (2010). Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer physics communications* **181**(2), 259–270.
- SALTELLI, ANDREA, RATTO, MARCO, TARANTOLA, STEFANO AND CAMPOLONGO, FRANCESCA. (2005). Sensitivity analysis for chemical models. *Chemical reviews* **105**(7), 2811–2828.
- SALTELLI, ANDREA, TARANTOLA, STEFANO AND CAMPOLONGO, FRANCESCA. (2000). Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 377–395.
- SANTRA, TAPESH. (2018). Fitting mathematical models of biochemical pathways to steady state perturbation response data without simulating perturbation experiments. *Scientific Reports* **8**(1), 1–14.

REFERENCES

19

- SHINAR, GUY, ALON, URI AND FEINBERG, MARTIN. (2009). Sensitivity and robustness in chemical reaction networks. *Siam J. Appl. Math.* **69**(4), 977–998.
- SOMMARIVA, SARA, CAVIGLIA, GIACOMO AND PIANA, MICHELE. (2021*a*). Gain and loss of function mutations in biological chemical reaction networks: a mathematical model with application to colorectal cancer cells. *Journal of Mathematical Biology* **82**(6), 1–25.
- SOMMARIVA, SARA, CAVIGLIA, GIACOMO, RAVERA, SILVIA, FRASSONI, FRANCESCO, BENVENUTO, FEDERICO, TORTOLINA, LORENZO, CASTAGNINO, NICOLETTA, PARODI, SILVIO AND PIANA, MICHELE. (2021*b*). Computational quantification of global effects induced by mutations and drugs in signaling networks of colorectal cancer cells. *Scientific reports* **11**(1), 1–13.
- TORTOLINA, LORENZO, DUFFY, DAVID J, MAFFEI, MASSIMO, CASTAGNINO, NICOLETTA, CARMODY, AIMÉE M, KOLCH, WALTER, KHOLODENKO, BORIS N, DE AMBROSI, CRISTINA, BARLA, ANNALISA, BIGANZOLI, ELIA M *and others*. (2015). Advances in dynamic modeling of colorectal cancer signaling-network regions, a path toward targeted therapies. *Oncotarget* **6**(7), 5041.
- VARMA, ARVIND AND MORBIDELLI, MASSIMO. (1999). *Parametric sensitivity in chemical systems*. Cambridge University Press.
- YU, POLLY Y AND CRACIUN, GHEORGHE. (2018). Mathematical analysis of chemical reaction systems. *Isr. J. Chem.* **58**(6-7), 733–741.
- ZI, ZHIKE. (2011). Sensitivity analysis approaches applied to systems biology models. *IET systems biology* **5**(6), 336–346.