scientific reports

Check for updates

OPEN

PCA-based synthetic sensitivity coefficients for chemical reaction network in cancer

Giorgia Biddau[1]✉, Giacomo Caviglia[1], Michele Piana[1,2] & Sara Sommariva[1]

Chemical reaction networks are powerful tools for modeling cell signaling and its disruptions in diseases like cancer. Realistic chemical reaction networks involve hundreds of proteins and reactions, resulting in a model depending on a consistently large number of kinetic parameters. Since finely calibrating all the parameters would require an unrealistic amount of data, proper sensitivity analysis is required to identify a subset of parameters for which fine tuning is needed and thus provide a fundamental tool for the qualitative analysis of the network. We present a multidisciplinary approach for computing a set of synthetic sensitivity indices. These indices rank the kinetic parameters, based on the impact that errors in their values would have on the protein concentration profile at equilibrium. Our tests on a chemical reaction network devised for colorectal cells demonstrate the effectiveness of the considered sensitivity indices in different scenarios including in-silico drug dosage and novel therapeutic target discovery. The Matlab code for computing the synthetic sensitivity indices and the data concerning the network for colorectal cells are available at https://github.com/theMIDAgroup/CRN_sensitivity.

Keywords Chemical reaction networks, In-silico simulation, Colorectal cancer, Principal component analysis, Sensitivity analysis

The mathematical modeling of chemical reactions networks (CRNs) may involve a large number of differential equations and parameters[1]. By instance, a model recently proposed for the G1-S phase of a colorectal (CR) cell incorporates the available biological knowledge into a system of 419 proteins and protein complexes interacting through 850 chemical reactions with as many rate constants[2–4]. In this model, under the assumption of the validity of 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[4]. 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[1], to put into evidence feedback effects[3,5], and to determine the impact of the introduction of targeted drugs[3,5,6].

Previous works[5–7] have demonstrated the strength of mathematical models grounded on CRNs in providing insights on the complex mechanism behind cancer onset and progression. In fact, most human cancers are highly heterogeneous and dynamic diseases that appear after accumulation of several genetic mutations in the cell[8,9] . For example, it has been estimated that five to ten tumor-specific driver mutations usually concur in colorectal cancer (CRC), and the most frequent and dangerous alterations usually pertain the genes TP53, APC, KRAS, PTEN, SMAD4, PIK3CA, BRAF, and AKT[2,10] . Modern CRC therapies often include the use of one or multiple targeted drugs which aim at facing cancer progression by inhibiting or expanding the effects of specific genes somehow involved in the genetic alteration underlying cancer growth[11,12] . From a mathematical point of view, it has been shown that global effects induced on the network by partial or complete mutations and by (combined) targeted therapies can be quantified by properly modifying the CR-CRN and then comparing the equilibrium states of the original and the mutated networks[5–7] . Furthermore, a fast optimization approach has been introduced[13] for efficiently computing such equilibria without simulating the whole network dynamics, resulting in a fast algorithm which is a first crucial ingredient towards the ultimate aim of tailoring the network to a personalized therapy. Focusing on the equilibrium states rather than on the whole network dynamics has obvious limitations. Despite this the analysis of steady-state perturbation response data has gained an increasing role in the design of

[1]MIDA, Dipartimento di Matematica, Dipartimento di Eccellenza 2023-2027, Università di Genova, Genova, Italy. [2]IRCCS Ospedale Policlinico San Martino, LISCOMP, Genova, Italy. ✉email: giorgia.biddau@dima.unige.it

novel targeted therapies[3,14,15] . In the context of simulating and comparing the effects of different mutations and different targeted therapies for cancer, the choice of focusing on the steady states of the network is motivated by the timescale of the considered biological processes. In fact, it is reasonable to assume that the involved mechanisms of cell signaling are almost instantaneous when compared to other cellular processes concerning, e.g., the production of large proteins or cell duplication. As a consequence, the concentration of the proteins involved in these signaling networks can be regarded as at or near equilibrium[16] .

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 CR-CRN in the G1-S phase[17], were fixed on the basis of literature data[2]. This fact opens a crucial issue about how to quantitatively assess the impact of the parameters' uncertainty on the equilibrium state of a CRN. A second, somehow associated problem concerns how to properly design the numerical simulations and interpret the corresponding (parameter dependent) results in order to identify possible molecular targets of drug's action. The present paper is devoted to a general analysis of these problems, and illustrates a multidisciplinary approach for their solution arising from sensitivity analysis. Since we are interested in the network equilibrium, we focused on a local approach specifically tailored to static conditions in order to avoid the complexity of global dynamical approaches[1,18–21], most of which cannot be easily adapted to our model, which includes several hundreds of state variables. Specifically, our approach relies on two steps: an application of the implicit function theorem in order to analytically compute the local sensitivity matrix of the equilibrium states with respect to the model parameters[22]; 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 set of synthetic indices, henceforth denoted as statistical sensitivity indices (SSIs), weighing the relevance of either a specific constant or of the corresponding reaction. This two–step approach was inspired by a previous work[23] , where a local sensitivity matrix was computed at each time-point of the network dynamics by solving through numerical approximation a dedicated ODEs system. Then PCA was performed on these time–varying sensitivity matrices. Two are the main differences of our approach with respect to that proposed by Liu and colleagues: (i) as we shall see later in the manuscript we provide a more statistically accurate definition of the sensitivity indices that is not affected by the ambiguity of the sign of the eigenvectors computed through PCA; (ii) by focusing on the equilibrium state rather than on the whole network dynamics we are able to analytically compute the sensitivity matrix thus providing more accurate results while keeping low the computational cost of the algorithm. This made it possible to apply our approach to a more complex network than the EGF signaling network considered by Liu and colleagues[24] which involves only 94 species interacting in 125 reversible reactions. The application of our analysis to CR-CRN shows that the SSIs provide 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. Further, the SSIs may play a crucial role in the identification of novel therapeutic targets and in the design of in-silico simulations for drug dosage.

Results

Synthetic sensitivity indices for CRNs

The present work focuses on a CRN devised for modeling signal transduction in a colorectal cell during its G1–S phase[2,5]. The network, henceforth denoted with CR-CRN, comprises $n = 419$ proteins and protein complexes, denoted as $A_1, \ldots, A_n$, involved in $r = 850$ chemical reactions. The list of the abbreviations used for some of the proteins relevant for this work is provided in Supplementary Table S2 . As detailed in the Methods section, by applying the law of mass action, the dynamic of the network can be described by the following system of 419 ordinary differential equations (ODEs)

$$\dot{\boldsymbol{x}} = \boldsymbol{S}\boldsymbol{v}(\boldsymbol{x}, \boldsymbol{k}) \tag{1}$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)^T \in R_+^n$ are the molar concentration (nM) of the chemical species, $\boldsymbol{k} = (k_1, \ldots, k_r)^T \in R_+^r$ is the vector of reaction rate constants, $\boldsymbol{S} \in Z^{n \times r}$ is the stoichiometric matrix, $\boldsymbol{v}(\boldsymbol{x}, \boldsymbol{k}) \in R_+^r$ is the vector of reaction fluxes defined through the law of mass action, and the superimposed dot denotes time derivative[25–27]. By studying the null space of $\boldsymbol{S}$, it can be shown that the proteins within the CR-CRN are involved in $p = 81$ independent semi-positive conservation laws[28], . Each conservation law identifies a group of proteins whose combined concentrations remain fixed during the network dynamics, while the amount of the single proteins may change (see Supplementary Note S1 for further details). In the following, we will denote with $\boldsymbol{c} = (c_1, \ldots, c_p)^T$ the set of total combined concentrations associated with the conservation laws. In a previous work[4] it has been conjectured the CR-CRN satisfies the global stability condition, that is the network has a unique asymptotically stable equilibrium $\boldsymbol{x}^*$ for any fixed $\boldsymbol{k}^*$ and $\boldsymbol{c}^*$. Furthermore, it has been shown that the effects of two classes of mutations commonly found in cancer and resulting in the gain or loss of function of some of the network proteins can be simulated by changing the values of some of the rate constants or by reducing the total combined concentration within some conservation laws, respectively.

Local sensitivity analysis consists in determining how the small variations of a parameter input influence the output of a certain model[29,30]. In this particular case, we are interested in studying the local sensitivity of an equilibrium/stationary state $\boldsymbol{x}^*$ with respect to perturbation of the rate constants $\boldsymbol{k}^*$ and the conservation laws' constants $\boldsymbol{c}^*$. Towards this end, we propose an algorithm based on the following two steps thoroughly described in the Methods section and summarized in Fig. 1.

1. Analytic computation of the local sensitivity matrix

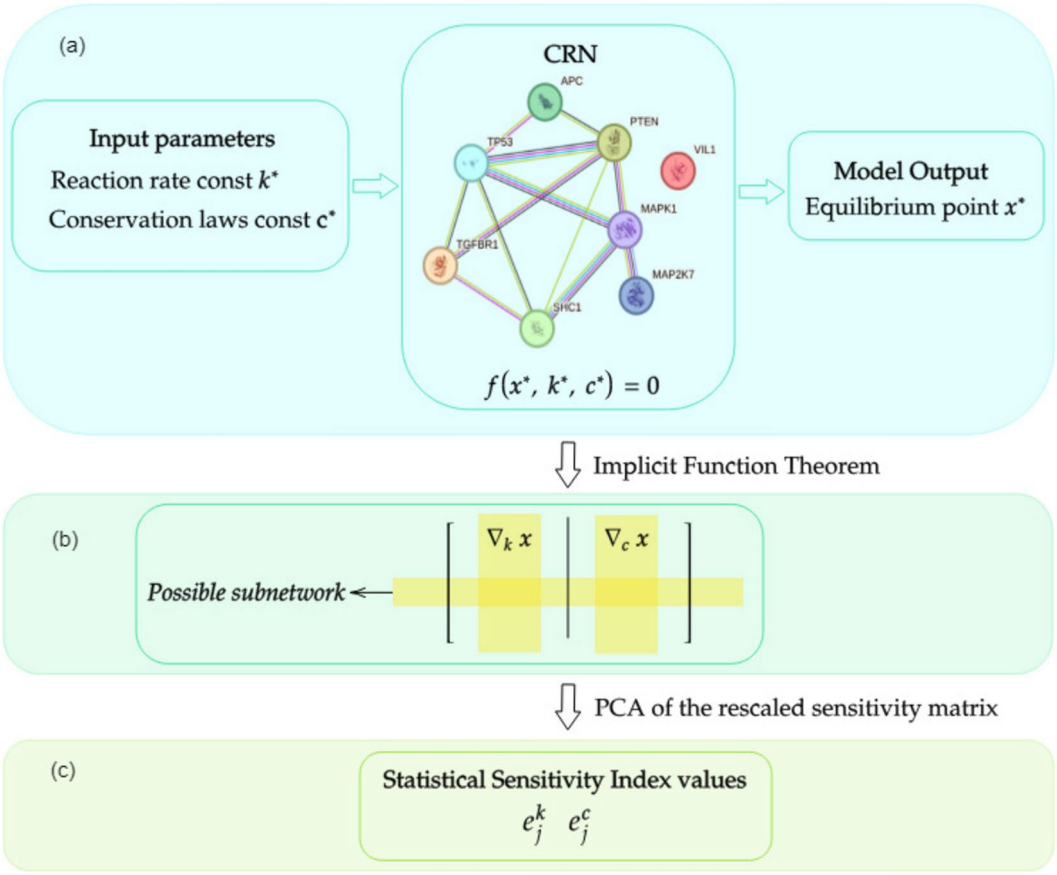

Figure 1. Schematic representation of the workflow for computing the SSIs. (**a**) Sensitivity analysis allows the quantification of how small changes in the input parameters, namely the rate constants $\boldsymbol{k}^*$ and the conservation laws' constants $\boldsymbol{c}^*$, impact the protein concentrations $\boldsymbol{x}^*$ at equilibrium. (**b**) By applying the implicit function theorem, the sensitivity matrix $\nabla_{\boldsymbol{h}}\boldsymbol{x} = [\nabla_{\boldsymbol{k}}\boldsymbol{x}, \nabla_{\boldsymbol{c}}\boldsymbol{x}]$ 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.

$$\nabla_{\boldsymbol{h}}\boldsymbol{x}(\boldsymbol{h}^*) := \begin{bmatrix} \frac{\partial x_1}{\partial k_1}(\boldsymbol{h}^*) & \dots & \frac{\partial x_1}{\partial k_r}(\boldsymbol{h}^*) & \frac{\partial x_1}{\partial c_1}(\boldsymbol{h}^*) & \dots & \frac{\partial x_1}{\partial c_p}(\boldsymbol{h}^*) \\ \vdots & & \vdots & & & \\ \frac{\partial x_n}{\partial k_1}(\boldsymbol{h}^*) & \dots & \frac{\partial x_n}{\partial k_r}(\boldsymbol{h}^*) & \frac{\partial x_n}{\partial c_1}(\boldsymbol{h}^*) & \dots & \frac{\partial x_n}{\partial c_p}(\boldsymbol{h}^*) \end{bmatrix}, \tag{2}$$

where $\boldsymbol{h}^* = (\boldsymbol{k}^*, \boldsymbol{c}^*)$. We observe that $\nabla_{\boldsymbol{h}}\boldsymbol{x}(\boldsymbol{h}^*)$ is a matrix of size $n \times (r+p)$ where the first r columns define the sensitivity matrix with respect to $\boldsymbol{k}^*$, while the last p columns define the sensitivity matrix with respect to $\boldsymbol{c}^*$.

2. Computation of the SSIs by applying PCA to a scaled version of $\nabla_{\boldsymbol{h}}\boldsymbol{x}(\boldsymbol{h}^*)$. PCA can be performed on the whole matrix, thus accounting for the effects of all the parameters on the overall equilibrium point of the network. However, in many practical scenarios, one may be interested in investigating the effects concerning specific proteins, acting e.g. as tumor suppressors, or groups of proteins involved in specific pathways, such as e.g. the mitogen-activated protein kinase (MAPK) pathway which plays a relevant role in the onset and progression of many cancers, including colorectal cancer[7,31] . In this case, our approach can be applied to a proper submatrixof $\nabla_{\boldsymbol{h}}\boldsymbol{x}(\boldsymbol{h}^*)$ defined by a subset of species $\mathcal{A} = \{A_{i_1}, ..., A_{i_{\tilde{n}}}\}$ with $\tilde{n} \leq n$, and a subset of parameters, $\widetilde{\boldsymbol{h}} = \left(h_{j_1}, \dots, h_{j_\tau}\right)$ with $\tau \leq r+p$, of interest. Here $\tilde{n}$ and τ are the number of selected species and network parameters, respectively, and can be arbitrary chosen by the user based on his/her research question. In the next sections, we will illustrate some possible choices for $\mathcal{A}$ and $\widetilde{\boldsymbol{h}}$, and thus for $\tilde{n}$ and τ, in four different scenarios.

Effectiveness of the SSIs

As a first experiment we ranked the impact of a perturbation in the value of each reaction and each conservation laws' constant within the CR-CRN. Specifically , we applied the algorithm sketched in Fig. 1 twice by defining the set of the considered parameters, $\tilde{\boldsymbol{h}}$, equal to $\boldsymbol{k}$ (hence $\tau = r$) and $\boldsymbol{c}$ (hence $\tau = p$), respectively, while keeping the whole set of species (hence $\tilde{n} = n = 419$) . As can be seen from Fig. 2, only 131 out of the 850 rate constants are associated with a SSI greater than 0.001, while for 462 rate constants the SSI ranges between 10^{-3} and 10^{-4} and for the remaining ones the SSI is lower than 10^{-4}. Instead, as far as the conservation laws' constants are concerned, 24 out of the 81 constants have a SSI grater than 10^{-2}, while 43 constants have a SSI that ranges between 10^{-2} and 10^{-3}, and the value of the lowest SSI is equal to 5.6×10^{-4} .

To verify the effectiveness of our definition of the SSIs, we quantified the effect induced on the equilibrium of the network by changing, one at the time, the components of $\boldsymbol{k}$ and $\boldsymbol{c}$ characterized by the maximum, median, and minimal SSI values, respectively. The results of this computation are presented in the lower panels of Fig. 2 and clearly show that these components are, respectively, the ones with highest, median, and lowest sensitivity, this latter being assessed as the Euclidean norm $||\hat{\Delta}\boldsymbol{x}||$ of the relative difference, defined in Eq. (15), between the equilibrium of the original network and that reached by the CRN containing the modified parameter. As thoroughly described in the Methods section, the squared norm $||\hat{\Delta}\boldsymbol{x}||^2$ has been denoted as $Q(\hat{\Delta}\boldsymbol{h})$ and quantifies the overall change on the protein concentrations at equilibrium induced by modifying the values of the network parameters as just described.

SSIs for identifying therapeutic targets

This second experiment aims at illustrating how the SSIs could support the development of targeted therapies. In detail , we analysed the sensitivity of the equilibrium concentration of p53 with respect to the conservation laws' constants by running the algorithm sketched in Fig. 1 with $\mathcal{A} = \{\text{p53}\}$ and $\tilde{\boldsymbol{h}} = \boldsymbol{c}$, and thus $\tilde{n} = 1$ and $\tau = p$. Since our aim is to illustrate the feasibility of our approach, we focused on the tumor suppressor p53 as mutations of the corresponding gene TP53 are among the most frequent mutations found in colorectal cancer and thus a rich literature is available to validate our results[32–34]. Furthermore, we focused on the SSIs for the conservation laws because we aim at identifying possible molecular targets, that are proteins whose concentration may be increased or reduced for indirectly acting on p53 concentration without modifying the network. Different experiments may be conducted by the interested readers by exploiting the code and data accompanying the manuscript.

As shown in Table 1, the SSI values allow ranking the conservation laws based on the impact that a change in the value of the corresponding constants would have on the concentration of p53. Possible therapeutic targets

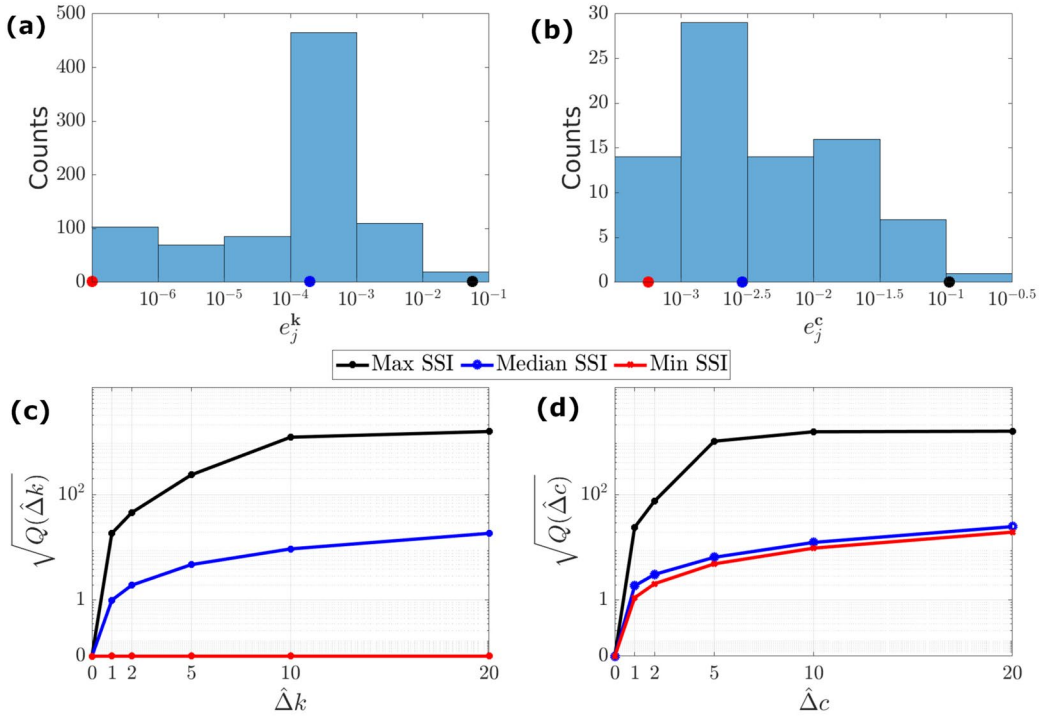

Figure 2. Sensitivity analysis for the CRN modeling a healthy colorectal cell. (**a**, **b**): distribution of the SSIs associated to the rate constants $\boldsymbol{k}$ (left) and to the conservation laws' constants $\boldsymbol{c}$ (right). In both panels, colored dots represent the values of the minimum (red), median (blue), and maximum (black) SSI. (**c**): $\sqrt{Q(\hat{\Delta}k)}$ as a function of the relative variation of the rate constants with minimum, median, and maximum SSI values. (**d**): $\sqrt{Q(\hat{\Delta}c)}$ as a function of the relative variation of the conservation laws' constants with minimum, median, and maximum SSI values.

e_j^c	j—th elemental species
0.455	MDM2
0.351	ARF
0.024	TP53_generator
0.022	Pho13
0.017	Pase2
0.016	AKT
0.016	Pase1

Table 1. SSI for the conservation laws' constants when only the concentration of TP53 is considered. For each of the SSIs in the first column, the second column shows one representative protein (elemental species) of the corresponding conservation law. Only SSI higher than 0.015 have been displayed.

may be identified by looking at the elemental species involved in the conservation laws associated with the highest SSI. As an example, in this case the conservation law associated with the maximum SSI value is the one associated to MDM2, a gene whose encoded protein can promote tumor formation by targeting tumor suppressors such as p53[5,35,36]. Coherently, a further analysis not shown here demonstrated that p53 concentration is increased by reducing the component of $\boldsymbol{c}$ that represents the total combined concentration available within the conservation law involving the elemental species MDM2. This may be achieved, for example, by using proper inhibitory drugs acting directly on MDM2 or on the proteins involved in its production .

SSIs for highlighting the impact of mutations

A greater significance of our approach to sensitivity analysis is concerned with applications to mutated CRNs. As illustrative examples, we considered four different mutated networks obtained incorporating into the CR-CRN four among the mutations more commonly found in colorectal cancer, namely the loss of function (LoF) of APC, SMAD4, and TP53 and the gain of functon (GoF) of KRAS. For each mutated network, we computed two distinct sets of SSIs, associated to the rate constants $\boldsymbol{k}$ and the conservation laws' constants $\boldsymbol{c}$, respectively, by applying the algorithm sketched in Fig. 1 as described in the Methods subsection *Sensitivity analysis of a mutated CRN*. The computed SSIs were then compared with the corresponding values in the physiological CR-CRN.

In this context, the SSIs play two roles. On the one hand, they allow a rigorous evaluation of the overall effects of each mutation on the network. This role is clearly shown by Figs. 3 and 4, which represent the two computed

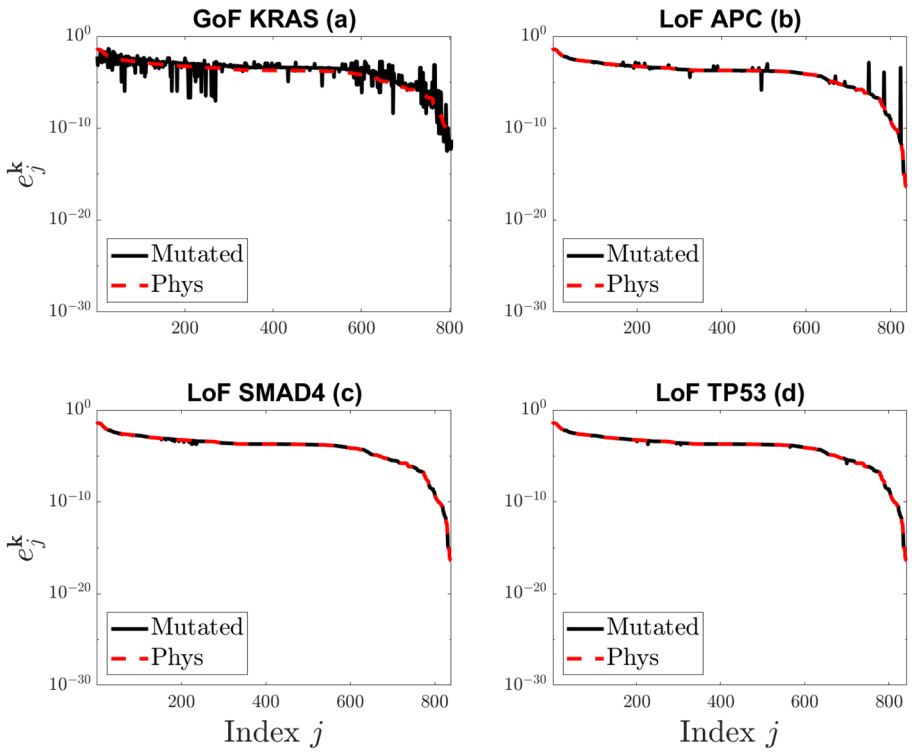

Figure 3. SSI values for all rate constants in the case of four mutated networks (black solid line) with respect to the SSI values of the physiological CRN (red dashed line).

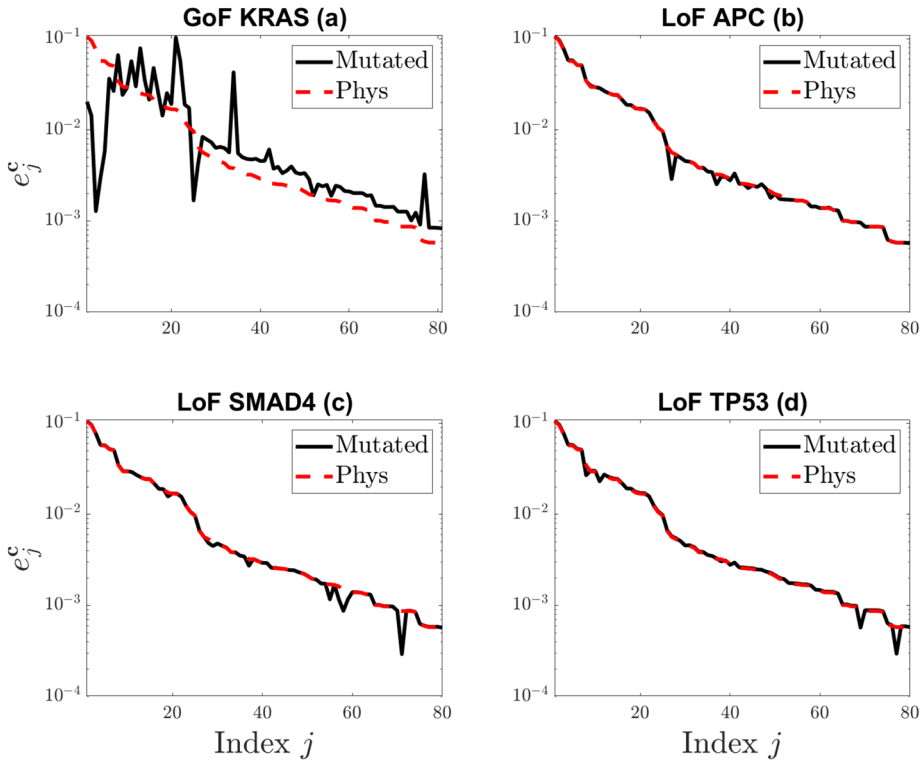

Figure 4. SSI values for all conservation laws' constants in the case of four mutated networks (black solid line) with respect to the SSI values of the physiological CRN (red dashed line).

sets of SSI values for all the considered colorectal cancer mutated networks. These figures indicate that the KRAS GoF mutation strongly modifies the SSI values of most indices 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[5].

On the other hand, and more importantly, the SSI values also allow the identification of the specific reactions in the network that were mostly affected by each mutation. This second role of SSIs is pointed out by Table 2 that contains a list of reactions ranked according to the values of the relative difference between the SSI values in the physiological and mutated cases. 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[9,37]. Specifically, via R41, free Ras is freezed in the inactive chemical compound *Ras_GDP*,

Reference #	Reaction	δe_j^k
R41	Ras + GDP -> Ras_GDP	1.91×10^7
R294	ERBP_ShP_G_S -> ERBP_ShP_G + SOS	2.91×10^4
R49	RP_ShP_G_S_Ras + GTP -> RP_ShP_G_S_Ras_GTP	2.88×10^4
R412	ERB3P_Sh -> ERB3P + Shc	2.86×10^4
R415	ShP + ERB3P -> ERB3P_ShP	1.23×10^4
R297	ERBP_ShP + G_S -> ERBP_ShP_G_S	8.64×10^3
R302	ERBP_ShP_G_S_Ras + GDP -> ERBP_ShP_G_S_Ras_GDP	5.78×10^2
R57	RP_G_S_Ras_GDP -> RP_G_S_Ras + GDP	5.78×10^2
R420	ERB3P_ShP_G + SOS -> ERB3P_ShP_G_S	5.78×10^2
R52	RP_ShP_G_S + Ras_GTP -> RP_ShP_G_S_Ras_GTP	4.95×10^2

Table 2. Reactions whose sensitivity indices are mostly effected by the GoF mutation of KRAS. For each reaction listed in the second column, the first column shows the reference number attributed to the reaction in the supplementary material of Tortolina and colleagues[2], while the third column contains the relative differences δe_j^k between the value of the SSI e_j^k in the original and in the mutated network. Only the first 10 reactions with highest absolute value of δe_j^k have been displayed.

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 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 the paper by Tortolina and colleagues[2]).

SSIs for in silico-based drug dosage

One of the main role of the SSIs is to identify which parameters require a fine calibration when, e.g., devising novel networks or when an existing network is employed for the in-silico comparison of different therapies. To illustrate this application, we considered the mutated CR-CRN incorporating a GoF mutation of KRAS and, following the model proposed in a previous work[7], we enlarged it with two groups of reactions mimicking the effect of two drugs: Dabrafenib (DBF) and Trametinib (TMT) targeting B-Raf and MEK, respectively. The action of 40 nM of DBF and 240 nM of TMT was considered since these concentrations were found to minimize the difference between the equilibrium of the modified and the original healthy network[7].

Figure 5 shows the SSI values for the conservation laws' constants obtained by applying the algorithm sketched in Fig. 1 with the parameters of the novel network. In particular, the SSI associated to the conservation law involving DBF is equal to 0.247 and results much higher than the SSI associated to the conservation law involving TMT, which is equal to 7.98×10^{-4}. This means that when the developed network is used for the in-silico comparison and evaluation of different drug dosages a finer set of values should be used for DBF because small variations in its concentration may have a great impact on the equilibrium reached by the network.

Discussion

CRNs are typically characterized by a huge number of variables and, although static formulations reduce the number of parameters in the game, calibration of a CRN is often a challenging issue. Sensitivity analysis of both physiological and mutated CRNs is an effective way to identify the parameters that mostly affect the overall behavior of the network and thus need to be finely calibrated.

In our view, the sensitivity index introduced in this paper has several advantages. It is specifically tailored to the static case and is based on PCA, which makes its computation rather straightforward. From a technical viewpoint, it is not affected by the intrinsic ambiguity of the sign of the eigenvectors, as it occurs in the case of the index introduced in previous works[3,23]. 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 CR-CRN, this index has a more general value, and can be utilized for assessing the equilibrium condition of any mutated signalling network.

We reckon that one of the main limitation of the present work is that it only focuses on a local approach to sensitivity analysis for which an initial guess of the values of the parameters is required and defined based on

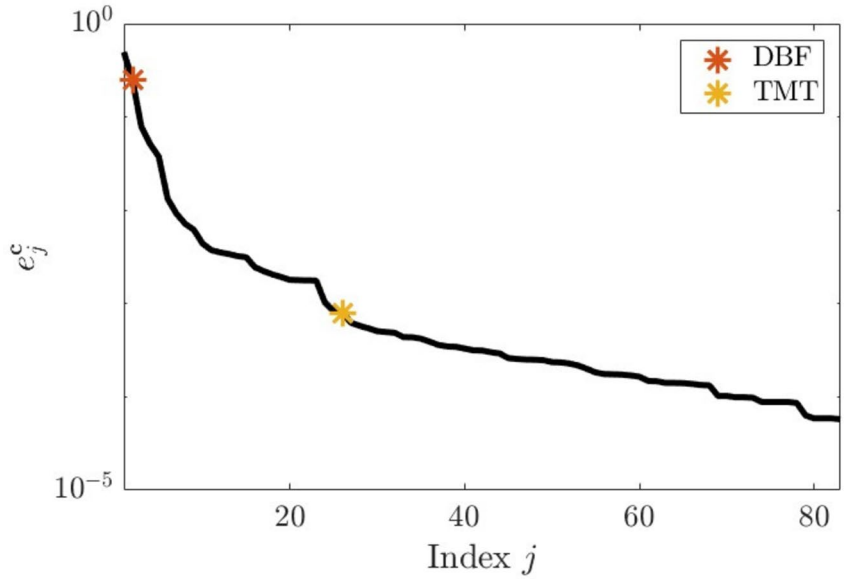

Figure 5. SSI values for all conservation laws' constants for the network incorporating a GoF mutation of KRAS and the combined effect of two targeted drugs, namely DBF and TMT. The red and yellow stars highlight the value of the SSI associated to the conservation law involving DBF and TMT, respectively.

literature data. This choice was mainly motivated by the fact that many of the existing global approaches, based e.g. on Monte Carlo sampling or polynomial chaos expansions, can currently tackle only small problems[38,39]. However, future effort will be devoted to implement a global sensitivity approach that do not suffer of such scalability issue and could be applied to the CR-CRN[40].

Another limitation concerns the fact that we focused our sensitivity analysis at the steady state of the network, after assuming that a unique steady state exists if the values of all the parameters of the network are fixed. Our definition of the SSI based on the PCA of the scaled local sensitivity matrix (second step of our approach) easily generalizes to the dynamic case. On the contrary, working at the equilibrium, either stable or not, is a crucial assumption in order to analytically compute the local sensitivity matrix through the implicit function theorem (first step of our approach). Future effort will be devoted in developing an algorithm capable of efficiently computing the local sensitivity matrix at each transition state of a solution of the ODEs system (1) also for complex networks involving hundreds of proteins and reactions such as the CR-CRN considered in this work. Furthermore, a second study has been planned aimed at exploiting algebraic tools, based e.g. on Gröbner basis, for identifying possible values of the kinetic parameters for which the equilibrium of the network is not unique[41,42]. If the identified values of the parameters correspond to biologically plausible scenarios, then we could use our approach for performing sensitivity analysis at each equilibrium reached by the network. In principle, the SSIs computed for distinct equilibrium points are different, thus a dedicated algorithm has to be developed for comparing and integrating them, for example by statistically comparing the PCA of the local sensitivity matrices at different locations[43].

Another possible next step will be to use the SSIs, possibly coupled with more advanced model reduction techniques based e.g. on information-geometry and sloppy modeling[44], to select the sensitive pathways within the global CR-CRN. Then, computational algorithms based on inverse problems theory[45] will be used to finally calibrate the parameters associated with a high value of the SSI using experimental measurements of the proteins' concentrations at equilibrium. The values of the other parameters will be more coarsely approximated by using literature data. It is worth noticing that in practice a perfect estimation of the network parameters is actually impossible since any CRN model is obviously incomplete with many missing proteins and reactions[46]. As a consequence, the considered parameters often account for numerous unmodelled effects. However, our final aim is not to perfectly estimate the value of the network parameters but rather to provide a valid tool for quantitatively simulating the effects of different therapies and thus, e.g., helping in prioritizing actual clinical trials.

Methods

Background: computation of CRNs equilibrium

We consider a CRN comprising r reactions involving n well-mixed chemical species and assume that the law of mass action kinetics holds. Throughout the paper, 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) as in Eq. (1).

We further assume the CRN to be weakly elemented[4], that is $p = n - \text{rank}(\boldsymbol{S})$ independent conservation laws exist described by the linear system

$$\boldsymbol{N}\boldsymbol{x} = \boldsymbol{c}, \tag{3}$$

where $\boldsymbol{c} \in R_+^p$ is the vector of conservations laws' constants whose values can be defined as $\boldsymbol{c} := \boldsymbol{N}\boldsymbol{x}_0$, being $\boldsymbol{x}_0$ the initial values of the species concentrations , and the matrix $\boldsymbol{N} \in N^{p\times n}$ is determined by computing the kernel of $\boldsymbol{S}^T$ 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 $\boldsymbol{x}$, we can consider the decompositions

$$\boldsymbol{N} = \begin{bmatrix} \boldsymbol{I}_p & \boldsymbol{N}_2^{p\times(n-p)} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_1^{p\times r} \\ \boldsymbol{S}_2^{(n-p)\times r} \end{bmatrix}, \tag{4}$$

where $\boldsymbol{x}_1$ is a p-vector of the so called elemental species[4,] and $\boldsymbol{S}_2 \in R^{(n-p)\times r}$ has rank $(n-p)$ and is the submatrix of $\boldsymbol{S}$ corresponding to the non-elemental species. An illustrative example on how to compute and interpret the conservation laws and the elemental species for a simpler network is provided in the supplementary material (see Supplementary Note S1).

It follows from the definition of conservation laws that $\boldsymbol{N}\boldsymbol{S} = 0$, whence $\boldsymbol{S}_1 = -\boldsymbol{N}_2\boldsymbol{S}_2$. It can be shown[13,26] that the equilibrium (stationary) solution of the Cauchy problem is determined by the algebraic system

$$\begin{cases} \boldsymbol{S}_2\boldsymbol{v}(\boldsymbol{x},\boldsymbol{k}) = 0 \\ \boldsymbol{N}\boldsymbol{x} - \boldsymbol{c} = 0 \end{cases}. \tag{5}$$

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

$$\boldsymbol{f}(\boldsymbol{x},\boldsymbol{k},\boldsymbol{c}) = 0, \tag{6}$$

where $f : R^n \times R^r \times R^p \longrightarrow R^n$ is a continuously differentiable function defined as

$$\boldsymbol{f}(\boldsymbol{x},\boldsymbol{k},\boldsymbol{c}) = \begin{bmatrix} \boldsymbol{S}_2\boldsymbol{v}(\boldsymbol{x},\boldsymbol{k}) \\ \boldsymbol{N}\boldsymbol{x} - \boldsymbol{c} \end{bmatrix}. \tag{7}$$

Since each reaction involves up to two reactants, from the law of mass action it follows that the components of $\boldsymbol{f}$ are polynomials of degree two or lower.

A CRN is said to satisfy the global stability condition[4] if to any fixed $\boldsymbol{k}^*$ and $\boldsymbol{c}^*$ there corresponds a unique asymptotically stable equilibrium $\boldsymbol{x}^*$, which may be determined by either a dynamic simulation[4] or the solution of (5)[13].

It has been demonstrated that the action of two particular classes of mutations can be simulated by modifying the original Cauchy problem associated to the CRN through proper projectors acting on the initial conditions, and thus on the conservation laws' constants $\boldsymbol{c}$, or on the stoichiometric matrix $\boldsymbol{S}$[4,5]. In detail, a class of mutations resulting in the loss of function (LoF) of an elemental species can be implemented by setting to zero the constant of the conservation law where it is involved. It follows that the concentration of all the species involved in such a conservation law will remain null over time. A class of mutations resulting in the gain of function (GoF) of a species is implemented by removing from the network the reactions involved in its deactivation.

Local sensitivity matrix and statistical sensitivity indices

In this section we thoroughly describe the two steps that form the proposed algorithm for computing the SSIs.

Step 1. Analytic computation of the sensitivity matrix

We consider a weakly elemented CRN and we assume that

$$\det[(\nabla_{\boldsymbol{x}}\boldsymbol{f})(\boldsymbol{x}^*,\boldsymbol{h}^*)] \neq 0 \ . \tag{8}$$

On account of the implicit function theorem[47], equations (5) and (8) imply that there exists one and only one $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{h})$ such that

$$\boldsymbol{f}(\boldsymbol{x}(\boldsymbol{h}),\boldsymbol{h}) = 0 \tag{9}$$

for $\boldsymbol{h}$ in a neighborhood of $\boldsymbol{h}^*$, meaning that $\boldsymbol{x}(\boldsymbol{h})$ is the equilibrium point of the CRN defined by the set of parameters $\boldsymbol{h}$. Further,

$$\nabla_{\boldsymbol{h}}\boldsymbol{x}(\boldsymbol{h}^*) = -\left[\nabla_{\boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}^*,\boldsymbol{h}^*)\right]^{-1}\nabla_{\boldsymbol{h}}\boldsymbol{f}(\boldsymbol{x}^*,\boldsymbol{h}^*)\,, \tag{10}$$

where, according to the definition of $\boldsymbol{f}$ in (7),

$$\nabla_{\boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}^*,\boldsymbol{h}^*) = \begin{bmatrix} \boldsymbol{S}_2\nabla_{\boldsymbol{x}}\boldsymbol{v}(\boldsymbol{x}^*,\boldsymbol{k}^*) \\ \boldsymbol{N} \end{bmatrix} \tag{11}$$

and

$$\nabla_{\boldsymbol{h}}\boldsymbol{f}(\boldsymbol{x}^*,\boldsymbol{h}^*) = \begin{bmatrix} \boldsymbol{S}_2\nabla_{\boldsymbol{k}}\boldsymbol{v}(\boldsymbol{x}^*,\boldsymbol{k}^*) & \boldsymbol{0} \\ \boldsymbol{0} & -\mathbf{I}_p \end{bmatrix}. \tag{12}$$

Step 2. Computation of the SSI

Following the work by Liu and colleagues[23], to quantify the overall effect that a relative change in the parameters $\boldsymbol{h}$ induces on the protein concentrations at equilibrium, we define the positive semi-definite quadratic form

$$Q(\hat{\Delta}\boldsymbol{h}) = \hat{\Delta}\boldsymbol{x}^T\hat{\Delta}\boldsymbol{x} \tag{13}$$

where

$$\hat{\Delta}\boldsymbol{h} = \left[\frac{h_1 - h_1^*}{h_1^*},\cdots,\frac{h_{r+p} - h_{r+p}^*}{h_{r+p}^*}\right]^T \tag{14}$$

and

$$\hat{\Delta}\boldsymbol{x} = \left[\frac{x_1(\boldsymbol{h}) - x_1^*}{x_1^*},\cdots,\frac{x_n(\boldsymbol{h}) - x_n^*}{x_n^*}\right]^T \tag{15}$$

$\boldsymbol{x}(\boldsymbol{h}) = (x_1(\boldsymbol{h}),\ldots,x_n(\boldsymbol{h}))^T$ being the equilibrium reached by the network with parameters $\boldsymbol{h}$. We observe that, in order to compute $\hat{\Delta}\boldsymbol{x}$, we shall assume that $x_i^* \neq 0$ for all $i = 1,\ldots,n$, which is true, e.g., for the original CR-CRN. However, as we shall see in the next sections, if some of the species have a null equilibrium concentration our approach can be applied by simply removing the components of $\boldsymbol{x}$ and the rows of the sensitivity matrix corresponding to such species. By applying some straightforward algebraic computation to the first order Taylor's polynomial of the implicit function $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{h})$, it follows that in a neighbour of $\boldsymbol{h}^*$ it holds

$$\hat{\Delta}\boldsymbol{x} = \boldsymbol{\mathcal{S}}\,\hat{\Delta}\boldsymbol{h}\ , \tag{16}$$

where we have introduced the relative local sensitivity matrix[29]

$$\boldsymbol{\mathcal{S}} = (\boldsymbol{D}_{\boldsymbol{x}^*})^{-1}\nabla_{\boldsymbol{h}}\boldsymbol{x}(\boldsymbol{h}^*)\,\boldsymbol{D}_{\boldsymbol{h}^*}\ , \tag{17}$$

being $\boldsymbol{D}_{\boldsymbol{x}^*} = \text{diag}[\text{x}_1^*, ..., \text{x}_\text{n}^*]$ and $\boldsymbol{D}_{\boldsymbol{h}^*} = \text{diag}[\text{h}_1^*, ..., \text{h}_{\text{r+p}}^*]$. Replacing equation (16) into (13) leads to

$$Q(\hat{\Delta}\boldsymbol{h}) = \hat{\Delta}\boldsymbol{h}^T \boldsymbol{\mathcal{S}}^T \boldsymbol{\mathcal{S}} \hat{\Delta}\boldsymbol{h} = \hat{\Delta}\boldsymbol{h}^T \, \boldsymbol{U} \, \boldsymbol{\Lambda} \, \boldsymbol{U}^T \, \hat{\Delta}\boldsymbol{h} \,, \tag{18}$$

where $\boldsymbol{\Lambda} = \text{diag}[\lambda_1, ..., \lambda_{\text{r+p}}]$, $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{r+p} \geq 0$, are the eigenvalues of the matrix $\boldsymbol{\mathcal{S}}^T\boldsymbol{\mathcal{S}}$ and $\boldsymbol{U} = [\boldsymbol{u}_1, ..., \boldsymbol{u}_{r+p}]$ collects a set of independent normalized eigenvectors.

As a special case, let us assume that only the j-th parameter is perturbed and thus $\hat{\Delta}\boldsymbol{h}$ has $\hat{\Delta}h_j$ as a unique non-vanishing component. Then, equation (18) reads as

$$Q(\hat{\Delta}h_j) = \hat{\Delta}h_j^2 \sum_{a=1}^{r+p} \lambda_a \, u_{ja}^2. \tag{19}$$

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

$$e_j^{\boldsymbol{h}} = \frac{\sum_{a=1}^{r+p} \lambda_a \, u_{ja}^2}{\sum_{a=1}^{r+p} \lambda_a} \quad j = 1, \dots r+p \,. \tag{20}$$

It can be shown that $e_j^{\boldsymbol{h}} \in [0, 1]$ and $\sum_{j=1}^{r+p} e_j^{\boldsymbol{h}} = 1$. Further, getting back to the original parameters $\boldsymbol{k}$ and $\boldsymbol{c}$, the first r SSIs in (20), also denoted as $e_j^{\boldsymbol{k}}$, $j = 1, \dots, r$, provide a synthetic and compact estimate of the weight of perturbation in reaction j on the whole CRN, while the latter p SSIs in (20), also denoted as $e_j^{\boldsymbol{c}}$, $j = 1, \dots, p$, refers to the overall sensitivity of the network to the j-th component of vector $\boldsymbol{c}$.

We observe that the definition of the SSIs in Eq. (20) is similar to that of Liu and colleagues[23] , the only difference being that we used the squared version, u_{ja}^2, of the eigenvector components instead of simply using u_{ja}. The motivation of this choice is twofold. On the one hand, it ensures that all the terms of the sums in (20) are positive thus avoiding cancellation effects. On the other hand, it accounts for the fact that the sign of the eigenvectors is ambiguous. In fact, for any eigenvector $\boldsymbol{u}_j$, $j = 1, \dots, r+p$, also $-\boldsymbol{u}_j$ is an eigenvector of $\boldsymbol{\mathcal{S}}^T\boldsymbol{\mathcal{S}}$ and different Matlab routines may actually return eigenvectors with different sign (see Supplementary Note S2 for further details) .

Sensitivity analysis of subnetworks

Since the SSIs in (20) are defined based on a PCA of the whole relative local sensitivity matrix $\boldsymbol{\mathcal{S}}$, their values allow discriminating among all the parameters whose variation mainly affects the overall equilibrium point reached by the network. However, it is possible to focus only on a subset of chemical species $\mathcal{A} = \{A_{i_1}, ..., A_{i_{\tilde{n}}}\}$, $\tilde{n} < n$, and on a subset of parameters $\boldsymbol{h} = \left(h_{j_1}, \dots, h_{j_\tau}\right)$, $\tau < r + p$, by applying the procedure just described on the submatrix $\widetilde{\boldsymbol{\mathcal{S}}}$ of $\boldsymbol{\mathcal{S}}$, formed by the $\tilde{n}$ rows referred to $\mathcal{A}$ and the τ columns referred to $\widetilde{\boldsymbol{h}}$. $\widetilde{\boldsymbol{\mathcal{S}}}$ represents the relative local sensitivity matrix of the equilibrium concentrations $(x_{i_1}, \dots, x_{i_{\tilde{n}}})$ of the protein in $\mathcal{A}$ with respect to the parameters in $\widetilde{\boldsymbol{h}}$. Hence the SSIs computed by using in (20) the eigenvalues and eigenvectors of $\widetilde{\boldsymbol{\mathcal{S}}}^T\widetilde{\boldsymbol{\mathcal{S}}}$ quantify the impact of each parameters only on the considered proteins.

As an example, let use assume $\widetilde{\boldsymbol{h}} = \boldsymbol{h}$, $\tilde{n} = 1$, and $\mathcal{A} = \{A_i\}$. Then $\widetilde{\boldsymbol{\mathcal{S}}}$ is the i-th row of $\boldsymbol{\mathcal{S}}$, denoted as $\boldsymbol{\mathcal{S}}_{i:}$, and the matrix $\widetilde{\boldsymbol{\mathcal{S}}}^T\widetilde{\boldsymbol{\mathcal{S}}}$ has a unique non-vanishing eigenvalue equal to $||\boldsymbol{\mathcal{S}}_{i:}||^2$ with $\frac{\boldsymbol{\mathcal{S}}_{i:}}{||\boldsymbol{\mathcal{S}}_{i:}||}$ as a normalized eigenvector. From the definition of the SSIs in (20) it follows that

$$e_j^{\boldsymbol{h}} = \frac{\mathcal{S}_{ij}^2}{||\boldsymbol{\mathcal{S}}_{i:}||} \,, \tag{21}$$

i.e., the SSIs coincide with the squared normalized component of the i-th row of $\boldsymbol{\mathcal{S}}$. This confirm our interpretation of the SSIs. Indeed, the i-th row of the rescaled sensitivity matrix $\boldsymbol{\mathcal{S}}$ describes the relative sensitivity of the equilibrium concentrations of the species $\mathcal{A}$ with respect to the parameters of the network[29].

Sensitivity analysis of mutated CRNs

As described in the *Background* section, a mutation resulting in the LoF of an elemental species can be implemented by setting to zero the constant of the conservation law involving such a species[4,5]. Coherently, in this scenario the algorithm sketched in Fig. 1 can be applied by setting to zero the component of $\boldsymbol{h}^*$ corresponding to the modified conservation law and then computing $\boldsymbol{x}^*$ as the equilibrium point of the CRN with the novel set of parameters. Furthermore, the species involved in the modified conservation law are removed from the set $\mathcal{A}$ because their equilibrium concentration will clearly be zero.

A mutation resulting in the GoF of a species implies that a set of reactions involving it no longer occurs in the network[5]. Hence, a reduced network is defined by removing the columns of $\boldsymbol{S}_2$, the rows of $\boldsymbol{v}$, and the rate constants within $\boldsymbol{h}^*$ associated to these reactions. The algorithm sketched in Fig. 1 can be applied to the novel network after computing the equilibrium $\boldsymbol{x}^*$ and by defining $\mathcal{A}$ as the set of species with an equilibrium concentration greater than 10^{-15} so as to neglect those species whose concentration is null or close to zero.

Data availability

The datasets and Matlab codes generated and analysed during the current study are available in the GitHub repository https://github.com/theMIDAgroup/CRN_sensitivity.

Received: 17 March 2024; Accepted: 16 July 2024
Published online: 31 July 2024

References

1. Zi, Z. Sensitivity analysis approaches applied to systems biology models. *IET Syst. Biol.* **5**, 336–346 (2011).
2. Tortolina, L. *et al.* Advances in dynamic modeling of colorectal cancer signaling-network regions, a path toward targeted therapies. *Oncotarget* **6**, 5041 (2015).
3. Santra, T. Fitting mathematical models of biochemical pathways to steady state perturbation response data without simulating perturbation experiments. *Sci. Rep.* **8**, 1–14 (2018).
4. Sommariva, S., Caviglia, G. & Piana, M. Gain and loss of function mutations in biological chemical reaction networks: A mathematical model with application to colorectal cancer cells. *J. Math. Biol.* **82**, 1–25 (2021).
5. Sommariva, S. *et al.* Computational quantification of global effects induced by mutations and drugs in signaling networks of colorectal cancer cells. *Sci. Rep.* **11**, 1–13 (2021).
6. Krishnan, J., Torabi, R., Schuppert, A. & Napoli, E. D. A modified ising model of barabási-albert network with gene-type spins. *J. Math. Biol.* **81**, 769–798 (2020).
7. Sommariva, S. *et al.* In-silico modelling of the mitogen-activated protein kinase (MAPK) pathway in colorectal cancer: Mutations and targeted therapy. *Front. Syst. Biol.*https://doi.org/10.1101/2023.04.18.537359 *(2023).*
8. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
9. Levine, A. J., Jenkins, N. A. & Copeland, N. G. The roles of initiating truncal mutations in human cancers: The order of mutations and tumor cell type matters. *Cancer Cell* **35**, 10–15 (2019).
10. Tariq, K. & Ghias, K. Colorectal cancer carcinogenesis: A review of mechanisms. *Cancer Biol. Med.* **13**, 120 (2016).
11. Lee, Y. T., Tan, Y. J. & Oon, C. E. Molecular targeted therapy: Treating cancer with specificity. *Eur. J. Pharmacol.* **834**, 188–196 (2018).
12. Bedard, P. L., Hyman, D. M., Davids, M. S. & Siu, L. L. Small molecules, big impact: 20 years of targeted therapy in oncology. *Lancet* **395**, 1078–1088 (2020).
13. Berra, S., La Torraca, A., Benvenuto, F. & Sommariva, S. Combined newton-gradient method for constrained root-finding in chemical reaction networks. *J. Optim. Theory Appl.* **200**, 404–427 (2024).
14. Klinger, B. *et al.* Network quantification of EGFR signaling unveils potential for targeted combination therapy. *Mol. Syst. Biol.* **9**, 673 (2013).
15. Sahin, Ö. *et al.* Modeling ERBB receptor-regulated G1/S transition to find novel targets for de novo trastuzumab resistance. *BMC Syst. Biol.* **3**, 1–20 (2009).
16. Shinar, G., Alon, U. & Feinberg, M. Sensitivity and robustness in chemical reaction networks. *SIAM J. Appl. Math.* **69**, 977–998 (2009).
17. Bartek, J. & Lukas, J. Pathways governing g1/s transition and their response to DNA damage. *FEBS Lett.* **490**, 117–122 (2001).
18. Saltelli, A., Tarantola, S. & Campolongo, F. Sensitivity analysis as an ingredient of modeling. *Stat. Sci.* **15**, 377–395 (2000).
19. Saltelli, A., Ratto, M., Tarantola, S. & Campolongo, F. Sensitivity analysis for chemical models. *Chem. Rev.* **105**, 2811–2828 (2005).
20. Saltelli, A. *et al.* Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Comput. Phys. Commun.* **181**, 259–270 (2010).
21. Quaiser, T. & Mönnigmann, M. Systematic identifiability testing for unambiguous mechanistic modeling-application to jak-stat, map kinase, and nf-κ b signaling pathway models. *BMC Syst. Biol.* **3**, 1–21 (2009).
22. Chen, H. & Heitjan, D. F. Analysis of local sensitivity to nonignorability with missing outcomes and predictors. *Biometrics* **78**, 1342–1352 (2022).
23. Liu, G., Swihart, M. T. & Neelamegham, S. Sensitivity, principal component and flux analysis applied to signal transduction: The case of epidermal growth factor mediated signaling. *Bioinformatics* **21**, 1194–1202 (2005).
24. Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D. & Müller, G. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat. Biotechnol.* **20**, 370–375 (2002).
25. Feinberg, M. Chemical reaction network structure and the stability of complex isothermal reactors-I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.* **42**, 2229–2268 (1987).
26. Chellaboina, V., Bhat, S. P., Haddad, W. M. & Bernstein, D. S. Modeling and analysis of mass-action kinetics. *IEEE Control Syst. Mag.* **29**, 60–78 (2009).
27. Yu, P. Y. & Craciun, G. Mathematical analysis of chemical reaction systems. *Israel J. Chem.* **58**, 733–741 (2018).
28. De Martino, A., De Martino, D., Mulet, R. & Pagnani, A. Identifying all moiety conservation laws in genome-scale metabolic networks. *PloS one* **9**(7), e100750 (2014).
29. Goulet, D. Modeling, simulating, and parameter fitting of biochemical kinetic experiments. *Siam Rev.* **58**, 331–353 (2016).
30. Varma, A. & Morbidelli, M. *Parametric Sensitivity in Chemical Systems* (Cambridge University Press, Cambridge, 1999).
31. Fang, J. Y. & Richardson, B. C. The MAPK signalling pathways and colorectal cancer. *Lancet Oncol.* **6**, 322–327 (2005).
32. De Roock, W., De Vriendt, V., Normanno, N., Ciardiello, F. & Tejpar, S. kras, braf, pik3ca, and pten mutations: Implications for targeted therapies in metastatic colorectal cancer. *Lancet Oncol.* **12**, 594–603 (2011).
33. Rivlin, N., Brosh, R., Oren, M. & Rotter, V. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. *Genes Cancer* **2**, 466–474 (2011).
34. Tsilimigras, D. I. *et al.* Clinical significance and prognostic relevance of kras, braf, pi3k and tp53 genetic mutation analysis for resectable and unresectable colorectal liver metastases: A systematic review of the current evidence. *Surg. Oncol.* **27**, 280–288 (2018).
35. Freedman, D., Wu, L. & Levine, A. Functions of the mdm2 oncoprotein. *Cell. Mol. Life Sci. CMLS* **55**, 96–107 (1999).
36. Li, J. & Kurokawa, M. Regulation of mdm2 stability after DNA damage. *J. Cell. Physiol.* **230**, 2318–2327 (2015).
37. Orton, R. J. *et al.* Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem. J.* **392**, 249–261 (2005).
38. Villaverde, A. F., Raimúndez, E., Hasenauer, J. & Banga, J. R. Assessment of prediction uncertainty quantification methods in systems biology. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **20**(3), 1725–1736 (2022).
39. Kaltenbach, H.-M., Dimopoulos, S. & Stelling, J. Systems analysis of cellular networks under uncertainty. *FEBS Lett.* **583**, 3923–3930 (2009).
40. Pianosi, F. *et al.* Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environ. Model. Softw.* **79**, 214–232 (2016).
41. Feliu, E. Injectivity, multiple zeros and multistationarity in reaction networks. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **471**, 20140530 (2015).
42. Sadeghimanesh, A. & Feliu, E. Gröebner bases of reaction networks with intermediate species. *Adv. Appl. Math.* **107**, 74–101 (2019).
43. Schott, J. R. Common principal component subspaces in two groups. *Biometrika* **75**, 229–236 (1988).
44. Transtrum, M. K. *et al.* Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *J. Chem. Phys.* **143**(1), 010901 (2015).

45. Bertero, M. & Piana, M. Inverse problems in biomedical imaging: Modeling and methods of solution. *Complex Syst. Biomed.*https://doi.org/10.1007/88-470-0396-2_1 *(2006).*
46. Nicolaou, Z. G. & Motter, A. E. Missing links as a source of seemingly variable constants in complex reaction networks. *Phys. Rev. Res.* **2**, 043135 (2020).
47. Krantz, S. G. & Parks, H. R. *The implicit function theorem: history, theory, and applications* (Springer Science & Business Media, 2002).

Acknowledgements

M. P. and S. S. kindly acknowledge the financial support of the "Hub Life Science - Digital Health (LSH-DH) PNC-E3-2022-23683267 - Progetto DHEAL-COM - CUP: D33C22001980001", granted by the Italian Ministero della Salute within the framework of the Piano Nazionale Complementare to the "PNRR Ecosistema Innovativo della Salute - Codice univoco investimento: PNC-E.3". G. B., M. P, and S.S. are member of the INdAM-GNCS group. S. S. has been partially supported by the INdAM-GNCS Project *Metodi computazionali per la costruzione e l'analisi di modelli matematici in biomedicina*, CUP_E55F22000270001. All the authors were partially supported by the MIUR Excellence Department Project awarded to Dipartimento di Matematica, Universitá di Genova, CUP D33C23001110001.

Author contributions

G.C., M.P. and S.S. conceived the study. G.B. implemented the method and performed the simulations. G.B., G.C. and S.S. analyzed the results. All the authors wrote the manuscript and read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-67862-5.

Correspondence and requests for materials should be addressed to G.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.