






Article

On Exact and Approximate Approaches for Stochastic Receptor-Ligand Competition Dynamics—An Ecological Perspective

Polly-Anne Jeffrey ¹, Martín López-García ^{1,*}, Mario Castro ² and Grant Lythe ¹
and Carmen Molina-París ¹

¹ Department of Applied Mathematics, School of Mathematics, University of Leeds, Leeds LS2 9JT, UK; mm13paj@leeds.ac.uk (P.-A.J.); grant@maths.leeds.ac.uk (G.L.); C.MolinaParis@leeds.ac.uk (C.M.-P.)

² Instituto de Investigación Tecnológica (IIT) and Grupo Interdisciplinar de Sistemas Complejos (GISC), Universidad Pontificia Comillas, E-28015 Madrid, Spain; marioc@comillas.edu

* Correspondence: m.lopezgarcia@leeds.ac.uk

Received: 15 May 2020; Accepted: 16 June 2020; Published: 20 June 2020



Abstract: Cellular receptors on the cell membrane can bind ligand molecules in the extra-cellular medium to form ligand-bound monomers. These interactions ultimately determine the fate of a cell through the resulting intra-cellular signalling cascades. Often, several receptor types can bind a shared ligand leading to the formation of different monomeric complexes, and in turn to competition for the common ligand. Here, we describe competition between two receptors which bind a common ligand in terms of a bi-variate stochastic process. The stochastic description is important to account for fluctuations in the number of molecules. Our interest is in computing two summary statistics—the steady-state distribution of the number of bound monomers and the time to reach a threshold number of monomers of a given kind. The matrix-analytic approach developed in this manuscript is exact, but becomes impractical as the number of molecules in the system increases. Thus, we present novel approximations which can work under low-to-moderate competition scenarios. Our results apply to systems with a larger number of population species (i.e., receptors) competing for a common resource (i.e., ligands), and to competition systems outside the area of molecular dynamics, such as Mathematical Ecology.

Keywords: receptor-ligand interaction; continuous-time Markov chain; summary statistics; steady-state; first-passage time; approximation

1. Introduction

Receptors play a significant role in determining the fate of cells through the initiation of intra-cellular phosphorylation (signalling) cascades [1–5]. For receptors on the cellular membrane, these cascades can be initiated by the binding of a ligand molecule, in the extra-cellular medium, to the extra-cellular domain of the receptor, forming a monomer. Tyrosine residues on the intra-cellular tail of the bound receptor can become phosphorylated, in turn, allowing the receptor to interact with downstream signalling proteins in the cellular cytoplasm. The fate of the cell, for example division, death or migration, will be determined by the specific receptor-ligand interaction and by the number of phosphorylated monomers. Often, ligand molecules are capable of binding to more than one kind of receptor [4,6]. These receptors may be similar in structure, yet different receptor-ligand interactions might lead to strikingly diverse cellular outcomes [7]. The strength of the signalling determines cellular fate, and in some systems this strength can be assumed to be proportional to the number of monomers formed [8,9]. It is therefore important to quantify the number, and timescales of formation, of the different ligand-bound monomers on the cell membrane.

When a ligand can bind two or more receptor types, competition for a shared resource will arise. For example, a ligand which is shared by two receptors is the vascular endothelial growth factor A (VEGF-A), which can bind two different vascular endothelial growth factor receptors (VEGFRs), namely VEGFR1 and VEGFR2 [10]. Although both receptors bind the same ligand, they have been shown to have different functions in normal and tumour vasculature [11,12]. Shibuya [11] explains how, during embryogenesis, these receptors have opposing functions upon VEGF-A stimulation; VEGFR1 is a negative regulator for angiogenesis (the formation of new blood cells), whereas VEGFR2 is a positive regulator. In the same reference, it is argued that the receptors have differing roles in pathology and the author suggests that it may be useful, in the control of disease, to be able to inhibit the receptors individually. In the development of such inhibitors it is crucial to be able to quantify the number, and timescales to formation, of ligand-bound receptor monomers.

There have been many attempts in the literature to model, mostly deterministically, the interaction between specific receptors and their corresponding ligands [13–18]. Stochastic models, which are generally more complex to analyse, can be especially relevant when there are low copy numbers of some molecular species (ligands or receptors), or when the dynamics of a small number of molecules regulates cellular fate. Stochastic descriptions have been useful to quantify the timescales that characterise cellular responses, where the time to reach a threshold number of ligand-bound monomers has been shown to correlate with the time to signal initiation and cellular response [19,20]. A range of stochastic methods has been used to analyse these receptor-ligand systems [18,21–24]. Gurevich et al. propose a stochastic model for receptor-ligand binding, in particular, a continuous-time birth-and-death Markov process [23]. They argue that such a stochastic model can be relevant for some ligand molecules, such as pharmaceuticals or naturally produced cytokines (for example, IL-2), which can produce a cellular response at a very low dose. A continuous-time birth-and-death Markov process [23], as used by Gurevich et al., is a stochastic process in one variable, so that in a sufficiently small time interval Δt , at most one change in the state of the variable can occur, either a birth, increasing the state variable by one, or a death, decreasing the state variable by one [25]. In 1961 Reuter [26] generalised these processes to two dimensions, framed within the theory of population dynamics in Mathematical Ecology, by considering Markov processes in two variables, for which the only possible transitions increase or decrease the value of a single variable by one, or simultaneously increase one variable whilst decreasing the other, both by one. Some of these processes were shown to be useful to represent competition dynamics between the two species for a shared resource [26]. In 1964 Iglehart [27] further generalised this idea to a multi-variate competition process with two or more variables (i.e., species).

Here we focus on a mathematical model for monomer formation, in which two different receptor species are competing for a shared ligand. We show how ligand competition leads to a bi-variate process of the kind considered by Reuter [26]. Our aim is to compute the steady-state distribution of the number of monomers of each type, and to quantify the timescales of formation of a particular kind of monomers. We also explore the timescales for monomer formation where bound monomers are only considered to be productive if they remain bound for longer than a given dwell time, τ . Attaining a threshold number of productive monomers has been proposed as a criterion of signal initiation for different receptor tyrosine kinases [18–20]. Since the two-dimensional process considered in this work leads to a quasi-birth-and-death process (QBD), a type of continuous-time Markov processes, it is possible to make use of matrix-oriented algorithmic methods [28]. These methods, used in queuing theory [28], have been used to analyse receptor-ligand models [29,30] and other systems in Mathematical Epidemiology [31,32].

The limitation of matrix-oriented approaches is their practical computational implementation. The space of states of a bi-variate stochastic process increases with the number of molecules in the system; large numbers are typical in experimental and physiological settings [10,11]. When more than two receptor species are present, or where intra-cellular dynamics is considered in the mathematical model, the dimensionality and number of states are further increased. Thus, in this paper, we propose

analytical approximations for the summary statistics of interest. Our approximations, for the two-dimensional system of interest (i.e., two kinds of receptor competing for a common ligand), are motivated by the fact that, under excess of ligand, the processes representing ligand-bound monomer formation for the two different receptors are approximately independent. The structure of this paper is as follows. In Section 2 the bi-variate stochastic model for two kinds of receptor, R_1 and R_2 , with a shared ligand, L , is introduced. In Section 3 we describe how to analyse two summary statistics: (i) the steady-state distribution of the number of monomers and (ii) the time to reach N monomers for receptor type 2. We carry out this analysis by means of both (exact) matrix-oriented approaches and novel approximations. Having proposed these approximations, we then test their accuracy for a wide range of biologically relevant parameter values and numbers of molecules in Section 4. We also consider a different summary statistic: the time to reach a threshold number N of *productive* monomers (i.e., monomers that remain bound for longer than a given dwell time τ). Finally, Section 5 provides a discussion of our results. Details of the matrix-oriented algorithms used in this paper are given in Appendices A and B. Appendix C has some additional numerical results, while Appendix D provides a comparison of the efficacy of the different computational methods.

2. A Stochastic Competition Model for Receptor-Ligand Interactions

We consider two receptor types, R_1 and R_2 , on the cell membrane, and a shared ligand, L , in the extra-cellular medium. Both receptors can associate with the ligand to form a ligand-bound monomer, which we denote by M_1 or M_2 , depending on the receptor type involved. We assume a spatially homogeneous distribution of molecules on the cell membrane and mass-action kinetics [33]. We also assume that these reactions occur with forward association rates $k_{f,1}$ and $k_{f,2}$, respectively. We allow for the monomers, M_1 and M_2 , to dissociate into their constituent receptor and ligand, with rates $k_{r,1}$ and $k_{r,2}$, respectively (see Figure 1a).

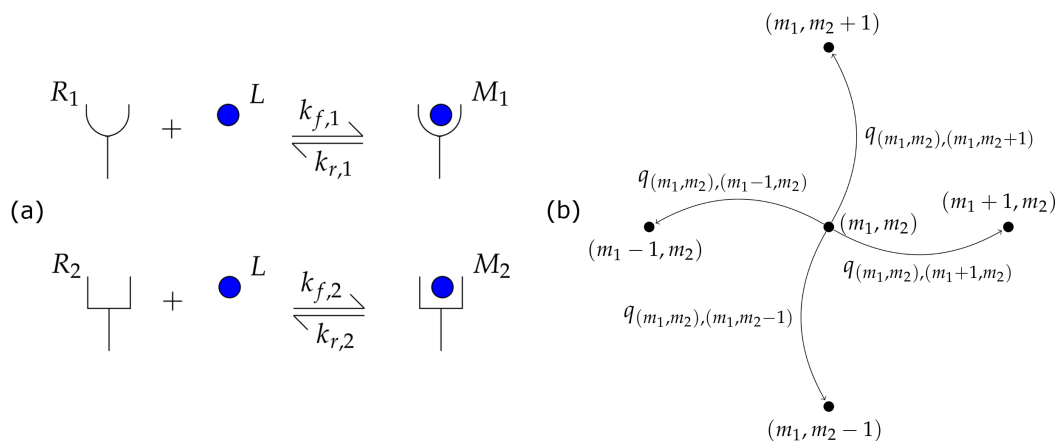


Figure 1. (a) Model reactions with two receptor types and one shared ligand. (b) Transition diagram for adjacent states in the continuous-time Markov chain (CTMC) \mathcal{X} .

If we denote by $n_{R,1}$ and $n_{R,2}$ the total number of receptors available in the system, and by n_L the total number of ligands, we can write $n_{R,1} = R_1(t) + M_1(t)$ and $n_{R,2} = R_2(t) + M_2(t)$, for all $t \geq 0$, where $R_1(t)$ and $R_2(t)$ represent the number of free receptors of type 1 and 2, respectively, at time t , and $M_1(t)$ and $M_2(t)$ represent the number of monomers of type 1 and 2, respectively, at time t . Then, the process can be described as a bi-variate continuous-time Markov chain (CTMC) $\mathcal{X} = \{\mathbb{X}(t) = (M_1(t), M_2(t)) : t \geq 0\}$, with $M_1(t), M_2(t) \geq 0$ for all $t \geq 0$, and the process \mathcal{X} evolves over the state space $\mathcal{S}_{\mathcal{X}} = \{(m_1, m_2) \in (\mathbb{N} \cup \{0\})^2 : m_1 \leq n_{R,1}, m_2 \leq n_{R,2}, m_1 + m_2 \leq n_L\}$, where m_1 and m_2 are the number of type 1 and type 2 monomers, respectively, at any given time.

The dynamics of monomer formation and dissociation are then represented by *jumps*, or transitions, between states in $\mathcal{S}_{\mathcal{X}}$, $(m_1, m_2) \rightarrow (m'_1, m'_2)$. The transition diagrams are shown in

Figures 1b and 2. We note that the boundaries of $\mathcal{S}_{\mathcal{X}}$, which are not explicitly depicted in Figure 2, are determined by the conditions $m_1 \leq n_{R,1}$, $m_2 \leq n_{R,2}$ and $m_1 + m_2 \leq n_L$. The infinitesimal transition rate from state (m_1, m_2) to state (m'_1, m'_2) , by assuming mass-action kinetics, is given by

$$q_{(m_1, m_2), (m'_1, m'_2)} = \begin{cases} k_{f,1}(n_{R,1} - m_1)(n_L - m_1 - m_2), & \text{if } (m'_1, m'_2) = (m_1 + 1, m_2), \\ k_{r,1}m_1, & \text{if } (m'_1, m'_2) = (m_1 - 1, m_2), \\ k_{f,2}(n_{R,2} - m_2)(n_L - m_1 - m_2), & \text{if } (m'_1, m'_2) = (m_1, m_2 + 1), \\ k_{r,2}m_2, & \text{if } (m'_1, m'_2) = (m_1, m_2 - 1), \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

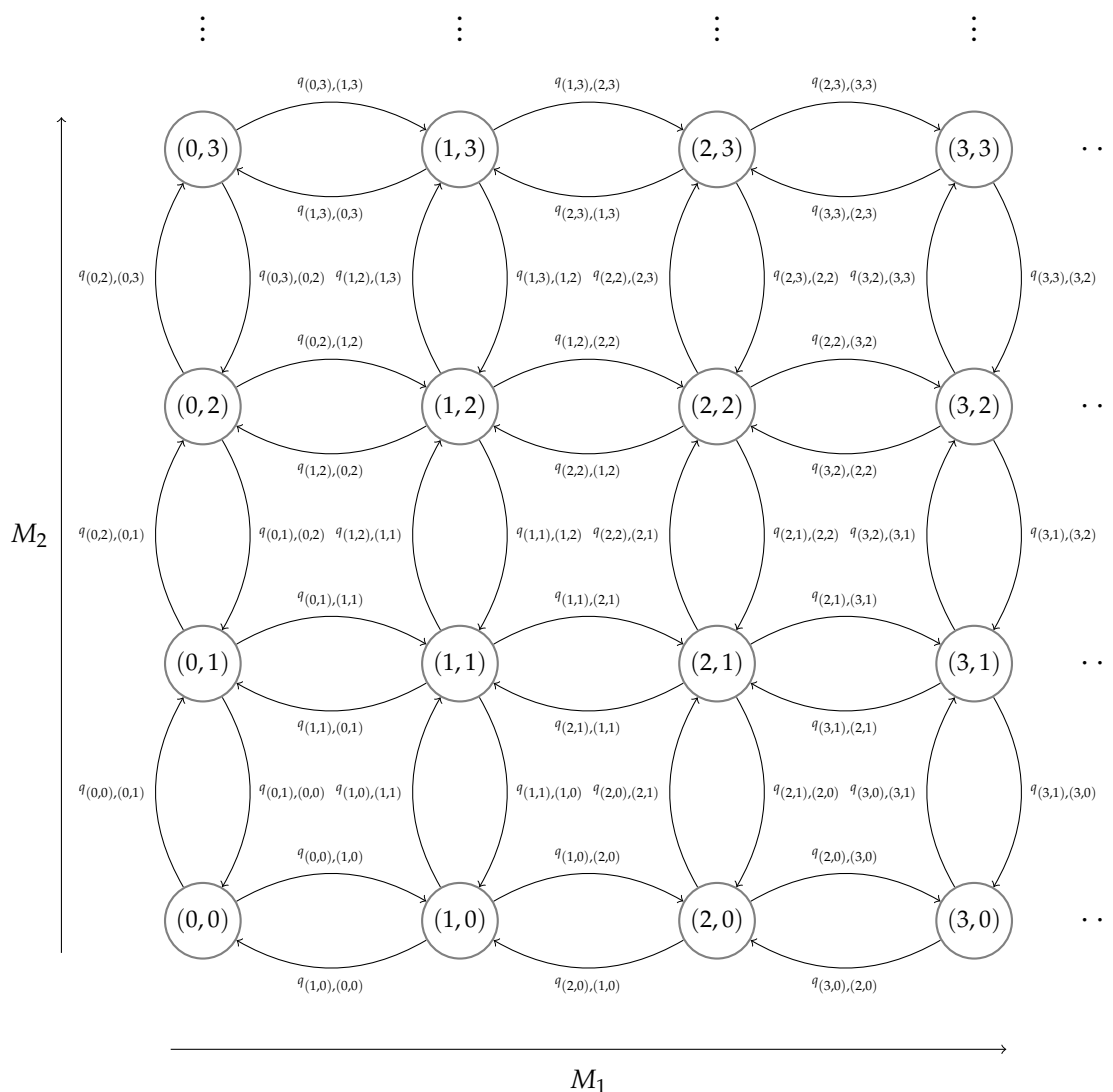


Figure 2. A diagram of the two-dimensional Markov process \mathcal{X} . Each state in the process represents a number of monomers of each type (m_1, m_2) , and the transitions between states represent formation and dissociation of type 1 and type 2 monomers over time. The maximal values of m_1 and m_2 will depend on the parameters $n_{R,1}$, $n_{R,2}$ and n_L , and will determine the overall shape of $\mathcal{S}_{\mathcal{X}}$.

The process shown in Figure 2 is an example of a two-dimensional competition process, as introduced by Reuter [26] in 1961. In this type of process, jumps can be made from a given state to their nearest neighbours, as seen by the connecting arrows between states in Figure 2. Here we

have two axes, each of which corresponds to a different receptor type, that is, the x -axis represents the number of type 1 monomers formed and the y -axis represents the number of type 2 monomers formed. In principle, as introduced by Iglehart [27] in his generalisation of Reuter’s work [26], one could add additional axes to represent further receptor variants competing for the same ligand; thus, expanding our model in Figure 1 to a multi-variate competition process (see the discussion in Section 5).

To illustrate the competition dynamics generated by the reactions in Figure 1a, we simulate in Figure 3 the stochastic model by means of the Gillespie algorithm [34]. We simulate the process for different numbers of molecules and K_d values of the reactions, where $K_{d,i} = \frac{k_{r,i}}{k_{f,i}}$ for $i \in \{1, 2\}$. We vary these parameter values, as well as the number of available receptors of type 2, to see the effect on the competition dynamics. In all plots in Figure 3, $n_L = 10^2$, $k_{r,1} = k_{r,2} = 10^{-3} \text{ s}^{-1}$ and $n_{R,1} = 10^2$.

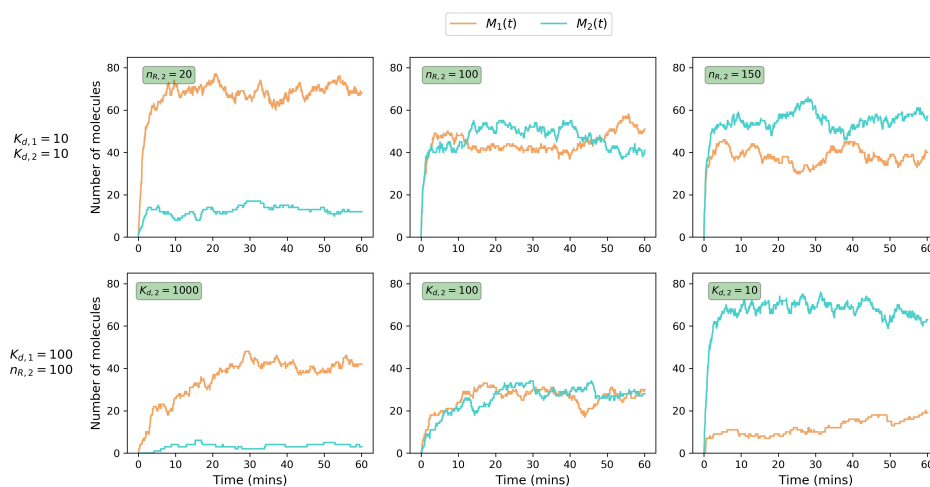


Figure 3. Gillespie simulations of the competition process showing the effect of varying rate constants and number of molecules on the time evolution of the monomers, $M_1(t)$ and $M_2(t)$. In the first row, for equal K_d values we vary the number, $n_{R,2}$, of type 2 receptors. In the second row, for equal number of receptors we vary $K_{d,2}$.

In the first plots of each row in Figure 3, we see that the process favours the formation of M_1 over M_2 . In this case, receptor type 1 out-competes receptor type 2 for ligand binding. In the second plot of each row, the number of molecules and rate constants governing the formation of M_1 and M_2 are identical and hence we see that the time courses for the two monomers are at a similar level, and are sometimes overlapping. Finally in the third plots of each row, we see that the receptors of type 2 are out-competing the receptors of type 1, and hence the time courses for M_2 are above those of M_1 . We also notice from Figure 3 that in some of the plots the process appears to reach a steady state by the end of the time course. One of the aims of this paper therefore, is to examine the effect of the rate constants and the number of molecules on the steady-state distribution of the process.

In Section 3 we compute two summary statistics of interest, namely the steady-state distribution of monomers on the cell surface [29,35] and the mean time to reach N monomers of one receptor type [29,36]. For each summary statistic, we present a computationally efficient matrix-analytic approach for its exact analysis, which can be implemented for low receptor and ligand numbers. Throughout this paper we will refer to this method of analysis as the exact matrix-analytic approach (EMA). Along with the exact analysis, and for scenarios where the large number of molecules leads to a huge number of states in \mathcal{S}_X , in turn making the EMA computationally intractable, we propose novel approximations to compute the proposed summary statistics. Finally, in Section 4 we compare the exact methods with the approximations, in order to quantify the goodness of the new approximations under different parameter regimes (i.e., in terms of the parameter values $(k_{f,1}, k_{r,1}, k_{f,2}, k_{r,2})$ of the model and the number of molecules).

3. Summary Statistics

3.1. Steady-State Distribution

In this section we analyse the steady-state distribution for the mathematical model defined in Section 2. We show how this distribution can be analytically computed by solving matrix equations in a computationally efficient way. Since this methodology is computationally limited by the number of molecules in the system, we propose as well approximations, which are based on the analysis of the process \mathcal{X} when $n_L \rightarrow +\infty$ or, equivalently, $n_L \gg n_{R,1}, n_{R,2}$. That is, instances when the competition for ligand is low. This analysis, when adjusting the number of available ligands to an *effective* number of ligands, leads to an approximation that can be used in low-to-moderate competition scenarios.

3.1.1. Exact Matrix-Analytic Approach (EMA)

The steady-state distribution for the model in Figure 1 is described in terms of the following probabilities

$$\pi_{(m_1, m_2)}^{EMA} = \lim_{t \rightarrow +\infty} \mathbb{P}\{(M_1(t), M_2(t)) = (m_1, m_2)\}, \quad (m_1, m_2) \in \mathcal{S}_{\mathcal{X}},$$

which can be stored in a row vector $\pi = (\pi_{(m_1, m_2)}, (m_1, m_2) \in \mathcal{S}_{\mathcal{X}})$ for any given order of states in $\mathcal{S}_{\mathcal{X}}$. These probabilities correspond to the number of monomers, of type 1 and type 2, respectively, found on the cellular surface at late times. They are known to satisfy [28] the system of equations

$$\begin{aligned} \pi \cdot Q &= \mathbf{0}_{\#\mathcal{S}_{\mathcal{X}}}, \\ \pi \cdot \mathbf{1}_{\#\mathcal{S}_{\mathcal{X}}} &= 1, \end{aligned} \tag{2}$$

where Q is the infinitesimal generator of the CTMC \mathcal{X} , $\#\mathcal{S}_{\mathcal{X}}$ is the number of states in the state space, $\mathbf{0}_a$ is a row vector of zeros with length a and $\mathbf{1}_a$ is a column vector of ones with length a .

The infinitesimal generator Q encodes the infinitesimal transition rates. Since the transitions described in Figure 2 only occur between adjacent states (see also Figure 1b), it is possible to order the states in $\mathcal{S}_{\mathcal{X}}$ so that Q is tri-diagonal by blocks, as shown in Appendix A. This means that the process \mathcal{X} can be described as a level-dependent quasi-birth-and-death (LD-QBD) process [28]. For this type of process, efficient computational methods exist to solve Equation (2). In particular we use Algorithm A1, as presented in Appendix A, to compute the probabilities $\pi_{(m_1, m_2)}^{EMA}$ for all $(m_1, m_2) \in \mathcal{S}_{\mathcal{X}}$. Once these probabilities are in hand, one can compute the mean number of monomers at steady-state as

$$\begin{aligned} \mathbb{E}^{EMA}[M_1] &= \sum_{m_1=0}^{N_1} m_1 \left(\sum_{m_2=0}^{\min(N_2, n_L - m_1)} \pi_{(m_1, m_2)}^{EMA} \right), \\ \mathbb{E}^{EMA}[M_2] &= \sum_{m_2=0}^{N_2} m_2 \left(\sum_{m_1=0}^{\min(N_1, n_L - m_2)} \pi_{(m_1, m_2)}^{EMA} \right), \end{aligned} \tag{3}$$

with $N_1 = \min(n_{R,1}, n_L)$ and $N_2 = \min(n_{R,2}, n_L)$. However, as the number of molecules increases (see Appendix A), Algorithm A1 becomes computationally expensive, and an alternative approach needs to be considered.

3.1.2. No-Competition Approximation (NCA) under Excess of Ligand

We seek to find an approximation to the steady-state distribution which can be computationally tractable even for a large state space. To this aim, we note that as $n_L \rightarrow +\infty$ (or, equivalently, $n_L \gg n_{R,1}, n_{R,2}$), the two populations of monomers behave independently of one another given the excess of ligand. We therefore hypothesised that we could approximate the steady-state distribution in the large n_L limit by considering two independent Markov chains for the variables $M_1(t)$ and $M_2(t)$. In the limit $n_L \rightarrow +\infty$, given that the monomers are linked only by the number of ligands present, the steady-state probabilities for the two-dimensional process will be precisely equal to the

product of the steady-state probabilities for the one-dimensional processes modelling $M_1(t)$ and $M_2(t)$. This product, in the limit $n_L \rightarrow +\infty$, comprises what we denote by the no-competition approximation (NCA).

In particular, we define two independent Markov chains

$$\mathcal{X}_1 = \{M_1(t) : t \geq 0\}, \quad \mathcal{X}_2 = \{M_2(t) : t \geq 0\},$$

for the monomers M_1 and M_2 , respectively, with state spaces

$$\mathcal{S}_{\mathcal{X}_1} = \{m_1 \in \mathbb{N}_0 = \mathbb{N} \cup \{0\} : m_1 \leq N_1\}, \quad \mathcal{S}_{\mathcal{X}_2} = \{m_2 \in \mathbb{N}_0 : m_2 \leq N_2\}.$$

We note that, since the number of monomers in each CTMC can only increase or decrease by one unit in every transition (by means of monomer formation or dissociation), each CTMC \mathcal{X}_j , for $j \in \{1, 2\}$, is a birth-and-death process as depicted in Figure 4, with $\lambda_n = k_{f,j}(n_{R,j} - n)(n_L - n)$ and $\mu_n = k_{r,j}n$.

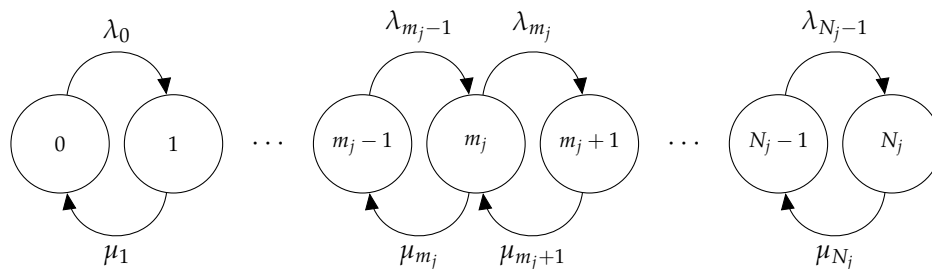


Figure 4. Diagram for the birth-and-death process $\mathcal{X}_j, j \in \{1, 2\}$.

In the limit $n_L \rightarrow +\infty$, one can consider a constant number of ligands (not depleted by monomer formation); the binding rates then become $\lambda_n = k_{f,j}n_L(n_{R,j} - n)$. For these rates, Equation (2) can be easily solved [25], as follows

$$\pi_{m_1} = \lim_{t \rightarrow +\infty} \mathbb{P}(M_1(t) = m_1) = \binom{n_{R,1}}{m_1} \left(\frac{k_{f,1}n_L}{k_{r,1} + k_{f,1}n_L} \right)^{m_1} \left(\frac{k_{r,1}}{k_{r,1} + k_{f,1}n_L} \right)^{n_{R,1}-m_1}, \quad 0 \leq m_1 \leq N_1, \quad (4)$$

$$\pi_{m_2} = \lim_{t \rightarrow +\infty} \mathbb{P}(M_2(t) = m_2) = \binom{n_{R,2}}{m_2} \left(\frac{k_{f,2}n_L}{k_{r,2} + k_{f,2}n_L} \right)^{m_2} \left(\frac{k_{r,2}}{k_{r,2} + k_{f,2}n_L} \right)^{n_{R,2}-m_2}, \quad 0 \leq m_2 \leq N_2, \quad (5)$$

which can be identified as binomial distributions. For these one-dimensional birth-and-death processes, \mathcal{X}_1 and \mathcal{X}_2 , one can determine the expected values of monomers in steady-state as

$$\mathbb{E}^{NCA}[M_1] = \sum_{m_1=0}^{N_1} m_1 \pi_{m_1} = \frac{k_{f,1}n_{R,1}n_L}{k_{r,1} + k_{f,1}n_L}, \quad \mathbb{E}^{NCA}[M_2] = \sum_{m_2=0}^{N_2} m_2 \pi_{m_2} = \frac{k_{f,2}n_{R,2}n_L}{k_{r,2} + k_{f,2}n_L}. \quad (6)$$

In the limit $n_L \rightarrow +\infty$, the steady-state distribution of the process \mathcal{X} is given by

$$\pi_{(m_1, m_2)}^{NCA} = \pi_{m_1} \times \pi_{m_2}, \quad (m_1, m_2) \in \mathcal{S}_{\mathcal{X}}. \quad (7)$$

We note that, when $n_L \rightarrow +\infty$, one gets $N_1 = n_{R,1}$, $N_2 = n_{R,2}$ and $\mathcal{S}_{\mathcal{X}} = \{(m_1, m_2) \in \mathbb{N}_0^2 : 0 \leq m_1 \leq N_1, 0 \leq m_2 \leq N_2\} = \mathcal{S}_{\mathcal{X}_1} \times \mathcal{S}_{\mathcal{X}_2}$, so that Equations (4)–(7) are consistent.

It is clear that Equation (7) will also provide a reasonable approximation to the steady-state distribution when the number of ligands is much greater than the number of receptors, that is, $n_L \gg n_{R,1}, n_{R,2}$, since we are then in a situation in which there is very little competition between the receptors for available ligand, and hence, ligand depletion can be neglected, leading to two independent birth-and-death processes for the two different monomer populations. On the other hand, when n_L is comparable in magnitude to $n_{R,1} + n_{R,2}$, Equation (7) overestimates the probabilities to

reach higher numbers of monomers of each type, since it does not take into account the fact that ligands are being depleted by the formation of monomers in the competing Markov chain. To effectively approximate the steady-state distribution in scenarios where $n_L \approx n_{R,1} + n_{R,2}$, we propose to employ an *effective* number of ligands available to each birth-and-death process, n_L^* , which is lower than the total number of ligands in the two-dimensional process, n_L . We propose that this effective number of ligands, and therefore an approximation to the steady-state distribution, can be reached iteratively using an algorithmic approach.

3.1.3. Moderate-Competition Approximation (MCA)

Here we propose Algorithm 1 to compute an approximate steady-state distribution, $\pi_{(m_1, m_2)}^{MCA}$, which can be used when $n_L \approx n_{R,1} + n_{R,2}$. In Algorithm 1, we iteratively reduce the total number of ligands available in each independent birth-and-death process \mathcal{X}_1 and \mathcal{X}_2 , by subtracting from n_L the expected number of each monomer type in steady-state computed from the previous iteration. In each iteration i , the mean number of monomers in steady-state under the NCA approximation, $\mathbb{E}^{NCA,(i)}[M_1]$ and $\mathbb{E}^{NCA,(i)}[M_2]$, are computed, and they are used to compute an effective number of free ligands available $N_L^{(i+1)}$, at iteration $i + 1$. The iterative scheme stops once these mean values are close enough for two consecutive iterations, as determined by a threshold value ϵ . We note that the final number of ligands considered, $n_L^* = N_L^{(i)}$ in the last iteration, can be seen as an *effective* number of ligands that reflects the competition for shared resources.

Algorithm 1 Iterative approximation for the steady-state distribution, $(\pi_{(m_1, m_2)}^{MCA} : (m_1, m_2) \in \mathcal{S}_{\mathcal{X}})$.

- 1: $i = 0, N_L^{(i)} = n_L$;
 - 2: Compute $\mathbb{E}^{NCA,(i)}[M_1]$ and $\mathbb{E}^{NCA,(i)}[M_2]$ from Equation (6), using $N_L^{(i)}$ as the number of ligands in the system, instead of n_L ;
 - 3: **while** $|\mathbb{E}^{NCA,(i)}[M_1] - \mathbb{E}^{NCA,(i-1)}[M_1]| > \epsilon$ or $|\mathbb{E}^{NCA,(i)}[M_2] - \mathbb{E}^{NCA,(i-1)}[M_2]| > \epsilon$ or $i < 1$ **do**:
 - 4: $i = i + 1$;
 - 5: $N_L^{(i)} = (n_L - \mathbb{E}^{NCA,(i-1)}[M_1] - \mathbb{E}^{NCA,(i-1)}[M_2])\alpha + N_L^{(i-1)}(1 - \alpha)$;
 - 6: Compute $\mathbb{E}^{NCA,(i)}[M_1]$ and $\mathbb{E}^{NCA,(i)}[M_2]$ from Equation (6), using $N_L^{(i)}$ instead of n_L ;
 - 7: **end while**.
 - 8: Compute $\pi_{m_1}^{NCA,(i)}$, for $0 \leq m_1 \leq N_1$, and $\pi_{m_2}^{NCA,(i)}$, for $0 \leq m_2 \leq N_2$, from Equations (4)–(5) using $N_L^{(i)}$ instead of n_L ;
 - 9: Compute $\pi_{(m_1, m_2)}^{MCA} = \pi_{m_1}^{NCA,(i)} \times \pi_{m_2}^{NCA,(i)}$, $(m_1, m_2) \in \mathcal{S}_{\mathcal{X}}$;
 - 10: **return** $\pi_{(m_1, m_2)}^{MCA} = \frac{\pi_{(m_1, m_2)}^{MCA}}{\sum_{(m'_1, m'_2) \in \mathcal{S}_{\mathcal{X}}} \pi_{(m'_1, m'_2)}^{MCA}}$, $(m_1, m_2) \in \mathcal{S}_{\mathcal{X}}$; ▷ MCA steady-state distribution
 - 11: **return** $\mathbb{E}^{MCA}[M_1] = \sum_{m_1=0}^{N_1} m_1 \left(\sum_{m_2=0}^{\min(N_2, n_L - m_1)} \pi_{(m_1, m_2)}^{MCA} \right)$; ▷ Mean number of type 1 monomers
 - 12: **return** $\mathbb{E}^{MCA}[M_2] = \sum_{m_2=0}^{N_2} m_2 \left(\sum_{m_1=0}^{\min(N_1, n_L - m_2)} \pi_{(m_1, m_2)}^{MCA} \right)$; ▷ Mean number of type 2 monomers
-

The parameter α ($\alpha \in [0, 1]$) is a tuning parameter needed in situations where $n_L < n_{R,1} + n_{R,2}$. It modulates the convergence speed of the algorithm. We set $\alpha = 1$ when $n_L \geq n_{R,1} + n_{R,2}$, since in this case it is impossible for $n_L - \mathbb{E}^{NCA,(i-1)}[M_1] - \mathbb{E}^{NCA,(i-1)}[M_2]$ to be negative at any iteration. Otherwise, when $n_L < n_{R,1} + n_{R,2}$, the parameter α is chosen between 0 and 1 such that the effective

number of ligands at every iteration in the algorithm is positive. Our numerical exploration indicates that the choice of α does not affect the resulting steady-state probabilities $\pi_{(m_1, m_2)}^{MCA}$. When $\alpha = 1$, the expected number of monomers and ligands in steady-state exhibit damped oscillations as they converge. An example of this can be found on the left plot of Figure 5. When $\alpha < 1$, for values of α for which the algorithm converges, the expected number of monomers and ligands in steady-state either exhibit damped oscillations as they converge (for larger values of α), or they decrease monotonically in convergence (for smaller values of α). An example of monotonic convergence can be seen in the right plot of Figure 5. We found that for values of α for which there is monotonic convergence, the smaller the value of α , the more iterations the algorithm requires to converge. We note that in Figure 5, the mean number of free ligands in steady-state, as predicted by the EMA, is given by $\mathbb{E}^{EMA}[L] = n_L - \mathbb{E}^{EMA}[M_1] - \mathbb{E}^{EMA}[M_2]$. Interestingly, we note how the values of $N_L^{(i)}$ tend, in the limit $i \rightarrow +\infty$, to this mean number of free ligands in steady-state.

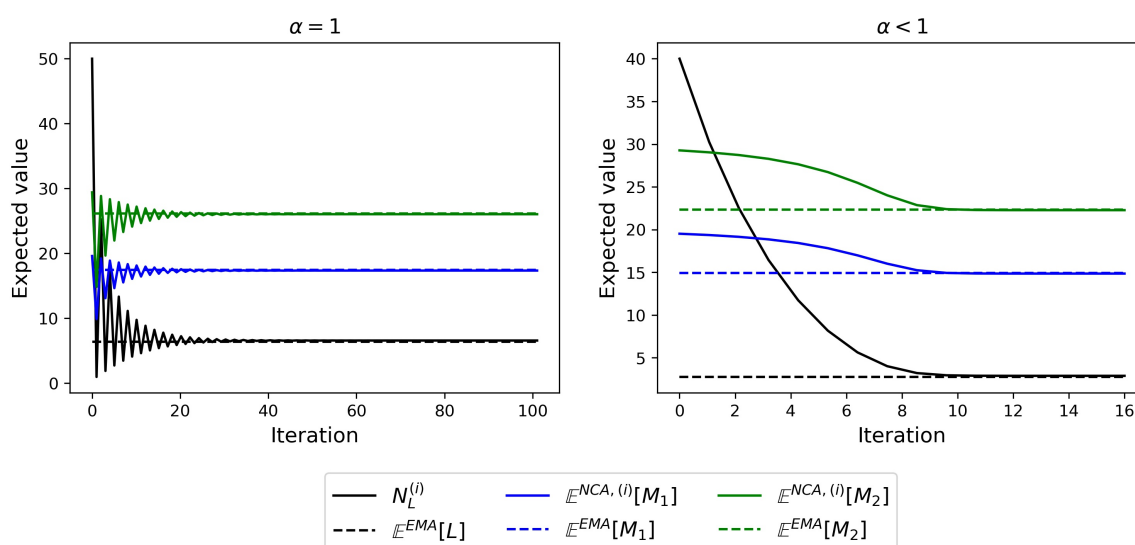


Figure 5. Examples of how Algorithm 1 converges when $n_L \geq n_{R,1} + n_{R,2}$ (left plot, $n_L = 50$) and when $n_L < n_{R,1} + n_{R,2}$ (right plot, $n_L = 40$). Iterative mean values $\mathbb{E}^{NCA, (i)}[M_1]$, $\mathbb{E}^{NCA, (i)}[M_2]$ and $N_L^{(i)}$ converge to the exact values $\mathbb{E}^{EMA}[M_1]$, $\mathbb{E}^{EMA}[M_2]$ and $\mathbb{E}^{EMA}[L]$. In both plots, $\epsilon = 10^{-5}$, $n_{R,1} = 20$, $n_{R,2} = 30$, $K_{d,1} = 1$, $K_{d,2} = 1$, and $\alpha = 0.2$ for the right plot.

In Section 4.1, we conduct a comprehensive numerical exploration of the accuracy of the MCA calculated with Algorithm 1 when compared to the exact steady-state distribution (EMA).

3.2. Timescales of Monomer Formation

In this section we study a second summary statistic of interest; namely the timescales of monomer formation. For the bi-variate process \mathcal{X} , we define

$$T_{(m_1, m_2)}(N) = \inf\{t \geq 0 : M_2(t) = N \mid (M_1(0), M_2(0)) = (m_1, m_2)\}, \quad (m_1, m_2) \in \mathcal{S}_{\mathcal{X}},$$

to be the time to reach N monomers of type 2 (without loss of generality, since our arguments would similarly apply to type 1 monomers) given the initial state $(m_1, m_2) \in \mathcal{S}_{\mathcal{X}}$. N can be any arbitrary value $M_2(0) < N \leq N_2$, since $T_{(M_1(0), M_2(0))}(M_2(0)) \equiv 0$.

Following the same strategy as that of Section 3.1, we first analyse $T_{(m_1, m_2)}(N)$ by means of an exact matrix-analytic approach. This implies solving a matrix equation for the two-dimensional process \mathcal{X} described in Figure 2. This EMA is once again computationally limited by the number of molecules in the system. Thus, we analyse $T_{(m_1, m_2)}(N)$ in the limit as $n_L \rightarrow +\infty$ (i.e., under effectively no competition for the shared ligand). We call this approximation the NCA. Based on this approximation,

we propose the MCA, where an effective number of ligands is considered when computing this summary statistic.

3.2.1. Exact Matrix-Analytic Approach (EMA)

In this case, we can efficiently analyse $T_{(m_1, m_2)}(N)$, for $(m_1, m_2) \in \mathcal{S}_{\mathcal{X}}$ with $m_2 \leq N$, by means of first-step arguments. We omit the index N from now on to simplify our notation. One can, according to arguments described in Appendix B, compute the l -th order moment of $T_{(m_1, m_2)}$, $\mathbb{E} \left[T_{(m_1, m_2)}^l \right]$, by solving the system of linear equations:

$$\Delta_{(m_1, m_2)} \mathbb{E} \left[T_{(m_1, m_2)}^l \right] = k_{f,1}(n_{R,1} - m_1)(n_L - m_1 - m_2) \mathbb{E} \left[T_{(m_1+1, m_2)}^l \right] + k_{r,1} m_1 \mathbb{E} \left[T_{(m_1-1, m_2)}^l \right] + k_{f,2}(n_{R,2} - m_2)(n_L - m_1 - m_2) \mathbb{E} \left[T_{(m_1, m_2+1)}^l \right] + k_{r,2} m_2 \mathbb{E} \left[T_{(m_1, m_2-1)}^l \right] + l \mathbb{E} \left[T_{(m_1, m_2)}^{l-1} \right], \tag{8}$$

with $\Delta_{(m_1, m_2)} = k_{f,1}(n_{R,1} - m_1)(n_L - m_1 - m_2) + k_{r,1} m_1 + k_{f,2}(n_{R,2} - m_2)(n_L - m_1 - m_2) + k_{r,2} m_2$. This system of equations can be efficiently solved by arranging the states of the bi-variate Markov process \mathcal{X} in levels $L(k)$, $0 \leq k \leq N_2$, as done for the steady-state distribution and detailed in Appendices A and B. We note that the system of equations given by Equation (8) can be solved in an iterative manner, for increasing values of l , where moments of order l can be computed once moments of order $l - 1$ have been calculated (see Algorithm A2 in Appendix B). For $l = 1$, we can then compute the mean time to reach N type 2 monomers, $\mathbb{E}^{EMA} [T_{(m_1, m_2)}(N)]$, making use of the exact matrix-analytic methodology.

3.2.2. No-Competition Approximation (NCA) under Excess of Ligand

As an approximation under no competition ($n_L \rightarrow +\infty$), we can compute the mean time to reach N type 2 monomers by considering the independent one-dimensional birth-and-death process \mathcal{X}_2 . In particular, we can analyse $T_{m_2}(N) = \inf \{ t \geq 0 : M_2(t) = N \mid M_2(0) = m_2 \}$ for the process \mathcal{X}_2 .

Let us focus on the mean value $\mathbb{E}[T_{m_2}(N)]$ for $m_2 = 0$ (that is, no type 2 monomers initially in the system), although we note that our arguments can be easily modified for different initial states $0 < m_2 < N$, or for higher order moments of $T_{m_2}(N)$. It is clear that, for $N = 1$, $\mathbb{E}[T_0(1)] = \frac{1}{\lambda_0}$. It is also clear from a first-step argument [25] that

$$\mathbb{E}[T_i(i + 1)] = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i} \mathbb{E}[T_{i-1}(i)].$$

This allows one to recursively obtain

$$\mathbb{E}[T_0(N)] = \frac{1}{\lambda_0} + \sum_{j=1}^{N-1} \left[\frac{\mu_1 \dots \mu_j}{\lambda_0 \dots \lambda_j} \left(\sum_{i=1}^j \frac{\lambda_0 \dots \lambda_{i-1}}{\mu_1 \dots \mu_i} + 1 \right) \right],$$

which is similar to the well-known extinction time expressions for a one-dimensional birth-and-death process [25]. Thus, we propose the following approximation

$$\mathbb{E}^{NCA} [T_{(0,0)}(N)] = \frac{1}{\lambda_0} + \sum_{j=1}^{N-1} \left[\frac{\mu_1 \dots \mu_j}{\lambda_0 \dots \lambda_j} \left(\sum_{i=1}^j \frac{\lambda_0 \dots \lambda_{i-1}}{\mu_1 \dots \mu_i} + 1 \right) \right], \tag{9}$$

when $n_L \gg n_{R,1} + n_{R,2}$.

In the case when n_L is comparable to $n_{R,1} + n_{R,2}$, it is clear that $\mathbb{E}^{NCA} [T_{(0,0)}(N)]$ will underestimate the exact mean time to reach N type 2 monomers, $\mathbb{E}^{EMA} [T_{(0,0)}(N)]$, since it does not take into account the ligands being depleted by the formation of monomers in the competing Markov chain. Thus,

we propose next, an approximate method (MCA) to calculate $\mathbb{E}[T_{(0,0)}(N)]$, which can be used under low-to-moderate competition scenarios (i.e., when $n_L \approx n_{R,1} + n_{R,2}$).

3.2.3. Moderate-Competition Approximation (MCA)

Since the NCA underestimates the time to reach N type 2 monomers, we now define the MCA, in order to approximately account for the depletion of ligand. In particular, we propose to compute $\mathbb{E}^{MCA}[T_{(0,0)}(N)]$ by implementing Equation (9), but with an effective number of free ligands

$$n_L^* = n_L - \mathbb{E}[M_1], \tag{10}$$

where $\mathbb{E}[M_1]$ is the mean number of type 1 monomers in steady-state. We note here that, ideally, one would use $\mathbb{E}^{EMA}[M_1]$ in Equation (10). Alternatively, if this is computationally not feasible, one can use its approximation $\mathbb{E}^{MCA}[M_1]$ instead. From a practical perspective this does not seem crucial, since as we show in Section 4, $\mathbb{E}^{EMA}[M_1]$ and $\mathbb{E}^{MCA}[M_1]$ are almost equal (to integer value) in most of the scenarios considered.

We expect that the EMA will be approximated best by the NCA or the MCA depending on the number of molecules $n_{R,1}$, $n_{R,2}$ and n_L and the parameter values $k_{f,1}$, $k_{f,2}$, $k_{r,1}$ and $k_{r,2}$. We explore the relationship between the EMA, NCA and MCA in terms of the parameters values and numbers of molecules in Section 4.2.

4. Results

In this section we analyse the behaviour of the different approximations described in Section 3 and compare them with the exact matrix-analytic approach. We do so numerically and for a wide range of parameter values and numbers of molecules, so as to identify regions of the parameter space for which the approximations are most accurate. In order to reduce the complexity of the numerical exploration we consider only the K_d value for each reaction in Figure 1, where $K_{d,i} = k_{r,i}/k_{f,i}$ for $i \in \{1, 2\}$, and not the individual forward and backward rate constants, by fixing $k_{r,1} = k_{r,2} = 10^{-3} \text{ s}^{-1}$ [6,37]. The value chosen for these rates is the common rate at which VEGFR1 and VEGFR2 dissociate their shared ligand, VEGF-A. The K_d values and numbers of each receptor type are varied within the ranges given in Table 1, inspired from Reference [24] (Table 3), and which correspond to VEGFR1, VEGFR2 and VEGF-A. Finally, the number of ligands, n_L , is varied to consider different competition regimes between the receptors, so that the comparisons between the EMA, NCA and MCA can be studied.

Table 1. Ranges for the parameter values and numbers of molecules in the model.

Parameter	Range	Unit
$K_{d,1}, K_{d,2}$	$[1 \times 10^1 - 1 \times 10^4]$	molecules
$n_{R,1}, n_{R,2}$	$[2 \times 10^1 - 2 \times 10^2]$	molecules

4.1. Steady-State Distribution

In this section we wish to compare the steady-state distributions computed via the EMA and the MCA, to determine how well the approximation works under different combinations of parameter values and number of molecules. We first present (see the first three columns of Figure 6) a comparison of the steady-state distributions computed via the EMA, NCA and MCA, for different numbers of ligands, n_L , and where the number of receptors and other parameter values remain constant (see Table 1). In these heatmaps, the colour of each state indicates its steady-state probability. In the fourth and fifth columns, the colour represents the absolute difference $|\pi_{(m_1, m_2)}^{EMA} - \pi_{(m_1, m_2)}^{NCA}|$ or $|\pi_{(m_1, m_2)}^{EMA} - \pi_{(m_1, m_2)}^{MCA}|$, respectively, for each state $(m_1, m_2) \in \mathcal{S}_X$. As the number of ligands, n_L , decreases and becomes comparable to $n_{R,1} + n_{R,2}$, the NCA approximation worsens since the number of monomers of each type is overestimated. In contrast, we see that the MCA approximation remains

similar to the EMA distribution, even for the lowest value of n_L . It also better captures the mean number of monomers in steady-state than the NCA.

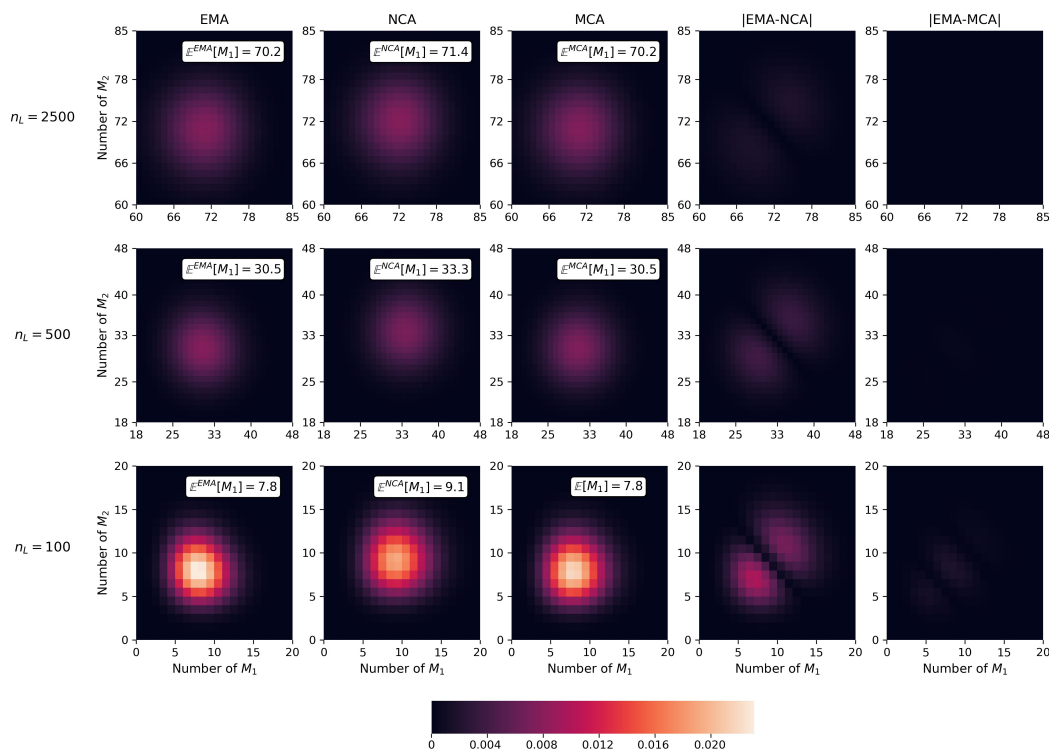


Figure 6. Comparison of the steady-state distributions for the EMA, NCA and MCA. The number of receptors are $n_{R,1} = n_{R,2} = 10^2$, and $K_{d,1} = K_{d,2} = 10^3$. The fourth and fifth columns of plots represent the absolute differences between probabilities, $|\pi_{(m_1,m_2)}^{EMA} - \pi_{(m_1,m_2)}^{NCA}|$ and $|\pi_{(m_1,m_2)}^{EMA} - \pi_{(m_1,m_2)}^{MCA}|$, respectively.

Having seen in Figure 6 that the MCA is able to accurately capture the steady-state distribution computed with the EMA for the specific number of molecules and parameter values chosen, we now seek to quantify the accuracy of the MCA for a wide range of parameter values and numbers of molecules. The *Hellinger distance*

$$H(\{\pi_{(m_1,m_2)}^{EMA}\}_{(m_1,m_2) \in \mathcal{S}_X}, \{\pi_{(m_1,m_2)}^{MCA}\}_{(m_1,m_2) \in \mathcal{S}_X}) \equiv \frac{1}{\sqrt{2}} \sqrt{\sum_{(m_1,m_2) \in \mathcal{S}_X} \left(\sqrt{\pi_{(m_1,m_2)}^{EMA}} - \sqrt{\pi_{(m_1,m_2)}^{MCA}} \right)^2}, \quad (11)$$

is a way to measure the similarity between two discrete probability distributions. The Hellinger distance (HD) can take values between 0 and 1, where lower values indicate that the probability distributions are more similar.

Analysis of the Hellinger distances between the steady-state distributions can be found in Appendix C. Here, we focus on *competing strength* parameters. In particular, in Figure 7 we display scatter plots of the Hellinger distance as a function of the parameters $n_{R,j}/K_{d,j}$, $j \in \{1, 2\}$. We note that the summary parameter $n_{R,j}/K_{d,j}$ is an indicator of the competing strength of receptor type j , which will be large whenever there are many of these receptors in the system and/or they have high affinity for the common ligand. In Figure 7 we sample 10^3 parameter values $(K_{d,1}, K_{d,2}, n_{R,1}, n_{R,2})$ within the ranges of Table 1, from uniform distributions $n_{R,j} \sim U(20, 200)$ and $K_{d,j} = 10^x$ with $x \sim U(1, 4)$, for $j \in \{1, 2\}$. We observe that the HD decreases with increasing n_L , so that in the third plot where $n_L = 500$, the HD is relatively small, even for the largest sampled values of $n_{R,1}/K_{d,1}$ and $n_{R,2}/K_{d,2}$. From the first plot in which $n_L = 10^2$, representing a high competition scenario, we observe larger values of the HD, especially when $n_{R,1}/K_{d,1}$ and $n_{R,2}/K_{d,2}$ are large. On the other hand, in the

$n_L = 250$ plot, smaller values of the competition strength parameters still lead to relatively small distances, indicating that even when n_L is small comparable to $n_{R,1} + n_{R,2}$ the approximation can still be reasonable if $n_{R,1}/K_{d,1}$ and $n_{R,2}/K_{d,2}$ are relatively small.

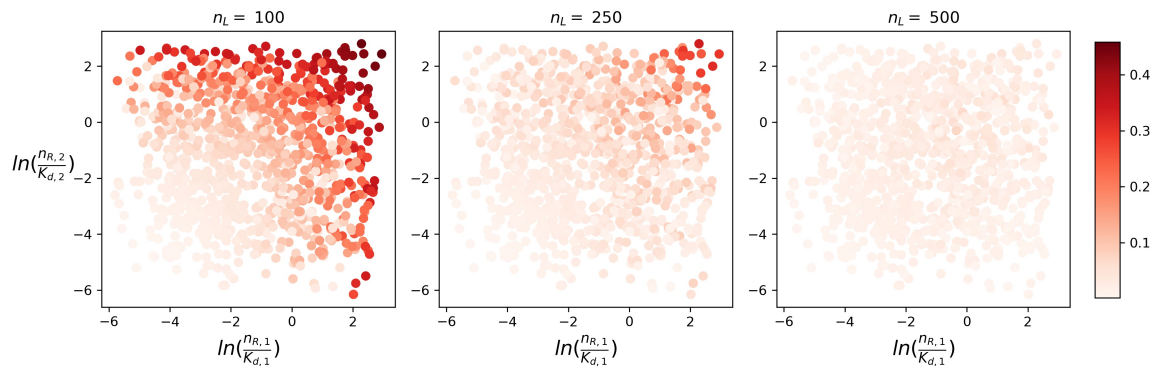


Figure 7. Hellinger distance between the steady-state distributions $\{\pi_{(m_1, m_2)}^{EMA}\}_{(m_1, m_2) \in \mathcal{S}_X}$ and $\{\pi_{(m_1, m_2)}^{MCA}\}_{(m_1, m_2) \in \mathcal{S}_X}$ plotted for sampled values $n_{R,j} \sim U(20, 200)$ and $K_{d,j} = 10^x$ with $x \sim U(1, 4)$, for $j \in \{1, 2\}$, and for different numbers of ligand $n_L \in \{100, 250, 500\}$. We set the threshold parameter $\epsilon = 10^{-5}$ for the Moderate-Competition Approximation (MCA).

Although it is clear that there is some disparity between the probability distributions, especially when the competition is the highest, we note that the expected number of monomers in steady-state (not reported here) are almost identical between the two methods of computation. In particular, for all the sets $(K_{d,1}, K_{d,2}, n_{R,1}, n_{R,2}, n_L)$ of parameter values considered in Figure 7, the percentage error $100 \cdot |1 - \mathbb{E}^{EMA}[M_j]/\mathbb{E}^{MCA}[M_j]|$, $j \in \{1, 2\}$, when computing the mean values through the MCA, instead of the EMA, is less than 5% in all 10^3 cases. In fact, the overwhelming majority of parameter choices (991 out of 10^3) lead to the same mean values to integer precision, when choosing $\epsilon = 10^{-5}$ in Algorithm 1, even if the resulting Hellinger distance is relatively large.

4.2. Timescales of Monomer Formation

In this section we carry out a similar numerical exploration to that of Section 4.1, of the accuracy of the NCA and MCA for the expected time to reach N monomers of type 2. We note here that, while the NCA for the steady-state distribution can be seen as a necessary step to then propose the MCA in Algorithm 1, but where the MCA is always expected to behave better than the NCA, the NCA and the MCA can both lead to reasonable and complementary approximations to the EMA when studying the time to reach N type 2 monomers. Thus, our objective in this section is to compare the expected value of $T_{(0,0)}(N)$ derived via the EMA and the NCA/MCA. We plot in Figure 8 the mean time to reach N monomers, where N is represented as a percentage P of the steady-state number of monomers, $N = \frac{P}{100} \mathbb{E}^{EMA}[M_2]$, by means of the EMA, NCA and MCA. In low competition scenarios (e.g., large n_L or when $K_{d,1} \gg K_{d,2}$ so that the dynamics of formation of type 1 monomers does not significantly affect the timescales of formation of type 2 monomers), the three approaches lead to almost indistinguishable results. For low-to-moderate competition settings, the MCA leads to a reasonable approximation. It is worth noting that the MCA and NCA seem to lead to lower and upper bounds for the mean time under analysis: the NCA always underestimates the timescales of type 2 monomer formation, by considering that there are n_L ligands available, neglecting ligand depletion due to type 1 monomer formation competition; while the MCA tends to overestimate the timescales of type 2 monomer formation because one considers only $n_L - \mathbb{E}[M_1]$ ligands available (that is, we remove from the beginning all the ligands occupied by steady-state type 1 monomers, even if these monomers may take some time to form). Since this steady-state amount of type 1 monomers takes some time to form, some of these ligands might still be able to contribute to the formation of type 2 monomers during early times, leading to the observed overestimation. Still, it is striking how well both approximations work when $K_{d,1} \gg K_{d,2}$,

even when the number of ligands is even less than the total number of receptors available in the system ($n_L = 100 < 200 = n_{R,1} + n_{R,2}$).

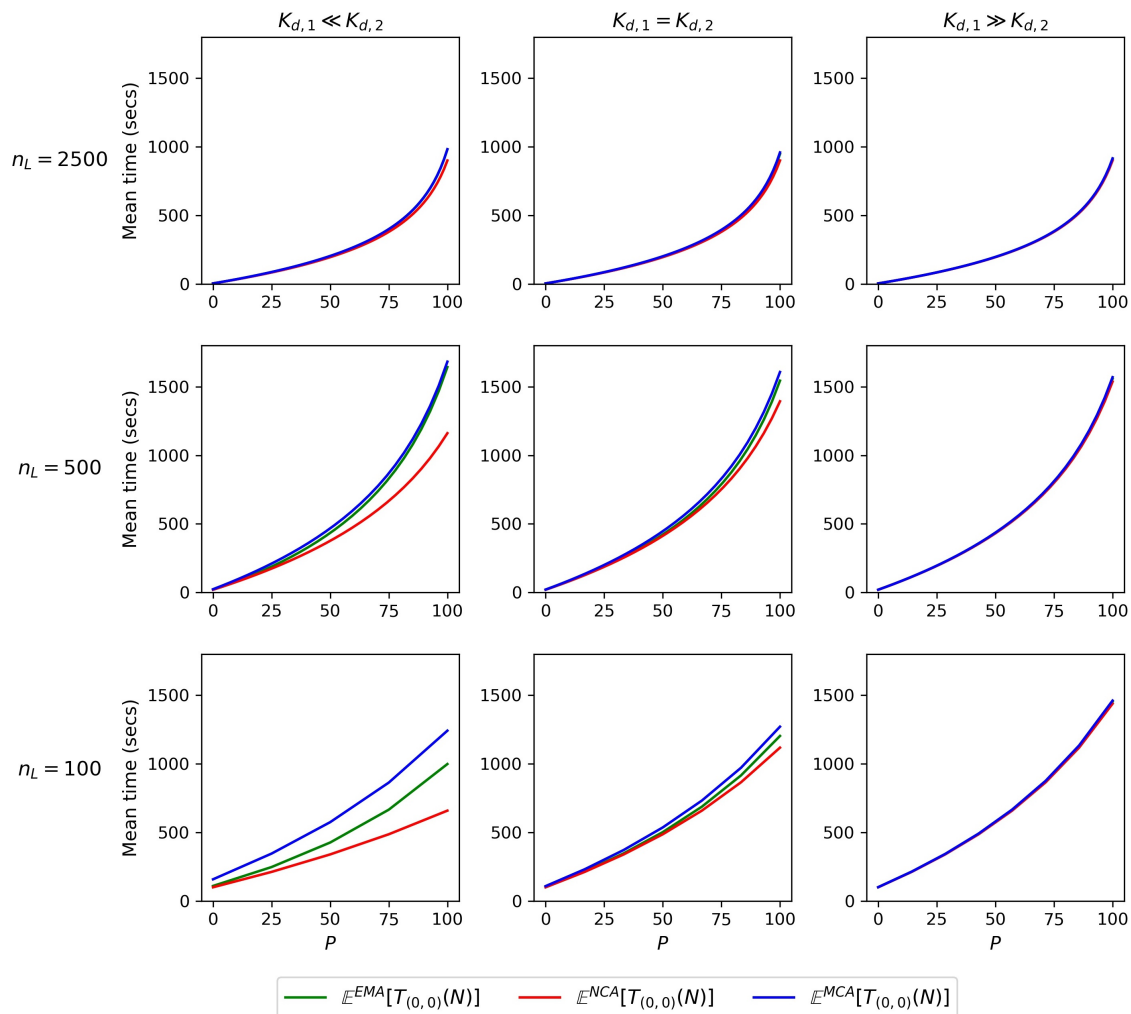


Figure 8. Comparison between the expected times to reach N monomers of type 2, computed using the Exact Matrix-Analytic Approach (EMA), No-Competition Approximation (NCA) and the MCA, for $K_{d,2} = 10^3$ and different values of $n_L \in \{100, 500, 2500\}$ and $K_{d,1} \in \{10^2, 10^3, 10^4\}$ (i.e., $K_{d,1} \ll K_{d,2}$, $K_{d,1} = K_{d,2}$ or $K_{d,1} \gg K_{d,2}$). For all plots the number of receptors are $n_{R,1} = n_{R,2} = 10^2$, and initial state $(m_1, m_2) = (0, 0)$. P represents a percentage of the mean steady-state value, so that $N = \frac{P}{100} \mathbb{E}^{EMA}[M_2]$.

Analysis of the two approximations over a range of parameter values and number of molecules can be found in Appendix C. Here, and in order to better illustrate the effect of competition on the goodness of the NCA and MCA, we plot in Figures 9 and 10 similar scatter plots to those of Figure 7, where we plot the relative differences between the mean times instead of the HD. From the first plot in Figure 9 we see that the relative difference increases with the competing strength parameter $\frac{n_{R,1}}{K_{d,1}}$. However, similarly to the case of the steady-state distributions, as we increase n_L , the competition from R_1 decreases and the NCA improves. For the sampled parameter sets, in general it seems that the MCA outperforms once again the NCA. Similarly to the NCA, the performance of the MCA tends to improve with large enough values of n_L , where in the third plot of Figure 10 in which $n_L = 500$, more than half of the sampled parameter sets lead to an approximation with a relative error smaller than 5%. It is clear that the MCA will behave rather well in situations where n_L is large enough and, in particular, of a different order of magnitude to $n_{R,1} + n_{R,2}$; for example, for $n_L = 2000$ (data not shown here; note that $n_{R,1} + n_{R,2}$ ranges from 40 to 400), an overwhelming majority of scenarios (more than 90%) in

Figure 10 lead to MCA predictions with relative errors smaller than 5%. Encouragingly also, we see that in Figure 9, all of the sampled points have relative differences greater than 0, which implies that the NCA always underestimates the EMA. This is expected given that the formation of monomers of type 2 will happen faster if there is no competition from another receptor, as assumed in the NCA. A similar pattern can be seen in Figure 10, in which the MCA tends to overestimate the EMA, leading to negative values of the relative difference. This is true for the majority of the points sampled. However, it is not always true as 6 of the 10^3 points sampled lead to very slightly positive relative differences (for $n_L = 100$ and $n_L = 250$ only).

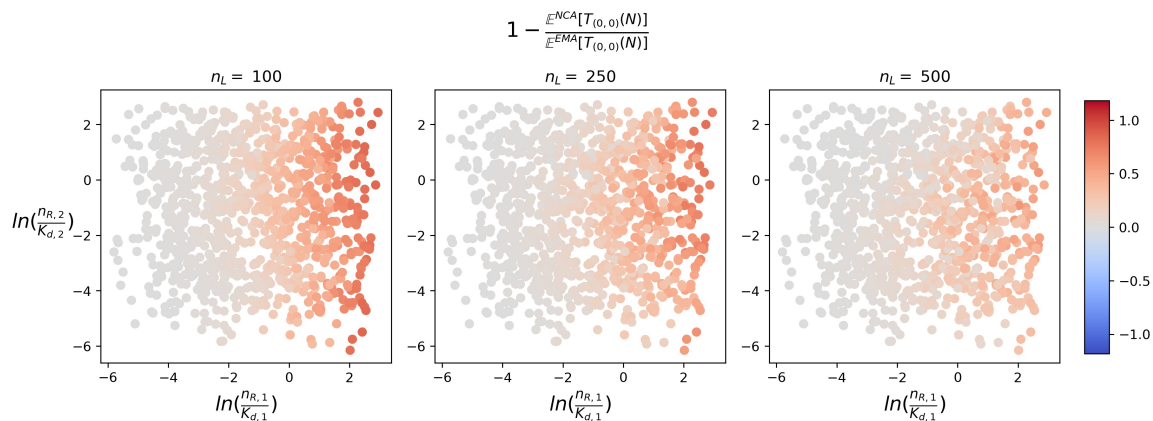


Figure 9. Relative difference between the EMA and NCA mean times to reach N monomers of type 2, plotted for 10^3 sampled parameter values in Figure 7, and for $n_L \in \{100, 250, 500\}$.

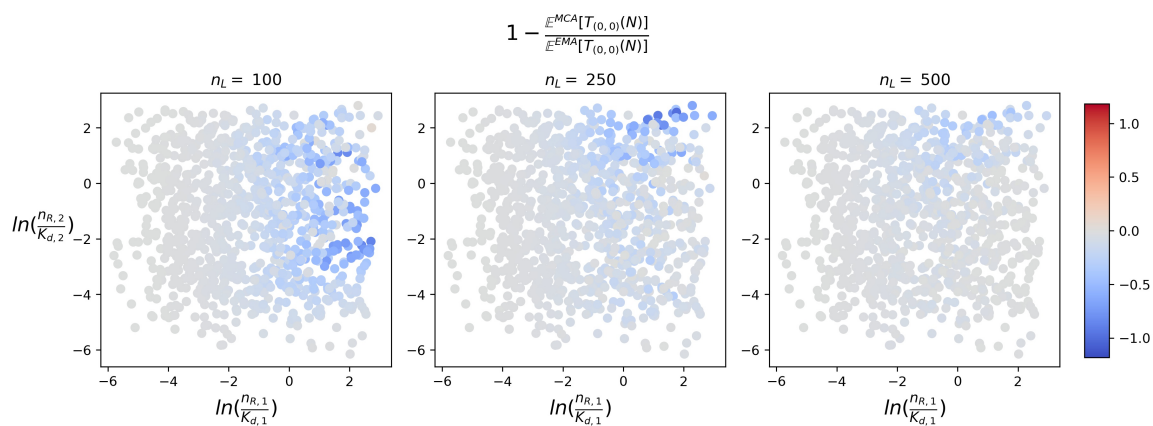


Figure 10. Relative difference between the EMA and MCA mean times to reach N monomers of type 2, plotted for 10^3 sampled parameter values in Figure 7, and for $n_L \in \{100, 250, 500\}$.

We finally note that the NCA may behave better than the MCA in situations where the mean steady-state number of type 1 monomers is perhaps non-negligible, but the timescales of type 1 monomer formation are significantly slower than the timescales of type 2 monomer formation. In these situations, considering an effective number of ligands $n_L - \mathbb{E}[M_1]$ in Equation (10) might lead to worse predictions, since this ligand depletion might take very long to occur, and should, thus, be neglected, so that the NCA would prevail. An example of this behaviour can be seen in the left plot of Figure A3 in Appendix C.

4.3. Time to Signal Initiation: Reaching a Threshold Number of Productive Monomers

We illustrate an extension of our methodology by considering a stochastic descriptor that is linked to signal initiation for some receptor tyrosine kinases (RTKs)—the time to reach a threshold number N

of productive monomers on the cell surface. In References [19,20], the authors hypothesise that signal initiation of T cells through the T cell receptor (TCR) is determined by the time a threshold number, N , of productive monomers is reached. A productive monomer is considered to be a receptor that remains bound to the ligand, for at least a dwell time τ . In References [19,20] the authors compute the mean time, $T(N, \tau)$ to reach N productive monomers. Note that $T_{(0,0)}(N)$ represents the time to reach a threshold number N of simultaneously bound monomers in the system, while $T(N, \tau)$ does not require these events to be simultaneous, but instead requires the corresponding monomers to be productive, based on the hypothesis that “counting devices are at work to allow signal accumulation, decoding and translation into biological responses”, as expressed in Reference [19].

The model in Figure 1a is proposed in Reference [19] when analysing $T(N, \tau)$ with a single receptor type (the TCR). Under ligand excess, authors in Reference [19] propose the approximation:

$$T(N, \tau) = \tau + \frac{1}{k_f n_L} \left(\frac{N'}{n_R} + \frac{1}{2!} \left(\frac{N'}{n_R} \right)^2 + \dots \right), \tag{12}$$

where k_f is the forward binding rate of the receptor, n_R is the number of receptors available in the system, and $N' = \exp(k_r \tau) N$ represents the average number of binding events required for N productive ones to be reached. Recall that dissociation of a monomer occurs after an exponentially-distributed random time $\text{Exp}(k_r)$, where k_r is the dissociation rate of the receptor.

The authors in Reference [19] make use of experimental data to test two different hypotheses: that (a) the timescale of a T cell response correlates with the time it takes to have had N receptor-ligand complexes bound for at least a threshold dwell time, τ , each; or that (b) the timescale of a T cell response correlates with the time a threshold number, N , of TCRs must be occupied at equilibrium. Their conclusion is that experimental data supports hypotheses (a), but not (b). The descriptor $T(N, \tau)$ has been proposed in References [19,20] for the T cell receptor, and a similar hypothesis has been considered for other RTKs [18,29]. We make use of the methods introduced in previous sections, to compute $T(N, \tau)$ for a model as described in Figure 1a, where the receptor of interest for signal initiation (receptor 1) has a competitor (receptor 2) for binding the ligand.

We note that, by following similar arguments to those above, the NCA approximation for $T(N, \tau)$ consists of applying Equation (12) with the total number of ligands in the system n_L , thus neglecting competition from receptor 2. On the other hand, as described in the MCA approximation, we may use Equation (12) with n_L replaced by the effective number of ligands $n_L^* = n_L - E[M_2]$.

Figure 11 shows a comparison between the mean time $T(N, \tau)$, computed making use of the NCA and MCA approximations, and numerical simulations ($T^{SIM}(N, \tau)$). The matrix-analytic approach for this descriptor is not applicable since τ is fixed. We note that the time $T(N, \tau)$ to reach N productive type 1 monomers is not well estimated by the NCA (using Equation (12) with n_L ligands) in situations where n_L is not large enough (e.g., $n_L = 100$ in Figure 11). The worst estimates are obtained in scenarios where the competing strength of receptor 1 is low, and that of receptor 2 is high, as one would expect. On the other hand, the MCA approach, using an effective number of ligands n_L^* , is better, even for a small number of ligands. In general, we observe that the MCA approximation gives good estimates of $T(N, \tau)$ over a wide range of parameter values: N and τ .

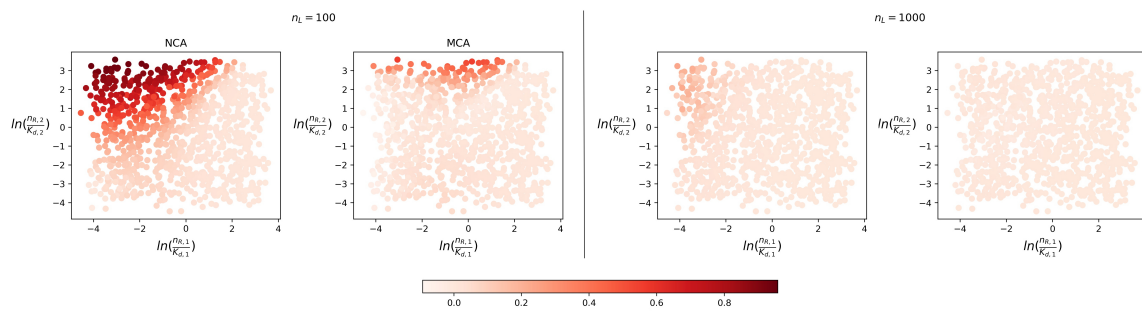


Figure 11. Relative difference $1 - \frac{T^j(N, \tau)}{T^{SIM}(N, \tau)}$, $j \in \{NCA, MCA\}$, between the mean time $T(N, \tau)$ computed through the NCA and MCA approaches, and the time computed through stochastic simulations, for $n_L \in \{10^2, 10^3\}$. In these scenarios, 10^3 parameter sets have been sampled by varying the number of receptors n_{R1} and n_{R2} between 100 and 400, $K_{d,1}$ and $K_{d,2}$ rates vary between 10^1 and 10^4 , and setting $k_{r,1} = k_{r,2} = 10^{-3} \text{ s}^{-1}$. In these examples, $N = 10$ and $\tau = k_{r,1}^{-1}$, so that a monomer is considered to be productive if it lasts for longer than its average lifetime.

5. Discussion

In this paper, our focus has been on the modelling of receptor competition for a shared ligand. We have expressed this process as a multi-variate (bi-variate, in this case) competition process, as defined by Reuter [26] and Iglehart [27] in the area of Mathematical Ecology. We have shown how this leads to the study of a bi-dimensional structured continuous-time Markov chain; in particular, a level-dependent quasi-birth-and-death-process. For this process, our interest is in the analysis of two summary statistics related to late time (the steady-state distribution of monomers) and transient (the timescales of type 2 monomer formation, or of productive type 1 monomer formation) dynamics. For these summary statistics, it is possible to implement first-step arguments which lead to the study of systems of linear equations, where the number of equations is equal to (for the steady-state distribution), or potentially less than (for the timescales of monomer formation), the number of states in the CTMC, $\#\mathcal{S}_X$. Since we are dealing with a LD-QBD process, there are matrix-oriented algorithms available in the literature [28], which we have exploited, that allow one to efficiently solve these systems.

The main limitation of this matrix-oriented approach is that the number of equations increases with the number of states, which turns out to increase with the number of molecules in the system, typically large in *in vivo* and *in vitro* settings [24]. Thus, we have proposed several approximations based on the analysis of the process in low-to-moderate competition scenarios (e.g., when the number of competing receptors is small, there is excess of ligand, or the competing receptor has a relatively low affinity compared to the other receptor). These approximations lead to the analysis of independent one-dimensional birth-and-death processes, for which analogous computations can be carried out much more efficiently. Our numerical results suggest that these approximations could be exploited in a wide range of parameter regimes, which have been explored inspired from values corresponding to the VEGF1 and VEGF2 receptors, which can bind the shared ligand, VEGF-A.

A striking advantage of using our approximate methods to compute the stochastic descriptors presented here is that the computational cost is much less than the cost of computing the same descriptors in an exact fashion (EMA). In Appendix D, Figures A4 and A5, we present heatmaps of the central processing unit (CPU) times to run each method for both descriptors and a range of the numbers of molecules, $n_{R,1}$ and $n_{R,2}$. From this analysis we find that, for the maximal values of $n_{R,1}$ and $n_{R,2}$ considered, the EMA takes over 40 min to compute the steady-state distribution, whereas the NCA and MCA can compute the distribution in a matter of seconds. Similarly, the EMA takes over 25 min to compute the mean time to reach $N = \frac{n_{R,2}}{2}$ monomers of type 2 for the maximal values of $n_{R,1}$ and $n_{R,2}$ considered, whereas the NCA and MCA again only take a few seconds. We note that as well as the CPU times, the EMA for both descriptors requires a large amount of memory for large values of $n_{R,1}$ and $n_{R,2}$, whereas for the NCA and MCA the memory required is minimal. Results in Figures A4 and A5 (see Appendix D) suggest that, due to both computational memory and time

constrains, the MCA and NCA would allow us to analyse the summary statistics of interest in this system when dealing with numbers of molecules significantly beyond those that can be considered when implementing matrix-analytic approaches.

In this paper we have considered a system of two receptors which share a common ligand. We note, however, that the modelling framework, and the analytic and approximate methods of analysis of the summary statistics, could be extended and applied to any other system in which two species compete for a shared resource. For example, our model here could be reversed to represent two ligand molecules competing for a shared receptor, such as the competition by several ligands for the epidermal growth factor receptor (EGFR) (see Roepstorff et al. [38]). For example, two major ligands for the EGFR are the epidermal growth factor (EGF) and transforming growth factor α (TGF- α). There are many other examples of multiple ligands stimulating the same receptor system to which our methodology could be applied [4,39,40]. Moreover, since our methods have been framed within processes originally devised in Mathematical Ecology, it is to be expected that they could be extended to more general competition processes in this area [41–43], not restricted to molecular dynamics.

Similarly, although our methods have been presented for bi-variate processes, they could in principle be extended to higher dimensions, where the competition of M species for a shared resource (a structured M -dimensional continuous-time Markov chain) could be approximated as M independent birth-and-death processes, or as a different combination of lower-dimensional LD-QBDs if some competing sub-groups of species need to be analysed jointly. This can be especially useful since, while the systems of equations we are dealing with in the original bi-variate process are, when $n_{R,1} + n_{R,2} \leq n_L$, of size $(n_{R,1} + 1) \cdot (n_{R,2} + 1)$ (see Appendix A), they would be of size $(n_{R,1} + 1) \cdot (n_{R,2} + 1) \dots (n_{R,M} + 1)$ for a system with M receptors, making the matrix-oriented approach not feasible in practice. On the other hand, our approximations would be dealing instead with M independent birth-and-death processes, each of them of size $n_{R,j} + 1$, $j \in \{1, \dots, M\}$. To illustrate the applicability of our approach in these higher-dimensional processes, we consider a situation with four different receptors competing for a common ligand. We plot in Figure 12 the absolute difference between the mean number of monomers in steady-state computed via the MCA and the mean number computed via stochastic simulations, for a wide range of parameter values, thus characterising different competition strengths for each receptor. We note that for the parameter regimes considered here (e.g., number of receptors ranging up to 2×10^2 for each species), the matrix-analytic approach is clearly unfeasible since it would require the analysis of CTMCs with more than 10^9 states. Thus, our NCA and MCA approximations allow one to analyse several descriptors of interest without having to implement these, numerically expensive, analytical methods. In particular, Figure 12 shows that the MCA approximation is able to capture, with high accuracy, these mean values of interest even for a small number of ligands (i.e., potentially high competition scenarios), unless the competing strengths of several receptors are too large. Even in these large competition scenarios, the MCA approach is able to estimate well these mean quantities for moderate ligand numbers (e.g., $n_L = 10^3$ in Figure 12).

Finally, a clear limitation of our approach is that it has focused on receptor-ligand binding dynamics without taking into account additional reactions such as receptor synthesis, degradation or internalisation. Our analysis can be generalised to include those scenarios; for example, when modelling the surface and intra-cellular dynamics of two kinds of receptor competing (on the surface) for a common ligand, one could use our techniques to disentangle this competition, by the consideration of an effective number of ligands, so that the two intra-cellular processes can be separately studied. In the same way, our methodology has the potential to be of use in other settings. For example, since links between Queuing Theory and receptor-ligand interactions have been already established [44,45] (e.g., the two receptors in our system can be considered as two types of customers, waiting to be served by the common ligand (i.e., server)), one could explore the applicability of our NCA and MCA approximations to Queueing Systems (e.g., by modelling a multidimensional M/M/c system with k different types of arrivals being served by c common servers, as k independent M/M/c

systems with an effective number of servers c^*). This is out of the scope of this manuscript and remains an area for future work.

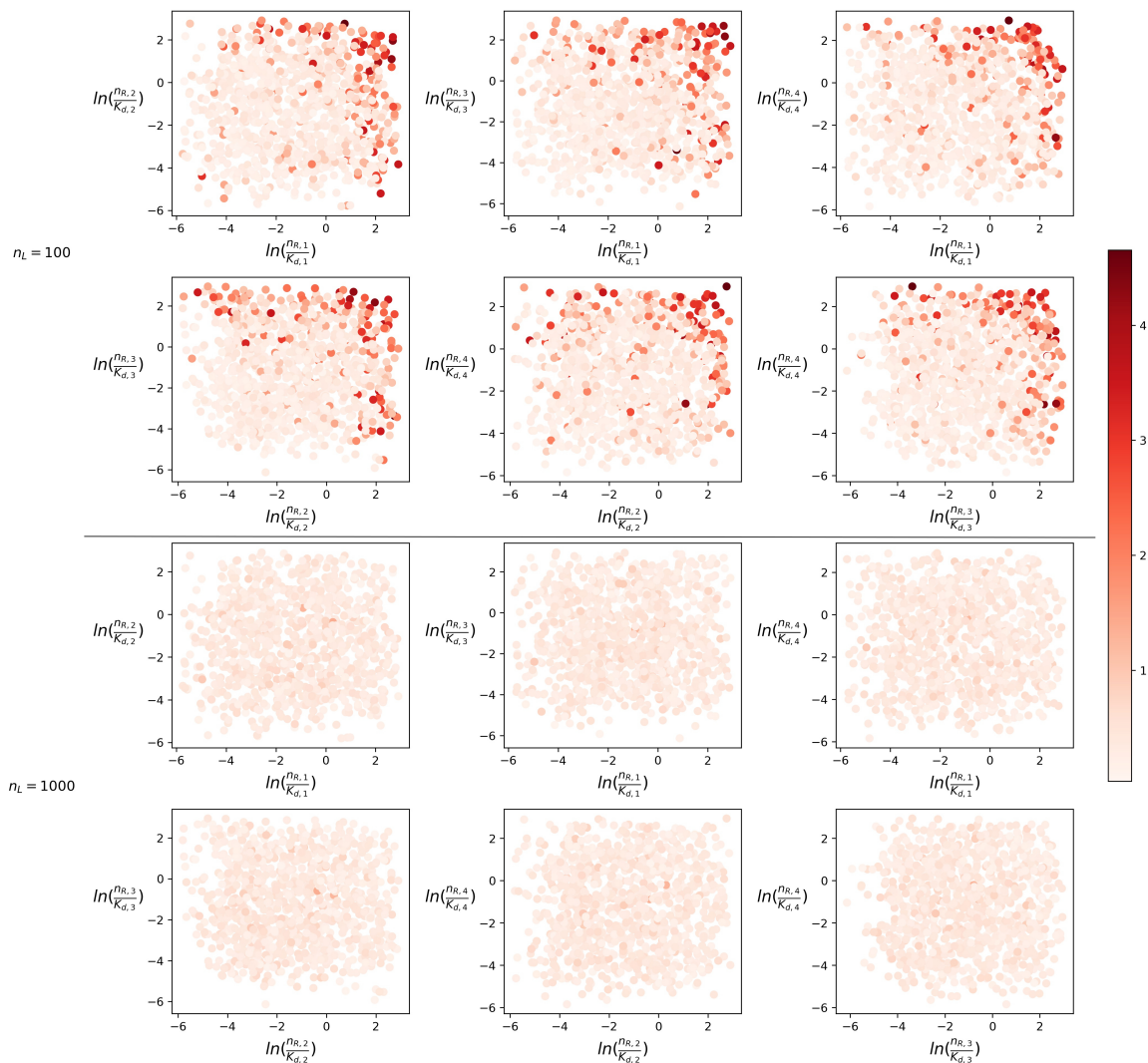


Figure 12. Absolute difference $\sum_{i=1}^4 |E^{MCA}[M_i] - E^{SIM}[M_i]|$ between the mean number of monomers in steady-state computed through simulations and the MCA approach, for $n_L \in \{10^2, 10^3\}$. 10^3 parameter sets are sampled by considering the number of receptors $n_{R,j} \sim U(20, 200)$ and $K_{d,j} = 10^x$ with $x \sim U(1, 4)$, $j \in \{1, 2, 3, 4\}$. We set $k_{r,j} = 10^{-3} \text{ s}^{-1}$ for $j \in \{1, 2, 3, 4\}$.

Author Contributions: Conceptualization, P.-A.J., M.L.-G., M.C., G.L. and C.M.-P.; Data curation, P.-A.J.; Formal analysis, P.-A.J., M.L.-G. and M.C.; Funding acquisition, C.M.-P.; Investigation, P.-A.J. and M.L.-G.; Methodology, P.-A.J., M.L.-G. and M.C.; Software, P.-A.J.; Supervision, M.L.-G., G.L. and C.M.-P.; Visualization, P.-A.J.; Writing—Original draft, P.-A.J. and M.L.-G.; Writing—Review & editing, P.-A.J., M.L.-G., M.C., G.L. and C.M.-P. All authors have read and agreed to the published version of the manuscript.

Funding: PJ is supported by the EPSRC, AstraZeneca and Smith Institute (Smith Institute CASE studentship, award reference 1969354). MC is partially supported by grants FIS2016-78883-C2-2-P and PID2019-106339GB-I00 (Ministerio de Economía, Industria y Competitividad - Agencia Estatal de Investigación).

Acknowledgments: Numerical work was undertaken on ARC3, which is part of the High Performance Computing facilities at the University of Leeds, UK. We acknowledge and are grateful to the International Centre for Mathematical Sciences (ICMS), Edinburgh, where early discussions of this work took place during a Research in Groups programme (MLG, MC, GL and CMP).

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability: Python (version 3.7) codes for computing the steady-state distribution and expected times to reach N monomers of type 2, via the EMA, NCA and MCA can be found in the public repository, "<https://github.com/PollyJeffrey/Competition-modelling-paper>".

Abbreviations

The following abbreviations are used in this manuscript:

- EMA Exact matrix-analytic approach
- NCA No-competition approximation
- MCA Moderate-competition approximation

Appendix A

In this appendix we explain how the tri-diagonal by blocks structure of the infinitesimal generator matrix Q in Section 3 results from a particular ordering of the states in $\mathcal{S}_{\mathcal{X}}$. We note first that the number of states in $\mathcal{S}_{\mathcal{X}}$ is

$$\#\mathcal{S}_{\mathcal{X}} = \sum_{k=0}^{N_2} (\min(N_1, n_L - k) + 1),$$

where $N_1 = \min(n_{R,1}, n_L)$ and $N_2 = \min(n_{R,2}, n_L)$. In fact, it is possible to obtain explicit formulae for the number of states in $\mathcal{S}_{\mathcal{X}}$ depending on the values of $n_{R,1}$, $n_{R,2}$ and n_L , as

$$\#\mathcal{S}_{\mathcal{X}} = \begin{cases} (n_{R,1} + 1)(n_{R,2} + 1), & \text{if } n_{R,1} + n_{R,2} \leq n_L, \\ (n_L - n_{R,1} + 1)(n_{R,1} + 1) + \frac{(n_{R,1} + n_{R,2} - n_L)(n_{R,1} + n_L - n_{R,2} + 1)}{2}, & \text{if } n_{R,1} \leq n_{R,2} \leq n_L \\ & \text{and } n_{R,1} + n_{R,2} > n_L, \\ (n_L - n_{R,2} + 1)(n_{R,2} + 1) + \frac{(n_{R,1} + n_{R,2} - n_L)(n_{R,2} + n_L - n_{R,1} + 1)}{2}, & \text{if } n_{R,2} \leq n_{R,1} \leq n_L \\ & \text{and } n_{R,1} + n_{R,2} > n_L, \\ \frac{(n_{R,2} + 1)(2n_L - n_{R,2} + 2)}{2}, & \text{if } n_{R,1} > n_L \text{ and } n_{R,2} < n_L, \\ \frac{(n_{R,1} + 1)(2n_L - n_{R,1} + 2)}{2}, & \text{if } n_{R,1} < n_L \text{ and } n_{R,2} > n_L, \\ \frac{(n_L + 1)(n_L + 2)}{2}, & \text{if } n_{R,1} \geq n_L \text{ and } n_{R,2} \geq n_L. \end{cases}$$

This state space can be organised into levels, as follows,

$$\mathcal{S}_{\mathcal{X}} = \bigcup_{k=0}^{N_2} L(k), \quad L(k) = \{(m_1, m_2) \in \mathcal{S}_{\mathcal{X}} : m_2 = k\}, \quad 0 \leq k \leq N_2.$$

Thus, level $L(k) = \{(0, k), (1, k), \dots, (\min(N_1, n_L - k), k)\}$ contains those states representing k type 2 monomers present on the cell surface at any given time. If, when constructing the matrix Q , one orders the states by levels with $L(0) \prec L(1) \prec \dots \prec L(N_2)$, and states within each level $L(k)$ as $(0, k) \prec (1, k) \prec \dots \prec (\min(N_1, n_L - k), k)$, this leads to a tri-diagonal by blocks infinitesimal generator. In particular, from the transition diagram in Figure 2, it is clear that only transitions between adjacent states are allowed. This means that from any state $(m_1, m_2) \in L(m_2)$, the next event in the Markov chain can take the process to either another state in level $L(m_2)$, a state in level $L(m_2 + 1)$, or a state in level $L(m_2 - 1)$, by means of an association or dissociation reaction involving M_1 , an association to form M_2 or a dissociation of M_2 , respectively. Thus, since the process only moves up or down by a

maximum of one level after each transition, the CTMC \mathcal{X} is a level-dependent quasi-birth-and-death process (LD-QBD) [46], and the infinitesimal generator matrix Q is tri-diagonal by blocks. Q is given by

$$Q = \begin{pmatrix} Q_{0,0} & Q_{0,1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ Q_{1,0} & Q_{1,1} & Q_{1,2} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Q_{2,1} & Q_{2,2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & Q_{N_2-1,N_2-1} & Q_{N_2-1,N_2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & Q_{N_2,N_2-1} & Q_{N_2,N_2} \end{pmatrix}.$$

We note that each level $L(k)$ contains $J(k) = \#L(k) = \min(N_1, n_L - k) + 1$ states, so that each sub-matrix $Q_{k,k'}$ has dimensions $J(k) \times J(k')$. Sub-matrices $Q_{k,k'}$ are given as follows:

- For $1 \leq k \leq N_2$,

$$(Q_{k,k-1})_{i,j} = \begin{cases} k_{r,2}k, & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

for $0 \leq i \leq J(k)$ and $0 \leq j \leq J(k-1)$.

- For $0 \leq k \leq N_2 - 1$,

$$(Q_{k,k+1})_{i,j} = \begin{cases} k_{f,2}(n_{R,2} - k)(n_L - i - k), & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

for $0 \leq i \leq J(k)$ and $0 \leq j \leq J(k+1)$.

- For $0 \leq k \leq N_2$,

$$(Q_{k,k})_{i,j} = \begin{cases} k_{f,1}(n_{R,1} - i)(n_L - i - k), & \text{if } j = i + 1, \\ k_{r,1}i, & \text{if } j = i - 1, \\ -(k_{f,1}(n_{R,1} - i)(n_L - i - k) + k_{r,1}i + k_{f,2}(n_{R,2} - k)(n_L - i - k) + k_{r,2}k), & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

for $0 \leq i \leq J(k)$ and $0 \leq j \leq J(k)$.

This tri-diagonal by blocks form of the infinitesimal generator allows us to solve the steady-state matrix equations, Equation (2), with Algorithm A1, where π is comprised of row vectors $\pi_0, \pi_1, \dots, \pi_{N_2}$ which contain the ordered probabilities $\pi_{(m_1,m_2)}$ for the states in the corresponding level, $\pi_k = \left(\pi_{(0,k)}^{EMA}, \pi_{(1,k)}^{EMA}, \dots, \pi_{(\min(N_1, n_L - k), k)}^{EMA} \right), 0 \leq k \leq N_2$. We note that Algorithm A1 is just an adapted version of the linear level reduction algorithm [47].

Algorithm A1 Computation of the steady-state distribution $\{\pi_{(m_1, m_2)}^{EMA}, (m_1, m_2) \in \mathcal{S}_{\mathcal{X}}\}$ [47,48].

- 1: $\mathbf{H}_0 = \mathbf{Q}_{0,0}$
 - 2: **for** $k = 1, \dots, N_2$ **do**
 - 3: $\mathbf{H}_k = \mathbf{Q}_{k,k} - \mathbf{Q}_{k,k-1} \mathbf{H}_{k-1}^{-1} \mathbf{Q}_{k-1,k}$
 - 4: **end for**
 - 5: Evaluate $\pi_{N_2}^*$ by solving $\pi_{N_2}^* \mathbf{H}_{N_2} = \mathbf{0}_{J(N_2)}$ with $\pi_{N_2}^* \mathbf{1}_{J(N_2)} = 1$
 - 6: **for** $k = N_2 - 1, \dots, 0$ **do**
 - 7: $\pi_k^* = -\pi_{k+1}^* \mathbf{Q}_{k+1,k} \mathbf{H}_k^{-1}$
 - 8: **end for**
 - 9: **for** $k = 0, \dots, N_2$ **do**
 - 10: $\pi_k = \frac{1}{\sum_{k=0}^{N_2} \pi_k^* \mathbf{1}_{J(k)}} \pi_k^*$
 - 11: **end for**
 - 12: **return** $\pi = (\pi_0, \dots, \pi_{N_2})$ ▷ EMA steady-state distribution
-

Appendix B

In this Appendix we present the details of the calculation of the moment of order l of the time to reach N monomers of type 2, starting from any initial state (m_1, m_2) , as introduced in Section 3. These moments can be obtained from the Laplace-Stieltjes transform of $T_{(m_1, m_2)}(N)$, given by

$$\phi_{(m_1, m_2)}^N(z) = \mathbb{E} \left[e^{-zT_{(m_1, m_2)}(N)} \right], \quad \Re(z) \geq 0.$$

The index N will be omitted from now on to simplify notation. Using a first-step argument, one finds the following system of linear equations

$$\begin{aligned} \phi_{(m_1, m_2)}(z) = & \frac{k_{f,1}(n_{R,1} - m_1)(n_L - m_1 - m_2)}{z + \Delta_{(m_1, m_2)}} \phi_{(m_1+1, m_2)}(z) + \frac{k_{r,1}m_1}{z + \Delta_{(m_1, m_2)}} \phi_{(m_1-1, m_2)}(z) \\ & + \frac{k_{f,2}(n_{R,2} - m_2)(n_L - m_1 - m_2)}{z + \Delta_{(m_1, m_2)}} \phi_{(m_1, m_2+1)}(z) - \frac{k_{r,2}m_2}{z + \Delta_{(m_1, m_2)}} \phi_{(m_1, m_2-1)}(z), \end{aligned} \quad (\text{A1})$$

for $(m_1, m_2) \in \mathcal{S}_{\mathcal{X}}$ with $m_2 < N$, and where $\Delta_{(m_1, m_2)} = k_{f,1}(n_{R,1} - m_1)(n_L - m_1 - m_2) + k_{r,1}m_1 + k_{f,2}(n_{R,2} - m_2)(n_L - m_1 - m_2) + k_{r,2}m_2$. Boundary conditions are given by $\phi_{(m_1, m_2)}(z) \equiv 1$ if $m_2 = N$, since $T_{(m_1, m_2)}(m_2) \equiv 0$.

We note that $E[T_{(m_1, m_2)}^l] = (-1)^l \frac{d^l}{dz^l} \phi_{(m_1, m_2)}(z) \Big|_{z=0}$. Thus, once the Laplace-Stieltjes transforms are in hand, if we multiply Equation (A1) by $(-1)^l$, differentiate it with respect to z l times, and set $z = 0$, we obtain the system of equations (8) for $\mathbb{E} \left[T_{(m_1, m_2)}^l \right]$, for any $l \geq 1$.

Moreover, by arranging the states in $\mathcal{S}_{\mathcal{X}}$ in levels as in Appendix A, we note that each equation in the system (8) corresponds to an initial state $(m_1, m_2) \in \cup_{k=0}^{N-1} L(k)$, so that one can rewrite Equation (8) in matrix form as follows

$$\mathbf{m}^{(l)} = \mathbf{A} \mathbf{m}^{(l)} + \mathbf{b}^{(l)} \quad (\text{A2})$$

where

$$\mathbf{m}^{(l)} = \begin{pmatrix} \mathbf{m}_0^{(l)} \\ \mathbf{m}_1^{(l)} \\ \vdots \\ \mathbf{m}_{N-1}^{(l)} \end{pmatrix}, \quad \mathbf{m}_k^{(l)} = \begin{pmatrix} \mathbb{E}[T_{(0,k)}^l] \\ \mathbb{E}[T_{(1,k)}^l] \\ \vdots \\ \mathbb{E}[T_{(\min(N_1, n_L-k), k)}^l] \end{pmatrix}, \quad 0 \leq k \leq N-1.$$

Vector $\mathbf{b}^{(l)}$ is also organised in sub-vectors as

$$\mathbf{b}^{(l)} = \begin{pmatrix} \mathbf{b}_0^{(l)} \\ \mathbf{b}_1^{(l)} \\ \vdots \\ \mathbf{b}_{N-1}^{(l)} \end{pmatrix}$$

which are obtained from the $(l-1)^{th}$ order moments, as $(\mathbf{b}_k^{(l)})_{(m_1, m_2)} = \frac{l}{\Delta_{(m_1, m_2)}} (\mathbf{m}_k^{(l-1)})_{(m_1, m_2)}$, for all $(m_1, m_2) \in L(k)$; we note that sub-vectors $\mathbf{m}_k^{(0)}$ are just column vectors of ones. Matrix \mathbf{A} is given by

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} & \mathbf{0}_{J(0) \times J(2)} & \cdots & \mathbf{0}_{J(0) \times J(N-2)} & \mathbf{0}_{J(0) \times J(N-1)} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{0}_{J(1) \times J(N-2)} & \mathbf{0}_{J(1) \times J(N-1)} \\ \mathbf{0}_{J(2) \times J(0)} & \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{0}_{J(2) \times J(N-2)} & \mathbf{0}_{J(2) \times J(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{J(N-2) \times J(0)} & \mathbf{0}_{J(N-2) \times J(1)} & \mathbf{0}_{J(N-2) \times J(2)} & \cdots & \mathbf{A}_{N-2, N-2} & \mathbf{A}_{N-2, N-1} \\ \mathbf{0}_{J(N-1) \times J(0)} & \mathbf{0}_{J(N-1) \times J(1)} & \mathbf{0}_{J(N-1) \times J(2)} & \cdots & \mathbf{A}_{N-1, N-2} & \mathbf{A}_{N-1, N-1} \end{pmatrix}.$$

The sub-matrices $\mathbf{A}_{k,k'}$ are defined as follows:

- For $1 \leq k \leq N-1$,

$$(\mathbf{A}_{k,k-1})_{i,j} = \begin{cases} \frac{k_{r,2k}}{\Delta_{(i,k)}}, & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq J(k)$ and $0 \leq j \leq J(k-1)$.

- For $0 \leq k \leq N-2$,

$$(\mathbf{A}_{k,k+1})_{i,j} = \begin{cases} \frac{k_{f,2(n_{R,2}-k)(n_L-i-k)}}{\Delta_{(i,k)}}, & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq J(k)$ and $0 \leq j \leq J(k+1)$.

- For $0 \leq k \leq N-1$,

$$(\mathbf{A}_{k,k})_{i,j} = \begin{cases} \frac{k_{f,1(n_{R,1}-i)(n_L-i-k)}}{\Delta_{(i,k)}}, & \text{if } j = i+1, \\ \frac{k_{r,1i}}{\Delta_{(i,k)}}, & \text{if } j = i-1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq J(k)$ and $0 \leq j \leq J(k)$.

Equation (A2) can then be solved efficiently using Algorithm A2 to obtain the l^{th} order moments of the random variable $T_{(m_1, m_2)}$. In this algorithm, \mathbf{I}_a denotes the $a \times a$ identity matrix.

Algorithm A2 Moments of the random variable $T_{(m_1, m_2)}$.

- 1: $(\mathbf{b}_k^{(1)})_{(m_1, m_2)} = \frac{1}{\Delta_{(m_1, m_2)}}$ for all $(m_1, m_2) \in L(k)$, for $k = 0, \dots, N - 1$
 - 2: $\mathbf{R}_0 = \mathbf{I}_{J(0)} - \mathbf{A}_{0,0}$
 - 3: **for** $p = 1, \dots, l$ **do**
 - 4: $\mathbf{S}_0 = \mathbf{R}_0^{-1} \cdot \mathbf{b}_0^{(p)}$
 - 5: **for** $k = 1, \dots, N - 1$ **do**
 - 6: $\mathbf{R}_k = \mathbf{I}_{J(k)} - \mathbf{A}_{k,k} - \mathbf{A}_{k,k-1} \mathbf{R}_{k-1}^{-1} \mathbf{A}_{k-1,k}$
 - 7: $\mathbf{S}_k = \mathbf{R}_k^{-1} \mathbf{A}_{k,k-1} \mathbf{S}_{k-1} + \mathbf{R}_k^{-1} \mathbf{b}_k^{(p)}$
 - 8: **end for**
 - 9: $\mathbf{m}_{N-1}^{(p)} = \mathbf{S}_{N-1}$
 - 10: **for** $k = N - 2, \dots, 0$ **do**
 - 11: $\mathbf{m}_k^{(p)} = \mathbf{S}_k + \mathbf{R}_k^{-1} \mathbf{A}_{k,k+1} \mathbf{m}_{k+1}^{(p)}$
 - 12: **end for**
 - 13: $(\mathbf{b}_k^{(p+1)})_{(m_1, m_2)} = \frac{p+1}{\Delta_{(m_1, m_2)}} (\mathbf{m}_k^{(p)})_{(m_1, m_2)}$, for all $(m_1, m_2) \in L(k)$, for $k = 0, \dots, N - 1$
 - 14: **end for**
 - 15: **return** $\mathbf{m}^{(l)} = (\mathbf{m}_0^T, \dots, \mathbf{m}_{N-1}^T)^T$ ▷ l^{th} order moment of $T_{(m_1, m_2)}$
-

Appendix C

In this Appendix we analyse how similar the distributions $\{\pi_{(m_1, m_2)}^{EMA}\}_{(m_1, m_2) \in \mathcal{S}_X}$ and $\{\pi_{(m_1, m_2)}^{MCA}\}_{(m_1, m_2) \in \mathcal{S}_X}$ are for different parameter values $(K_{d,1}, K_{d,2}, n_{R,1}, n_{R,2}, n_L)$, by plotting the corresponding Hellinger distance (HD) in Figure A1. The range of parameter values considered here allows us to consider scenarios of low to moderate and high competition between the two receptors for the shared ligand. The HD seems to be lower than 0.16 for all of the scenarios considered, even though in some of them, the number of ligands is significantly smaller than the total number of receptors (e.g., $n_L = 250 < 400 = 200 + 200 = n_{R,1} + n_{R,2}$). The number of molecules in the system does not, by itself, explain the goodness of the MCA; the dissociation constants $K_{d,1}$ and $K_{d,2}$ need also to be taken into account. Settings where both $K_{d,1}$ and $K_{d,2}$ are large (indicating the lowest affinity for ligand) correspond to low competition, since a small number of monomers of each type is formed, and then the baseline number of ligands considered, $n_L = 250$, is sufficient for the competition to be negligible. On the other hand, scenarios where for example, $K_{d,1}$ is small, can still lead to low competition if n_L is sufficiently increased, or alternatively if $n_{R,1}$ decreased.

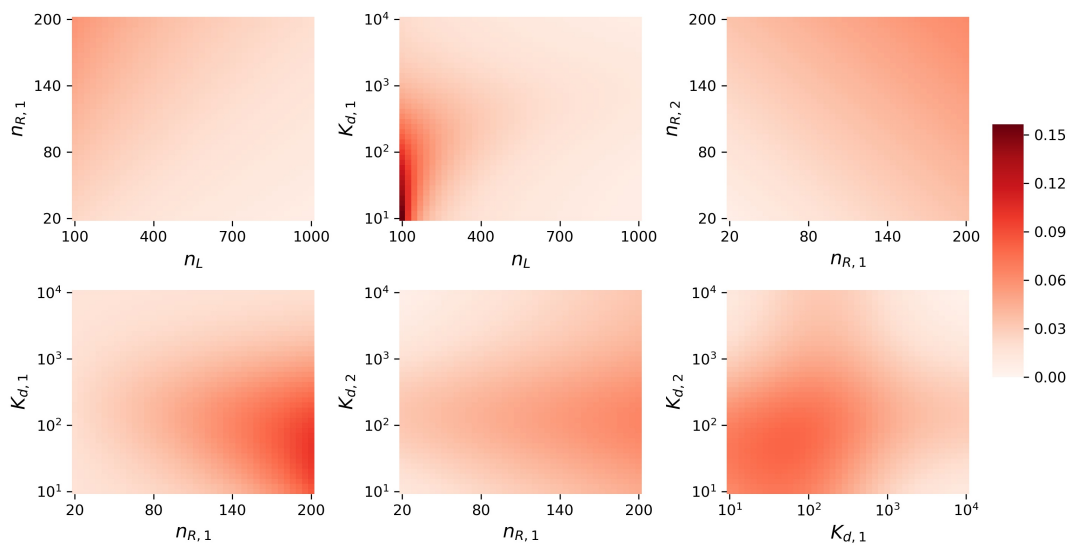


Figure A1. Hellinger distance between the steady-state distributions $\{\pi_{(m_1,m_2)}^{EMA}\}_{(m_1,m_2)\in\mathcal{S}_X}$ and $\{\pi_{(m_1,m_2)}^{MCA}\}_{(m_1,m_2)\in\mathcal{S}_X}$. Baseline parameter values (that is, the value chosen in each plot for any parameter that has been fixed) are $n_{R,1} = 10^2$, $n_{R,2} = 10^2$, $K_{d,1} = 10^3$, $K_{d,2} = 10^3$ and $n_L = 250$. We set the threshold parameter $\epsilon = 10^{-5}$ for the MCA.

When focusing on the time to reach N monomers of type 2, and similarly to the comprehensive sensitivity analysis carried out in Figure A1 for the steady-state distributions, we vary in Figure A2 the parameters $(K_{d,1}, K_{d,2}, n_L, n_{R,1}, n_{R,2})$ in pairs, and plot heatmaps of the relative differences,

$$1 - \frac{\mathbb{E}^{NCA}[T_{(0,0)}(N)]}{\mathbb{E}^{EMA}[T_{(0,0)}(N)]} \quad \text{and} \quad 1 - \frac{\mathbb{E}^{MCA}[T_{(0,0)}(N)]}{\mathbb{E}^{EMA}[T_{(0,0)}(N)]},$$

for $N = \mathbb{E}^{EMA}[M_2]$, so that $T_{(0,0)}(N)$ encodes the timescales for type 2 monomers to reach steady-state. We note that the MCA seems to better approximate the exact mean time in almost all parameter regimes explored in Figure A2. We expect however that the NCA might perform better in some scenarios (some of them can be identified in Figure A2), for example, when N is small (see Figure 8). Still, we observe in Figure A2 (left) that the NCA performs relatively well in parameter regimes representing low natural competition from R_1 , for example when $K_{d,1}$ is relatively large (indicative of low affinity of R_1 for L) and/or $n_{R,1}$ is small. In Figure A2 (right), this effect of competition from receptors of type 1 is reduced, and the MCA in general performs much better thanks to the effective ligand numbers considered in Equation (10), which takes into account ligand depletion due to receptor competition. We note that the MCA and the NCA lead to very similar, and almost exact, results under very low competition scenarios, since when $E[M_1] \approx 0$, $n_L^* \approx n_L$ in Equation (10) and thus, the two approximations are effectively the same.

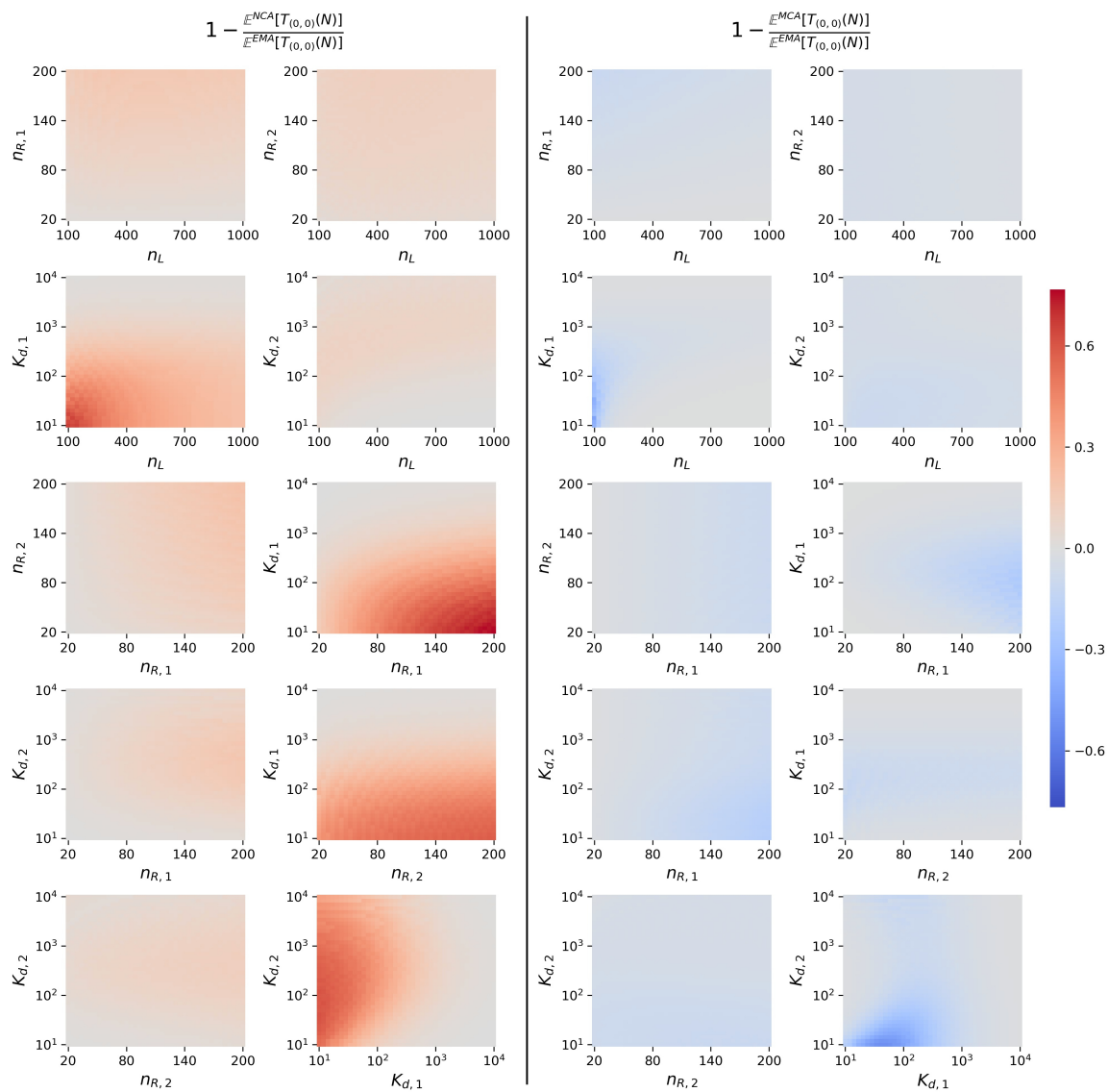


Figure A2. Relative difference between the EMA and NCA (left) and MCA (right) expected times to reach N monomers of type 2. Baseline parameter values are $n_{R,1} = 10^2$, $n_{R,2} = 10^2$, $K_{d,1} = 10^3$, $K_{d,2} = 10^3$ and $n_L = 250$.

As mentioned in Section 4.2, there are some situations in which the NCA behaves better than the MCA. For example, the NCA seems to lead to better predictions than the MCA in Figure A3 (left), where the formation of type 2 monomers occurs more rapidly than that of type 1 monomers. In this figure, by carrying out a single stochastic (Gillespie) simulation of the process, we get a particular realisation of the time $T_{(0,0)}(N)$ to reach $N = \frac{P}{100} \mathbb{E}^{EMA}[M_2]$ type 2 monomers, for different values of P ; in particular, for $P = 20, 40, 60, 80$ and 100% of the average number of monomers in steady-state, we get the dots plotted in Figure A3 (top). On the other hand, in situations where the type 1 receptor is a strong competitor, such as in Figure A3 (right), the MCA outperforms the NCA.

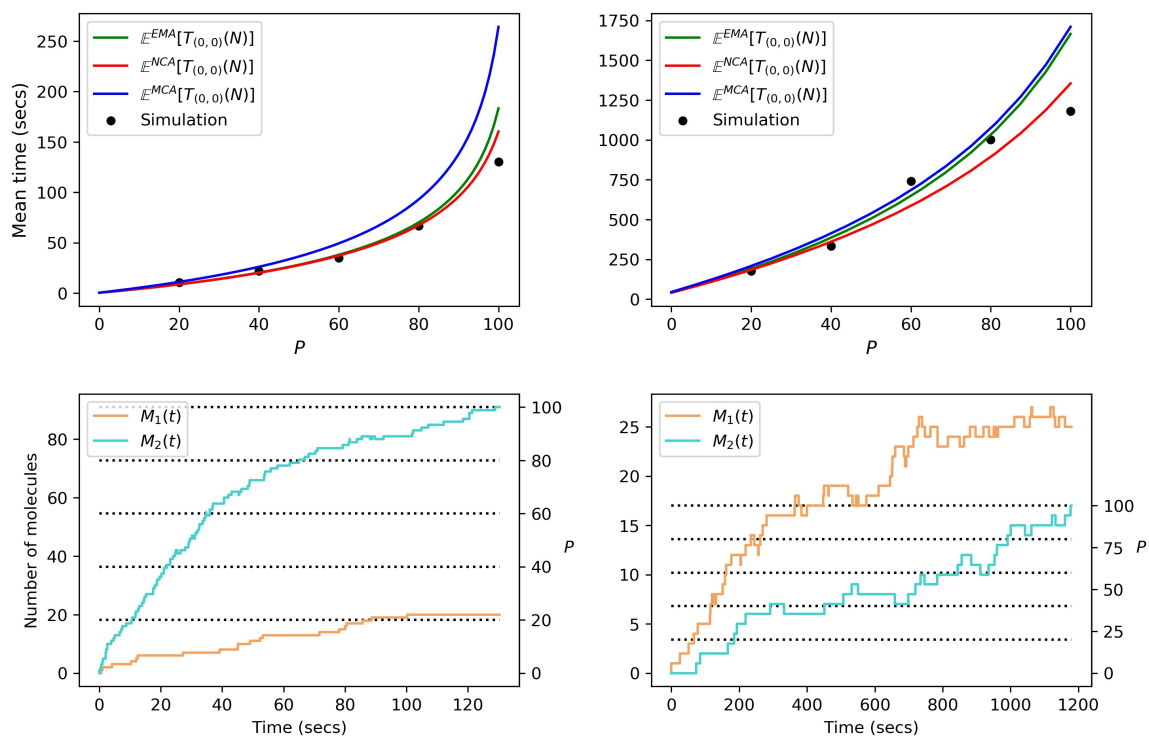


Figure A3. Top: Comparison between the mean times computed using the EMA, NCA and MCA. P represents a percentage of the mean steady-state value, as in Figure 8, so that $N = \frac{P}{100} \mathbb{E}^{EMA}[M_2]$. **Bottom:** Stochastic realisations of the processes analysed in the top scenarios, leading to stochastic realisations (plotted as black dots in top plots) of random variable $T_{(0,0)}(N)$ for $N = \frac{P}{100} \mathbb{E}^{EMA}[M_2]$ and different values of P . **Left:** $n_{R,1} = n_{R,2} = 10^2$, $K_{d,1} = 10^2$ and $K_{d,2} = 10$. **Right:** $n_{R,1} = 50$, $n_{R,2} = 10^2$, $K_{d,1} = 10$ and $K_{d,2} = 10^3$. Number of ligands $n_L = 250$.

Appendix D

In this appendix we illustrate the computational advantage of using our approximations. In Figures A4 and A5, a simple computational comparison between the NCA/MCA and the EMA is presented. In particular, we plot heatmaps of the central processing unit (CPU) times to run each method, for a range of numbers of the molecules $n_{R,1}$ and $n_{R,2}$. In both figures, we set $n_L = 10,000$ and $K_{d,1} = K_{d,2} = 1000$. We note that, for a large number of ligands such as $n_L = 10,000$, the computational cost of the EMA just depends on the values of $n_{R,1}$ and $n_{R,2}$ in Figures A4 and A5 (see Appendix A, Algorithms A1 and A2), and is independent of $K_{d,1}$ and $K_{d,2}$. The colour bar on each figure represents the CPU time for the computation of the summary statistic, with units minutes. We notice that, as expected, the EMA requires the largest amount of time to compute both descriptors since, especially when $n_{R,1}$ and $n_{R,2}$ are large, the method involves the computation of relatively large matrix inverses. There is a similar trend for both the NCA and MCA approaches to compute the steady-state distribution, that is, the larger $n_{R,1}$ and $n_{R,2}$, the greater the CPU time. We see that, whereas the CPU time under the EMA takes up to a maximum of approximately 45 min, the NCA and MCA take less than 1 min for all numbers of molecules considered. For the second descriptor, we observe from Figure A5 that the EMA takes roughly half the CPU time as the EMA for the steady state distribution in Figure A4. This is since we are considering a value of $N = \frac{n_{R,2}}{2}$, for the expected times and hence we are only inverting half the matrices as required in Figure A4. Again we see for the expected times in Figure A5, that the NCA and MCA are strikingly faster to compute than the EMA. As expected, the NCA in this case does not depend on the value of $n_{R,1}$ since this quantity does not feature in the calculation. For the MCA however, we must first compute $\mathbb{E}^{MCA}[M_1]$ using the MCA for the steady-state distribution and hence the CPU times here do depend on $n_{R,1}$.

In Figure A6, we explore in more detail specific pixels from Figures A4 and A5. In particular we choose 3 values of $n_{R,2}$, corresponding to the 3 columns in Figure A6 and vary on the x -axis of each plot, the value of $n_{R,1}$. We are thus considering in each plot of Figure A6, a cross-section of the heatmaps in Figures A4 and A5. The top row of Figure A6 relates to the steady-state distribution descriptor and hence the y -axis of the plots in this row is the Hellinger distance. The second row in the figure relates to the expected times to reach N monomers of type 2 descriptor and hence the y -axis of the plots in this row is the relative difference as plotted in Figures 9 and 10. Finally the colour bar represents the CPU time saved with units *minutes*, by computing the descriptors using the approximations (NCA and MCA) instead of the analytic (EMA). For the steady state descriptor in the first row we see that as expected the Hellinger distance between the EMA and the NCA increases as $n_{R,1}$ increases, whereas the Hellinger distance between the EMA and the MCA remains almost constant, indicating that this approximation works very well for these numbers of molecules. As we increase $n_{R,1}$ and $n_{R,2}$ we find that the time saved by using the NCA or MCA increases. In the second row we note that both the NCA and MCA worsen as approximations as we increase $n_{R,1}$, but this worsening is only very slight (the relative differences are always $\ll 1$). Again we see that as we increase $n_{R,1}$ and $n_{R,2}$ the time saved by using the NCA or MCA increases. From this figure we see that there is a “playoff” between the Hellinger distance or relative difference and the CPU time saved, but for the numbers of molecules and parameter values considered here, the time saved clearly outweighs the slight deviation from the EMA result caused by using the MCA for the steady state distribution and both the NCA and MCA for the expected times descriptor.

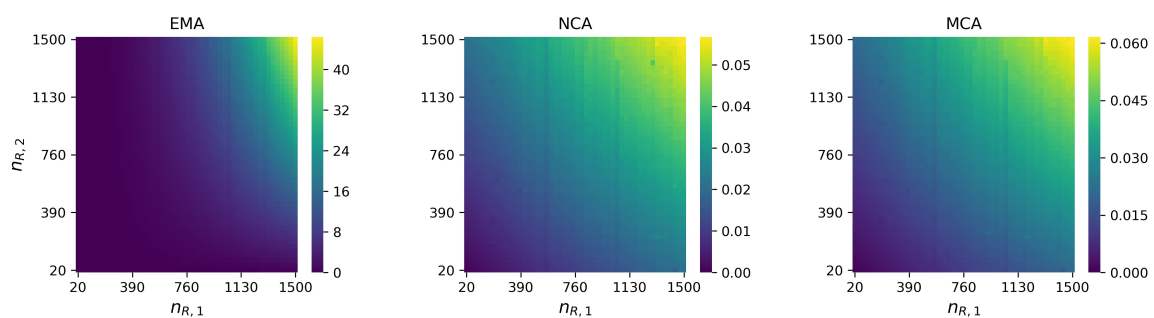


Figure A4. Comparison of CPU time (in *minutes*) required to compute the steady-state distribution for the EMA, NCA and MCA for different receptor numbers.

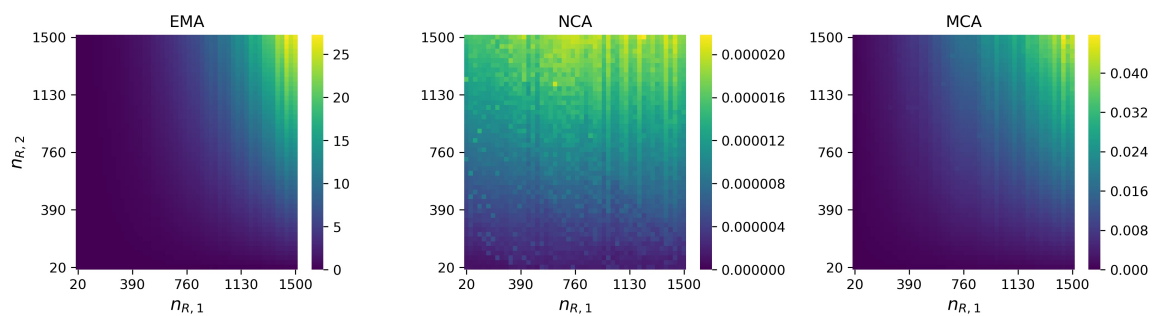


Figure A5. Comparison of CPU time (in *minutes*) required to compute the mean time to reach $N = \frac{n_{R,2}}{2}$ monomers of type 2 for the EMA, NCA and MCA for different receptor numbers.

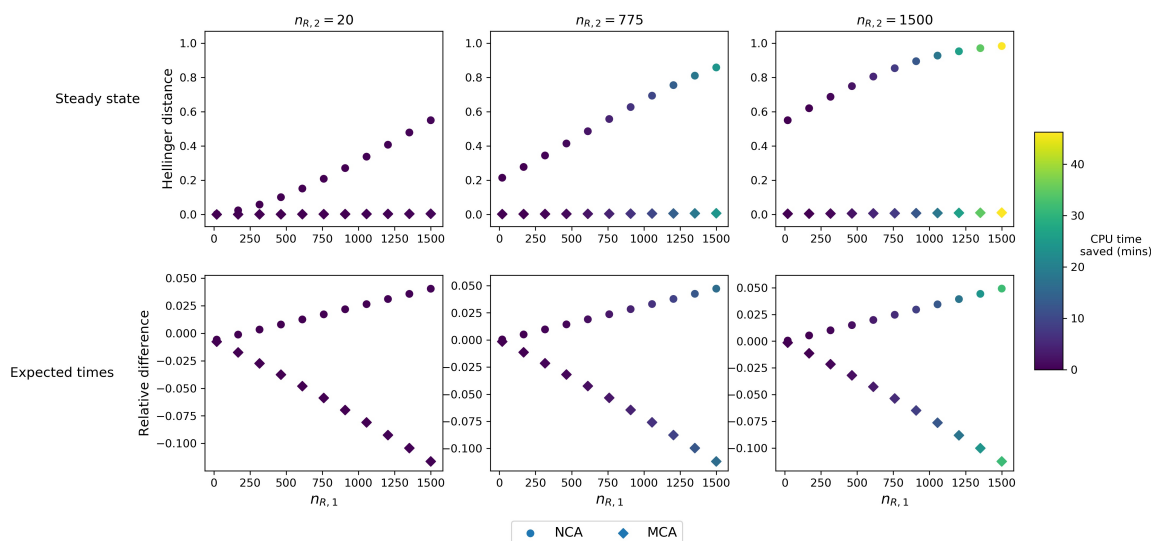


Figure A6. Top: Playoff between the Hellinger distance, the value of $n_{R,1}$ and the CPU time saved by considering either the NCA or MCA instead of the EMA when computing the steady state distribution. **Bottom:** Playoff between the relative difference, the value of $n_{R,1}$ and the CPU time saved by considering either the NCA or MCA instead of the EMA when computing the expected time to reach N monomers of type 2. In all plots, $n_L = 10,000$ and $K_{d,1} = K_{d,2} = 1000$.

References

- Santos, S.D.; Verveer, P.J.; Bastiaens, P.I. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat. Cell Biol.* **2007**, *9*, 324. [[CrossRef](#)] [[PubMed](#)]
- Katoh, M.; Katoh, M. FGF signaling network in the gastrointestinal tract. *Int. J. Oncol.* **2006**, *29*, 163–168. [[CrossRef](#)] [[PubMed](#)]
- Ernst, M.; Jenkins, B.J. Acquiring signalling specificity from the cytokine receptor gp130. *Trends Genet.* **2004**, *20*, 23–32. [[CrossRef](#)] [[PubMed](#)]
- Chen, S.; Guo, X.; Imarenezor, O.; Imoukhuede, P. Quantification of VEGFRs, NRP1, and PDGFRs on endothelial cells and fibroblasts reveals serum, intra-family ligand, and cross-family ligand regulation. *Cell. Mol. Bioeng.* **2015**, *8*, 383–403. [[CrossRef](#)]
- Hackel, P.O.; Zwick, E.; Prenzel, N.; Ullrich, A. Epidermal growth factor receptors: critical mediators of multiple receptor pathways. *Curr. Opin. Cell Biol.* **1999**, *11*, 184–189. [[CrossRef](#)]
- Weddell, J.C.; Imoukhuede, P.I. Integrative meta-modeling identifies endocytic vesicles, late endosome and the nucleus as the cellular compartments primarily directing RTK signaling. *Integr. Biol.* **2017**, *9*, 464–484. [[CrossRef](#)]
- Riese, D.J.; Stern, D.F. Specificity within the EGF family/ErbB receptor family signaling network. *Bioessays* **1998**, *20*, 41–48. [[CrossRef](#)]
- López-García, M.; Nowicka, M.; Fearnley, G.W.; Ponnambalam, S.; Lythe, G.; Molina-París, C. Performance measures in stochastic processes and the matrix-analytic approach. In *Quantitative Biology: Theory, Computational Methods, and Models*; Springer: Cham, Switzerland, 2018; pp. 379–402.
- Starbuck, C.; Lauffenburger, D.A. Mathematical model for the effects of epidermal growth factor receptor trafficking dynamics on fibroblast proliferation responses. *Biotechnol. Progress* **1992**, *8*, 132–143. [[CrossRef](#)]
- Cross, M.J.; Dixelius, J.; Matsumoto, T.; Claesson-Welsh, L. VEGF-receptor signal transduction. *Trends Biochem. Sci.* **2003**, *28*, 488–494. [[CrossRef](#)]
- Shibuya, M. Differential roles of vascular endothelial growth factor receptor-1 and receptor-2 in angiogenesis. *BMB Rep.* **2006**, *39*, 469–478. [[CrossRef](#)]
- Dikov, M.M.; Ohm, J.E.; Ray, N.; Tchekneva, E.E.; Burlison, J.; Moghanaki, D.; Nadaf, S.; Carbone, D.P. Differential roles of vascular endothelial growth factor receptors 1 and 2 in dendritic cell differentiation. *J. Immunol.* **2005**, *174*, 215–222. [[CrossRef](#)] [[PubMed](#)]

13. Kozer, N.; Barua, D.; Orchard, S.; Nice, E.C.; Burgess, A.W.; Hlavacek, W.S.; Clayton, A.H. Exploring higher-order EGFR oligomerisation and phosphorylation—A combined experimental and theoretical approach. *Mol. bioSyst.* **2013**, *9*, 1849–1863. [[CrossRef](#)] [[PubMed](#)]
14. Lauffenburger, D.A.; Linderman, J.J. *Receptors: Models for Binding, Trafficking, and Signaling*; Oxford University Press on Demand: Oxford, UK, 1996.
15. Waters, C.M.; Oberg, K.C.; Carpenter, G.; Overholser, K.A. Rate constants for binding, dissociation, and internalization of EGF: Effect of receptor occupancy and ligand concentration. *Biochemistry* **1990**, *29*, 3563–3569. [[CrossRef](#)] [[PubMed](#)]
16. Klein, P.; Mattoon, D.; Lemmon, M.A.; Schlessinger, J. A structure-based model for ligand binding and dimerization of EGF receptors. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 929–934. [[CrossRef](#)]
17. Blinov, M.L.; Faeder, J.R.; Goldstein, B.; Hlavacek, W.S. A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. *Biosystems* **2006**, *83*, 136–151. [[CrossRef](#)]
18. Alarcón, T.; Page, K.M. Stochastic models of receptor oligomerization by bivalent ligand. *J. R. Soc. Interface* **2006**, *3*, 545–559. [[CrossRef](#)]
19. Currie, J.; Castro, M.; Lythe, G.; Palmer, E.; Molina-París, C. A stochastic T cell response criterion. *J. R. Soc. Interface* **2012**, *9*, 2856–2870. [[CrossRef](#)]
20. Castro, M.; van Santen, H.M.; Férez, M.; Alarcón, B.; Lythe, G.; Molina-París, C. Receptor pre-clustering and T cell responses: Insights into molecular mechanisms. *Front. Immunol.* **2014**, *5*, 132. [[CrossRef](#)]
21. Bel, G.; Munsky, B.; Nemenman, I. The simplicity of completion time distributions for common complex biochemical processes. *Phys. Biol.* **2009**, *7*, 016003. [[CrossRef](#)]
22. Resat, H.; Ewald, J.A.; Dixon, D.A.; Wiley, H.S. An integrated model of epidermal growth factor receptor trafficking and signal transduction. *Biophys. J.* **2003**, *85*, 730–743. [[CrossRef](#)]
23. Gurevich, K.G.; Agutter, P.S.; Wheatley, D.N. Stochastic description of the ligand—Receptor interaction of biologically active substances at extremely low doses. *Cell. Signal.* **2003**, *15*, 447–453. [[CrossRef](#)]
24. López-García, M.; Nowicka, M.; Bendtsen, C.; Lythe, G.; Ponnambalam, S.; Molina-París, C. Stochastic models of the binding kinetics of VEGF-A to VEGFR1 and VEGFR2 in endothelial cells. *arXiv* **2016**, arXiv:1606.07269.
25. Allen, L.J. *An Introduction to Stochastic Processes with Applications to Biology*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2010.
26. Reuter, G. Competition processes. In Proceedings of the 4th Berkeley Symposium Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1961, Volume 2, pp. 421–430.
27. Iglehart, D.L. Multivariate competition processes. *Ann. Math. Stat.* **1964**, *35*, 350–361. [[CrossRef](#)]
28. Latouche, G.; Ramaswami, V.; Kulkarni, V. Introduction to matrix analytic methods in stochastic modeling. *J. Appl. Mathe. Stoch. Anal.* **1999**, *12*, 435–436.
29. López-García, M.; Nowicka, M.; Bendtsen, C.; Lythe, G.; Ponnambalam, S.; Molina-París, C. Quantifying the phosphorylation timescales of receptor–ligand complexes: A Markovian matrix-analytic approach. *Open Biol.* **2018**, *8*, 180126. [[CrossRef](#)]
30. De la Higuera, L.; López-García, M.; Lythe, G.; Molina-París, C. IL-2 Stimulation of Regulatory T Cells: A Stochastic and Algorithmic Approach. In *Modeling Cellular Systems*; Springer: Cham, Switzerland, 2017; pp. 81–105.
31. Economou, A.; Gómez-Corral, A.; López-García, M. A stochastic SIS epidemic model with heterogeneous contacts. *Physica A Stat. Mech. Appl.* **2015**, *421*, 78–97. [[CrossRef](#)]
32. López-García, M. Stochastic descriptors in an SIR epidemic model for heterogeneous individuals in small networks. *Math. Biosci.* **2016**, *271*, 42–61. [[CrossRef](#)]
33. Steinfeld, J.I.; Francisco, J.S.; Hase, W.L. *Chemical Kinetics and Dynamics*; Prentice Hall: Englewood Cliffs, NJ, USA, 1989; Volume 3.
34. Gillespie, D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **1976**, *22*, 403–434. [[CrossRef](#)]
35. Stroock, D.W. *An Introduction to Markov Processes*; Springer Science & Business Media: Berlin, Germany, 2013; Volume 230.
36. Karlin, S. *A First Course in Stochastic Processes*; Academic Press: Cambridge, MA, USA, 2014.
37. Weddell, J.C.; Imoukhuede, P. Quantitative characterization of cellular membrane-receptor heterogeneity through statistical and computational modeling. *PLOS ONE* **2014**, *9*, e97271. [[CrossRef](#)]

38. Roepstorff, K.; Grandal, M.V.; Henriksen, L.; Knudsen, S.L.J.; Lerdrup, M.; Grøvdal, L.; Willumsen, B.M.; Van Deurs, B. Differential effects of EGFR ligands on endocytic sorting of the receptor. *Traffic* **2009**, *10*, 1115–1127. [[CrossRef](#)]
39. Schmidt, A.M.; Du Yan, S.; Yan, S.F.; Stern, D.M. The biology of the receptor for advanced glycation end products and its ligands. *Biochim. Biophys. Acta (BBA) Mol. Cell Res.* **2000**, *1498*, 99–111. [[CrossRef](#)]
40. Ma, B.; Shatsky, M.; Wolfson, H.J.; Nussinov, R. Multiple diverse ligands binding at a single protein site: A matter of pre-existing populations. *Protein Sci.* **2002**, *11*, 184–197. [[CrossRef](#)] [[PubMed](#)]
41. Gómez-Corral, A.; López-García, M. Extinction times and size of the surviving species in a two-species competition process. *J. Math. Biol.* **2012**, *64*, 255–289. [[CrossRef](#)]
42. Gómez-Corral, A.; López-García, M. On the number of births and deaths during an extinction cycle, and the survival of a certain individual in a competition process. *Comput. Math. Appl.* **2012**, *64*, 236–259. [[CrossRef](#)]
43. Gómez-Corral, A.; López-García, M. Lifetime and reproduction of a marked individual in a two-species competition process. *App. Math. Compu.* **2015**, *264*, 223–245. [[CrossRef](#)]
44. Wedagedera, J.R.; Burroughs, N.J. T-cell activation: A queuing theory analysis at low agonist density. *Biophys. J.* **2006**, *91*, 1604–1618. [[CrossRef](#)]
45. Wedagedara, J. Some large deviation estimates for an Erlang queue of phosphorylated T-cell receptors. *J. Natl. Sci. Found. Sri Lanka* **2011**, *39*, 03–12. [[CrossRef](#)]
46. Kulkarni, V.G. *Modeling and Analysis of Stochastic Systems*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2016.
47. Gaver, D.; Jacobs, P.; Latouche, G. Finite birth-and-death models in randomly changing environments. *Adv. Appl. Probab.* **1984**, *16*, 715–731. [[CrossRef](#)]
48. Gómez-Corral, A.; López-García, M. Perturbation analysis in finite LD-QBD processes and applications to epidemic models. *Numer. Linear Algebra Appl.* **2018**, *25*, e2160. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).