SIAM J. APPL. MATH.
Vol. 84, No. 3, pp. S105–S150

ALGEBRAIC STUDY OF RECEPTOR-LIGAND SYSTEMS: A DOSE-RESPONSE ANALYSIS*

LÉA STA†, MICHAEL F. ADAMER‡, AND CARMEN MOLINA-PARÍS†§

Abstract. The study of a receptor-ligand system generally relies on the analysis of its dose-response (or concentration-effect) curve, which quantifies the relation between ligand concentration and the biological effect (or cellular response) induced when binding its specific cell surface receptor. Mathematical models of receptor-ligand systems have been developed to compute a dose-response curve under the assumption that the biological effect is proportional to the number of ligand-bound receptors. Given a dose-response curve, two quantities (or metrics) have been defined to characterize the properties of the ligand-receptor system under consideration: amplitude and potency (or half-maximal effective concentration, and denoted by EC_{50}). Both the amplitude and the EC_{50} are key quantities commonly used in pharmaco-dynamic modeling, yet a comprehensive mathematical investigation of the behavior of these two metrics is still outstanding; for a large (and important) family of receptors, called cytokine receptors, we still do not know how amplitude and EC_{50} depend on receptor copy numbers. Here we make use of algebraic approaches (Gröbner basis) to study these metrics for a large class of receptor-ligand models, with a focus on cytokine receptors. In particular, we introduce a method, making use of two motivating examples based on the interleukin-7 (IL-7) receptor, to compute analytic expressions for the amplitude and the EC_{50}. We then extend the method to a wider class of receptor-ligand systems, sequential receptor-ligand systems with extrinsic kinase, and provide some examples. The algebraic methods developed in this paper not only reduce computational costs and numerical errors, but allow us to explicitly identify key molecular parameters and rates which determine the behavior of the dose-response curve. Thus, the proposed methods provide a novel and useful approach to perform model validation, assay design and parameter exploration of receptor-ligand systems.

Key words. cytokine, receptor, kinase, signaling, cellular fate, dose-response, steady state, Gröbner basis, amplitude, half-maximal effective concentration

MSC codes. 92C40, 92C42, 92C45, 13P10

DOI. 10.1137/22M1506262

1. Introduction.

The human body consists of more than 3×10^{13} cells [3], each of them receiving, at any given time, hundreds of signals from extracellular molecules when these bind their specific membrane receptors. These signals are integrated, translated, and read by a small number of intracellular molecules to generate appropriate cellular responses [28]. Surface receptors specifically bind to extracellular molecules called ligands. The binding of a ligand to its receptor induces an intracellular cascade of signaling events which regulate a cell's fate, such as migration, proliferation, death, or differentiation [39, 60]. Receptor-ligand interactions are essential in cell-to-cell communication, as is the case for immune cell populations [17]

*Received by the editors June 29, 2022; accepted for publication (in revised form) January 5, 2023; published electronically July 12, 2023.
https://doi.org/10.1137/22M1506262
Funding: The first and third authors received funding from the European Union's Horizon 2020 research and innovation programme under the Skłodowska-Curie grant agreement 764698.

†School of Mathematics, University of Leeds, Leeds LS2 9JT, UK (mmls@leeds.ac.uk).
‡Department of Biosystems Science and Engineering, ETH Zürich, Mattenstrasse 26, 4058 Basel, Switzerland and Swiss Institute of Bioinformatics, Quartier Sorge, Batiment Amphipole, 1015 Lausanne, Switzerland (michael.adamer@bsse.ethz.ch).
§Corresponding author. Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos 87545, NM, USA (molina-paris@lanl.gov).

and, thus, a large body of literature has been devoted to the experimental and theoretical study of cell signaling dynamics [64, 49, 23, 50, 32, 31, 46, 41, 20]. Exploiting the controlled environment of in vitro experiments, most cell signaling studies focus on the estimation of the affinity constant for a given receptor-ligand system, and the quantification of biochemical on and off rates for the binding and unbinding, respectively, of receptor and ligand molecules. Recent single-cell studies have shown that cells have heterogeneous expression levels of receptor copy numbers. Not only does the copy number depend on the cell type, but receptor copy numbers vary strongly between isogenic cells of one cell type [20, 8, 17]. Given the heterogeneity of receptor copy numbers across and within cell types, it is timely to understand how a cell's response to a given ligand depends on the expression levels of its receptor. This quantification will be a first step to account for the variability of receptor expression levels when designing and studying receptor-ligand models (both from an experimental and mathematical perspective) [17, 8, 23, 49].

The study of a receptor-ligand system generally relies on the analysis of its dose-response (or concentration-effect) curve, which describes the relation between ligand concentration and the biological effect (or cellular response) it generates when binding its specific receptor [39, 30, 31]. Mathematical models of receptor-ligand systems have been developed to compute a dose-response curve, under the assumption that a biological effect is proportional to the number of ligand-bound receptors [41, 8, 17, 64]. Given a dose-response curve, two quantities (or metrics) have been defined to characterize the properties of the ligand-receptor system under consideration. These metrics are the amplitude, which is defined as the difference between the maximal and minimal response, and the half-maximal effective concentration (or EC_{50}), which is the concentration of ligand required to induce an effect corresponding to 50% of the amplitude [39, 30, 31]. The amplitude is a measure of the efficacy of the ligand, and the EC_{50} is a measure of the potency (or sensitivity) of the ligand (for a given receptor) [39, 30, 14]. Both amplitude and EC_{50} are key quantities commonly used in pharmacodynamic modeling, yet a comprehensive mathematical investigation of the behavior of these two metrics is still outstanding for most receptor-ligand systems. For instance, for a large (and important) family of receptors, called cytokine receptors [48, 62, 2], we still do not know how amplitude and EC_{50} depend on receptor copy numbers (for a given concentration of ligand) [20, 8]. In this paper we bridge this gap by deriving closed-form expressions for a class of cytokine-receptor models. We further highlight how tools from computational algebra can be used to facilitate the calculation of both the amplitude and the EC_{50} for this family of models.

Previous work has shown that the estimation of the amplitude and the EC_{50} from experimental data is often possible, although strong inductive biases might be introduced [30, 39]. Usually one starts with a data set where the number (or concentration) of receptor-ligand signaling complexes formed (see section 2.2) is measured for different values of the ligand concentration. Then, the estimation of the amplitude and the EC_{50} is turned into a regression problem by assuming a functional relationship in the data set and fitting a parametric curve. A simple first approach is to plot experimental values (corresponding to a measurable variable which quantifies cellular response) as a function of ligand concentration. The amplitude and the EC_{50} are then read directly from a curve formed by interpolation of the data points. Since the EC_{50} is likely to fall between two data points, a geometrical method [1] can be used for an accurate determination. Nowadays many software packages can compute the amplitude and the EC_{50} from the data set making use of statistical methods, which consist in finding the "best-fit" equation to the dose-response curve. The most

common shape of the dose-response curve is a sigmoid and, thus, can be fitted with the famous Hill equation [22, 24]. However, other functions are also possible, such as a logistic equation [6, 33], a log-logistic equation [29, 58], or the Emax model [37, 59]. An asymmetrical sigmoid equation is sometimes needed for better precision [6, 58]. The amplitude and the EC_{50} are parameters of these equations and can, thus, be directly inferred from the fitting process. When a data set does not follow the strictly increasing pattern of these Hill-like functions, then more complex functions, such as bell-shaped curves [51], or multiphasic curves [10] can be used. It is important to note that even though these empirical regression methods allow one to quantify the two key receptor-ligand metrics, amplitude and EC_{50}, they do not offer any mechanistic insights for the receptor-ligand system under consideration. To this end, mathematical models can be used to describe the receptor-ligand system at a molecular level; that is, mathematical models consider the biochemical reactions which initiate a cellular response [15, 64, 31]. The challenge in such models is finding analytical, ideally closed-form, expressions for the amplitude and the EC_{50}. Due to the nonlinear nature of the biochemical reactions involved, this poses a significant and practical challenge.

Cytokine-receptor systems are of great relevance in immunology [48, 62, 2], and here we want to address this challenge in the context of this family of receptors [34, 2]. The advantages of having analytical (or closed-form) expressions of the amplitude and the EC_{50} for a large class of receptor-ligand systems are many: (i) they allow one to quantify their dependence on receptor copy numbers, (ii) they facilitate mathematical model validation and parameter exploration, and (iii) they reduce computational cost. To the best of our knowledge such expressions have been obtained in a few instances: closed or open bimolecular receptor-ligand systems [21], monomeric receptors [38], or ternary complexes [13]. More complicated receptor-ligand models have been studied with chemical reaction network theory (CRNT) [18, 44, 55], but CRNT has thus far, been focused on the analysis of the steady state of the system (i.e., existence and number of steady states and their stability). Yet, we believe CRNT is an essential and useful framework to start any mathematical investigation of the amplitude and the EC_{50}.

Another aspect which can be effectively addressed by mechanistic mathematical modeling is the effect of internal or external perturbations to the state of a cell. For example, in single-cell experiments or even repetitions of bulk experiments [8, 20], the experimental conditions can never be replicated exactly. This leads to noise not only in the measured quantities, but also in the reaction mechanisms themselves. This variation can be captured in mathematical models which encode parameters such as affinity constants or total copy number of constituent molecular species. An analytical study of the dependency of pharmacologically relevant quantities, such as amplitude and EC_{50}, on the reaction parameters can facilitate in silico drug design [43]. While amplitude and EC_{50} are widely employed to characterize biological phenomena, the manner in which they depend on the parameters of the receptor-ligand model is not fully understood. Thus, improved understanding of these relationships could provide novel biological and computational insights.

Motivated by the previous challenges and making use of methods from CRNT and algebraic geometry, such as the Gröbner basis, in this paper we propose a new method to obtain analytic expressions of the amplitude and the EC_{50} for a large class of receptor-ligand models, with a focus on cytokine receptors. The paper is organized as follows. In section 2 we introduce the mathematical background and essential notions of CRNT used in the following sections. With the IL-7 cytokine receptor as a paradigm, in section 3 we propose a general method to calculate the amplitude and the

EC_{50} of the dose-response curve for a class of receptor-ligand systems. In section 4 we generalize the previous results to a wider class of receptor-ligand systems, sequential receptor-ligand systems with extrinsic kinase, and provide a few biological examples of these systems. Finally, we discuss and summarize our results in section 5. We have included an appendix to provide additional details of our methods (perturbation theory) and our algebraic computations.

2. Mathematical background. In this section we briefly summarize the relevant notions of CRNT and formally define amplitude, EC_{50}, and signaling function. A very short introduction to the use of Gröbner bases is also given.

2.1. A brief introduction to CRNT. In this paper we view a chemical reaction network (CRN), $\mathcal{N}$, as a multiset $\mathcal{N} = \{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$, where $\mathcal{S}$ is the set of species, $\mathcal{C}$ the set of complexes, and $\mathcal{R}$ the set of reactions. We note that in the context of CRN, a "complex" is a linear combination of species and need not be a "biological functional unit," which we refer to as a *biological complex*. We denote, whenever useful, a biological complex formed by species X and Y as $X : Y$, where the colon denotes the physical bond between X and Y. The order of species in the biological complex is irrelevant, i.e., $X : Y = Y : X$.

Example 2.1 (heterodimeric receptor tyrosine kinase). A simple heterodimeric receptor tyrosine kinase (RTK) model has a species set $\mathcal{S} = \{X_1, X_2, Y_1, Y_2\}$, a complex set $\mathcal{C} = \{X_1 + X_2, Y_1, Y_2\}$, and a reaction set $\mathcal{R} = \{X_1 + X_2 \to Y_1, Y_1 \to X_1 + X_2, Y_1 \to Y_2, Y_2 \to Y_1\}$. Ligand binding induces dimerization of these receptors resulting in autophosphorylation of their cytoplasmic domains (tyrosine autophosphorylation sites) [54]. X_1 and X_2 are the two components of the heterodimeric RTK. The biological complexes $Y_1 = X_1 : X_2$ and $Y_2 = L : Y_1$ are the heterodimeric receptor with intrinsic kinase activity and the heterodimeric receptor bound to the ligand, respectively. In this paper the ligand concentration (L) is taken to be an input parameter and, hence, it does not feature as a separate chemical species in the species set $\mathcal{S}$.

We can associate a reaction graph to every CRN $\mathcal{N}$, by letting the vertex set be $\mathcal{C}$ and the (directed) edge set $\mathcal{R}$. There exists a class of important CRNs defined by their network reversibility.

DEFINITION 2.2 (network reversibility). *Let $\mathcal{N}$ be a CRN with its associated reaction graph $\mathcal{G}(\mathcal{C}, \mathcal{R})$. An edge between C_i and $C_j \in \mathcal{C}$ exists if $C_i \to C_j \in \mathcal{R}$. If for every edge $C_i \to C_j \in \mathcal{R}$, the edge $C_j \to C_i \in \mathcal{R}$ also exists, then the network is called reversible. If for every edge, $C_i \to C_j \in \mathcal{R}$, a directed path exists going back from C_j to C_i, then the network is called weakly reversible. All reversible networks are also weakly reversible.*

A general reaction from complex C_i to complex C_j can be written as

$$\sum_{k=1}^{n} \alpha_{ik} X_k \to \sum_{k=1}^{n} \alpha_{jk} X_k, \tag{2.1}$$

where the sum is over the set of species $(X_1, X_2, \ldots, X_n)$, and $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{in})^T$ and $\alpha_j = (\alpha_{j1}, \ldots, \alpha_{jn})^T$ are nonnegative integer vectors. The corresponding reaction vector is given by $r = \alpha_j - \alpha_i$. For a CRN with n species and m reactions we can now define the $n \times m$ matrix of all reaction vectors, Γ, such that $\Gamma = (r_1, \ldots, r_m)$. This matrix is called the *stoichiometric matrix*.

Example 2.3 (heterodimeric RTK continued). The reaction graph of the heterodimeric RTK model is given by

$$X_1 + X_2 \rightleftharpoons Y_1 \rightleftharpoons Y_2.$$

The model is reversible with reaction vectors $r_1 = (-1,-1,1,0)^T$, $r_2 = (1,1,-1,0)^T$, $r_3 = (0,0,-1,1)^T$, and $r_4 = (0,0,1,-1)^T$.

To derive dynamical properties from the static description so far provided, we make use of the law of mass action kinetics [27]. First, we assign a rate constant $k_j \in \mathbb{R}_{>0}$ with $j \in \{1,\dots,m\}$, to each reaction in the network. Second, we denote the concentration of species X_i by x_i. With this notation, we then associate a monomial to every complex $C_i = \sum_k \alpha_{ik} X_k$, as follows,

$$x^{\alpha_i} = x_1^{\alpha_{i1}} \cdots x_n^{\alpha_{in}}, \tag{2.2}$$

where n is the number of species in the network. We define the reactant complex of a reaction as the complex on the left-hand side of reaction (2.1). The reaction rate of a reaction is the monomial of its reactant complex multiplied by the rate constant. The flux vector, $R(x)$, is the $m \times 1$ column vector of all reaction rates. The ordinary differential equations (ODEs) governing the dynamics of the reaction network are given by

$$\frac{dx}{dt} = \Gamma\, R(x), \tag{2.3}$$

where Γ is the stoichiometric matrix (defined above). We note that the reaction rate of the i^{th} reaction is the i^{th} row in $R(x)$, and similarly, the stoichiometry of ith reaction is given by the ith column of Γ.

From (2.3) we can also deduce the conserved quantities of the reaction network. That is, if a vector exists, $z \in \mathbb{Z}^n$, such that $d(z^T x)/dt = z^T \Gamma R(x) = 0$, the quantity $z^T x$ is conserved. Consequently, the left kernel of Γ defines a basis for the space of conserved quantities. In this way, conservations induce linear relations between the variables. Informally we say that a molecular species X_i is conserved if its total number of molecules, N_i, is constant. N_i is determined by the initial conditions.

Example 2.4 (heterodimeric RTK continued). The dynamical system associated with the heterodimeric RTK model is given by

$$\begin{aligned}\frac{dx}{dt} = \frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} &= \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} k_1 x_1 x_2 \\ q_1 y_1 \\ k_2 y_1 \\ q_2 y_2 \end{pmatrix} \\ &= \begin{pmatrix} -k_1 x_1 x_2 + q_1 y_1 \\ -k_1 x_1 x_2 + q_1 y_1 \\ k_1 x_1 x_2 - (q_1 + k_2) y_1 + q_2 y_2 \\ k_2 y_1 - q_2 y_2 \end{pmatrix}\end{aligned}$$

with k_i as the reaction constants of the forward chemical reactions ($\rightharpoonup$) and q_i the reaction constants of the backward reactions ($\leftharpoondown$). A basis for the conservation equations is given by the linear relations $X_1 + Y_1 + Y_2 = N_1$ and $X_2 + Y_1 + Y_2 = N_2$. These imply that the total amount of the species X_1 (X_2) is conserved by adding the amounts of the bound states of the molecule (Y_1 and Y_2) to the amount of free molecule X_1 (X_2).

We can now define the biologically relevant steady states of a CRN.

DEFINITION 2.5. *A vector x^* is a biologically relevant steady state if $\Gamma R(x^*) = 0$ and $x_i^* > 0 \, \forall i \in \{1, \ldots, n\}$.*

A useful connection between the static network structure (defined earlier) and the existence (and stability) of unique biologically relevant steady states can be made via deficiency theory [18].

DEFINITION 2.6 (deficiency). *Let $\mathcal{N}$ be a CRN with ℓ connected components in the reaction graph and $\eta = \dim \operatorname{span}(r_1, \ldots, r_m)$ be the dimension of the span of the reaction vectors. The deficiency of $\mathcal{N}$ is then given by*

$$\delta = |\mathcal{C}| - \ell - \eta.$$

The notion of network deficiency leads to one of the fundamental theorems of CRNT, the d*eficiency zero theorem* [18], which connects the network structure to the dynamics of a CRN.

THEOREM 2.7 (deficiency zero theorem). *Let $\mathcal{N}$ be a weakly reversible CRN with $\delta = 0$. Then the network has a unique biologically relevant steady state for every set of initial conditions, and this steady state is asymptotically stable.*

With certain additional conditions on the reaction rates (see [19, 12]), biologically relevant steady states are detailed balanced. This means that for every reaction of the form (2.1), the steady states satisfy

$$K x^{\alpha_i} = x^{\alpha_j},$$

where $K = k/q$, the ratio of the rate constants of the forward and backward reactions, is called the *affinity constant* of the reaction.

Example 2.8 (heterodimeric RTK continued). The heterodimeric RTK model has 3 complexes, 1 connected component and the dimension of the span of the reaction vectors is 2; hence, $\delta = 3 - 1 - 2 = 0$. Since the network is reversible, we know from Theorem 2.7 that there exists exactly one stable positive steady state for each set of initial conditions. One can show that in fact $y_1 = (k_1/q_1)x_1x_2$ and $y_2 = (k_2/q_2)y_1$.

2.2. Signaling function: Amplitude and half-maximal effective concentration. In this paper we want to closely investigate pharmacological properties of receptor-ligand systems, rather than the steady state structure of the models. In particular, we want to study the dose-response (or concentration-effect) curve of the system, which describes the relation between ligand concentration and the biological effect (or cellular response) it generates when binding its specific cell surface receptor. As mentioned in section 1 a lot of effort has been devoted to exploring the steady state structure of chemical reaction networks. In this paper we make use of algebraic methods to explore the dose-response of receptor-ligand systems. To do so we start with the definition of signaling complex. We note that in most biological instances the signaling complex is formed by all the subunit chains that make up the full receptor, intra-cellular kinases, and the specific ligand [28, 60, 23, 32, 20, 8, 14].

DEFINITION 2.9. *The signaling complex of a receptor-ligand system is the biological complex which induces a biological response.*

This leads to the following definitions of the signaling function and dose-response curve.

DEFINITION 2.10. *We define the signaling function,* $\sigma : \mathbb{R}_+ \to \mathbb{R}_+, L \mapsto \sigma(L)$, *as the univariate function which assigns to a given value of ligand concentration,* L, *the number (or concentration) of signaling complexes formed at steady state. The dose-response curve is the corresponding plot of the signaling function.*

We note that in what follows we will not distinguish between number (or concentration) of signaling complexes since one can be obtained from the other if we know the volume of the system and Avogadro's number.

The specific choice of σ will depend on the receptor-ligand system under consideration. In this paper we focus on the class of cytokine receptors and the signaling function will be defined in section 3. In our examples the signaling function will be a product of the steady state values (numbers) of subunit chains that make up the full receptor, intracellular kinases, affinity constants of the reactions involved, and ligand concentration. This together with (2.2) and (2.3) indicates that the signaling function will always be algebraic. Next, we define a central object of study in this paper, namely, the amplitude of the signaling function, often referred as efficacy in the pharmacology literature [39].

DEFINITION 2.11. *The amplitude of the signaling function,* A, *is the difference between the maximum and the minimum of* σ, *that is,* $A \equiv max(\sigma) - min(\sigma)$.

We note that when $\min(\sigma) = 0$, which is the case considered in this paper ($\min(\sigma) = \sigma(0) = 0$), the amplitude is given by the maximum of the signaling function. If, in addition, the dose-response curve attains its maximum at large concentrations (for instance, when the dose-response curve is sigmoid), we have

$$A = \lim_{L \to +\infty} \sigma(L). \tag{2.4}$$

The amplitude provides information about the magnitude of the intracellular response to the stimulus, L. The larger the amplitude is, the larger the response variability will be. The amplitude is always bounded by the number of molecules available. However, this bound is often not tight [47]. To quantify the sensitivity of the model to the stimulus, i.e., the potency of the ligand L, we introduce the *half-maximal effective concentration*, EC_{50}.

DEFINITION 2.12. *The half-maximal effective concentration, or* EC_{50}, *is the ligand concentration* L^* *which satisfies* $\sigma(L^*) = min(\sigma) + \frac{max(\sigma) - min(\sigma)}{2} = min(\sigma) + \frac{A}{2}$.

We say that the EC_{50} is inversely proportional to ligand potency, namely, the lower the EC_{50}, the higher the potency of the ligand. Figure 1 illustrates the amplitude and the EC_{50} of a sigmoid dose-response curve (A) when its minimum is zero: increasing the amplitude shifts up the maximum of the curve and results in greater efficacy (B), and decreasing the EC_{50} shifts the dose-response curve to the left and increases the potency of the ligand (C). We now review some algebraic and analytic tools which will enable us to compute the EC_{50} and the amplitude.

2.3. Gröbner bases. Since we assume the law of mass action, the models studied in this paper are systems of polynomial equations and, thus, we can use the techniques developed in the field of computational algebra and algebraic geometry [9]. Such methods have also been successfully applied to many topics in CRNT, see e.g., [26, 52, 11]. In particular, we make use of Gröbner bases. Informally speaking, a Gröbner basis is a nonlinear generalization of the concept of a basis in linear algebra and, therefore when a Gröbner basis for a polynomial system is calculated, many properties of the system can be investigated, such as the number of solutions and the

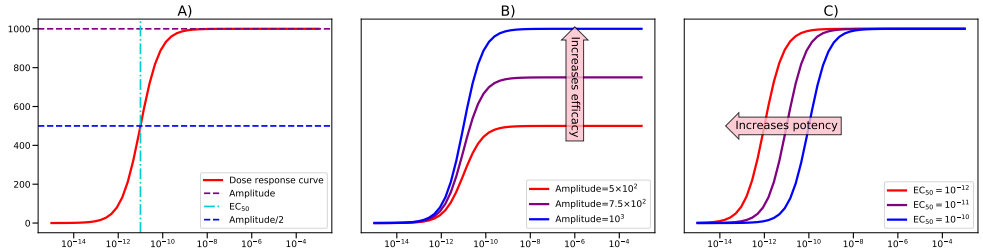


FIG. 1. *Sigmoid dose-response curve: number of signaling complexes formed, $\sigma(L)$, as a function of the concentration of ligand L (arbitrary units).* (A) *The maximum value defines the amplitude. The EC_{50} is the concentration of ligand which corresponds to half the amplitude.* (B) *Three dose-response curves with the same EC_{50} value and different amplitudes. Increasing the amplitude shifts up the maximum of the curve and increases the efficacy.* (C) *Three dose-response curves with the same amplitude and different EC_{50} values. Decreasing the EC_{50} shifts the dose-response curve to the left and increases the potency of the ligand.*

dimensionality of the space of solutions. Strictly speaking, however, a Gröbner basis is not a basis as it is not unique and it depends on the monomial ordering chosen (for instance, lexicographical (lex) order or graded reverse lex order). For more details we refer the reader to [9].

A lex Gröbner basis is a triangular polynomial system; that is, for a polynomial system (ideal) in $\mathbb{Q}[x_1,\dots,x_n]$ we obtain a polynomial system of the form

$$g_n(x_1,\dots,x_n) = g_{n-1}(x_1,\dots,x_{n-1}) = \cdots = g_1(x_1) = 0. \tag{2.5}$$

Other monomial orderings may result in triangular systems, however, there is no theoretical guarantee that this should be the case. When the solution space is positive dimensional, then $g_1,\dots,g_n$ are identically zero for an infinite number of points. For a given Gröbner basis with a zero-dimensional solution space we can now iteratively, and often numerically, solve the constituent polynomials to obtain a finite number of solutions (in $\mathbb{C}^n$) for the polynomial system. We can also find all real and, further, positive solutions, if there are any [9].

Remark 2.13. We note that computing the solution of the polynomial system iteratively only works if partial solutions of the polynomials forming a Gröbner basis are not zeros of the leading coefficients of any of the other polynomials in the Gröbner basis. We assume this to be the case in this paper. Indeed, for all the examples considered here, this is always the case.

Remark 2.14. Technically, in this paper, we consider Gröbner bases over a parametric coefficient field, e.g., $\mathrm{frac}(\mathbb{Q}[k_1,\dots,k_m])$, where $\mathrm{frac}(\cdot)$ denotes the fraction field of the ring $\mathbb{Q}[k_1\dots,k_m]$. If we restrict ourselves to a specific set of rate constants, we are not guaranteed that the "generic" Gröbner basis obtained is indeed a Gröbner basis for this restriction, i.e., the Gröbner basis may not be comprehensive [63]. One way to construct a comprehensive Gröbner basis is via the Gröbner cover (`GrobCov`) algorithm [42]. We assert, however, that for most practical cases this technicality will not be required.

3. Methods: analytical study of receptor-ligand systems. In this section we first outline the computation of the analytic expressions of the steady state, amplitude and EC_{50} for two IL-7 receptor (IL-7R) models. These two examples then allow us to introduce a more general method to analytically compute the amplitude and the EC_{50} of receptor-ligand systems under the following hypotheses:

1. The system is in steady state.
2. The ligand is in excess (we consider ligand concentration, L, as a parameter instead of a dynamic variable).
3. A unique biologically relevant solution exists for any given set of rate constants and initial conditions.

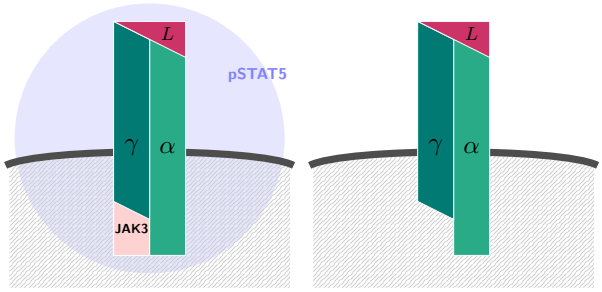
The IL-7R models we have chosen are simple enough to illustrate our method and, thus, to derive analytic expressions for the amplitude and the EC_{50}, yet complex enough to show its limitations.

3.1. Two motivating examples: IL-7 cytokine receptor as a paradigm. We now consider the cytokine interleukin-7 (IL-7) and its receptor (IL-7R) [45, 50, 23, 32, 8, 46] as a motivating receptor-ligand system. IL-7 is a cytokine involved in T cell development, survival, and homeostasis [36]. Its receptor, IL-7R, is displayed on the surface of T cells and is composed of two transmembrane chains: the common gamma chain (denoted by γ) and the specific high affinity chain IL-7Rα (denoted by α) [23, 46, 36, 41]. Cytokine receptors do not contain intrinsic kinase domains and, thus, make use of Janus family tyrosine kinases (JAKs) and signal in part by the activation of signal transducer and activator of transcription (STAT) proteins [35]. In the case of the gamma chain, it binds to the intracellular extrinsic Janus kinase molecule, JAK3. Binding of IL-7 to the dimeric JAK3-bound IL-7R, defined as $\alpha:\gamma:JAK3$, initiates a series of biochemical reactions from the membrane of the cell to its nucleus, which in turn lead to a cellular response. For the IL-7R system the STAT protein preferentially activated is STAT5 [35], so that the amount of phosphorylated STAT5 can be used as the experimental measure of the intracellular response generated by the IL-7 stimulus. The IL-7R receptor is illustrated in Figure 2(a), where the hatched area determines the intracellular environment.

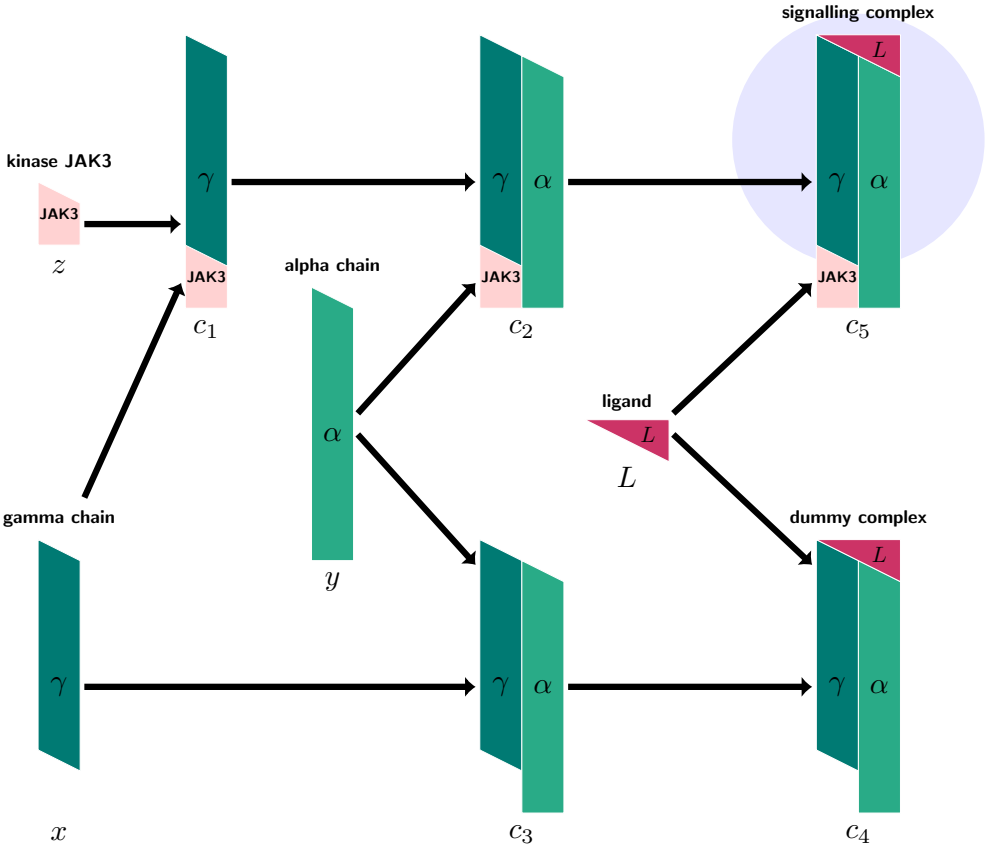
The first model we consider is shown in Figure 2(c). As discussed in [50], the gamma chain is shared by other cytokine receptors. This model does not include the competition for the gamma chain between different cytokine receptors, therefore, later in this section we introduce a second model to account for this competition. In this section we will provide an (algebraic) analytic treatment of both models. We consider the formation of dummy receptors, $\alpha:\gamma$, which are formed of the IL-7R devoid of JAK3 and, therefore, they cannot signal (see Figure 2(b)). We further assume no allostery, that is, the affinity constants of the biochemical reactions involved in the formation of the dummy complex, $L:\alpha:\gamma$, are the same as the affinity constants involved in the formation of the signaling complex, $L:\alpha:\gamma:JAK3$.

3.1.1. The IL-7 receptor-ligand system: Two receptor chains and a kinase. We first consider a model in which the IL-7R is formed sequentially, one molecule at a time; the γ chain binds to the kinase, JAK3, then the α chain binds to the complex formed by γ and JAK3. Finally, the ligand, IL-7, binds to the signaling receptor composed of γ, α, and $JAK3$. The model also includes the formation of dummy receptors, which do not involve the kinase JAK3. Figure 2(c) illustrates the sequential formation of the signaling and dummy complexes. The reaction scheme for this model is as follows,

$$
\begin{array}{lcll}
\gamma + JAK3 & \rightleftharpoons & \gamma:JAK3, & K_1,\\
\alpha+\gamma:JAK3 & \rightleftharpoons & \alpha:\gamma:JAK3, & K_2,\\
\alpha+\gamma & \rightleftharpoons & \alpha:\gamma, & K_2,\\
L+\alpha:\gamma & \rightleftharpoons & L:\alpha:\gamma, & K_3,\\
L+\alpha:\gamma:JAK3 & \rightleftharpoons & L:\alpha:\gamma:JAK3, & K_3,
\end{array}
\tag{3.1}
$$



(a) Signalling IL-7 complex. (b) Dummy IL-7 complex.



(c) IL-7R model: sequential chemical reaction scheme.

FIG. 2. *First IL-7R model:* (a) *The IL-7 receptor is composed of the transmembrane* γ *and* α *chains. The* γ *chain can bind the intracellular downstream kinase JAK3. When the ligand, IL-7, binds the full receptor, it phosphorylates STAT5.* (b) *The IL-7R model allows the formation of "dummy" complexes: IL-7 bound IL-7R complexes, devoid of JAK3, which are unable to induce intracellular signaling.* (c) *IL-7 bound IL-7R complexes with JAK3 are able to induce intracellular signaling and, thus, are called "signaling" complexes. IL-7R dummy and signaling complexes are formed sequentially. The mathematical notation used in this paper is shown below each molecule or complex.*

where for $i = 1, 2, 3$, K_i is the affinity constant of the appropriate reaction. One can show that this system has deficiency zero and is reversible (see section 2.1). Therefore, for every set of rate constants and initial conditions, there exists exactly one positive steady state. Moreover, this positive steady state is in detailed balance. We remind the reader that in this paper we assume mass action kinetics to determine reaction rates. We denote the concentration of γ, α, JAK3, and IL-7 by x, y, z, and L, respectively. The reaction rate for the forward/backward reaction ($\rightharpoonup/\leftharpoondown$) is given by k_i and q_i, respectively, for $i = 1, 2, 3$. We note that $K_i = k_i/q_i$. The concentrations of the product complexes of the forward reactions are denoted by c_i in order of appearance (see Figure 2(c)). We can now write down the ODEs associated with the system of reactions (3.1):

$$
\begin{aligned}
\frac{dx}{dt} &= -k_1xz + q_1c_1 - k_2xy + q_2c_3,\\
\frac{dy}{dt} &= -k_2yc_1 + q_2c_2 - k_2xy + q_2c_3,\\
\frac{dz}{dt} &= -k_1xz + q_1c_1,\\
\frac{dc_1}{dt} &= k_1xz - q_1c_1 - k_2yc_1 + q_2c_2,\\
\frac{dc_2}{dt} &= k_2yc_1 - q_2c_2 - k_3c_2L + q_3c_5,\\
\frac{dc_3}{dt} &= k_2xy - q_2c_3 - k_3c_3L + q_3c_4,\\
\frac{dc_4}{dt} &= k_3c_3L - q_3c_4,\\
\frac{dc_5}{dt} &= k_3c_2L - q_3c_5.
\end{aligned}
\tag{3.2}
$$

A suitable basis for the conservation equations is

$$
\begin{aligned}
N_x &= x + c_1 + c_2 + c_3 + c_4 + c_5,\\
N_y &= y + c_2 + c_3 + c_4 + c_5,\\
N_z &= z + c_1 + c_2 + c_5,
\end{aligned}
\tag{3.3}
$$

that is, single chain molecules are conserved since we do not consider the generation or degradation of molecules. The constants N_x, N_y, and N_z represent the total copy number of γ, α, and JAK3 molecules per cell, respectively. Detailed balance [19] leads to the following steady state equations:

$$
\begin{aligned}
c_1 &= K_1xz,\\
c_2 &= K_2yc_1,\\
c_3 &= K_2xy,\\
c_4 &= K_3Lc_3,\\
c_5 &= K_3Lc_2.
\end{aligned}
\tag{3.4}
$$

Substituting the steady state equations into the conservation equations, we obtain a system of polynomials:

$$
\begin{aligned}
0 &= -N_x + x + K_1xz + K_2K_1xyz + K_2xy + K_3K_2Lxy + K_3K_2K_1Lxyz,\\
0 &= -N_y + y + K_2K_1xyz + K_2xy + K_3K_2Lxy + K_3K_2K_1Lxyz,\\
0 &= -N_z + z + K_1xz + K_2K_1xyz + K_3K_2K_1Lxyz.
\end{aligned}
\tag{3.5}
$$

Analytic computation of the steady state. The polynomial system (3.5) can be solved numerically for a particular set of parameter values. However, an analytic solution will provide greater insight and will allow us to derive expressions for the amplitude and the EC_{50}. We make use of Macaulay2 [25] to compute a lex Gröbner basis for this model, which will lead to a triangular set of polynomials,[1] as follows:

$$0 = z^2 + \frac{[1 + K_1(N_x - N_z)]}{K_1} z - \frac{N_z}{K_1}, \tag{3.6a}$$

$$0 = y^2 + \frac{[1 + K_2(K_3L + 1)(N_x - N_y)]}{K_2(K_3L + 1)} y - \frac{N_y}{K_2(K_3L + 1)}, \tag{3.6b}$$

$$0 = x - \frac{1}{N_x} yz - \frac{(N_x - N_z)}{N_x} y - \frac{(N_x - N_y)}{N_x} z - \frac{N_x(N_x - N_y - N_z) + N_y N_z}{N_x}. \tag{3.6c}$$

We make use of the `GrobCov` algorithm to assert that (3.6) is a Gröbner basis for all positive parameter configurations. Equation (3.6c) gives

$$x = \frac{(N_x - N_y + y)(N_x - N_z + z)}{N_x} = \frac{N_x - N_y + y}{1 + K_1 z},$$

where the last equality follows from (3.6a). Solving the system (3.6) and selecting the biologically relevant solution, we obtain an analytic expression for the number of free (unbound) JAK3, α, and γ molecules at steady state,

$$z = \frac{-1 + K_1(N_z - N_x) + \sqrt{\Delta_1}}{2K_1}, \tag{3.7a}$$

$$y = \frac{-1 + K_2(N_y - N_x)(K_3L + 1) + \sqrt{\Delta_2}}{2K_2(K_3L + 1)}, \tag{3.7b}$$

$$x = \frac{N_x - N_y + y}{1 + K_1 z}, \tag{3.7c}$$

where we have introduced

$$\Delta_1 = 4K_1N_z + [K_1(N_x - N_z) + 1]^2,$$

and

$$\Delta_2 = 4K_2N_y(K_3L + 1) + [K_2(N_x - N_y)(K_3L + 1) + 1]^2.$$

We study the dose-response curve of this model given by the number of signaling complexes, $L : \gamma : \alpha : JAK$, per cell at steady state and as a function of L. The signaling function, $\sigma(L)$, is given by

$$\sigma(L) \equiv c_5 = K_3K_2K_1Lxyz. \tag{3.8}$$

Analytic computation of the amplitude. A simple inspection of the behavior of (3.7) shows that the dose-response curve is a sigmoid, such that $\sigma(0) = 0$. Therefore the amplitude A is given by the asymptotic behavior of the signaling function as follows:

$$A \equiv \lim_{L \to +\infty} \sigma(L). \tag{3.9}$$

[1] Example code is provided in Appendix C.

We will prove this result rigorously for a more general class of models in section 4.

We first notice that z is independent of L. We now compute the product xy (at steady state) as follows:

$$xy = \frac{(N_x - N_y)y + y^2}{1 + K_1 z}.$$

From (3.6b) we can replace the polynomial in y of degree two by an expression linear in y:

$$(N_x - N_y)y + y^2 = \frac{N_y - y}{K_2(K_3 L + 1)}.$$

Thus, we obtain the following analytic expression for the signaling function:

$$\sigma(L) = K_3 K_2 K_1 L x y z = \frac{K_1 z}{(1 + K_1 z)} \frac{K_3 L}{(K_3 L + 1)} (N_y - y). \tag{3.10}$$

Since $\frac{K_3 L}{1+K_3 L} \to 1$ when $L \to +\infty$, we need to study the expression $N_y - y$ in this limit. We have

$$N_y - y = \frac{(N_y + N_x)K_2(K_3 L + 1) + 1 - \sqrt{\Delta_2}}{2K_2(K_3 L + 1)}, \tag{3.11}$$

where

$$\Delta_2 = K_2^2 (K_3 L + 1)^2 (N_x - N_y)^2 + 2K_2(K_3 L + 1)(N_x + N_y) + 1.$$

Keeping to lowest order in $\mathcal{O}(\frac{1}{L})$ we obtain

$$\begin{aligned} N_y - y &= \frac{1 + (N_x + N_y)K_2(K_3 L + 1) - K_2(K_3 L + 1)|N_x - N_y|(1 + \mathcal{O}(\frac{1}{L}))}{2K_2(K_3 L + 1)} \\ &= \frac{N_x + N_y - |N_x - N_y|}{2} + \mathcal{O}\left(\frac{1}{L}\right). \end{aligned} \tag{3.12}$$

Finally, noticing that

$$\frac{N_x + N_y - |N_x - N_y|}{2} = \min(N_x, N_y),$$

we obtain the amplitude

$$A = \min(N_x, N_y) \frac{K_1 z}{1 + K_1 z}, \tag{3.13}$$

where z is the analytic expression obtained in (3.7). This result indicates that the amplitude of this model is the total number of the limiting transmembrane chain modulated by a factor, valued in the interval $[0, 1]$, which only depends on K_1, N_x, and N_z.

Analytic computation of the EC_{50}. We now determine the EC_{50} by finding the value of L_{50} such that

$$\sigma(L_{50}) = \frac{A}{2} = K_1 K_2 K_3 L_{50} x_{50} y_{50} z_{50}, \tag{3.14}$$

where x_{50}, y_{50}, and z_{50} are the steady state expressions found in (3.7) evaluated at $L = L_{50}$. Two expressions satisfy this equation but only one provides a relevant

biological solution with $L, x, y, z > 0$. The relevant analytic expression of the EC_{50} is given by

(3.15)
$$EC_{50} = M\,\frac{1 + K_2(N_x + N_y - M) + \sqrt{1 + K_2^2(N_y - N_x)^2 + 2K_2(N_x + N_y - M)}}{K_2K_3(M - 2N_x)(M - 2N_y)},$$

with $M = \min(N_x, N_y)$. The details of the computation can be found in Appendix B. This result shows that the EC_{50} value for this system is independent of the kinase, since the parameters K_1 and N_z are absent in the previous expression.

Alternatively, we now propose a more algebraic method to derive the analytic expression of the EC_{50}. We compute a lex Gröbner basis for the augmented system of polynomials consisting of the steady state equations (3.5) and

$$K_1K_2K_3Lxyz\,(1 + K_1z) - \frac{MK_1z}{2} = 0, \tag{3.16}$$

where this time x, y, z, and L are variables. The resulting triangular system describes directly the EC_{50} and x, y, z at $L = EC_{50}$:

$$0 = L^2 + \frac{2M[-1 + K_2(M - N_x - N_y)]}{K_2K_3(M - 2N_x)(M - 2N_y)}L + \frac{M^2}{K_3^2(M - 2N_x)(M - 2N_y)}, \tag{3.17a}$$

$$0 = z^2 + \frac{1 + K_1(N_x - N_z)}{K_1}z - \frac{N_z}{K_1}, \tag{3.17b}$$

$$0 = y - \frac{K_3(M - 2N_x)(M - 2N_y)}{2M}L + \frac{K_2(2N_x - M) + 2}{2K_2}, \tag{3.17c}$$

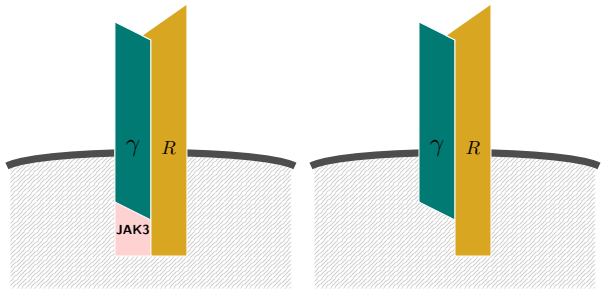
$$0 = x - \frac{K_3(M - 2N_x)(M - 2N_y)(N_x - N_z + z)}{2MN_x}L + \frac{[2 + K_2(2N_y - M)](N_x - N_z + z)}{2K_2N_x}. \tag{3.17d}$$

This set of polynomials is a Gröbner basis for all positive parameter configurations. Solving (3.17a) and selecting the solution for which y and x, given by (3.17c) and (3.17d), respectively, are positive yields the final result, in agreement with (3.15).

3.1.2. The IL-7 receptor-ligand system: An additional subunit receptor chain. The previous model described the IL-7 receptor system without any consideration for the fact that the γ chain is shared with other cytokine receptors [50]. We now account for this competition by including in the previous model an additional receptor chain, R, which can bind to the γ chain, or the complex $\gamma : JAK3$, to form decoy receptor complexes (see Figures 3(a) and 3(b), where the hatched area indicates the cytoplasmic region). The resulting reaction scheme (summarized in Figure 3(c)) is given by

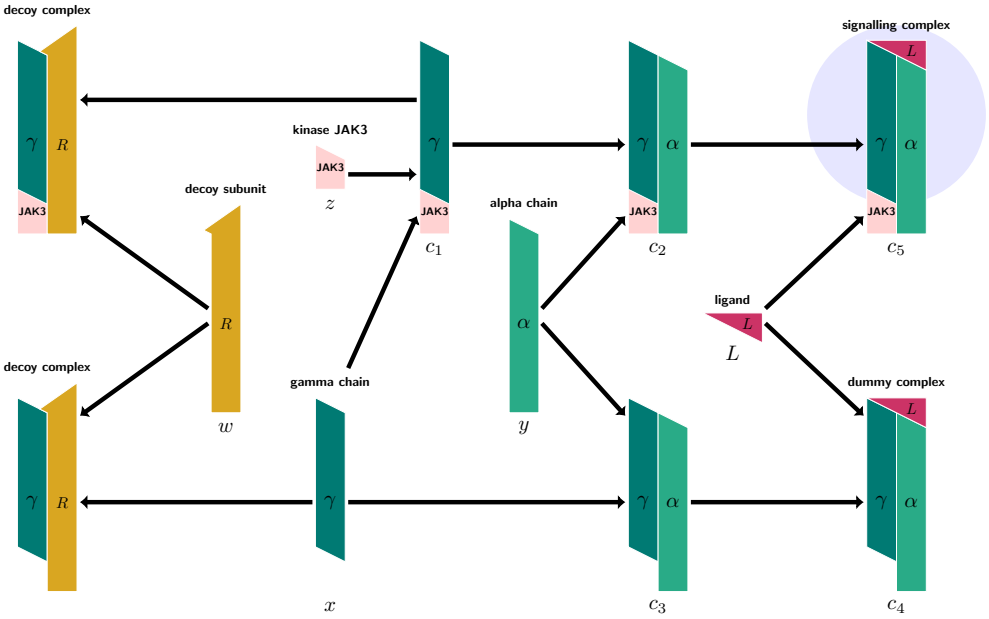
$$\begin{array}{lcll}
JAK3 + \gamma & \rightleftharpoons & JAK3 : \gamma, & K_1,\\
\alpha + JAK3 : \gamma & \rightleftharpoons & \alpha : \gamma : JAK3, & K_2,\\
\alpha + \gamma & \rightleftharpoons & \alpha : \gamma, & K_2,\\
L + \alpha : \gamma & \rightleftharpoons & L : \alpha : \gamma, & K_3,\\
L + \alpha : \gamma : JAK3 & \rightleftharpoons & L : \alpha : \gamma : JAK3, & K_3,\\
R + \gamma & \rightleftharpoons & R : \gamma, & K_4,\\
R + JAK3 : \gamma & \rightleftharpoons & R : \gamma : JAK3, & K_4.
\end{array}$$

We use w to describe the concentration of the additional chain R. Similarly to the previous model, we write the system of ODEs describing the time evolution for each



(a) Decoy complex with kinase.

(b) Decoy complex without kinase.



(c) Second IL-7R model: sequential chemical reaction scheme.

FIG. 3. *IL-7R model with an additional receptor subunit. The signaling and dummy complexes are the same as in the first IL-7R model. This second model allows the formation of decoy complexes:* (a) *with the kinase JAK3 or* (b) *without the kinase.* (c) *The IL-7R dummy and signaling complexes are formed sequentially. Decoy complexes can be formed to prevent the formation of signaling or dummy complexes. The mathematical notation used is annotated below each molecule or complex.*

complex and then derive (a basis for) the conservation and steady state equations. Combining them, we obtain the following polynomial system,

$$
\begin{aligned}
0 &= -N_x + x + K_2xy + K_1xz + K_2K_1xyz + K_3K_2Lxy \\
&\quad + K_3K_2K_1Lxyz + K_4xw + K_1K_4xwz, \\
0 &= -N_y + y + K_2xy + K_2K_1xyz + K_3K_2Lxy + K_3K_2K_1Lxyz, \\
0 &= -N_z + z + K_1xz + K_2K_1xyz + K_3K_2K_1Lxyz + K_1K_4xwz, \\
0 &= -N_w + w + K_4xw + K_4K_1xwz,
\end{aligned}
\tag{3.18}
$$

where N_w is the additional conserved quantity. Again, we compute a lex Gröbner basis for this set of polynomials to obtain the following triangular system,

$$0 = K_1 z^2 + z[1 + K_1(N_x - N_z)] - N_z, \tag{3.19a}$$
$$0 = Ay^3 + By^2 + Cy + D, \tag{3.19b}$$
$$0 = [K_2K_4(1 + K_3L)N_xN_y]\, x + (Ay^2 + By + C + K_4N_y)(N_x - N_z + z), \tag{3.19c}$$
$$0 = [K_2K_4(1 + K_3L)N_y]\, w + Ay^2 + By + [K_2(1 + K_3L) - K_4]N_y, \tag{3.19d}$$

where

$$A = -K_2(1 + K_3L)[K_2(1 + K_3L) - K_4],$$
$$B = K_4 - K_2(1 + K_3L)[1 + K_4(N_w - N_x + 2N_y) + K_2(1 + K_3L)(N_x - N_y)],$$
$$C = N_y[-2K_4 + K_2(1 + K_3L)(1 + K_4(N_w - N_x + N_y))],$$
$$D = K_4N_y^2.$$

Using the `GrobCov` algorithm we show that this set of polynomials is a Gröbner basis for all positive rate constants except when $LK_2K_3 + K_2 - K_4 = 0$. Since this relation defines a measure zero set in parameter space, we ignore it for now. Solving (3.19a) gives the number of free JAK3 molecules per cell at steady state, z; solving (3.19b) gives the number of free (unbound) α chains per cell, y; and substituting y and z into (3.19c) and (3.19d) gives the remaining steady states. We obtain the following implicit steady state expressions for the number of free (unbound) chains:

$$\begin{aligned} z &= \frac{-1 + K_1(N_z - N_x) + \sqrt{[1 + K_1(N_x - N_z)]^2 + 4N_zK_1}}{2K_1}, \\ x &= -\frac{(Ay^2 + By + C + K_4N_y)(N_x - N_z + z)}{K_2K_4(1 + K_3L)N_xN_y}, \\ w &= -\frac{Ay^2 + By + [K_2(1 + K_3L) - K_4]N_y}{K_2K_4(1 + K_3L)N_y}. \end{aligned} \tag{3.20}$$

The problem now reduces to finding the positive real roots of (3.19b). As (3.19b) is a polynomial of degree three, we could, in principle, find an exact analytic solution. However, such a solution might not be very informative. Instead, we show how perturbation theory can be used to obtain the amplitude of the dose response. In this model, the signaling complex is still $L : \alpha : \gamma : JAK3$. The signaling function is given by

$$\sigma(L) \equiv K_3K_2K_1Lxyz. \tag{3.21}$$

In section 4.3 we will show that, for this model, the maximum of σ is attained in the limit $L \to +\infty$. Hence, we have

$$A \equiv \lim_{L\to+\infty} \sigma(L). \tag{3.22}$$

Combining (3.19a), written as $N_x - N_z + z = \frac{N_x}{1+K_1z}$, and (3.19b), we obtain a reduced expression for the product xy

$$xy = \frac{N_y - y}{K_2(1 + K_3L)(1 + K_1z)}, \tag{3.23}$$

which allows us to rewrite the amplitude as

$$A = \lim_{L\to+\infty} \frac{K_1z}{(1 + K_1z)} \frac{K_3L}{(K_3L + 1)}(N_y - y). \tag{3.24}$$

We note that z is independent of L and, therefore, to compute the amplitude we only need to find the behavior of y as $L \to +\infty$.

Perturbation theory to determine y as $L \to +\infty$. We now apply the method described in [56] and summarized in Appendix A. Let $\epsilon = \frac{1}{L}$ and define the polynomial P_ϵ as follows,

$$P_\epsilon(y) \equiv P_2(y)\epsilon^2,$$

where P_2 is the polynomial (3.19b). We added a factor of ϵ^2 to remove any negative powers of ϵ in P_2. We obtain the polynomial

$$P_\epsilon(y) = A_\epsilon y^3 + B_\epsilon y^2 + C_\epsilon y + D_\epsilon, \tag{3.25}$$

where

$$\begin{aligned}
A_\epsilon &= -K_2(\epsilon + K_3)[K_2(\epsilon + K_3) - \epsilon K_4],\\
B_\epsilon &= K_2^2K_3^2(N_y - N_x) - \epsilon K_2K_3[1 + 2K_2N_x - 2K_2N_y + K_4(N_w - N_x + 2N_y)]\\
&\quad + \epsilon^2(K_4 - K_2(1 + K_2N_x - K_2N_y + K_4(N_w - N_x + 2N_y))),\\
C_\epsilon &= \epsilon N_y(K_2K_3(1 + K_4(N_w - N_x + N_y)) + \epsilon(K_2 - 2K_4 + K_2K_4(N_w - N_x + N_y))),\\
D_\epsilon &= K_4N_y^2\epsilon^2.
\end{aligned}$$

We now replace y by $\epsilon^p\omega(\epsilon)$ with $\omega(0) \neq 0$ according to theorem A.3. We obtain

$$P_\epsilon(\epsilon^p\omega(\epsilon)) = A_{p,\epsilon}\omega^3 + B_{p,\epsilon}\omega^2 + C_{p,\epsilon}\omega + D_{p,\epsilon}, \tag{3.26}$$

where

$$\begin{aligned}
A_{p,\epsilon} &= -\epsilon^{3p}K_2(\epsilon + K_3)(K_2(\epsilon + K_3) - \epsilon K_4),\\
B_{p,\epsilon} &= \epsilon^{2p}(K_2^2K_3^2(N_y - N_x) - \epsilon K_2K_3(1 + 2K_2(N_x - N_y) + K_4(N_w - N_x + 2N_y))\\
&\quad + \epsilon^2(K_4 - K_2(1 + K_2N_x - K_2N_y + K_4(N_w - N_x + 2N_y)))),\\
C_{p,\epsilon} &= \epsilon^{1+p}N_y(K_2K_3(1 + K_4(N_w - N_x + N_y))\\
&\quad + \epsilon(K_2 - 2K_4 + K_2K_4(N_w - N_x + N_y))),\\
D_{p,\epsilon} &= K_4N_y^2\epsilon^2.
\end{aligned}$$

The smallest exponents in the previous equation are

$$E = \{2, 1 + p, 2p, 3p\}.$$

We note that 0 is not in E because we multiplied P_2 by ϵ^2. Applying the graphical algorithm detailed in Appendix A, we find the proper values $(0, 0)$ and $(1, 2)$ (see Figure 4). We investigate these two branches.

Branch (0,0). We make use of the notation in Appendix A to define

$$T_\epsilon^{(1)}(\omega) \equiv \epsilon^0 P_\epsilon(\omega\epsilon^0).$$

The least common denominator of {2,1,0,0} is $q_1 = 1$. Therefore in accordance with the notation of Appendix A

$$\epsilon = \beta,$$

and the polynomial $R_\beta^{(1)}$ defined as

$$R_\beta^{(1)}(\omega) \equiv T_\epsilon^{(1)}(\omega),$$

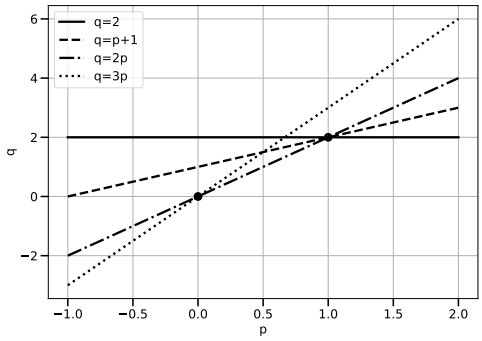


FIG. 4. *The lines defined in set E and the proper values (black dots) computed following the graphical algorithm described in Appendix* A.

is the polynomial P_ϵ. It means that we have $y = \omega$ and we can directly carry out a regular perturbation expansion.

Let us write the asymptotic expansion $y = y_0 + y_1\epsilon + y_2\epsilon^2 + \cdots$ and substitute it into $P_\epsilon(y)$. Since $P_\epsilon(y) = 0$, by the fundamental theorem of perturbation theory (Theorem A.2) we obtain a system of equations in y_0, $y_1, \ldots$, which can be solved. The first equation of the system is given by

$$-K_2^2 K_3^2 y_0^2 (N_x - N_y + y_0) = 0. \tag{3.27}$$

We are are only interested in nonnegative values of y_0, since we want y to be biologically relevant. We also require $\omega(0) = y(0) = y_0 \neq 0$ from Theorem A.3. Thus, solving (3.27), we obtain $y_0 = N_y - N_x$ if $N_y > N_x$ and $y_0 = 0$ otherwise. Assuming $y_0 = 0$ (i.e., $N_x \geq N_y$), we solve the next order equation

$$K_4 N_y^2 + K_2 K_3 N_y (1 + K_4(N_w - N_x + N_y))y_1 + K_2^2 K_3^2 (N_y - N_x) y_1^2 = 0. \tag{3.28}$$

We select the positive root of this polynomial and obtain an expression for y_1 when $y_0 = 0$. Thus, we have

$$y_1 = N_y \frac{1 + K_4(N_w - N_x + N_y) + \sqrt{\Delta_{y_1}}}{2K_2 K_3 (N_x - N_y)}, \tag{3.29}$$

where $\Delta_{y_1} = 1 + 2K_4(N_w + N_x - N_y) + K_4^2(N_w - N_x + N_y)^2$. Equation (3.29) shows that $y_1 > 0$ when $N_x \geq N_y$. Hence, $y \approx \epsilon y_1$ converges to zero from above and, therefore, represents a biologically relevant solution. We can conclude that

$$\lim_{\epsilon \to 0}(N_y - y) = \left\{ \begin{matrix} N_x, & \text{if } N_y > N_x, \\ N_y, & \text{otherwise} \end{matrix} \right\} = \min(N_x, N_y). \tag{3.30}$$

Branch (1, 2). On this branch, and again following the notation of Appendix A, we define

$$T_\epsilon^{(2)}(\omega) \equiv \epsilon^{-2} P_\epsilon(\epsilon^1 \omega).$$

The least common denominator of $\{2, 1+1, 0+1, 0+1\}$ is $q_2 = 1$, so $R_\beta^{(2)}$ is the same polynomial as $T_\epsilon^{(2)}$. Since $y = \omega\epsilon$, we have $N_y - y \underset{\epsilon \to 0}{\sim} N_y$. Furthermore, when replacing ω by an asymptotic expansion $\omega_0 + \omega_1\epsilon + \cdots$ in $T_\epsilon^{(2)}$ and applying the fundamental

theorem of perturbation theory (Theorem A.2), we obtain the same equation for w_0 as for y_1 in the previous branch (see (3.28)):

$$K_4N_y^2 + K_2K_3N_y(1 + K_4(N_w - N_x + N_y))\omega_0 + K_2^2K_3^2(N_y - N_x)\omega_0^2 = 0\,. \tag{3.31}$$

We have $y_1 = \omega_0$. In other words, at large, but finite $L = 1/\epsilon$, the convergence behavior of the two branches is identical. This agrees with Theorem 2.7 which states that there is only one positive solution for each set of reaction constants and initial conditions. In conclusion, we find that $N_y - y = \min(N_x, N_y)$, which gives the following expression for the amplitude,

$$A \equiv \frac{K_1 z}{1 + K_1 z}\min(N_x, N_y), \tag{3.32}$$

with z defined in (3.21). As the steady state concentration of JAK3, z, is the same in the IL-7R model with or without the extra chain R, the amplitude of both models has the exact same expression.

Computation of the EC_{50}. Since we did not compute analytic expressions for each steady state concentration, the EC_{50} expression has to be obtained by computing a Gröbner basis of the polynomial system (3.19) augmented by the polynomial

$$K_3K_2K_1Lxyz(1 + K_1z) - \frac{K_1zM}{2} = 0, \tag{3.33}$$

considering x, y, z, and L as variables with $M = \min(N_x, N_y)$. The lex Gröbner basis obtained for this system is

$$0 = K_1z^2 + (1 + K_1(N_x - N_z))z - N_z, \tag{3.34a}$$

$$0 = K_3aL^3 + A_LL^2 + B_LL + C_L, \tag{3.34b}$$

$$0 = y + \frac{-aL^2 + B_yL + C_y}{2K_2(K_2 - K_4)M^2}, \tag{3.34c}$$

$$0 = w + \frac{aL^2 + B_wL + C_w}{2K_4(K_2 - K_4)M^2}, \tag{3.34d}$$

$$0 = x + \frac{aL^2 + B_xL + C_x}{2K_2K_4M^2(1 + K_1z)}, \tag{3.34e}$$

where we wrote

$$\begin{aligned}
a &= K_2^2K_3^2(M - 2N_x)(M - 2N_y)^2,\\
A_L &= K_2K_3^2M(M - 2N_y)(-2 + 3K_2M - K_4(M + 2N_w - 2N_x) - 2K_2(2N_x + N_y)),\\
B_L &= K_3M^2(2K_4\\
&\quad + K_2(-2 + 3K_2M + 2K_4(-M - N_w + N_x + N_y) - 2K_2(N_x + 2N_y))),\\
C_L &= K_2(K_2 - K_4)M^3,\\
B_y &= -\frac{A_L}{K_3},\\
C_y &= M^2(-2K_4 + K_2(2 + K_4(M + 2N_w - 2N_x) - 2K_2(M - N_x - N_y))),\\
B_w &= -2K_2K_3M(M - 2N_y)(1 + K_4N_w + K_2(N_x + N_y - M)),\\
C_w &= K_2M^2(K_2(M - 2N_y) - 2K_4N_w),\\
B_x &= -K_2K_3M(M - 2N_y)(2 + K_4(M + 2N_w - 2N_x) - 2K_2(M - N_x - N_y)),\\
C_x &= M^2(2K_4 + K_2(K_2 - K_4)(M - 2N_y)).
\end{aligned}$$

Again, this polynomial system forms a Gröbner basis for most parameter choices. One obvious exception is the case $K_2 = K_4$. The polynomial (3.33a) is expected to be independent of the ligand concentration, L. The EC_{50} expression is the real positive root of polynomial (3.34b) at which x, y, and w (obtained via polynomials (3.34e), (3.34c), and (3.34d), respectively) are positive. The polynomial (3.34b) reflects the parameter dependency of the EC_{50}: since the parameters K_1 and N_z are not present in its coefficients, we can affirm that the EC_{50} is, once again, independent of the kinase. Thus, we reduced the problem of computing the EC_{50} to solving a univariate polynomial (3.34b). In comparison, before any algebraic manipulation was possible, the polynomial system (3.19) had to be solved multiple times to obtain the dose-response curve (a sigmoid), which was then fitted with a Hill equation. Finally, the EC_{50} was computed from the fitted parameters of the Hill curve. Alternatively, if one wanted to apply the Gröbner basis-free method of section 3.1.1, one would have to solve the polynomial (3.19b) in y (which is possible in theory), find its positive real solution (which is potentially hard), substitute the expression of y into $\sigma(L)$, and then solve for the EC_{50}.

3.2. Summary of proposed algebraic method to study the signaling function. From the two previous examples, we devise a general algorithm to compute analytic expressions of the steady state, the amplitude and the EC_{50} for some receptor-ligand systems when ligand is in excess.

Key steps:

1. Write the mass action kinetics set of ODEs for the system under consideration.
2. Obtain the polynomial system by combining the steady state and conservation equations.
3. Compute a lex Gröbner basis of the polynomial system obtained in step 2. (Assert that this is indeed a Gröbner basis for the desired choice of parameters.)
4. Define the signaling function $\sigma(L)$.
5. Compute the amplitude expression by finding the extreme values of σ:
$$\text{Amplitude} = \max(\sigma) - \min(\sigma).$$
6. Compute a lex Gröbner basis of the polynomial system obtained in step 2 augmented by the equation
$$\sigma(L) - \left[\frac{\text{Amplitude}}{2} + \min(\sigma)\right] = 0,$$
with the ligand concentration, L, considered as an additional variable. This additional equation corresponds to Definition 2.12 of the EC_{50}.
7. Find the positive roots of the univariate polynomial in L of the Gröbner basis obtained in step 6. The root which allows the other variables to be positive is the EC_{50}.

Remark 3.1. Computing a Gröbner basis is, in general, an expensive computation and the run time and memory requirements of the proposed method may vary seemingly erratically when different models are considered.

One of the crucial parts of the proposed algebraic algorithm is the amplitude computation. Usually, we have the simplification that $\min(\sigma) = \sigma(0) = 0$, however, finding $\max(\sigma)$ can be challenging. For certain classes of models we have

$$\lim_{L\to+\infty} \sigma(L) = \max(\sigma),$$

which greatly reduces the calculation. We can now either solve the Gröbner basis from step 3 directly, to obtain analytic expressions of the steady state concentrations of the single chains components, or use perturbation theory as outlined in section 3.1.2. In the final step, if an exact expression for the EC_{50} cannot be computed, i.e., the univariate polynomial in L has a large degree, one already reduces the cost of the EC_{50} computation compared to the naive approach. In summary, in this section we compute the lex Gröbner bases with the computer algebra package Macaulay2 [25] and provide a Macaulay2 code example in Appendix C.

4. Analytical study of general sequential receptor-ligand systems. In spite of the general applicability of the method outlined in the previous section, we still have to make the assumption that the computed limit of the signaling function coincides with its amplitude. In this section we show that this is indeed the case for a wider class of receptor-ligand systems. An analytic closed-form expression for the amplitude follows with little extra work. The EC_{50} can then be studied making use of key steps 6 and 7 in section 3.2. We start by giving an abstract generalization of the example from section 3.1.1.

Definition 4.1 (SRLK model). *We call a sequential receptor-ligand model with extrinsic kinase (SRLK) a receptor-ligand model with the following properties:*

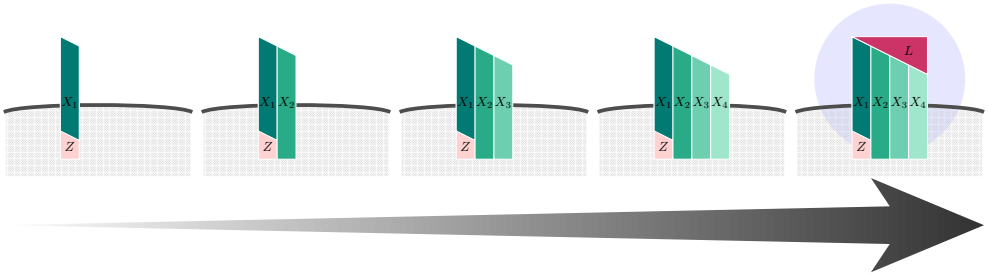
- *The receptor is composed of n different transmembrane chains, $X_1, \dots, X_n$, which bind sequentially,*

$$X_1 : \dots : X_{i-1} + X_i \rightleftharpoons X_1 : \dots : X_i \quad \text{for all } i \in \{2, \dots, n\}.$$

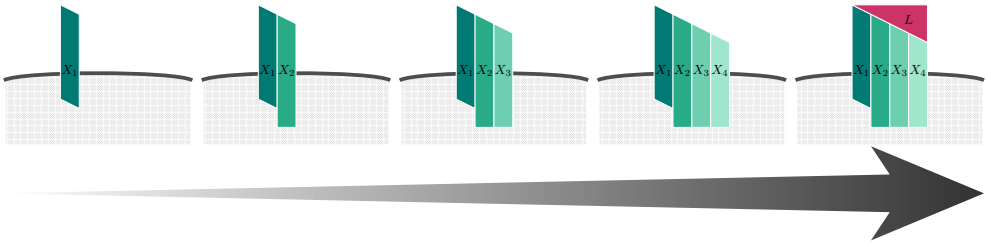
- *X_1 can bind reversibly to an intracellular extrinsic kinase Z.*
- *The signaling receptor is given by $Z : X_1 : \dots : X_n$ and the dummy receptor by $X_1 : \dots : X_n$.*
- *The extracellular ligand, L, binds reversibly to the signaling (or dummy) receptor, forming the signaling (or dummy) complex $Z : X_1 : \dots : X_n : L$ (or $X_1 : \dots : X_n : L$).*

The biochemical reaction network for a general SRLK model is given by

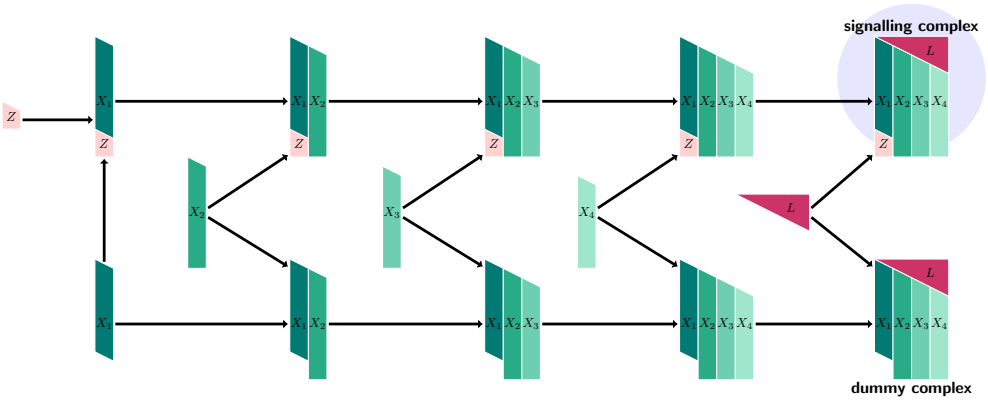
$$\begin{array}{llllll}
Z & + & X_1 & \rightleftharpoons & Z : X_1, & K_0, \\
X_2 & + & Z : X_1 & \rightleftharpoons & Z : X_1 : X_2, & K_1, \\
X_2 & + & X_1 & \rightleftharpoons & X_1 : X_2, & K_1', \\
\vdots & & \vdots & & \vdots & \vdots \\
X_{i+1} & + & Z : X_1 : \dots : X_i & \rightleftharpoons & Z : X_1 : \dots : X_{i+1}, & K_i, \\
X_{i+1} & + & X_1 : \dots : X_i & \rightleftharpoons & X_1 : \dots : X_{i+1}, & K_i', \\
\vdots & & \vdots & & \vdots & \vdots \\
X_n & + & Z : X_1 : \dots : X_{n-1} & \rightleftharpoons & Z : X_1 : \dots : X_n, & K_{n-1}, \\
X_n & + & X_1 : \dots : X_{n-1} & \rightleftharpoons & X_1 : \dots : X_n, & K_{n-1}', \\
L & + & Z : X_1 : \dots : X_n & \rightleftharpoons & Z : X_1 : \dots : X_n : L, & K_n, \\
L & + & X_1 : \dots : X_n & \rightleftharpoons & X_1 : \dots : X_n : L, & K_n',
\end{array} \tag{4.1}$$



(a) Sequential formation of the signalling complex.



(b) Sequential formation of the dummy complex.



(c) SRLK model: sequential biochemical reaction scheme.

FIG. 5. *SRLK model with $n = 4$ transmembrane chains:* (a)/(b) *Sequential formation of the signaling/dummy complex.* (c) *Scheme of the sequential formation of the signaling/dummy complex: the chain X_1 binds first to the intracellular extrinsic kinase Z (only for the signaling complex). Next, the chain X_2 binds to the complex $Z : X_1$ (or X_1) and X_3 binds to $Z : X_1 : X_2$ (or $X_1 : X_2$). Then, X_4 binds to $Z : X_1 : X_2 : X_3$ (or $X_1 : X_2 : X_3$). Finally, the ligand L binds to the signaling receptor $Z : X_1 : X_2 : X_3 : X_4$ (or dummy receptor $X_1 : X_2 : X_3 : X_4$), thus forming the signaling (or dummy) complex.*

where the K_i (or K_i') are the affinity constants related to the formation of the signaling (or dummy) complex. Figure 5 illustrates the formation of the signaling and dummy complexes in an SRLK model with $n = 4$ transmembrane chains. We assume the system at steady state and ligand in excess. In what follows we refer to these two assumptions as the *experimental hypotheses*.

We write z (or x_i) for the steady state concentration of unbound chain Z (or X_i). We also use L to denote the ligand concentration. Finally, N_z (or N_i) denotes the total copy number per cell of the species Z (or X_i). An SRLK model satisfying the experimental hypotheses is then described by the following polynomial system:

$$N_z = z + K_0 z x_1 + K_0 K_1 z x_1 x_2 + \cdots + K_0 K_1 \dots K_{n-1} z x_1 \dots x_n + K_0 \dots K_n z x_1 \dots x_n L = z + K_0 z \left[x_1 + \sum_{j=2}^{n} \left(\prod_{l=1}^{j-1} K_l x_l x_j \right) + L \prod_{j=1}^{n} K_j x_j \right], \tag{4.2a}$$

$$N_1 = x_1 + K_1' x_1 x_2 + \cdots + K_1' \dots K_{n-1}' x_1 \dots x_n + K_1' \dots K_n' L x_1 \dots x_n + K_0 z (x_1 + K_1 x_1 x_2 + \cdots + K_1 \dots K_{n-1} x_1 \dots x_n + K_1 \dots K_n x_1 \dots x_n L) \tag{4.2b}$$

for $i = 2, \dots, n-1$:

(4.2c)

$$N_i = x_i + K_1' \dots K_{i-1}' x_1 \dots x_i + \cdots + K_1' \dots K_{n-1}' x_1 \dots x_n + K_1' \dots K_n' x_1 \dots x_n L + K_0 z (K_1 \dots K_{i-1} x_1 \dots x_i + \cdots + K_1 \dots K_{n-1} x_1 \dots x_n + K_1 \dots K_n x_1 \dots x_n L) = x_i + \sum_{j=i}^{n} \left(\prod_{l=1}^{j-1} K_l' x_l x_j + K_0 z \prod_{l=1}^{j-1} K_l x_l x_j \right) + L \prod_{j=1}^{n} K_j' x_j + K_0 z L \prod_{j=1}^{n} K_j x_j,$$

(4.2d)

$$N_n = x_n + K_0 \dots K_{n-1} z x_1 \dots x_n + K_1' \dots K_{n-1}' x_1 \dots x_n + K_1' \dots K_n' L x_1 \dots x_n + K_0 \dots K_n z L x_1 \dots x_n .$$

We note that many results in this section can be further simplified under the additional hypothesis of no allostery.

DEFINITION 4.2. *There is no allostery in an SRLK model if $K_i = K_i'$ for all $i = 1, \dots, n$.*

Finally, we formally define the signaling and dummy functions for this class of models.

DEFINITION 4.3. *For an SRLK model under the experimental hypotheses the signaling function, $\sigma(L)$, is the number of signaling complexes formed as a function of the ligand concentration, L, and can be written as follows:*

$$\sigma(L) = K_0 z L \prod_{i=1}^{n} K_i x_i.$$

Similarly, the dummy function, $\delta(L)$, is the number of dummy complexes formed as a function of the ligand concentration, L, and can be written as follows:

$$\delta(L) = L \prod_{i=1}^{n} K_i' x_i.$$

Note that the IL-7R model of section 3.1.1 is one example of an SRLK model and the definition of signaling function given in section 2.2 is equivalent. We now introduce the notion of a limiting component.

DEFINITION 4.4. *The species, X_j, which has the smallest total copy number of molecules*

$$0 < N_j < N_i \ \forall i \neq j,$$

is the limiting component of the system. If there are multiple limiting components, $X_{j_1}, \dots, X_{j_r}$, then

$$0 < N_{j_1} = \cdots = N_{j_r} < N_i \ \forall i \notin \{j_1, \dots, j_r\}.$$

If the signaling function attains its maximum for large values of the ligand concentration, then, since by definition $\sigma(0) = 0$, the amplitude of such a model is given by

$$A \equiv \lim_{L \to +\infty} \sigma(L).$$

In this section we present some general results for $\lim_{L\to+\infty} \delta(L)$ and $\lim_{L\to+\infty} \sigma(L)$ applicable to SRLK models. The proofs of the lemmas and theorems can be found in Appendix D.

4.1. Asymptotic study of the steady states. While it is difficult to find closed-form expressions of the steady states for general receptor-ligand systems, in what follows we show that considerable progress can be made for the specific case of SRLK models. In this section we describe the behavior of the concentrations, x_i, in the limit $L \to +\infty$. The proofs of our results can be found in Appendix D. First, we recall the definition and a property of algebraic functions.

DEFINITION 4.5. *A univariate function $y = f(x)$ is said to be algebraic if it satisfies the polynomial equation*

$$(*) \qquad y^m + R_{m-1}(x)y^{m-1} + \cdots + R_0(x) = 0,$$

where the $R_i(x)$ are rational functions of x, i.e., are of the form $\frac{p(x)}{q(x)}$, where p and q are polynomial functions and $q(x) \neq 0$ for all $x \in \mathbb{R}$.

Remark 4.6. Note that the polynomial $(*)$ has m solutions. These solutions are called the branches of an algebraic function and one often specifies a particular branch.

Since we are interested in the limit behavior, the following lemma proves useful.

LEMMA 4.7. *Any bounded, continuous solution of $(*)$ defined on $\mathbb{R}$ has a finite limit at $+\infty$ (and $-\infty$).*

With this background in place, we can now proceed to study SRLK models in detail. We start by showing that in steady state the signaling and the dummy functions have a positive limit when L tends to $+\infty$.

LEMMA 4.8. *The signaling and the dummy functions of an SRLK model satisfying the experimental hypotheses admit a finite limit when $L \to +\infty$ and this limit is positive.*

An equivalent result holds for the steady state concentration of the kinase.

LEMMA 4.9. *In an SRLK model under the experimental hypotheses, the concentration of the extrinsic intracellular kinase Z admits a positive finite limit, $c_z > 0$, when $L \to +\infty$.*

In the particular case of no allostery, we can write an explicit expression of the limit of z, c_z.

LEMMA 4.10. *Consider an SRLK model which satisfies the experimental hypotheses. If we assume no allostery, then the steady state value of the extrinsic intracellular kinase, z, is given by*

$$z = \frac{-1 + K_0(N_z - N_1) + \sqrt{\Delta_z}}{2K_0}, \tag{4.3}$$

where

$$\Delta_z = (1 + K_0(N_1 - N_z))^2 + 4K_0N_z.$$

By Lemma 4.10, z is independent of L (thus, $c_z = z$) and only depends on K_1, N_1, and N_z. Note that this result is equivalent to the one obtained in section 3.1.1 for the IL-7R model. Finally, we study the behavior of the concentration x_i in the limit $L \to +\infty$. We first give bounds to the asymptotic dependency of x_i on L.

LEMMA 4.11. *Let us consider an SRLK model which satisfies the experimental hypotheses. Then no concentration x_i behaves proportionally to L^q, $q > 0$, or $\frac{1}{L^p}$, $p > 1$, when $L \to +\infty$.*

We can now state the main theorem of this section.

THEOREM 4.12. *We consider an SRLK model which satisfies the experimental hypotheses. If there exists a unique limiting component X_{i_0}, then*

$$x_{i_0} \underset{L\to+\infty}{\sim} \frac{c_{i_0}}{L},$$

and for all $i = 1, \dots, n, i \neq i_0$,

$$x_i \underset{L\to+\infty}{\sim} c_i,$$

where c_{i_0} and c_i are positive constants.

COROLLARY 4.13. *If an SRLK model, which satisfies the experimental hypotheses, has multiple limiting components, $X_{i_1}, \dots, X_{i_r}$, $i_1 < \dots < i_r$, then*

$$x_{i_1} \underset{L\to+\infty}{\sim} \frac{c_{i_1}}{L^{p_1}}, \dots, x_{i_r} \underset{L\to+\infty}{\sim} \frac{c_{i_r}}{L^{p_r}},$$

where $c_{i_1}, \dots c_{i_r}$ are positive constants and $p_1 = \dots = p_r = \frac{1}{r}$. The concentrations of the nonlimiting components, x_i (for $i \notin \{i_1, \dots, i_r\}$) tend to positive constants, $c_i > 0$.

4.2. Asymptotic study of the signaling and dummy functions. The previous section presented numerous small results which give insight into the steady state behavior of SRLK receptor-ligand systems. We are now in a position to combine these results to state and prove our main theorem, which gives closed-form formulas for the limits of the signaling and dummy functions.

THEOREM 4.14. *Consider an SRLK model which satisfies the experimental hypotheses. Let us write $X_{i_1}, \dots, X_{i_r}$ as the limiting components and denote N_{i_0} as the total amount of any limiting component, i.e., $N_{i_0} \equiv N_{i_1} = \dots = N_{i_r}$. The limit of the signaling function is given by*

$$\lim_{L\to+\infty} \sigma(L) = \frac{\prod_{i=1}^n K_i K_0 c_z}{\prod_{i=1}^n K_i' + \prod_{i=1}^n K_i K_0 c_z} N_{i_0},$$

and the limit of the dummy function is

$$\lim_{L\to+\infty} \delta(L) = \frac{\prod_{i=1}^n K_i'}{\prod_{i=1}^n K_i' + \prod_{i=1}^n K_i K_0 c_z} N_{i_0},$$

where

$$c_z = \lim_{L\to+\infty} z.$$

Under the assumption of no allostery, these expressions can be further simplified.

COROLLARY 4.15. *Consider the system of Theorem* 4.14 *and assume there is no allostery. Denote the limiting components by* $X_{i_1}, \dots, X_{i_r}$ *and and let* N_{i_0} *be the total amount of any limiting component, i.e.,* $N_{i_0} \equiv N_{i_1} = \cdots = N_{i_r}$. *The limit of the signaling function is*

$$\lim_{L\to+\infty} \sigma(L) = \frac{K_0 z}{1 + K_0 z} N_{i_0},$$

and the limit of the dummy function is

$$\lim_{L\to+\infty} \delta(L) = \frac{N_{i_0}}{1 + K_0 z},$$

with z *given by* (4.3) *in Lemma* 4.10.

From the previous expressions we observe that the limit of the signaling and dummy functions are equal to the total copy number of the limiting component, N_{i_0}, multiplied by a term which is bounded between 0 and 1. This term only depends on the affinity constant K_0 and the steady state concentration of the kinase. In order to relate the above limits back to biologically meaningful quantities, all there is left to show is that the explicit expression of the limit of σ is in fact the amplitude of the system. Since $\sigma(0) = 0$, let us first note that the amplitude is equal to the maximum of σ. Under the no allostery assumption, we can show mathematically that this maximum is the limit of σ when $L \to +\infty$. To this end, the following lemma is needed.

LEMMA 4.16. *Consider an SRLK model under the experimental hypotheses. If there is no allostery, then we have*

$$\sup \sigma(L) = \lim_{L\to+\infty} \sigma(L).$$

The supremum here is attained and is a maximum. Thus, the amplitude for an SRLK receptor-ligand system when there is no allostery is the limit of σ described in Corollary 4.15. This result is the generalization of the example discussed in section 3.1.1. We note that the amplitude of the IL-7R model of section 3.1.1 can be recovered by setting $N_{i_0} = \min(N_x, N_y)$. We now have also rigorously shown that the limit of the signaling function is indeed the amplitude. The EC_{50} can now be found as outlined in section 3.1.1.

Remark 4.17. Theorem 4.14 provides a Gröbner basis-free result to find the amplitude of an SRLK system. This is crucial due to the fact that finding Gröbner bases is a costly process, and no a priori algorithmic complexity bounds have been derived for SRLK systems.

4.3. SRLK models with additional subunit receptor chains. As hinted at in section 3.1.2, the IL-7R model with the additional subunit receptor chain is part of a larger group of models which are an extension of the SRLK family. Therefore, our previous results can be extended to these types of models. Again, we start by giving an abstract definition of the extended SRLK family of models.

DEFINITION 4.18 (extended SRLK model). *An* extended SRLK model *is an SRLK model where we assume that each intermediate complex,* $Z : X_1 : \ldots : X_i$ *(or* $X_1 : \ldots : X_i$*), for* $i = 1, \dots, n$ *can bind to an extra chain,* Y_i*, with an affinity constant* K_{y_i} *(or* K'_{y_i}*), to form a decoy complex* $Z : X_1 : \ldots X_i : Y_i$ *(or* $X_1 : \ldots X_i : Y_i$*). The*

addition of a subunit chain of the kind Y_i prevents the binding of ligand to the receptor, and thus, does not allow the formation of signaling or dummy complexes.

The chemical reaction network for an extended SRLK model is given by:

$$
\begin{array}{lcll}
Z & + & X_1 & \rightleftharpoons Z:X_1, \quad K_0,\\
Y_1 & + & X_1 & \rightleftharpoons X_1:Y_1, \quad K'_{y_1},\\
Y_1 & + & Z:X_1 & \rightleftharpoons Z:X_1:Y_1, \quad K_{y_1},\\
X_2 & + & Z:X_1 & \rightleftharpoons Z:X_1:X_2, \quad K_1,\\
X_2 & + & X_1 & \rightleftharpoons X_1:X_2, \quad K'_1,\\
Y_2 & + & X_1:X_2 & \rightleftharpoons X_1:X_2:Y_2, \quad K'_{y_2},\\
Y_2 & + & Z:X_1:X_2 & \rightleftharpoons Z:X_1:X_2:Y_2, \quad K_{y_2},\\
\vdots & & \vdots & \vdots\\
X_{i+1} & + & Z:X_1:\ldots:X_i & \rightleftharpoons Z:X_1:\ldots:X_{i+1}, \quad K_i,\\
X_{i+1} & + & X_1:\ldots:X_i & \rightleftharpoons X_1:\ldots:X_{i+1}, \quad K'_i,\\
Y_i & + & X_1:\ldots:X_i & \rightleftharpoons X_1:\ldots:X_i:Y_i, \quad K'_{y_i},\\
Y_i & + & Z:X_1:\ldots:X_i & \rightleftharpoons Z:X_1:\ldots:X_i:Y_i, \quad K_{y_i},\\
\vdots & + & \vdots & \vdots\\
Y_n & + & X_1:\ldots:X_n & \rightleftharpoons X_1:\ldots:X_n:Y_n, \quad K'_{y_n},\\
Y_n & + & Z:X_1:\ldots:X_n & \rightleftharpoons Z:X_1:\ldots:X_n:Y_n, \quad K_{y_n},\\
L & + & Z:X_1:\ldots.X_n & \rightleftharpoons Z:X_1:\ldots:X_n:L, \quad K_n,\\
L, & + & X_1:\ldots.X_n & \rightleftharpoons X_1:\ldots:X_n:L, \quad K'_n,
\end{array}
$$

where K_i, K'_i, K_{y_i}, and K'_{y_i} denote the affinity constants. Figures 6(a) and 6(b) show the decoy complexes of an extended SRLK receptor-ligand system with $n = 4$ transmembrane chains. The signaling and dummy complexes are built sequentially similarly to the classic SRLK model (see Figures 5 and 6(c)).

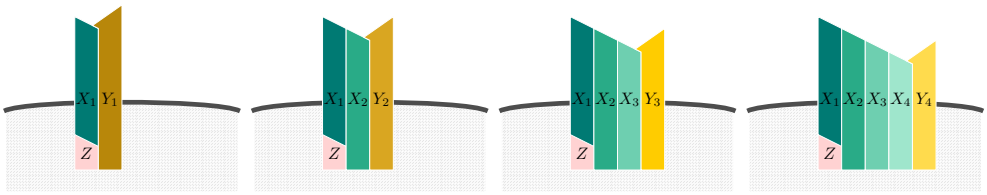
We note that while we assume all the X_i to be different species, we allow that $Y_i = Y_j$ or $Y_i = \emptyset$, as long as for $i = 1, \ldots, n$, $Y_i \notin \{X_1, \ldots, X_n, Z, L\}$. We assume that the receptor-ligand system is in a steady state and the ligand is in excess. We further assume that the concentrations of the species Y_i (which we write y_i) are all bounded. We could consider the case when the Y_i are in excess, and thus, treat their concentration as a parameter of the model, or assume that the number of Y_i molecules is conserved. We refer to these assumptions as the *extended experimental hypotheses*.

The signaling and dummy functions of classic and extended SRLK receptor-ligand systems are equivalent (see Definition 4.3). The polynomial system describing an extended SRLK model under the extended experimental hypotheses is given by

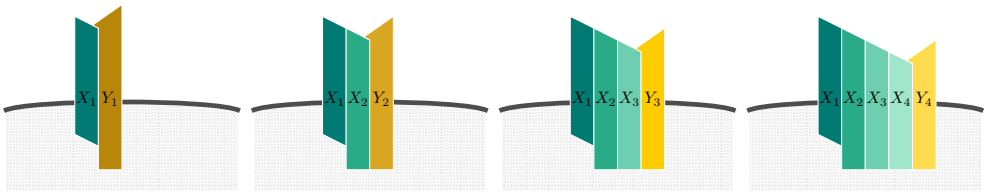
$$
N_z = z + K_0 z \left(x_1(1 + K_{y_1} y_1) + \sum_{j=2}^{n} \left((1 + K_{y_j} y_j) \prod_{l=1}^{j-1} K_l x_l x_j \right) \right) + \sigma(L), \tag{4.4a}
$$

$$
\begin{aligned}
N_1 = {} & x_1(1 + K'_{y_1} y_1) + \sum_{j=2}^{n} \left(\left(1 + K'_{y_j} y_j\right) \prod_{l=1}^{j-1} K'_l x_l x_j \right) + \delta(L)\\
& + K_0 z \left(x_1 \left(1 + K_{y_1} y_1\right) + \sum_{j=2}^{n} \left(\left(1 + K_{y_j} y_j\right) \prod_{l=1}^{j-1} K_l x_l x_j \right) \right) + \sigma(L),
\end{aligned} \tag{4.4b}
$$

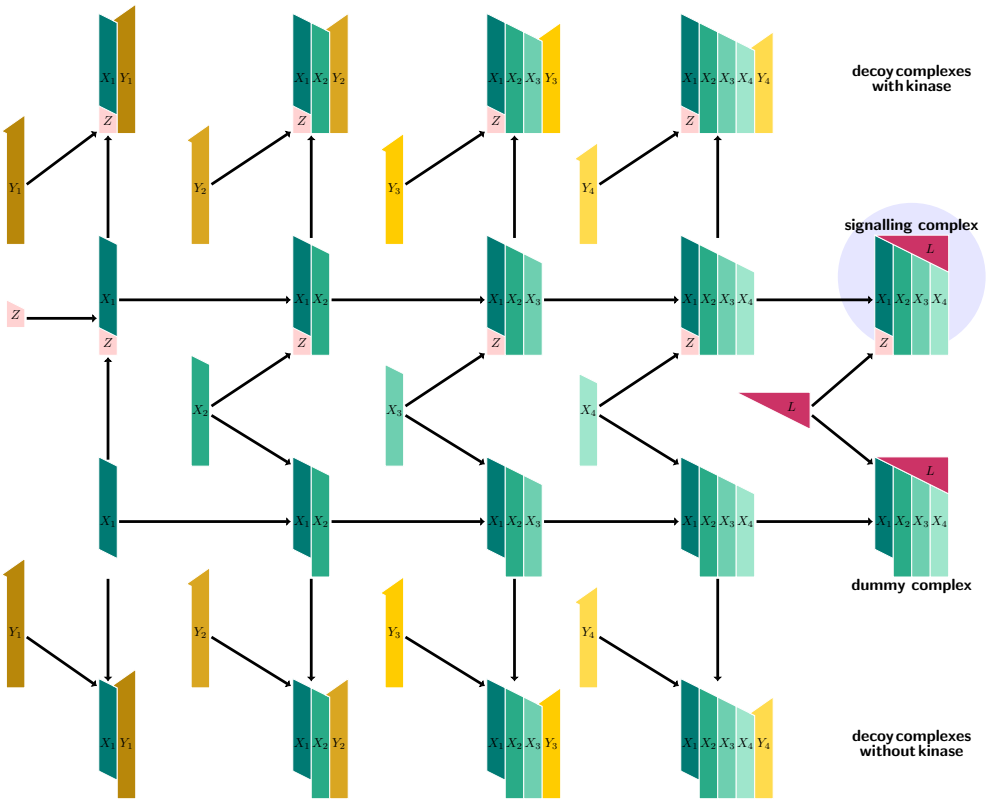
$$
\vdots
$$



(a) Decoy complexes with kinase.



(b) Decoy complexes without kinase.



(c) Extended SRLK model: sequential biochemical reaction scheme.

FIG. 6. *Extended SRLK model with $n = 4$ transmembrane chains:* (a) *An additional subunit chain, Y_i, can bind to each intermediate signaling complex $Z : X_1 : \ldots : X_i$, to form decoy complexes with kinase.* (b) *The subunit chain Y_i can also bind to the intermediate dummy complexes $X_1 : X_2 : \ldots : X_i$, forming decoy complexes without kinase.* (a) *Scheme of the sequential formation of the signaling and dummy complexes. At each step their formation can be interrupted by the binding of a subunit chain, Y_i, to the intermediate complex, forming a decoy complex.*

and for $i = 2, \ldots, n-1$,

(4.4c)
$$N_i = x_i + \sum_{j=i}^{n} \left(\left(1 + K'_{y_j} y_j\right) \prod_{l=1}^{j-1} K'_l x_l x_j + K_0 z (1 + K_{y_j} y_j) \prod_{l=1}^{j-1} K_l x_l x_j \right) + \delta(L) + \sigma(L),$$

$\vdots$

(4.4d)
$$N_n = x_n + K_0 z (1 + K_{y_n} y_n) \prod_{j=1}^{n-1} K_j x_j x_n + (1 + K'_{y_n} y_n) \prod_{j=1}^{n-1} K'_j x_j x_n + \delta(L) + \sigma(L).$$

This system of polynomials is completed by the conservation equations of the species Y_i, for $i = 1, \ldots, n$, if we assume they are conserved.

We can extend the notion of no allostery to the extended models.

DEFINITION 4.19. *An extended SRLK model is said to be under the assumption of no allostery if for each* $i = 1, \ldots, n$, $K_i = K'_i$, *and* $K_{y_i} = K'_{y_i}$.

With these expanded definitions, we can extend the results previously obtained for the SRLK receptor-ligand systems.

THEOREM 4.20. *The theorems and lemmas previously true for the SRLK models are true for the extended SRLK models under the same (extended) hypotheses.*

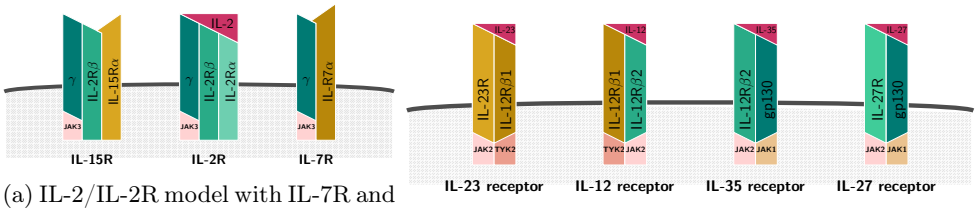
4.4. A few examples of (extended) SRLK models. In spite of some presumably strong modeling assumptions, the (extended) SRLK models can describe a broad range of cytokine-receptor systems. The IL-7R models described in sections 3.1.1 and 3.1.2 are, respectively, an SRLK and an extended SRLK model. In this section, we provide examples of other interleukin-signaling systems which are part of the SRLK family.

Example 4.21 (SRLK models: IL-2R and IL-15R). The interleukin-2 (IL-2) receptor is composed of three transmembrane subunit chains: the common gamma chain, γ, the IL-2Rα chain, and the IL-2Rβ chain. Additionally, γ binds to the intracellular extrinsic kinase JAK3. This IL-2 receptor-ligand system can be considered an SRLK model with $\{Z, X_1, X_2, X_3, L\} = \{\text{JAK3}, \gamma, \text{IL-2R}\beta, \text{IL-2R}\alpha, \text{IL-2}\}$. Similarly, the interleukin-15 (IL-15) receptor is composed of three transmembrane subunit chains, γ, IL-2Rβ, and IL-15Rα, as well as the kinase JAK3. It can be considered an SRLK model with $\{Z, X_1, X_2, X_3, L\} = \{\text{JAK3}, \gamma, \text{IL-2R}\beta, \text{IL-15R}\alpha, \text{IL-15}\}$.

A number of interleukin receptors share different molecular components. For instance, cytokine receptors of the common gamma family (comprising the receptors for IL-2,4,7,9,15, and 21 [50]) share the common gamma chain, γ. In addition the IL-2 and IL-15 receptors share the subunit chain, IL-2Rβ. The competition for these subunit chains can be mathematically described with an extended SRLK model, as follows.

Example 4.22 (extended SRLK model: IL-2/IL-2R model with formation of IL-7R and IL-15R). Let us suppose we want to study the formation of IL-2/IL-2R complexes taking into account the competition for the γ chain between IL-2Rβ and IL-7Rα, and the competition for the complex γ :IL-2Rβ between the subunits IL-2Rα and IL-15Rα. We can use an extended SRLK model with

$$\{Z, X_1, X_2, X_3, L\} = \{\text{JAK3}, \gamma, \text{IL-2R}\beta, \text{IL-2}\alpha, \text{IL-2}\}$$



(a) IL-2/IL-2R model with IL-7R and IL-15R competition.

(b) IL-12 receptor family.

FIG. 7. (a) *Illustration of example* 4.22*: IL-2R, IL-7R, and IL-15R competing for the common gamma chain and IL-2Rβ. The IL-2R is composed of three subunit chains: the gamma chain, IL-2Rβ, and IL-2Rα. IL-15R is composed of the gamma chain, IL-2Rβ, and the specific chain IL-15Rα. The IL-7R is composed of the gamma chain and IL-7Rα. All these receptors signal through the Janus kinase JAK3.* (b) *Illustration of example* 4.23*:* (i) *models the competition for IL-12Rβ1 between the IL-12 and the IL-23 receptors. We assume that IL-23R and IL-12Rβ2 are already bound to their associated extrinsic kinase JAK2;* (ii) *models the competition for IL-12Rβ2 between the IL-12 and IL-35 receptors. We consider the complexes IL-12Rβ1:TYK2 and gp130:JAK1 already preformed;* (iii) *models the competition for gp130 between the IL-35 and the IL-27 receptors. We consider the complexes IL-12Rβ2:JAK2 and IL-27R:JAK2 already preformed.*

and

$$\{Y_1, Y_2, Y_3\} = \{\text{IL-7R}\alpha, \text{IL-15R}\alpha, \emptyset\}.$$

This example is illustrated in Figure 7a.

A further extended SRLK example is that of the IL-12 family of receptors, which share multiple components [61], and each of which is composed of two transmembrane subunit chains. The IL-12 receptor is composed of the subunit chains IL-12Rβ1 and IL-12Rβ2. The IL-23 receptor signals via the IL-23R chain and the IL-12Rβ2 chain. The IL-27R (also known as WSX-1) and glycoprotein 130 (gp130) form the IL-27 receptor. Finally, IL-12Rβ2 and gp130 form the IL-35 receptor. The subunit chains gp130, IL-12Rβ1, and IL-12Rβ2 bind to a kinase from the JAK family (JAK1, TYK2, and JAK2, respectively). This competition can be described with extended SRLK models as follows.

Example 4.23 (extended SRLK models: IL-12R family). We provide three examples of extended SRLK systems which characterize the competition for receptor subunits between receptors of the IL-12 family (see Figure 7(b)).

1. Suppose we want to study the IL-12-induced signaling process taking into account the competition for IL-12Rβ1. We can use an extended SRLK model with

$$\{Z, X_1, X_2, L, Y_1, Y_2\} = \{\text{TYK2}, \text{IL-12R}\beta1, \text{IL-12R}\beta2^*, \text{IL-12}, \text{IL-23R}^*, \emptyset\}.$$

2. To study IL-35-induced signaling taking into account the competition for IL-12Rβ2, we can use an extended SRLK model with

$$\{Z, X_1, X_2, L, Y_1, Y_2\} = \{\text{JAK2}, \text{IL-12R}\beta2, \text{gp130}^*, \text{IL-35}, \text{IL-12R}\beta1^*, \emptyset\}.$$

3. An extended SRLK model with

$$\{Z, X_1, X_2, L, Y_1, Y_2\} = \{\text{JAK1}, \text{gp130}, \text{IL-27R}^*, \text{IL-27}, \text{IL-12R}\beta2^*, \emptyset\}$$

can describe the IL-27-induced signaling, when there is competition for the subunit chain gp130 with the IL-35 receptor.

Above we have made use of the notation X^* to denote the preformed complex composed of the receptor chain X and its intracellular extrinsic kinase (TYK2 for IL-12Rβ1, JAK1 for gp130, and JAK2 for all the others).

5. Conclusion. In the first part of this paper we propose a method to compute analytic expressions for two relevant pharmacodynamic metrics, the amplitude and the EC_{50} for receptor-ligand systems, based on two (simple) IL-7 receptor models. Our method starts with the computation of a Gröbner basis for the polynomial system of the receptor-ligand system in a steady state. As shown in our IL-7R models from sections 3.1.1 and 3.1.2, the derivation of the amplitude is easier when the maximum of the dose-response curve is attained at large ligand concentration (for instance when the dose-response curve is a sigmoid). In that case, the amplitude is the limit of the signaling function when the ligand concentration tends to infinity. When the model is simple enough, as is the case of the first IL-7R model, the polynomial system, simplified by the computation of the Gröbner basis, can be solved iteratively to obtain an analytic expression for the steady state. From these expressions, it is then relatively straightforward to compute the amplitude (i.e., the limit of the signaling function at large values of the ligand concentration) and the EC_{50}. For more complex models, such as our second IL-7R model, getting such steady state expressions can be more challenging. However, perturbation theory can be used to derive the expression for the amplitude. Computing another Gröbner basis can dramatically simplify the calculation of the EC_{50}, and in turn display how it depends on the parameters of the model. Analytic expressions for the amplitude and the EC_{50} offer mechanistic insight for the receptor-ligand systems under consideration, allow one to quantify the parameter dependency of these two key variables, and can facilitate model validation and parameter exploration. Indeed, for both IL-7R models, we noticed that the affinity constant of the association of the gamma chain to the kinase JAK3, K_1, was the only constant involved in the expression for the amplitude. As a consequence, and if conducting parameter inference to fit the model to experimental data, K_1 would be the only parameter that could be inferred by comparison of the theoretical to the experimental amplitude. On the contrary, this constant was absent from both EC_{50} expressions and, thus, its value would be impossible to infer by only comparing the experimental to the theoretical EC_{50}. Our exact analysis also showed that both models have the same amplitude. Finally, the application of our method no longer requires the numerical computation of the dose-response curve, finding its maximum to then obtain the amplitude and fit the curve to derive the EC_{50}. This reduces dramatically the computational cost and numerical errors. However, our method requires models simple enough to be able to compute a lex Gröbner basis, which is known to be computationally intensive [9, 40]. Additionally, computing the amplitude when the maximum response is not the asymptotic behavior of the dose-response curve can be tricky. For instance the computation of the maximum for bell-shaped dose-response curves (which has been done for simple models in [38, 13]) may involve the computation of the derivative of the signaling function. This computation can be laborious even with the use of symbolic software. Finally, our method often requires additional mathematical tools or knowledge, such as perturbation theory in section 3.1.2, which makes it rather a challenge to be used by those who are not mathematically trained. In spite of the (sometimes complicated) calculations that our method requires, we believe that analytic expressions of the pharmacological metrics characterizing simple receptor-ligand systems may provide significant advantages when studying such biological systems.

In the second part of this paper, we introduced a family of receptor-ligand systems, called SRLK, in which the signaling complex, composed of a kinase, a ligand, and n transmembrane subunit chains, is built sequentially. These models could also form dummy and decoy complexes, similarly to the IL-7R models which the SRLK family encompasses. By manipulating the polynomials describing the SRLK models, we are able to derive an analytic expression of the amplitude under the no allostery assumption. We also show that the maximum of the dose-response curve for both our IL-7R models was indeed the amplitude of the models. Despite relatively strong assumptions, we believe that the SRLK approach can be used to model a broad range of biochemical systems, such as receptor competition in interleukin signaling. The analytic expressions obtained for the amplitude could improve our understanding of biological mechanisms requiring a fine tuning of cytokine signaling such as cancer treatment [57] or cytokine storm control [16, 53]. We showed in section 4.4 how our SRLK models can account for the competition for the gamma chain between the IL-2 family of receptors and the competition for receptor components between the IL-12 family of receptors. However, many receptors signal through different configurations. IL-35, for instance, can signal through homodimerization of gp130 or IL-12Rβ2 [7]. It has been shown that IL-6, a cytokine implied in cytokine storms [5, 16], signals through a hexameric structure composed of two IL-6Rα chains and two gp130 molecules [4]. Furthermore, it seems that the ligand IL-6 first binds to the IL-6Rα chain before any association with gp130 [4]. Thus, one could imagine other general receptor models that may involve any of the following: (1) homooligomerization (when two transmembrane chains X_i are identical); (2) other orders of signaling complex formation (nonsequential orders or, for instance, if the ligand is not the final subunit to be bound); (3) thermodynamic cycles (when there are several ways to form the signaling complex); (4) multiple kinases (including kinases binding to other subunit chains, such as JAK1 which binds to IL-7Rα [46]); or (5) a more detailed JAK-STAT pathway (most cytokine receptors activate multiple STAT molecules, whose copy numbers tune the immune response elicited [35]).

With this paper we hope to have initiated, or renewed, an interest for the algebraic analysis of receptor-ligand systems. Finally, we believe the results presented in this paper are a first step to account for the variability of receptor expression levels when designing and studying receptor-ligand models (both from an experimental and mathematical perspective) [17, 8, 23, 49].

Appendix A. Perturbation theory. A well-known difficulty with the lex Gröbner basis method, and polynomial equations in general, is that there is usually no analytic solution when the degree of a univariate polynomial is greater than four. This result is known as the Abel–Ruffini theorem [65]. Therefore, in order to make progress, we need to resort to either numerical computations or analytic approximations. Since receptor-ligand systems are often characterized by a sigmoidal dose-response curve, at least to calculate the amplitude, the only quantity of interest is the limit of the signaling function at infinity. In order to calculate this limit (where possible, analytically, otherwise numerically) we make use of perturbation theory for polynomial equations.

Greatly inspired by the Dover book written by Simmonds and Mann [56], this section reviews some notions of perturbation theory and justifies the steps of the method used to compute the analytic amplitude expression in section 3.1.2. We start by defining an asymptotic expansion.

DEFINITION A.1 (asymptotic expansion). *We say that*

$$\sum_{n=0}^{N} c_n\, f_n(\epsilon)$$

is an asymptotic expansion of f in ϵ if

- $\{f_n\}_{n=0,\dots,N+1}$ *is a gauge sequence, i.e.,* $f_n(\epsilon) = o(f_{n-1}(\epsilon))$ *as* $\epsilon \to 0$, *for* $n = 0, \dots, N+1$, *and*
- $f(\epsilon) - \sum_{n=0}^{N} c_n\, f_n(\epsilon) = \mathcal{O}(\epsilon^{N+1})$ *as* $\epsilon \to 0$.

The core of perturbation theory is the notion of asymptotic expansion and the following fundamental theorem.

THEOREM A.2 (fundamental theorem of perturbation theory). *If an asymptotic expansion satisfies*

$$A_0 + A_1\epsilon + A_2\epsilon^2 + \cdots + A_n\epsilon^N + \mathcal{O}(\epsilon^{N+1}) = 0$$

for any sufficiently small ϵ, and the coefficients A_i are independent of ϵ, then we have

$$A_0 = A_1 = \cdots = A_N = 0.$$

We are now ready to study the behavior of the root of a univariate polynomial. Let $n \in \mathbb{N}^*$, where $\mathbb{N}^*$ is the set of natural numbers without zero. We consider a univariate polynomial, $P_\epsilon(x)$, of degree n, in the variable x, with coefficients which depend on the parameter ϵ. We are interested in the behavior of the roots of $P_\epsilon(x)$ when $\epsilon \to 0$. This polynomial can be rewritten in the following form,

$$\begin{aligned} P_\epsilon(x) = {} & (1 + b_0\epsilon + c_0\epsilon^2 + \cdots) + A_1\epsilon^{\alpha_1}(1 + b_1\epsilon + c_1\epsilon^2 + \cdots)x + \cdots \\ & + A_n\epsilon^{\alpha_n}(1 + b_n\epsilon + c_n\epsilon^2 + \cdots)x^n, \end{aligned} \tag{A.1}$$

where for each i α_i is a rational number, b_i, c_i, ... are real constants, and $(1+b_i\epsilon+\cdots)$ is a regular asymptotic expansion of the general form

$$a_0 + a_1\epsilon + \cdots + a_N\epsilon^N + \mathcal{O}(\epsilon^{N+1}).$$

For such a polynomial, $P_\epsilon(x)$, we have the following result.

THEOREM A.3. *Each root of a polynomial, such as* (A.1) *is of the form*

$$x(\epsilon) = \epsilon^p \omega(\epsilon), \quad \omega(0) \neq 0, \tag{A.2}$$

where ω is a continuous function of ϵ for ϵ sufficiently small and $p \in \mathbb{Q}$.

The proof of this theorem (see [56]) gives a method to study the asymptotic behavior of the roots of polynomial (A.1).

Method. Let $P_\epsilon(x)$ be a polynomial that can be written as in (A.1). Let p be a rational and x a root of P_ϵ. Let us replace x by $\epsilon^p\omega(\epsilon)$ in P_ϵ. We can rewrite the polynomial as follows,

$$P_\epsilon(\epsilon^p\omega(\epsilon)) = Q_\epsilon(\omega) + \epsilon(b_0 + b_1A_1\epsilon^{\alpha_1+p}\omega(\epsilon) + \cdots + b_nA_n\epsilon^{\alpha_n+np}\omega(\epsilon)^n) + \cdots, \tag{A.3}$$

where

$$Q_\epsilon(\omega) = 1 + A_1\omega(\epsilon)\epsilon^{\alpha_1+p} + \cdots + A_n\omega(\epsilon)^n\epsilon^{\alpha_n+np}.$$

As $\epsilon \to 0$, the dominant term in P_ϵ is the term with the smallest exponent in Q_ϵ, i.e., the smallest element of

$$E = \{0, \alpha_1 + p, \ldots, \alpha_n + np\}. \tag{A.4}$$

However, the set E must have two identical values. Indeed, if $\alpha_k + kp$ is the smallest value of E, then we have

$$\epsilon^{-(\alpha_k + kp)} P_\epsilon(\epsilon^p \omega(\epsilon)) \underset{\epsilon \to 0}{\sim} A_k \omega(0).$$

Since $\omega(0) \neq 0$ and $P_\epsilon(\epsilon^p \omega(\epsilon)) = 0$ by hypothesis, we have $A_k = 0$ which is a contradiction with the fact that $\alpha_k + kp \in E$. To select the proper value of p, we follow a graphical algorithm which indicates when two or more components of E have equal minimal values:

1. On a plane (p, q), draw the lines $q = \alpha_j + jp$, $j = 1, \ldots, n$, and the line $q = 0$.
2. From the right, for p sufficiently large, the smallest exponent is 0. As p decreases (one can imagine a fictive vertical line moving from right to left), there will be a first point where at least two lines intersect ($q = 0$ and another one). Let us call this point $(p_1, 0)$. One line will have the largest slope, n_1.
3. Let the fictive vertical line keep moving to the left and follow this line of slope n_1 until the next intersection (p_2, e_2). Find the new intersected line with the largest slope n_2.
4. Continue until there is no other intersection. The last intersection involves the line with the largest slope of all the lines n.

We apply this method on an example and illustrate the algorithm in section 3.1.2. This algorithm finds all the intersection points of the lines of equation $q = \alpha_j + jp$, $j = 0, \ldots, n$, and $q = 0$ that are on the lower envelope of these lines. In this way, we have generated a set of pairs $\{(p_j, e_j)\}_{j=1,\ldots,m}$ corresponding to each intersection we encountered. Each of these intersection points corresponds to an asymptotic behavior of one branch of the roots of our original polynomial P_ϵ. Now let us define for each branch j, the scaled polynomial $T_\epsilon^{(j)}$, as follows:

$$T_\epsilon^{(j)}(\omega) = \epsilon^{-e_j} P_\epsilon(\epsilon^{p_j} \omega). \tag{A.5}$$

We can rewrite $T_\epsilon^{(j)}$ as a sum of two polynomials

$$T_\epsilon^{(j)}(\omega) = T_0^{(j)}(\omega) + E_\epsilon^{(j)}(\omega),$$

where $E_0^{(j)} = 0$ and $T_0^{(j)}$ do not depend on ϵ. The nonzero roots of $T_\epsilon^{(j)}$ (approached by the roots of $T_0^{(j)}$ as $\epsilon \to 0$) need to be regular but this is not necessarily the case. Indeed, α_j or (p_j, e_j) may be noninteger rationals or $T_0^{(j)}$ may have repeated roots. To obtain regular expansions, we introduce the new variable β such that:

$$\epsilon = \beta^{q_j}, \tag{A.6}$$

where q_j is the least common denominator of the set $\{0, \alpha_1 + p_j, \ldots, \alpha_n + np_j\}$. Finally, we define

$$R_\beta^{(j)}(\omega) = T_j(\omega, \beta^{q_j}) = \beta^{-q_j e_j} P(\beta^{q_j p_j} \omega, \beta^{q_j}), \tag{A.7}$$

where $T_j(\omega,\epsilon) = T_\epsilon^{(j)}(\omega)$ and $P(\omega,\epsilon) = P_\epsilon(\omega)$. The polynomial $R_\beta^{(j)}$ has the same roots as the polynomial $T_\epsilon^{(j)}$ but its nonzero roots have a regular expansion in ω of the form

$$\omega(\beta) = a_0 + a_1\beta + \cdots + a_N\beta^N + \mathcal{O}(\beta^{N+1}).$$

By substituting this expansion into $R_\beta^{(j)}$ and applying the fundamental theorem of perturbation theory (Theorem A.2), we find an expression for $a_0, a_1, \ldots$. We then come back to x with the transformation $x = \beta^{q_j p_j}\omega(\beta)$ for each branch. In practice we explore each branch one by one and can eliminate those which are irrelevant (for instance, when we have a negative root, since in our case the roots of the polynomials are concentrations of species, or $\omega(0) = 0$).

The above discussion can be summarized algorithmically as follows.

1. Replace the variable x by $\epsilon^p\omega(\epsilon)$ in $P_\epsilon(x)$, assuming $\omega(0) \neq 0$. One obtains a polynomial of the form
$$P_\epsilon(\epsilon^p\omega(\epsilon)) = Q_\epsilon(\omega) + \epsilon(\ldots) + \cdots.$$
2. Write the set of exponents for Q_ϵ: $E = \{0, \alpha_1 + p, \ldots, \alpha_n + np\}$.
3. Determine the pairs, (p_j, e_j), of proper values and minimal exponents following the graphical algorithm described above. Each pair corresponds to an asymptotic branch to explore.
4. For each branch j:
   4.1. Define $T_\epsilon^{(j)}(\omega) = \epsilon^{-e_j}P_\epsilon(\epsilon^{p_j}\omega)$.
   4.2. Introduce β such that $\epsilon = \beta_j^q$, where $q_j = \mathrm{lcd}(0, \alpha_1 + p, \ldots, \alpha_n + np)$, and define $R_\beta^{(j)}(\omega) = T_{\beta_j^q}^{(j)}(\omega)$.
   4.3. In $R_\beta^{(j)}(\omega) = 0$, substitute ω by a regular expansion $\omega(\beta) = a_0 + a_1\beta + \cdots + a_N\beta^N + \mathcal{O}(\beta^{N+1})$.
   4.4. Apply the fundamental theorem of perturbation theory to obtain an analytic expression for $a_0, a_1, \ldots$. Usually at this step, we can discriminate whether this branch is relevant (see example 3.1.2).
   4.5. Find the asymptotic expansion of the root of the original polynomial, P_ϵ by $x = \beta^{q_j p_j}\omega(\beta)$.

In this paper we are mainly interested in the first nonzero coefficient of the regular expansion of ω since it drives the behavior of the root of P_ϵ in the limit $\epsilon \to 0$.

Appendix B. Computation of EC_{50} for the IL-7R model. We make use of the expression for $\sigma(L)$, the signaling function described in (3.10), and (3.11), to isolate the square root in (3.14),

$$\sqrt{\Delta_2} = K_2(K_3L_{50} + 1)(N_x + N_y) + 1 - \frac{K_2(K_3L_{50} + 1)^2M}{K_3L}, \tag{B.1}$$

with $M = \min(N_x, N_y)$. We square the equation to remove the root and simplify the expression to obtain

$$\begin{aligned} 0 = {} & 4K_2^2(K_3L_{50} + 1)^2N_xN_y + K_2^2(K_3L_{50} + 1)^4\frac{M^2}{K_3^2L_{50}^2} \\ & - 2\frac{K_2^2(K_3L_{50} + 1)^3M(N_x + N_y)}{K_3L_{50}} - 2\frac{K_2(K_3L_{50} + 1)^2M}{K_3L_{50}}. \end{aligned} \tag{B.2}$$

Since we are looking for a positive value of L_{50}, we divide by $K_2(K_3L_{50}+1)^2$ and rewrite the previous expression as follows:

$$0 = 4K_2K_3^2L_{50}^2N_xN_y + K_2(K_3L_{50}+1)^2M^2 - 2K_2(K_3L_{50}+1)M(N_x+N_y)K_3L_{50} - MK_3L_{50}. \tag{B.3}$$

It leads to a polynomial of degree 2 in L_{50}:

$$0 = M^2K_2 + 2K_3L_{50}M(-1+K_2(M-N_x-N_y)) + K_3^2K_2L_{50}^2(M-2N_x)(M-2N_y). \tag{B.4}$$

The discriminant of this polynomial is positive,

$$\Delta = [1+K_2^2(N_y-N_x)^2 + 2K_2(N_x+N_y-M)]4K_3^2M^2, \tag{B.5}$$

so that there are two potential solutions:

$$L_{50}^+ = M\frac{1+K_2(N_x+N_y-M)+\sqrt{1+K_2^2(N_y-N_x)^2+2K_2(N_x+N_y-M)}}{K_2K_3(M-2N_x)(M-2N_y)},$$
$$L_{50}^- = M\frac{1+K_2(N_x+N_y-M)-\sqrt{1+K_2^2(N_y-N_x)^2+2K_2(N_x+N_y-M)}}{K_2K_3(M-2N_x)(M-2N_y)}.$$

Two solutions exist since by squaring (B.1) we lose the positive steady state hypothesis. Substituting these expressions back into the steady state equations shows that only L_{50}^+ leads to a biologically relevant solution. The use of the algebraic method described at the end of section 3.1.1 is more elegant as it gives directly the correct EC_{50} expression.

Appendix C. Macaulay2 code to compute Gröbner bases. Every Gröbner basis of this paper has been computed making use of the software Macaulay2 [25]. We provide the code to compute the Gröebner basis of the IL-7R model described in section 3.1.1.

```
R = frac(QQ[Nx,Ny,Nz,L,K1,K2,K3])[x,y,z,MonomialOrder=>Lex]

I = ideal(-Nx + x + K2*x*y + K1*x*z + K2*K_1*x*y*z
      + K3*K2*L*x*y + K3*K2*K1*L*x*y*z,
- Ny + y + K2*x*y + K2*K_1*x*y*z + K3*K2*L*x*y
      + K3*K2*K1*L*x*y*z,
- Nz + z + K1*x*z + K2*K_1*x*y*z + K3*K2*K1*L*x*y*z)

g = gens gb I
```

Appendix D. Analytic study of general sequential receptor-ligand systems.

D.1. Asymptotic study of the steady states.

LEMMA 4.7. *Any bounded, continuous solution of* (∗∗) *defined on* $\mathbb{R}$ *has a finite limit at* $+\infty$ *(and* $-\infty$*).*

Proof. Multiply (∗∗) by the common denominator of the R_i and let $x = \epsilon^{-1}$ to obtain

$$\underbrace{\left[\prod_{i=0}^{m} \tilde{q}_i(\epsilon)\right]}_{\tilde{r}_m(\epsilon)} \tilde{y}^m + \tilde{r}_{m-1}(\epsilon)\tilde{y}^{m-1} + \cdots + \tilde{r}_0(\epsilon) = 0,$$

with $\tilde{y} = \tilde{f}(\epsilon)$. We have now recast the original problem into the form of (A.1). By Theorem A.3 we know, that an expansion for the roots exists and we note that the points of $f(x)$ as $x \to \infty$ correspond to the points of $\tilde{f}(\epsilon)$ as $\epsilon = 0$. Note that, since all real $f(x)$ are bounded, so are the real $\tilde{f}(\epsilon)$. Therefore all real $\tilde{f}(0)$ are finite and equal to the limits $\lim_{x\to\infty} f(x)$. A unique limit is chosen by specifying a branch of $f(x)$. The proof for $x \to -\infty$ follows mutatis mutandis. ∎

LEMMA 4.8. *The signaling and the dummy functions of an SRLK model satisfying the experimental hypotheses admit a finite limit when $L \to +\infty$ and this limit is positive.*

Proof. The function σ (or δ) is an algebraic function bounded on $\mathbb{R}$ between 0 and $\min(N_z, N_1, \ldots, N_n)$ (or $\min(N_1, \ldots, N_n)$) so they admit a finite limit when $L \to +\infty$. Let us denote this limit by c_σ (or c_δ). We know that c_σ and c_δ are nonnegative because σ and δ are products of nonnegative functions.

Consider $c_\delta = 0$. Then since $\sigma(L) = K_0 z \prod_{i=1}^{n} \frac{K_i}{K_i'} \delta(L)$, we have $c_\sigma = 0$ (we note that z being also an algebraic function, z also admits a finite limit when $+\infty$). Since δ converges to 0, we need

$$\prod_{i=1}^{n} x_i \underset{L\to+\infty}{\sim} \frac{C_n}{L^p}, \tag{D.1}$$

with C_n a positive constant and $p > 1$. We recall and rewrite polynomial (4.2d):

$$N_n = x_n + K_0 z \prod_{i=1}^{n-1} K_i \prod_{i=1}^{n} x_i + \prod_{i=1}^{n-1} K_i' \prod_{i=1}^{n} x_i + \delta(L) + \sigma(L). \tag{D.2}$$

Assuming (D.1) when $L \to +\infty$ in (D.2), we obtain

$$\lim_{L\to+\infty} x_n = N_n,$$

and so we must have

$$\prod_{i=1}^{n-1} x_i \underset{L\to+\infty}{\sim} \frac{C_{n-1}}{L^p},$$

with $p > 1$ and C_{n-1} a positive constant. Passing to the limit in polynomial (4.2c) for $i = n - 1$, we obtain

$$\lim_{L\to+\infty} x_{n-1} = N_{n-1}.$$

We repeat the process for every conservation equation (4.2c) of the species X_i and we obtain

$$\forall i = 1, \ldots, n, \ \lim_{L\to+\infty} x_i = N_i,$$

which is in contradiction to (D.1). So $c_\delta > 0$.

Now, consider $c_\sigma = 0$. Then since $\sigma(L) = K_0 z \prod_{i=1}^{n} \frac{K_i}{K_i'} \delta(L)$, z has to tend to 0. However, when passing to the limit $L \to +\infty$ in (4.2a), we obtain

$$N_z = \lim_{L \to +\infty} (z + K_0 z x_1 + \cdots + \sigma(L)) = 0,$$

which is a contradiction.

Conclusion: $c_\sigma > 0$ and $c_\delta > 0$. ∎

LEMMA 4.9. *In an SRLK model under the experimental hypotheses, the concentration of the extrinsic intracellular kinase Z admits a positive finite limit, $c_z > 0$, when $L \to +\infty$.*

Proof. The concentration of kinase z being an algebraic function bounded on $\mathbb{R}$ between 0 and N_z, it admits a finite limit c_z when $L \to +\infty$. We know that $c_z \geq 0$ because z is a concentration. We now prove that $c_z > 0$. Since δ converges to a positive constant when $L \to +\infty$, we must have

$$\prod_{i=1}^{n} x_i \underset{L \to +\infty}{\sim} \frac{c_d}{L},$$

where c_d is a positive constant. Since σ also admits a finite limit when $L \to +\infty$, it means that

$$z \prod_{i=1}^{n} x_i \underset{L \to +\infty}{\sim} \frac{c_s}{L},$$

where c_s is a positive constant. So z has to satisfy

$$z \underset{L \to +\infty}{\sim} c_z,$$

where $c_z = \frac{c_s}{c_d}$ is a positive constant. ∎

LEMMA 4.10. *Consider an SRLK model which satisfies the experimental hypotheses. If we assume no allostery, then the steady state value of the extrinsic intracellular kinase, z, is given by*

$$z = \frac{-1 + K_0(N_z - N_1) + \sqrt{\Delta_z}}{2K_0}, \tag{4.3}$$

where

$$\Delta_z = (1 + K_0(N_1 - N_z))^2 + 4K_0 N_z.$$

Proof. We assumed no allostery so $K_i = K_i'$ for all $i = 1, \ldots, n$. Equation (4.2a) gives

$$N_z - z = K_0 z \left(x_1 + \sum_{j=2}^{n} \left(\prod_{l=1}^{j-1} K_l x_l x_j \right) + L \prod_{j=1}^{n} K_j x_j \right).$$

By substituting this equality into (4.2b), we obtain

$$N_1 = N_z - z + \frac{N_z - z}{K_0 z},$$

so z is a positive root of the polynomial

$$-N_z + z(1 + K_0(N_1 - N_z)) + K_0 z^2$$

with L-independent coefficients. The two possibilities are

$$z_1 = \frac{-1 + K_0(N_z - N_1) + \sqrt{4K_0N_z + (1 + K_0(N_1 - N_z))^2}}{2K_0},$$
$$z_2 = \frac{-1 + K_0(N_z - N_1) - \sqrt{4K_0N_z + (1 + K_0(N_1 - N_z))^2}}{2K_0}.$$

The expression z_1 is always positive while z_2 is always negative. Hence z_1 is the steady state kinase concentration, z. ∎

LEMMA 4.11. *Let us consider an SRLK model which satisfies the experimental hypotheses. Then no concentration x_i behaves proportionally to L^q, $q > 0$, or $\frac{1}{L^p}$, $p > 1$, when $L \to +\infty$.*

Proof. Lemma 4.9 affirms that z tends to a positive constant when $L \to +\infty$. In order for σ or δ to converge to a positive constant as stated in lemma 4.8, we need

$$\prod_{i=1}^{n} x_i \underset{L\to+\infty}{\sim} \frac{c}{L}, \tag{D.3}$$

where c is a positive constant. Since the concentrations $x_1, \ldots, x_n$ are bounded functions (between 0 and their respective N_i), it is impossible to have for any $i = 1 \ldots n$, $x_i \underset{L\to+\infty}{\sim} c_i L^q$ with c_i constant and $q > 0$. From (D.3) it follows that it is impossible to have any $x_i \underset{L\to+\infty}{\sim} \frac{c_i}{L^p}$ for $p > 1$. ∎

THEOREM 4.12. *We consider an SRLK model which satisfies the experimental hypotheses. If there exists a unique limiting component X_{i_0}, then*

$$x_{i_0} \underset{L\to+\infty}{\sim} \frac{c_{i_0}}{L},$$

and for all $i = 1, \ldots, n, i \neq i_0$,

$$x_i \underset{L\to+\infty}{\sim} c_i,$$

where c_{i_0} and c_i are positive constants.

Proof. Since the concentrations x_i are algebraic functions (with coefficients in $\mathbb{R}$) bounded on $\mathbb{R}$, they admit a nonnegative limit when $L \to +\infty$.

We know that we need

$$\prod_{i=1}^{n} x_i \underset{L\to+\infty}{\sim} \frac{c}{L}, \tag{D.4}$$

with c a positive constant, so that σ and δ converge when $L \to +\infty$. Lemma 4.9 shows that z tends to a positive constant when $L \to +\infty$. Thus, it follows from (D.4) and Lemma 4.11 that at least one of the x_i must tend to 0. We will prove that the only concentration that can tend to 0 is x_{i_0} and so $x_{i_0} \underset{L\to+\infty}{\sim} \frac{c_{i_0}}{L}$ with c_{i_0} a constant.

(1) There exists at least one chain X_j whose concentration tends to 0. The conservation equation of X_j described in (4.2c) is

$$N_j = x_j + \prod_{i=1}^{j} K_i' x_i + K_0 z \prod_{i=1}^{j} K_i x_i + \prod_{i=1}^{j+1} K_i' x_i + K_0 z \prod_{i=1}^{j+1} K_i x_i + \cdots + \delta(L) + \sigma(L).$$

When $L \to +\infty$, we obtain

$$N_j = \lim_{L\to+\infty} \delta(L) + \lim_{L\to+\infty} \sigma(L).$$

We cannot form more dummy or signaling complexes than the number of molecules available. Since X_{i_0} is the limiting component, we have

$$\delta(L) + \sigma(L) \le N_{i_0} \quad \forall L.$$

This yields, in the limit $L \to +\infty$, $N_j \le N_{i_0}$. By hypothesis this implies that $j = i_0$ and so X_j is our limiting component X_{i_0}.

(2) Reciprocally, if x_{i_0} tends to a positive constant when $L \to +\infty$, then there exists at least one x_j, $j \ne i_0$, such that $x_j \to 0$ when $L \to +\infty$. The limit when $L \to +\infty$ of (4.2c) when $i = j$ gives

$$\lim_{L\to+\infty} [\delta(L) + \sigma(L)] = N_j.$$

However, since we also have $\delta + \sigma \le N_{i_0}$, we obtain when taking the limit, $N_j \le N_{i_0}$, which is a contradiction with the fact that X_{i_0} is the only limiting component.

Conclusion: X_{i_0} is limiting if and only if its concentration tends to 0, and we have

$$x_{i_0} \underset{L\to+\infty}{\sim} \frac{c_{i_0}}{L},$$

and for $i \ne i_0$,

$$x_i \underset{L\to+\infty}{\sim} c_i,$$

where c_{i_0} and c_i are positive constants. ∎

COROLLARY 4.13. *If an SRLK model, which satisfies the experimental hypotheses, has multiple limiting components, $X_{i_1}, \ldots, X_{i_r}$, $i_1 < \cdots < i_r$, then*

$$x_{i_1} \underset{L\to+\infty}{\sim} \frac{c_{i_1}}{L^{p_1}}, \ldots, x_{i_r} \underset{L\to+\infty}{\sim} \frac{c_{i_r}}{L^{p_r}},$$

where $c_{i_1}, \ldots c_{i_r}$ are positive constants and $p_1 = \cdots = p_r = \frac{1}{r}$. The concentrations of the nonlimiting components, x_i (for $i \notin \{i_1, \ldots, i_r\}$), tend to positive constants, $c_i > 0$.

Proof. If X_{i_1} and X_{i_2} are limiting components, they are the only ones whose concentrations, x_{i_1} and x_{i_2}, tend to 0 when $L \to +\infty$. From (D.4) we can write

$$x_{i_1} \underset{L\to+\infty}{\sim} \frac{c_{i_1}}{L^{p_1}},$$
$$x_{i_2} \underset{L\to+\infty}{\sim} \frac{c_{i_2}}{L^{p_2}},$$

with c_{i_1} and c_{i_2} constants and $p_1, p_2 > 0$, such that $p_1 + p_2 = 1$.

From system (4.2), we have:

$$N_{i_1} = x_{i_1} + \sum_{j=i_1}^{n} \left(\prod_{l=1}^{j-1} K_l' x_l x_j + K_0 z \prod_{l=1}^{j-1} K_l x_l x_j \right) + L \prod_{j=1}^{n} K_j' x_j + K_0 z L \prod_{j=1}^{n} K_j x_j,$$
$$N_{i_2} = x_{i_2} + \sum_{j=i_2}^{n} \left(\prod_{l=1}^{j-1} K_l' x_l x_j + K_0 z \prod_{l=1}^{j-1} K_l x_l x_j \right) + L \prod_{j=1}^{n} K_j' x_j + K_0 z L \prod_{j=1}^{n} K_j x_j.$$

Since X_{i_1} and X_{i_2} are limiting components, we have $N_{i_1} = N_{i_2}$ and, if $i_1 < i_2$, we obtain

(D.5)
$$N_{i_1} = N_{i_2} \iff x_{i_1} \left(1 + \sum_{j=i_1}^{i_2-1} \left(\prod_{l=1}^{j-1} K_l' \prod_{l=1, l\neq i_1}^{j} x_l + K_0 z \prod_{l=1}^{j-1} K_l \prod_{l=1, l\neq i_1}^{j} x_l \right) \right) = x_{i_2}.$$

Since all the x_i with $i \neq i_1$, $i \neq i_2$, tend to a positive constant when $L \to +\infty$, we have

$$1 + \sum_{j=i_1}^{i_2-1} \left(\prod_{l=1}^{j-1} K_l' \prod_{l=1, l\neq i_1}^{j} x_l + K_0 z \prod_{l=1}^{j-1} K_l \prod_{l=1, l\neq i_1}^{j} x_l \right) \underset{L\to+\infty}{\sim} C,$$

where C is a positive constant. Thus, we obtain the behavior of the left side of (D.5),

$$x_{i_1} \left(1 + \sum_{j=i_1}^{i_2-1} \left(\prod_{l=1}^{j-1} K_l' \prod_{l=1, l\neq i_1}^{j} x_l + K_0 z \prod_{l=1}^{j-1} K_l \prod_{l=1, l\neq i_1}^{j} x_l \right) \right) \underset{L\to+\infty}{\sim} \frac{C c_{i_1}}{L^{p_1}}.$$

Since the right side is given by

$$x_{i_2} \underset{L\to+\infty}{\sim} \frac{c_{i_2}}{L^{p_2}},$$

it results in $p_1 = p_2 = \frac{1}{2}$.

If there are r limiting components $x_{i_1}, \ldots, x_{i_r}$, $i_1 < \cdots < i_r$, then we have

$$\begin{aligned} x_{i_1} &\underset{L\to+\infty}{\sim} \frac{c_{i_1}}{L^{p_1}}, \\ &\vdots \qquad\qquad \vdots \\ x_{i_r} &\underset{L\to+\infty}{\sim} \frac{c_{i_r}}{L^{p_r}}, \end{aligned}$$

with $c_{i_1}, \ldots, c_{i_r}$ positive constants, and $p_1, \ldots, p_r > 0$ such that $p_1 + \cdots + p_r = 1$. We proceed the same way as for the case of two limiting components and we obtain $p_1 = \cdots = p_r = \frac{1}{r}$. ∎

D.2. Asymptotic study of the signaling and dummy functions.

THEOREM 4.14. *Consider an SRLK model which satisfies the experimental hypotheses. Let us write $X_{i_1}, \ldots, X_{i_r}$ as the limiting components and denote N_{i_0} as the total amount of any limiting component, i.e., $N_{i_0} \equiv N_{i_1} = \cdots = N_{i_r}$. The limit of the signaling function is given by*

$$\lim_{L\to+\infty} \sigma(L) = \frac{\prod_{i=1}^{n} K_i K_0 c_z}{\prod_{i=1}^{n} K_i' + \prod_{i=1}^{n} K_i K_0 c_z} N_{i_0},$$

and the limit of the dummy function is

$$\lim_{L\to+\infty}\delta(L)=\frac{\prod_{i=1}^{n}K_i'}{\prod_{i=1}^{n}K_i'+\prod_{i=1}^{n}K_iK_0c_z}N_{i_0},$$

where

$$c_z=\lim_{L\to+\infty}z.$$

Proof. By the definition of σ and δ we have

$$\delta(L)+\sigma(L)=L\prod_{i=1}^{n}x_i\left(\prod_{i=1}^{n}K_i'+K_0z\prod_{i=1}^{n}K_i\right),$$

which implies that

$$\lim_{L\to+\infty}[\delta(L)+\sigma(L)]=\lim_{L\to+\infty}\left(L\prod_{i=1}^{n}x_i\left(\prod_{i=1}^{n}K_i'+K_0z\prod_{i=1}^{n}K_i\right)\right).$$

Using the limit properties and since everything converges, we obtain

$$\lim_{L\to+\infty}[\delta(L)+\sigma(L])=\lim_{L\to+\infty}\left(L\prod_{i=1}^{n}x_i\right)\left(\prod_{i=1}^{n}K_i'+K_0\lim_{L\to\infty}(z)\prod_{i=1}^{n}K_i\right).$$

However, Theorem 4.12 states that x_{i_0} tends to 0 when $L\to+\infty$. Thus, (4.2c) when $i=i_0$ at $L\to+\infty$ gives

$$N_{i_0}=\lim_{L\to+\infty}[\delta(L)+\sigma(L)].$$

Consequently since $z\to c_z>0$ from Lemma 4.9, we obtain

$$\lim_{L\to+\infty}\left(L\prod_{i=1}^{n}x_i\right)=\frac{N_{i_0}}{\prod_{i=1}^{n}K_i'+K_0c_z\prod_{i=1}^{n}K_i}.$$

We substitute this limit into the expression of σ and δ and obtain the desired expressions. ∎

COROLLARY 4.15. *Consider the system of Theorem* 4.14 *and assume there is no allostery. Denote the limiting components by* $X_{i_1},\dots,X_{i_r}$ *and and let* N_{i_0} *be the total amount of any limiting component, i.e.,* $N_{i_0}\equiv N_{i_1}=\cdots=N_{i_r}$. *The limit of the signaling function is*

$$\lim_{L\to+\infty}\sigma(L)=\frac{K_0z}{1+K_0z}N_{i_0},$$

and the limit of the dummy function is

$$\lim_{L\to+\infty}\delta(L)=\frac{N_{i_0}}{1+K_0z},$$

with z *given by* (4.3) *in Lemma* 4.10.

Proof. Since there is no allostery, we have $K_i = K'_i$ for all i. Lemma 4.10 states that z is independent of L, thus $c_z = z$. Applying these statements in the expressions of the previous theorem, we obtain the expressions of this corollary. ∎

LEMMA 4.16. *Consider an SRLK model under the experimental hypotheses. If there is no allostery, then we have*

$$\sup \sigma(L) = \lim_{L \to +\infty} \sigma(L).$$

Proof. Since X_{i_0} is the limiting component, we know from Theorem 4.12 that its concentration tends to 0 when $L \to +\infty$. We have

$$\delta + \sigma \leq N_{i_0} = \lim_{L \to +\infty} (\delta + \sigma). \tag{D.6}$$

In the no allostery case, z is independent of L and we have $\sigma = K_0 z \delta$. Thus, (D.6) gives

$$(1 + K_0 z)\delta \leq (1 + K_0 z) \lim_{L \to +\infty} \delta.$$

Hence, we can conclude

$$\lim_{L \to +\infty} \delta = \sup \delta,$$

and

$$\lim_{L \to +\infty} \sigma = \sup \sigma.$$ ∎

D.3. SRLK models with additional receptor sub-units.

THEOREM 4.20. *The theorems and lemmas previously true for the SRLK models are true for the extended SRLK models under the same (extended) hypotheses.*

Proof. The concentrations y_i are bounded ($0 \leq y_i \leq N_{y_i}$) algebraic functions on $\mathbb{R}$, and therefore admit a limit when $L \to +\infty$. As the expressions of σ and δ are not modified, the addition of the Y_i variables to an SRLK model, assuming the extended experimental hypotheses, does not change the proofs of the previous lemmas and theorems. ∎

Acknowledgments. We would like to thank Grégoire Altan-Bonnet for inspiring this work and for encouraging us to explore with analytical methods how the dose-response curve depends on receptor expression levels. We thank Elisenda Feliú and Alicia Dickenstein for carefully reading and providing detailed feedback to an earlier version of this manuscript. We also thank Grant Lythe and Martín López-García for their input in some of the early ideas behind this work and for supervising the doctoral project of one of us (LS). This manuscript has been internally reviewed at Los Alamos National Laboratory, and assigned the reference number LA-UR-22-25806 (CMP).

REFERENCES

[1] B. ALEXANDER, D. BROWSE, S. READING, AND I. BENJAMIN, *A simple and accurate mathematical method for calculation of the EC50*, J. Pharmacol. Toxicol. Methods, 41 (1999), pp. 55–58.

[2] G. Altan-Bonnet and R. Mukherjee, *Cytokine-mediated communication: A quantitative appraisal of immune complexity*, Nature Rev. Immunol., (2019), p. 1.
[3] E. Bianconi et al., *An estimation of the number of cells in the human body*, Ann. Human Biol., 40 (2013), pp. 463–471.
[4] M. J. Boulanger, D.-c. Chow, E. E. Brevnova, and K. C. Garcia, *Hexameric structure and assembly of the interleukin-6/IL-6 α-receptor/gp130 complex*, Science, 300 (2003), pp. 2101–2104.
[5] L. Y. Chen, C. M. Biggs, S. Jamal, S. Stukas, C. L. Wellington, and M. S. Sekhon, *Soluble interleukin-6 receptor in the covid-19 cytokine storm syndrome*, Cell Rep. Med., 2 (2021), 100269.
[6] Z. Chen, R. Bertin, and G. Froldi, *EC50 estimation of antioxidant activity in DPPH assay using several statistical programs*, Food Chem., 138 (2013), pp. 414–420.
[7] L. W. Collison et al., *The composition and signaling of the IL-35 receptor are unconventional*, Nat. Immunol., 13 (2012), pp. 290–299.
[8] J. W. Cotari, G. Voisinne, O. E. Dar, V. Karabacak, and G. Altan-Bonnet, *Cell-to-cell variability analysis dissects the plasticity of signaling of common γ chain cytokines in t cells*, Sci. Signal., 6 (2013), ra17.
[9] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Aalgorithms,* Undergrad. Texts Mathematics, Springer, New York,1997.
[10] G. Y. Di Veroli, C. Fornari, I. Goldlust, G. Mills, S. B. Koh, J. L. Bramhall, F. M. Richards, and D. I. Jodrell, *An automated fitting procedure and software for dose-response curves with multiphasic features*, Sci. Rep., 5 (2015), pp. 1–11.
[11] A. Dickenstein, M. P. Millan, A. Shiu, and X. Tang, *Multistationarity in structured reaction networks*, Bull. Math. Biol., 81 (2019), pp. 1527–1581.
[12] A. Dickenstein and M. Pérez Millán, *How far is complex balancing from detailed balancing?*, Bull. Math. Biol., 73 (2011), pp. 811–828.
[13] E. F. Douglass Jr, C. J. Miller, G. Sparer, H. Shapiro, and D. A. Spiegel, *A comprehensive mathematical model for three-body binding equilibria*, J. Amer. Chem. Soc., 135 (2013), pp. 6092–6099.
[14] O. Dushek, M. Aleksic, R. J. Wheeler, H. Zhang, S.-P. Cordoba, Y.-C. Peng, J.-L. Chen, V. Cerundolo, T. Dong, D. Coombs, and P. A. van der Merwe, *Antigen potency and maximal efficacy reveal a mechanism of efficient t cell activation*, Sci. Signal., 4 (2011), ra39.
[15] R. Eftimie, J. J. Gillard, and D. A. Cantrell, *Mathematical models for immunology: Current state of the art and future research directions*, Bull. Math. Biol., 78 (2016), pp. 2091–2134.
[16] D. C. Fajgenbaum and C. H. June, *Cytokine storm*, New England J. Med., 383 (2020), pp. 2255–2273.
[17] A. M. Farhat, A. C. Weiner, C. Posner, Z. S. Kim, B. Orcutt-Jahns, S. M. Carlson, and A. S. Meyer, *Modeling cell-specific dynamics and regulation of the common gamma chain cytokines*, Cell Rep., 35 (2021), 109044.
[18] M. Feinberg, *Chemical reaction network structure and the stability of complex isothermal reactors—*I. *The deficiency zero and deficiency one theorems*, Chem. Eng. Sci., 42 (1987), pp. 2229–2268.
[19] M. Feinberg, *Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity*, Chem. Eng. Sci., 44 (1989), pp. 1819–1827.
[20] O. Feinerman, G. Jentsch, K. E. Tkach, J. W. Coward, M. M. Hathorn, M. W. Sneddon, T. Emonet, K. A. Smith, and G. Altan-Bonnet, *Single-cell quantification of IL-2 response by effector and regulatory t cells reveals critical plasticity in immune response*, Mol. Syst. Biol., 6 (2010), 437.
[21] J. Gabrielsson, L. A. Peletier, and S. Hjorth, *Lost in translation: What's in an EC50? Innovative pk/pd reasoning in the drug development context*, Eur. J. Pharmacol., 835 (2018), pp. 154–161.
[22] R. Gesztelyi, J. Zsuga, A. Kemeny-Beke, B. Varga, B. Juhasz, and A. Tosaki, *The Hill equation and the origin of quantitative pharmacology*, Arch. Hist. Exact Sci., 66 (2012), pp. 427–438.
[23] P. Gonnord, B. R. Angermann, K. Sadtler, E. Gombos, P. Chappert, M. Meier-Schellersheim, and R. Varma, *A hierarchy of affinities between cytokine receptors and the common gamma chain leads to pathway cross-talk*, Sci. Signal., 11 (2018), eaal1253.
[24] S. Goutelle, M. Maurin, F. Rougier, X. Barbaut, L. Bourguignon, M. Ducher, and P. Maire, *The Hill equation: A review of its capabilities in pharmacological modelling*, Fundam. Clin. Pharmacol., 22 (2008), pp. 633–648.
[25] D. R. Grayson and M. E. Stillman, *Macaulay2,* http://www.math.uiuc.edu/Macaulay2/.

[26] E. Gross, H. A. Harrington, Z. Rosen, and B. Sturmfels, *Algebraic systems biology: A case study for the Wnt pathway*, Bull. Math. Biol., 78 (2016), pp. 21–51.
[27] F. Horn and R. Jackson, *General mass action kinetics*, Arch. Ration. Mech. Anal., 47 (1972), pp. 81–116.
[28] K. A. Janes and D. A. Lauffenburger, *Models of signalling networks-what cell biologists can gain from them and give to them*, J. Cell Sci., 126 (2013), pp. 1913–1921.
[29] X. Jiang and A. Kopp-Schneider, *Summarizing EC50 estimates from multiple dose-response experiments: A comparison of a meta-analysis strategy to a mixed-effects model approach*, Biomet. J., 56 (2014), pp. 493–512.
[30] D. Lambert, *Drugs and receptors*, Cont. Educ. Anaesthesia Crit. Care Pain, 4 (2004), pp. 181–184.
[31] D. A. Lauffenburger and J. Linderman, *Receptors: Models for Binding, Trafficking, and Signaling*, Oxford University Press, New York, 1996.
[32] W. J. Leonard, J.-X. Lin, and J. J. O'Shea, *The γc family of cytokines: Basic biology to therapeutic ramifications*, Immunity, 50 (2019), pp. 832–850.
[33] J.-L. Li, X.-Y. Liu, J.-T. Xie, Y.-L. Di, and F.-X. Zhu, *A comparison of different estimation methods for fungicide EC50 and EC95 values*, J. Phytopathol., 163 (2015), pp. 239–244.
[34] J.-X. Lin and W. J. Leonard, *The common cytokine receptor γ chain family of cytokines*, Cold Spring Harbor Perspect. Biol., 10 (2018), a028449.
[35] J.-X. Lin and W. J. Leonard, *Fine-tuning cytokine signals*, Annu. Rev. Immunol., 37 (2019), pp. 295–324.
[36] A. Ma, R. Koka, and P. Burkett, *Diverse functions of IL-2, IL-15, and ILil-7 in lymphoid homeostasis*, Annu. Rev. Immunol., 24 (2006), pp. 657–679.
[37] J. Macdougall, *Analysis of dose-response studies—e max model*, in *Dose Finding in Drug Development*, Springer, New York, 2006, pp. 127–145.
[38] E. T. Mack, R. Perez-Castillejos, Z. Suo, and G. M. Whitesides, *Exact analysis of ligand-induced dimerization of monomeric receptors*, Anal. Chem., 80 (2008), pp. 5550–5555.
[39] S. R. Maxwell and D. J. Webb, *Receptor functions*, Medicine, 36 (2008), pp. 344–349.
[40] V. Mityunin and E. V. Pankratiev, *Parallel algorithms for Groebner-basis construction*, J. Math. Sci., 142 (2007), pp. 2248–2266.
[41] C. Molina-Paris, J. Reynolds, G. Lythe, and M. C. Coles, *Mathematical model of naive t cell division and survival IL-7 thresholds*, Front. Immunol., 4 (2013), 434.
[42] A. Montes and M. Wibmer, *Gröbner bases for polynomial systems with parameters*, J. Symbolic Comput., 45 (2010), pp. 1391–1425.
[43] I. Moraga, J. B. Spangler, J. L. Mendoza, M. Gakovic, T. S. Wehrman, P. Krutzik, and K. C. Garcia, *Synthekines are surrogate cytokine and growth factor agonists that compel signaling through non-natural receptor dimers*, Elife, 6 (2017), e22882.
[44] I. Otero-Muras, P. Yordanov, and J. Stelling, *Chemical reaction network theory elucidates sources of multistability in interferon signaling*, Plos Comput. Biol., 13 (2017), e1005454.
[45] M. J. Palmer, V. S. Mahajan, L. C. Trajman, D. J. Irvine, D. A. Lauffenburger, and J. Chen, *Interleukin-7 receptor signaling network: An integrated systems perspective*, Cell Mol. Immunol., 5 (2008), pp. 79–89.
[46] J.-H. Park, A. T. Waickman, J. Reynolds, M. Castro, and C. Molina-París, *Il-7 receptor signaling in t cells: A mathematical modeling perspective*, Wiley Interdiscip. Rev. Syst. Biol. Med., 11 (2019), e1447.
[47] M. Pérez Millán and A. Dickenstein, *Implicit dose-response curves*, J. Math. Biol., 70 (2015), pp. 1669–1684.
[48] M. E. Raeber, Y. Zurbuchen, D. Impellizzieri, and O. Boyman, *The role of cytokines in t-cell memory in health and disease*, Immunol. Rev., 283 (2018), pp. 176–193.
[49] A. M. Ring et al., *Mechanistic and structural insight into the functional dichotomy between IL-2 and IL-15*, Nat. Immunol., 13 (2012), pp. 1187–1195.
[50] Y. Rochman, R. Spolski, and W. J. Leonard, *New insights into the regulation of t cells by γ c family cytokines*, Nat. Rev. Immunol., 9 (2009), pp. 480–490.
[51] G. E. Rovati and S. Nicosia, *Lower efficacy: Interaction with an inhibitory receptor or partial agonism?*, Trends Pharmacol. Sci., 15 (1994), pp. 140–144.
[52] A. Sadeghimanesh and E. Feliu, *Gröbner bases of reaction networks with intermediate species*, Adv. Appl. Math., 107 (2019), pp. 74–101.
[53] C. Savarin and C. C. Bergmann, *Fine tuning the cytokine storm by IFN and IL-10 following neurotropic coronavirus encephalomyelitis*, Front. Immunol., 9 (2018), 3022.
[54] J. Schlessinger, *Cell signaling by receptor tyrosine kinases*, Cell, 103 (2000), pp. 211–225.
[55] A. J. Shiu, *Algebraic Methods for Biochemical Reaction Network Theory*, University of California, Berkeley, Berkeley, CA, 2010.

[56] J. G. Simmonds and J. E. Mann, *A First Look at Perturbation Theory*, Dover, Mineola, NY, 2013.
[57] R. Spolski, D. Gromer, and W. J. Leonard (2017). *The γ c family of cytokines: Fine-tuning signals from IL-2 and IL-21 in the regulation of the immune response*, F1000 Research,6 (2017), 1872.
[58] R. Suriyatem, R. A. Auras, P. Intipunya, and P. Rachtanapun, *Predictive mathematical modeling for EC50 calculation of antioxidant activity and antibacterial ability of Thai bee products*, J. Appl. Pharm. Sci., 7 (2017), pp. 122–133.
[59] N. Thomas, *Hypothesis testing and Bayesian estimation using a sigmoid e max model applied to sparse dose-response designs*, J. Biopharm. Stat., 16 (2006), pp. 657–677.
[60] I. Uings and S. Farrow, *Cell receptors and cell signalling*, Mol. Pathol., 53 (2000), 295.
[61] D. A. Vignali and V. K. Kuchroo, *Il-12 family cytokines: Immunological playmakers*, Nature Immunol., 13 (2012), pp. 722–728.
[62] A. V. Villarino, Y. Kanno, J. R. Ferdinand, and J. J. O'Shea, *Mechanisms of JAK/STAT signaling in immunity and disease*, J. Immunol., 194 (2015), pp. 21–27.
[63] V. Weispfenning, *Comprehensive Gröbner bases*, J. Symbolic Comput., 14 (1992), pp. 1–29.
[64] H. S. Wiley, S. Y. Shvartsman, and D. A. Lauffenburger, *Computational modeling of the EGF-receptor system: A paradigm for systems biology*, Trends Cell Biol., 13 (2003), pp. 43–50.
[65] H. Żołdek, *The topological proof of Abel-Ruffini theorem*, Topol. Methods Nonlinear Anal., 16 (2000), pp. 253–265.