

Research



Cite this article: López-García M, Nowicka M, Bendtsen C, Lythe G, Ponnambalam S, Molina-París C. 2018 Quantifying the phosphorylation timescales of receptor–ligand complexes: a Markovian matrix-analytic approach. *Open Biol.* **8**: 180126.
<http://dx.doi.org/10.1098/rsob.180126>

Received: 19 July 2018

Accepted: 10 August 2018

Subject Area:

biochemistry/cellular biology/systems biology

Keywords:

stochastic model, receptor–ligand interaction, phosphorylation, time to signalling threshold, quasi-birth-and-death process

Author for correspondence:

C. Molina-París

e-mail: carmen@maths.leeds.ac.uk

[†]These authors contributed equally to this study.

Quantifying the phosphorylation timescales of receptor–ligand complexes: a Markovian matrix-analytic approach

M. López-García^{1,†}, M. Nowicka^{1,3,†}, C. Bendtsen⁴, G. Lythe¹,
S. Ponnambalam² and C. Molina-París¹

¹Department of Applied Mathematics, School of Mathematics, and ²Endothelial Cell Biology Unit, School of Molecular and Cellular Biology, University of Leeds, LS2 9JT Leeds, UK

³Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY, USA

⁴Quantitative Biology, Discovery Sciences, IMED Biotech Unit, AstraZeneca, Cambridge Science Park, Milton Road, CB4 0WG Cambridge, UK

ML-G, 0000-0003-3833-8595; CM-P, 0000-0001-9828-6737

Cells interact with the extracellular environment by means of receptor molecules on their surface. Receptors can bind different ligands, leading to the formation of receptor–ligand complexes. For a subset of receptors, called receptor tyrosine kinases, binding to ligand enables sequential phosphorylation of intra-cellular residues, which initiates a signalling cascade that regulates cellular function and fate. Most mathematical modelling approaches employed to analyse receptor signalling are deterministic, especially when studying scenarios of high ligand concentration or large receptor numbers. There exist, however, biological scenarios where low copy numbers of ligands and/or receptors need to be considered, or where signalling by a few bound receptor–ligand complexes is enough to initiate a cellular response. Under these conditions stochastic approaches are appropriate, and in fact, different attempts have been made in the literature to measure the timescales of receptor signalling initiation in receptor–ligand systems. However, these approaches have made use of numerical simulations or approximations, such as moment-closure techniques. In this paper, we study, from an analytical perspective, the stochastic times to reach a given signalling threshold for two receptor–ligand models. We identify this time as an extinction time for a conveniently defined auxiliary absorbing continuous time Markov process, since receptor–ligand association/dissociation events can be analysed in terms of quasi-birth-and-death processes. We implement algorithmic techniques to compute the different order moments of this time, as well as the steady-state probability distribution of the system. A novel feature of the approach introduced here is that it allows one to quantify the role played by each kinetic rate in the timescales of signal initiation, and in the steady-state probability distribution of the system. Finally, we illustrate our approach by carrying out numerical studies for the vascular endothelial growth factor and one of its receptors, the vascular endothelial growth factor receptor of human endothelial cells.

1. Introduction

Cells interact with the extracellular environment by means of molecules located on their surface, referred to as *receptors*. These receptors interact with extracellular molecules called *ligands*, so that bound receptor–ligand complexes are formed, which eventually phosphorylate, initiating downstream signalling in the cytoplasm, and leading to a cellular response. Phosphorylation of a particular class of receptors, receptor tyrosine kinases (RTKs), occurs upon sequential activation of tyrosine residues located in the intra-cellular tail of

the receptors. In order to model cell behaviour regulated by receptor–ligand interactions, initial cell surface binding events and subsequent intra-cellular processes must be first quantified. Once this foundation is established, cellular behaviour can be analysed based on the number, state, and location of the molecules and complexes involved. The receptor population is involved in binding to the ligand, cross-linking to other receptors or membrane associated molecules, internalization, recycling, degradation and synthesis, broadly termed ‘trafficking’ events [1].

Detailed analyses of receptor–ligand interactions and phosphorylation kinetics on the cell membrane usually make use of mathematical models which ignore endocytosis (or internalization) events, and focus on the biochemical reactions taking place on the cell surface. For example, Starbuck *et al.* [2] consider a particular RTK, the epidermal growth factor receptor (EGFR), to study the role of epidermal growth factor (EGF) on mammalian fibroblasts. They argue that the receptor signal is generated at a rate proportional to the number of activated receptors present, so that the amount of phosphorylated ligand-bound dimeric complexes is directly related to the initiation of signalling cascades. Tan *et al.* [3] consider a mathematical model of pre-formed RTK receptor dimers, with instantaneous phosphorylation of ligand-bound dimeric complexes. However, phosphorylation is in fact a multi-step process, in which the different tyrosine domains of each receptor transfer phosphate (from ATP) onto side chains of specific tyrosine residues of the partner receptor, i.e. trans-autophosphorylation [4]. In Alarcón & Page [5], stochastic models of receptor oligomerization by a bivalent ligand are introduced to study the role of ligand-induced receptor cross-linking in cell activation. A particular feature of this study is that a small number of receptors is considered, making a stochastic approach more suitable than a deterministic one (see [6] for a comparison between deterministic and stochastic approaches for models of vascular endothelial growth factor receptors). In order to relate receptor–ligand dynamics on the cellular membrane to cell activation, the authors [5] introduce a threshold number of bound oligomers that need to be formed before a cellular response can take place. Once the stochastic process reaches this threshold, they study (by means of Gillespie simulations) the probability of staying above this threshold for a given time, $T = 10 k_{\text{off}}^{-1}$, which is identified with the time required for the activation of kinases and for the signalling pathway to be initiated [5].

In this study, we analyse receptor–ligand interactions and phosphorylation dynamics on the cell surface, to compute the time to reach a given signalling threshold [7], and the late time probability distribution of the system. To this end, we first introduce a mathematical model (instantaneous phosphorylation (IP) model), in which receptor monomers can bind a bivalent ligand, which allows a second receptor monomer to cross-link. This model is similar to Model 1 of Alarcón & Page [5]. However, rather than assuming that a fixed time interval above the threshold leads to a cellular response, we consider phosphorylation an intrinsic characteristic of the ligand cross-linked receptor dimers. In the IP model, ligand-bound receptor dimers are assumed to be instantaneously phosphorylated, so that the time to initiate the signalling cascade is identified with the time to reach a given threshold number of ligand-bound phosphorylated receptor dimers. This results in the analysis of a first-passage time or an absorption time in the theory of continuous time Markov processes. In

the second model, the delayed phosphorylation (DP) model, phosphorylation of ligand-bound receptor dimers is considered as an additional reaction in the system, and we also consider the possibility of ligand-bound receptor dimer dephosphorylation. We then compute the time to reach a given threshold number of phosphorylated ligand-bound receptor dimers in the DP model. Finally, the late time behaviour of the system is studied by analysing its stationary probability distribution.

As stated in Alarcón & Page [5], the analytical treatment of the multi-variate stochastic processes describing these biological receptor–ligand systems is typically extremely difficult, and numerical approaches, such as Gillespie simulations, are normally used instead. However, it is still possible to carry out an analytical study of these processes without the need to solve the corresponding master equation. Here, we do so by making use of a matrix-analytic technique and by considering a number of stochastic descriptors, conveniently defined in the spirit of Alarcón & Page [5]. This matrix-analytic approach, which has its origins in the seminal work by Neuts [8], allows us to study the stochastic descriptors of interest for moderate numbers of ligands and receptors in an *exact* way, as discussed in §2. Matrix-analytic techniques have historically been developed in the context of Queueing Theory [9]. However, more recently, they have been applied in Mathematical Biology [10–12].

We illustrate our methods by considering a receptor–ligand interaction involving vascular endothelial growth factors (VEGFs) and receptors (VEGFRs) in human endothelial cells. VEGFs are a family of bivalent ligands consisting of mammalian and virus-encoded members. The first member discovered was VEGF-A [13], which occurs in different isoforms of varying lengths. Mounting evidence suggests that the various isoforms are involved in diverse cellular responses [4]. VEGFs specifically bind to three type V RTKs, VEGFR1, VEGFR2 and VEGFR3, as well as co-receptors, such as neuropilins. In physiological conditions, the vascular endothelium expresses VEGFR1 and VEGFR2, whereas the lymphatic endothelium expresses VEGFR2 and VEGFR3 [14]. Each receptor has an extracellular domain for binding ligand, a trans-membrane domain and an intra-cellular or cytoplasmic domain [1]. Like many other RTKs, VEGFRs normally require dimerization to become activated: once VEGF binds to VEGFRs, the intra-cellular domains become activated by auto-phosphorylation and start cascades of intra-cellular enzymatic reactions [4]. We aim to develop a new quantitative study of receptor–ligand interaction and phosphorylation kinetics to aid our understanding of processes such as angiogenesis and vasculogenesis.

The paper is organized as follows. In §2, two different stochastic models are introduced to describe the association and dissociation dynamics of ligand-bound receptor monomers and dimers on the cell surface. The models include instantaneous phosphorylation or phosphorylation as an additional reaction. Matrix-analytic techniques are applied (for further details about these techniques, see appendices B and C) to study a number of stochastic descriptors of interest to the system, making use of an auxiliary absorbing continuous time Markov process. A particular feature of this method is that a sensitivity analysis (described in appendix D) to quantify the effect of association, dissociation and phosphorylation rates on the stochastic descriptors can be carried out. In §3.1, parameter estimation is carried out following arguments first described in Lauffenburger & Linderman [1], and applied to obtain the

kinetic rates of the receptor–ligand system of interest (VEGFR2 and VEGF-A, respectively) from the physiological parameters given in §3.2. Finally, numerical results are presented in §3.3 and §3.4, followed by a discussion in §4. The notation used in the paper is introduced in appendix A.

2. Stochastic models

In this section, we introduce two different stochastic models for the interaction of a surface receptor and a bivalent ligand (see §3). The bivalent ligand can bind a receptor monomer, creating a bound monomeric complex. The free site of the ligand in a bound monomeric complex can then bind to a second receptor monomer, while these molecules diffuse on the cell surface. This leads to a bound dimeric complex, consisting of two receptors bound to a bivalent ligand.

In our models, receptor dimerization is ligand-induced, as the dimeric VEGF-A ligand binds and recruits two receptor monomers into a single complex (cross-linking). We thus assume that two monomeric and free receptors are not able to create a *pre-dimer* in the absence of ligand (ligand-induced dimerization or LID [15, LID model]). We note that the consideration of receptor pre-dimerization in the model does not significantly change the dynamics of the process, especially for low ligand concentrations [15], as considered here. In some instances, and for highly saturated situations, the existence of pre-dimers may alter the dynamics of the system (see, for example, MacGabhann & Popel [15, Figs. 2 and 3] for details). On the other hand, there is experimental support for the following hypothesis: free VEGFR2 is observed (electron microscopy) in monomeric form on the cell surface [16].

Once ligand-bound dimeric complexes are formed, their activation leads to the initiation of a signalling pathway. From a biological perspective, this activation is usually the result of a sequence of phosphorylation events, involving different tyrosine residues on the intra-cellular tails of the receptors forming the dimer. From a mathematical perspective, this sequence of events is usually neglected by considering instantaneous phosphorylation [5,13]. This is described in §2.1, where the IP model is described. However, we also consider an extension of this model in §2.2, the DP model, where the phosphorylation of ligand-bound dimeric complexes is considered as an additional reaction. We refer the reader to MacGabhann & Popel [15] for a brief discussion on the importance of including phosphorylation, and to Bel *et al.* [17] for a discussion of the conditions under which the sequence of phosphorylation events can be treated as a single reaction.

For the IP and DP models, the aim in §2.1 and 2.2, as well as appendices B and C, is to compute the time to reach a given signalling threshold, where the amount of signalling in the process is identified with the number of phosphorylated (either instantaneously or not) complexes at any given time. Moreover, the steady-state distribution of the system is also computed. Finally, a sensitivity analysis of both models is carried out in appendix D, to quantify how the association, dissociation, phosphorylation and de-phosphorylation rates affect the dynamics of the receptor–ligand system.

The study of the number of ligand-bound monomeric, non-phosphorylated and phosphorylated ligand-bound dimeric molecules on the cell surface over time can be viewed as the

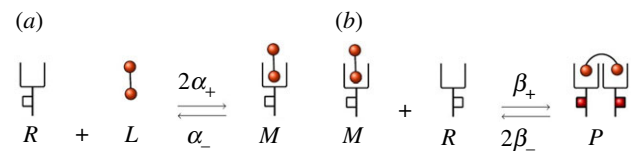


Figure 1. Reactions of the IP model. (a) Association and dissociation of bound monomers (M). (b) Association and dissociation of bound dimers (P), which instantaneously phosphorylate (represented by red squares as phosphorylated residues in the intra-cellular tail of the receptors).

analysis of the transient behaviour of a specific multi-variate Markov process, a problem which, in general, is not solvable in closed form [18]. Therefore, one typically carries out Gillespie simulations [19], or applies moment-closure techniques [20,21] to deal with the master equation of the Markov process under study. In this study, and for the models considered in §2.1 and 2.2, we apply alternative methods, which allow us to analyse, in an exact way, the quantities of interest mentioned above. In particular, by considering the time to reach a given signalling threshold as a continuous random variable, and by conveniently structuring the space of states of the continuous time Markov processes under study, we identify this time as the absorption time in an auxiliary absorbing continuous time Markov process. We compute the Laplace–Stieltjes transforms of this random variable, as well as the steady-state probabilities, by making use of first-step and matrix-analytic arguments. A novel local sensitivity analysis for the Markov processes considered is adapted and applied here by generalizing arguments from Caswell [22] (see also [23]). This analysis allows us to quantify how the stochastic descriptors considered in §3.3, time to signalling threshold and steady-state probability distribution, are affected by the association, dissociation, phosphorylation and de-phosphorylation rates.

2.1. IP model: instantaneous phosphorylation

In this section, we consider a model of a bivalent ligand that can bind a free receptor to form a bound monomer (or M complex). Receptors can diffuse on the cell surface, so that eventually a free receptor can bind an extracellular ligand to form a bound monomer M . This complex in turn can further engage a second receptor to form a ligand-bound and cross-linked receptor dimer (or P complex). Once a P complex is formed, it is instantaneously phosphorylated, so that P complexes on the plasma membrane initiate signalling, in the spirit of Starbuck *et al.* [2] and Alarcón & Page [5]. Ligand-bound monomers and dimers can dissociate. We assume that de-phosphorylation of P takes place when cross-linked receptor dimers also dissociate. In this scenario, four possible reactions can occur with different association and dissociation rates as shown in figure 1.

In what follows, we consider an environment with constant number, n_R and n_L , of receptors and ligands, spatially well-mixed on the cell surface and in the extracellular space, respectively. We are interested in the number of M and P complexes on the cell surface as a function of time, which we model using a stochastic approach: as a continuous time Markov chain (CTMC) $\mathcal{X} = \{\mathbf{X}(t) = (M(t), P(t)): t \geq 0\}$, where $M(t)$ and $P(t)$ represent the number of M and P complexes, respectively, at time t . We note that, if we define the random variables $R(t)$ and $L(t)$ as the numbers of free receptors and ligands,

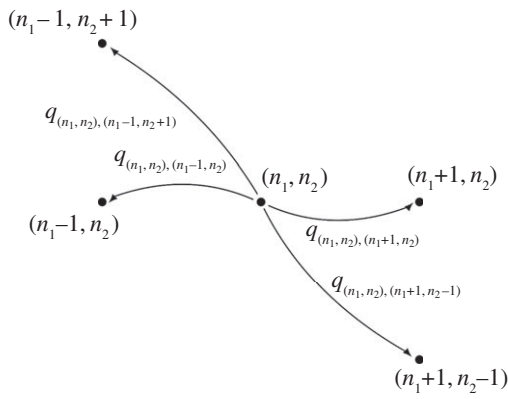


Figure 2. Transition diagram for the IP model (process \mathcal{X}).

respectively, at time $t \geq 0$, it is clear that $R(t) = n_R - M(t) - 2P(t)$ and $L(t) = n_L - M(t) - P(t)$, for all $t \geq 0$. Then, $R(t)$ and $L(t)$ are implicitly analysed in \mathcal{X} and do not need to be explicitly considered in the CTMC. We need to impose the conditions $M(t), P(t) \geq 0$ and, from the previous comments, we have

$$L(t) \geq 0 \Rightarrow M(t) + P(t) \leq n_L$$

and $R(t) \geq 0 \Rightarrow M(t) + 2P(t) \leq n_R$,

for all $t \geq 0$, which specify the state space \mathcal{S} of \mathcal{X} . Specifically, we note that given $(M(t), P(t)) = (n_1, n_2)$ at some time $t \geq 0$, then

- if $2n_L \leq n_R$: $n_1 + n_2 \leq n_L \Rightarrow n_1 + 2n_2 \leq n_R$ and
- if $n_R \leq n_L$: $n_1 + 2n_2 \leq n_R \Rightarrow n_1 + n_2 \leq n_L$

so that three different specifications of the state space \mathcal{S} are obtained, depending on the particular values of n_R and n_L . In particular:

- if $2n_L \leq n_R$, then $\mathcal{S} = \{(n_1, n_2) \in (\mathbb{N} \cup \{0\})^2: n_1 + n_2 \leq n_L\}$,
- if $n_R < 2n_L < 2n_R$, then $\mathcal{S} = \{(n_1, n_2) \in (\mathbb{N} \cup \{0\})^2: n_1 + n_2 \leq n_L, n_1 + 2n_2 \leq n_R\}$ and
- if $n_R \leq n_L$, then $\mathcal{S} = \{(n_1, n_2) \in (\mathbb{N} \cup \{0\})^2: n_1 + 2n_2 \leq n_R\}$.

Although we can deal with each of these cases, without loss of generality, we focus here on the first one, $2n_L \leq n_R$, since these are the physiological conditions for the receptor–ligand system analysed in §3. Thus, the stochastic process \mathcal{X} is defined over $\mathcal{S} = \{(n_1, n_2) \in (\mathbb{N} \cup \{0\})^2: n_1 + n_2 \leq n_L\}$. From figure 1, it is clear that transitions from states in the interior of \mathcal{S} , that is, from states $(n_1, n_2) \in \mathbb{N}^2$ with $n_1 + n_2 < n_L$, can take place to four adjacent states as shown in figure 2. Transitions for states within the boundary of \mathcal{S} are obtained in a similar way by discarding those transitions that leave \mathcal{S} .

Transitions between states in our CTMC are governed by the infinitesimal transition rates $q_{(n_1, n_2), (n'_1, n'_2)}$, with $(n_1, n_2), (n'_1, n'_2) \in \mathcal{S}$. These infinitesimal transition rates are obtained by mass action kinetics, and by the fact that if the process is in state (n_1, n_2) at a given time, there are $(n_L - n_1 - n_2)$ free ligands and $(n_R - n_1 - 2n_2)$ free receptors available. The formation of M complexes depends on the number of free receptors and ligands, and their dissociation only depends on the number of M complexes. A similar analysis can be made for P complexes. Finally, we note that the formation of M complexes and dissociation of P complexes can take place

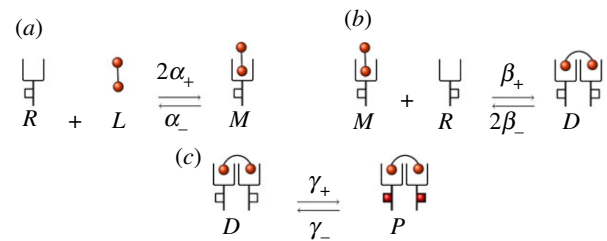


Figure 3. Reactions of the DP model. (a) Association and dissociation of bound monomers (M). (b) Association and dissociation of non-phosphorylated bound dimers (D). (c) Phosphorylation and de-phosphorylation of phosphorylated bound dimers (P).

with any of the two available binding sites of the ligand. Then, the specific values of the non-null infinitesimal transition rates are given by

$$q_{(n_1, n_2), (n'_1, n'_2)} = \begin{cases} 2\alpha_+(n_R - n_1 - 2n_2) & \text{if } (n'_1, n'_2) = (n_1 + 1, n_2), \\ \alpha_- n_1 & \text{if } (n'_1, n'_2) = (n_1 - 1, n_2), \\ \beta_+ n_1 (n_R - n_1 - 2n_2) & \text{if } (n'_1, n'_2) = (n_1 - 1, n_2 + 1), \\ 2\beta_- n_2 & \text{if } (n'_1, n'_2) = (n_1 + 1, n_2 - 1), \end{cases} \quad (2.1)$$

where α_+ , α_- , β_+ and β_- are positive constants representing the association and dissociation rates for M and P complexes, respectively.

For this model, the focus in §3.3 is on several summary statistics (or stochastic descriptors) that allow one to study the timescales for signal initiation on the cell membrane, as well as the late time behaviour of the system, and to carry out a local sensitivity analysis to test how these summary statistics depend on the different parameters (e.g. kinetic rates) of the model. An efficient matrix-oriented analysis of these summary statistics, for the IP model, can be found in appendix B.

2.2. DP model: delayed phosphorylation

In the previous section, the P complexes were instantaneously phosphorylated. Here we relax this requirement and include phosphorylation as an additional reaction (figure 3). We note that, in the DP model presented in figure 3, dissociation of phosphorylated receptors can only occur after their de-phosphorylation. One may alternatively consider that dissociation can occur due to ligand unbinding to one of the receptors, even if de-phosphorylation has not occurred yet. For this case, a similar analysis to the one carried out in this section could be developed, and bound phosphorylated monomers should be incorporated as a new molecular *species*. Numerical results for the VEGFR2 receptor and VEGF-A ligand system (§3), including this additional molecular species and not reported here, show similar qualitative dynamics to the simpler model considered in this section.

In what follows, we adapt the arguments of the previous section to the DP model. This not only allows us to evaluate the relevance of phosphorylation as an independent reaction (with numerical results presented in §3), but also serves as an example of how to include new reactions in this type of stochastic model, while adapting the matrix-analytic arguments.

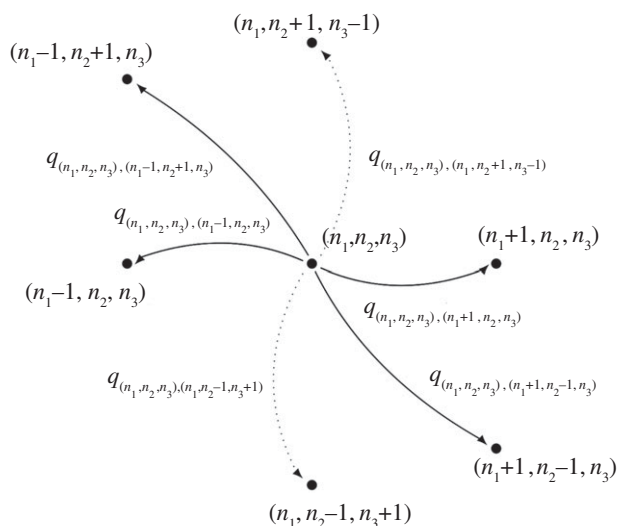


Figure 4. Transition diagram for the DP model (process $\hat{\mathcal{X}}$).

In brief, we consider the CTMC $\hat{\mathcal{X}} = \{\hat{\mathbf{X}}(t) = (\hat{M}(t), \hat{D}(t), \hat{P}(t)): t \geq 0\}$, where

$\hat{M}(t)$ = ‘number of M complexes at time t ’,

$\hat{D}(t)$ = ‘number of D complexes at time t ’

and $\hat{P}(t)$ = ‘number of P complexes at time t ’,

for all $t \geq 0$, where D complexes refer to non-phosphorylated bound dimers and P to phosphorylated ones. From the reactions in figure 3, it is clear that for all $t \geq 0$

$$\hat{M}(t) + \hat{D}(t) + \hat{P}(t) \leq n_L$$

$$\text{and } \hat{M}(t) + 2\hat{D}(t) + 2\hat{P}(t) \leq n_R,$$

and, by assuming as previously, that $2n_L \leq n_R$, it is easy to show that

$$\hat{M}(t) + \hat{D}(t) + \hat{P}(t) \leq n_L, \forall t \geq 0$$

$$\Rightarrow \hat{M}(t) + 2\hat{D}(t) + 2\hat{P}(t) \leq n_R, \forall t \geq 0,$$

so that $\hat{\mathcal{X}}$ is defined over $\hat{\mathcal{S}} = \{(n_1, n_2, n_3) \in (\mathbb{N} \cup \{0\})^3: n_1 + n_2 + n_3 \leq n_L\}$.

From figure 3, the transition diagram can be obtained (figure 4), where the non-null infinitesimal transition rates are obtained in a manner analogously to (2.1). In particular, we have

$$q_{(n_1, n_2, n_3), (n'_1, n'_2, n'_3)} = \begin{cases} 2\alpha_+(n_R - n_1 - 2n_2 - 2n_3)(n_L - n_1 - n_2 - n_3), & \text{if } (n'_1, n'_2, n'_3) = (n_1 + 1, n_2, n_3), \\ \alpha_- n_1, & \text{if } (n'_1, n'_2, n'_3) = (n_1 - 1, n_2, n_3), \\ \beta_+ n_1 (n_R - n_1 - 2n_2 - 2n_3), & \text{if } (n'_1, n'_2, n'_3) = (n_1 - 1, n_2 + 1, n_3), \\ 2\beta_- n_2, & \text{if } (n'_1, n'_2, n'_3) = (n_1 + 1, n_2 - 1, n_3), \\ \gamma_+ n_2, & \text{if } (n'_1, n'_2, n'_3) = (n_1, n_2 - 1, n_3 + 1), \\ \gamma_- n_3, & \text{if } (n'_1, n'_2, n'_3) = (n_1, n_2 + 1, n_3 - 1), \end{cases}$$

where α_+ , α_- , β_+ , β_- , γ_+ and γ_- are positive constants representing the association, dissociation and phosphorylation rates for the complexes in figure 3. Similar summary statistics to those studied for the IP model, and analysed in §3.3, are analysed for the DP model in appendix C, by following a matrix-oriented approach.

3. The vascular endothelial growth factor receptor–ligand system

In this section, we illustrate the analytical work developed in the previous ones and the appendices, and focus on the interaction between VEGFR2 receptors and VEGF-A ligands on the surface of human umbilical vein endothelial cells (HUVECs), an interaction initiating signalling cascades that can cause diverse cellular responses, such as cell motility, division or death (i.e. apoptosis). We first develop, in §3.1,

a method to estimate the parameters α_+ , α_- , β_+ and β_- for the interaction between the VEGFR2 receptor and the VEGF-A ligand molecule. We do so by making use of the methods proposed by Lauffenburger and Linderman [1], where the transport mechanism of free ligand or free receptor is modelled by molecular diffusion, and where diffusive transport dominates convective transport caused by fluid motion at cellular and sub-cellular length scales [1,24].

The rates estimated in §3.1 depend on several physiological parameters, which are presented in §3.2. In §3.3, we analyse a number of stochastic descriptors of interest when the IP or the DP models are considered for this interaction. This allows us to study the impact of phosphorylation as a separate reaction in the process (delayed phosphorylation), to quantify timescales for signalling initiation under different ligand concentrations and to analyse the impact that each kinetic rate has in these stochastic descriptors. Finally, we investigate in §3.4 the effect that synthesis of new free

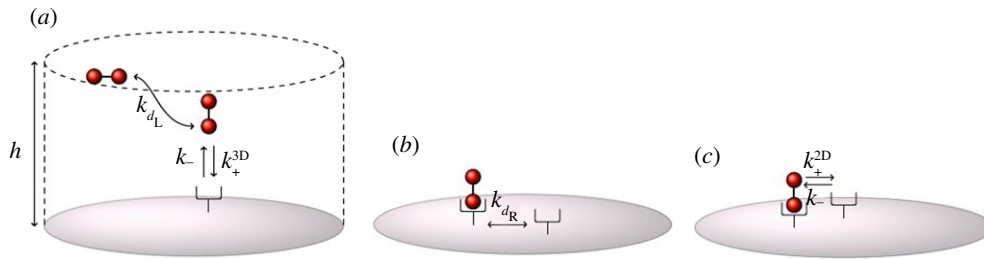


Figure 5. (a) Two-step binding and unbinding of receptor and ligand: k_{d_L} is the ligand transport rate, k_+^{3D} and k_- are the intrinsic binding and unbinding rates, respectively, and h is the characteristic length of the experimental volume. (b) Diffusive transport of surface receptor: k_{d_R} is the transport rate for both receptor R and bound monomer M . (c) Once in the reaction zone of M , R can bind with rate k_+^{2D} (which is a 2D version of k_+^{3D}) or unbind with rate k_- .

receptors on the cell surface, and internalization of bound complexes into endosomal compartments, can have on the molecular dynamics.

3.1. Estimation of association and dissociation rates

We estimate in §3.3 the parameters α_+ , α_- , β_+ and β_- (s^{-1}) for the binding and unbinding of the VEGFR2 receptor and its bivalent VEGF-A ligand. We consider a fraction, $0 < f < 1$, of a HUVEC, for computational reasons, and denote the receptor molecule by R and the ligand by L . Firstly, we set the dissociation rate $k_{off} = 1.32 \times 10^{-3} s^{-1}$ as reported in MacGabhann & Popel [15] for VEGFR2. From the equilibrium dissociation rate, K_d ($mm^{-3} mol$), given by $K_d = k_{off}/k_{on}$, it is possible then to obtain the biophysical binding rate, k_{on} ($mol^{-1} mm^3 s^{-1}$). Therefore, the transition rates α_+ and α_- of §2 are given by

$$\alpha_+ = \frac{k_{on}}{N_A f h s_c} \quad \text{and} \quad \alpha_- = k_{off},$$

where h (mm) is the characteristic length of the experimental volume, s_c (mm^2) is the total area of the cell surface and N_A (mol^{-1}) is Avogadro's number. In order to estimate the transition rates β_+ and β_- , we first note that the binding process between the receptor and the ligand, such as reaction (a) in figure 1, can be considered as a one-step process, with q_{on} ($mm^3 s^{-1}$) the association constant and q_{off} (s^{-1}) the dissociation constant. Constants q_{on} and q_{off} are related to the biophysical rates k_{on} and k_{off} as follows:

$$q_{on} = \frac{k_{on}}{N_A} \quad \text{and} \quad q_{off} = k_{off}.$$

However, these binding and unbinding events are in fact two-step processes [1,25–29]. In the first step, the ligand and the receptor simply encounter each other; that is, ligands diffuse into a sufficiently close proximity of the receptor to allow the chemical reaction step to occur. Let us define the ligand diffusion rate k_{d_L} ($mm^2 s^{-1}$), and the 3D reaction intrinsic rate k_+^{3D} ($mm^3 s^{-1}$). The mechanism of the reverse process is similar, so that the unbinding of the receptor and the ligand occurs with intrinsic dissociation rate k_- (s^{-1}) and the outward diffusion with transport rate k_{d_L} (figure 5a).

As mentioned earlier, we restrict our study to a fraction $0 < f < 1$ of the cell surface, so that the radius of this target surface is given by

$$r = \sqrt{\frac{n_R s_c}{\pi n_R^T}},$$

where n_R^T is the total number of receptors on the cell surface, and $n_R = f n_R^T$ is the number of receptors present on the target surface. We have assumed, thus, an homogeneous spatial

distribution of VEGFR2 on the cell surface [30,31], neglecting receptor clustering, which might be initiated upon ligand stimulation [32]. Under this assumption, the contributions of rates k_{d_L} , k_+^{3D} and k_- to the overall association and dissociation rates, q_{on} and q_{off} , respectively, are given by

$$q_{on} = \frac{k_{d_L} k_+^{3D}}{k_{d_L} + n_R k_+^{3D}} \quad \text{and} \quad q_{off} = \frac{k_{d_L} k_-}{k_{d_L} + n_R k_+^{3D}}, \quad (3.1)$$

where $k_{d_L} = 4\pi D_L r$, as shown elsewhere [1,25–28]. We note that q_{on} is a *per receptor* rate, as explained elsewhere [1,33]. A similar argument (figure 5b, c) applies when computing the rate of free receptor binding (k_c ($mm^2 s^{-1}$)) or unbinding (k_u (s^{-1})) to a monomer on the cell membrane [1], which occurs with rates

$$k_c = \frac{k_+^{2D} k_{d_R}}{k_{d_R} + k_+^{2D}} \quad \text{and} \quad k_u = \frac{k_- k_{d_R}}{k_{d_R} + k_+^{2D}}, \quad (3.2)$$

where the transport rate k_{d_R} ($mm^2 s^{-1}$) (figure 5b) is given by $k_{d_R} = 2\pi D / \log(w/b)$. The diffusion constant $D = D_R + D_M$ ($mm^2 s^{-1}$) is the sum of the diffusivities of the receptor and the bound monomer on the cell membrane (which are assumed to be the same $D_R = D_M$), b (mm) is the characteristic length of the receptor, and w (mm) is one-half the mean distance between receptors, given by

$$w = \sqrt{\frac{s_c}{\pi n_R^T}}.$$

We find k_+^{3D} and k_- from equation (3.1). Once k_+^{3D} is in hand, the intrinsic 3D binding rate allows to compute its 2D version, k_+^{2D} , as follows:

$$k_+^{2D} = \frac{k_+^{3D}}{\delta},$$

where δ (mm) is the cell membrane thickness, as suggested in Lauffenburger & Linderman [1]. Given k_+^{2D} , rate constants k_c and k_u can be found by means of equation (3.2). Finally, these rates, k_c and k_u , are related to β_+ and β_- , respectively, for the CTMCs considered in §2, as follows:

$$\beta_+ = \frac{k_c}{f s_c} \quad \text{and} \quad \beta_- = k_u.$$

3.2. Physiological parameters

All the rates of the IP and DP models (figures 1 and 3, respectively) used in §3.3 and §3.4 have been obtained following the approach described in §3.1, with physiological parameters taken from the literature. In particular, physiological parameters are given in table 1, and the specific rates for the IP and DP models are given in table 2. The equilibrium dissociation rate

Table 1. Physiological parameters.

physiological parameter	value	reference
endothelial cell surface area, s_c	10^{-3} mm^2	[15]
VEGF-A diffusion coefficient at 4°C , D_L	$5.2 \times 10^{-5} \text{ mm}^2 \text{ s}^{-1}$	[34]
VEGFR2 diffusion coefficient, D_R	$10^{-8} \text{ mm}^2 \text{ s}^{-1}$	[35]
VEGFR2 radius, b	$5 \times 10^{-7} \text{ mm}$	[5]
average membrane thickness of ECs, δ	10^{-4} mm	[36]
characteristic length of the experimental volume, h	1 mm	[15]
dissociation rate, k_{off}	$1.32 \times 10^{-3} \text{ s}^{-1}$	[15]
equilibrium dissociation rate, K_d for VEGFR2	$1.5 \times 10^{-16} \text{ mm}^{-3} \text{ mol}$	[15]
phosphorylation rate for D complexes, γ_+	$3.67 \times 10^{-3} \text{ s}^{-1}$	[1]
de-phosphorylation rate for P complexes, γ_-	$9.17 \times 10^{-4} \text{ s}^{-1}$	[1]

Table 2. Rates (in s^{-1}) for the IP and DP models, considering $f = 4\%$ of the cell surface. Note that parameters γ_+ and γ_- are not considered in the IP model.

reactions of the IP model	α_+	3.653×10^{-7}
$R + L \xrightleftharpoons[\alpha_-]{2\alpha_+} M$	α_-	1.320×10^{-3}
$M + R \xrightleftharpoons[2\beta_-]{\beta_+} P$	β_+	4.483×10^{-4}
reactions of the DP model	β_-	1.620×10^{-4}
$R + L \xrightleftharpoons[\alpha_-]{2\alpha_+} M$	γ_+	3.667×10^{-3}
$M + R \xrightleftharpoons[2\beta_-]{\beta_+} D$	γ_-	9.167×10^{-4}
$D \xrightleftharpoons[\gamma_-]{\gamma_+} P$		

for VEGF-A and VEGFR2 is equal to $K_d = 150 \text{ pM}$, as reported in MacGabhann & Popel [15]. This rate is consistent with previously reported values for *in silico* experiments [37], and agrees with experimentally determined ones [38–41].

We consider in this section the subset of endothelial cells, called human umbilical vein endothelial cells (HUVECs), which have been characterized to express (on average) 5800 VEGFR2s per cell [42]. We focus on 4% of the cell surface ($f = 0.04$) for computational reasons, so that in this area the total number of VEGFR2s is $n_R = 232$. For the IP and DP models, our numerical results should be considered exact, since they have been obtained making use of the analytical arguments described in the appendices.

3.3. Results

In this section (both for the IP and the DP models), we focus on two stochastic descriptors (or summary statistics) that allow one to study the timescales for signal initiation on the cell surface (in terms of phosphorylated dimers), as well as the late time behaviour of the system (in terms of the steady-state number of free receptors, monomers and dimers). In particular, we focus on

- (1) Starting in any state \mathbf{n} ($\mathbf{n} = (n_1, n_2) \in \mathcal{S}$ for the IP model, and $\mathbf{n} = (n_1, n_2, n_3) \in \hat{\mathcal{S}}$ for the DP model), the time $T_{\mathbf{n}}(N)$

to reach, for the first time, N phosphorylated dimers on the cell surface; that is, $T_{\mathbf{n}}(N) = \inf\{t \geq 0 : P(t) = N\}$ for the IP model, and $T_{\mathbf{n}}(N) = \inf\{t \geq 0 : \hat{P}(t) = N\}$ for the DP model.

- (2) The stationary probability distribution of the system, which does not depend on the initial conditions; that is, the probabilities $\pi_{(n_1, n_2)} = \lim_{t \rightarrow +\infty} P((M(t), P(t)) = (n_1, n_2))$ for the IP model and $\pi_{(n_1, n_2, n_3)} = \lim_{t \rightarrow +\infty} P((\hat{M}(t), \hat{D}(t), \hat{P}(t)) = (n_1, n_2, n_3))$ for the DP model.

We note that in this section we always consider initial states such that $(n_1, n_2) = (0, 0)$ and $(n_1, n_2, n_3) = (0, 0, 0)$. These initial conditions indicate that at time $t = 0$ (when ligand stimulation occurs), all receptors are in monomeric form. We report in appendices B and C a matrix-oriented approach to study these summary statistics for the IP and the DP model, respectively, and in appendix D, a matrix-oriented method to carry out a local sensitivity analysis of these summary statistics with respect to the model parameters (e.g. kinetic rates). This allows one to explore what the contribution is of each kinetic rate to a given stochastic descriptor.

3.3.1. Time to reach a signalling threshold

In figure 6, we plot $E[T_{(0,0)}(N)]$ (for the IP model) and $E[T_{(0,0,0)}(N)]$ (for the DP model), for values $0 \leq N \leq n_L$, where $n_L \in \{23, 58, 116\}$ is the number of ligands considered, which corresponds to the following ligand concentrations, $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}\}$. We note that these concentrations are similar to those reported in serum for healthy controls and cancer studies (see table I in [43]). The three different values of n_L correspond to 10%, 25% and 50% of n_R , the total number of VEGFR2 on the fraction of the cell surface considered. The number of ligands, thus, verifies the condition $2n_L \leq n_R$, assumed in the analysis of $T_{(0,0)}(N)$, as discussed in §2.1. $T_{(0,0)}(N)$ is the continuous random variable that represents the time to reach a total number, N , of phosphorylated bound dimers, P , given the initial state $(0, 0)$, in the IP model where instantaneous phosphorylation is considered (for details, see §2.1), while $T_{(0,0,0)}(N)$ is its DP model counterpart. Our results have been restricted to times up to 60 min, to describe the early time dynamics on the cell surface. The late time behaviour of the system will be analysed by means of its steady-state distribution. In figure 6, *solid* curves represent the values of $E[T_{(0,0)}(N)]$,

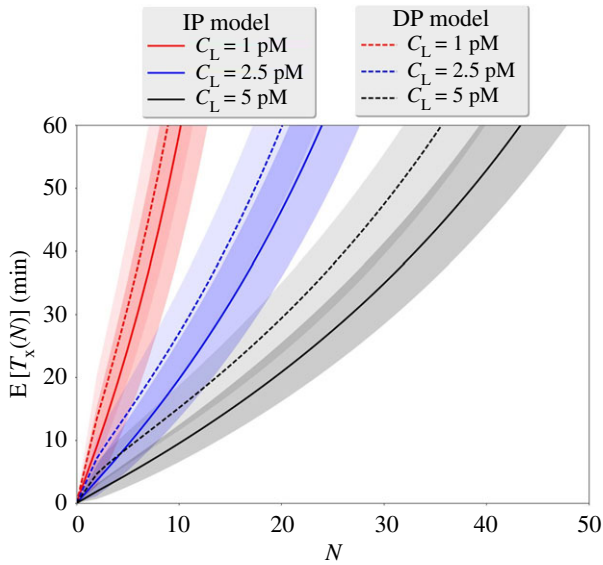


Figure 6. $E[T_x(N)]$ for (from left to right) ligand concentrations, $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}\}$, and for the IP model (solid curves) and the DP model (dashed curves). The initial state for the IP model is $\mathbf{x} = (0, 0)$ and for the DP model it is $\mathbf{x} = (0, 0, 0)$.

while dashed curves represent the values of $E[T_{(0,0)}(N)]$, obtained by means of algorithm 1 (see appendix B). Shaded areas have been obtained for both models by considering $E[T_x(N)] \pm SD[T_x(N)]$, where $SD[X]$ represents the standard deviation of the random variable X , also obtained from algorithm 1.

In figure 6, a monotonic behaviour is observed. For a fixed value of N in the IP model, $E[T_{(0,0)}(N)]$ is always smaller for larger ligand concentrations, c_L . Indeed, an increase in the amount of available ligand to bind receptors implies reaching the given signalling threshold (encoded by N) in a shorter time. The behaviour for $E[T_{(0,0)}(N)]$ is similar to that observed for $E[T_{(0,0)}(N)]$, so that the consideration of delayed phosphorylation in the DP model does not seem to qualitatively affect the main features of this descriptor. This can be explained as follows: the most likely fate of a bound monomer is to phosphorylate before its dissociation. However, the consideration of phosphorylation as an additional reaction delays the time to reach a given threshold N and every curve for the DP model is displaced to the left of its corresponding one for the IP model. For example, for $c_L = 1 \text{ pM}$, the mean time $E[T_{(0,0)}(N)]$ to reach a threshold $N = 5$ (20% of n_L) of phosphorylated bound dimers is approximately 25 min under the IP model. When the phosphorylation of bound complexes is explicitly considered (DP model), this mean time increases approximately up to 31 min.

3.3.2. Stationary probability distribution

The asymptotic behaviour of the curves shown in figure 6 is directly related to the maximum signalling threshold that is, in fact, reached by the process in short and intermediate time-scales. From a purely mathematical perspective, any state within \mathcal{S} (or $\hat{\mathcal{S}}$ in the DP model) is reached in the IP model (DP model) as $t \rightarrow +\infty$, since \mathcal{S} ($\hat{\mathcal{S}}$) is an irreducible finite class of states for the process \mathcal{X} ($\hat{\mathcal{X}}$). However, according to our numerical results, there exists a subset of (high) signalling thresholds that is not reached in practice by \mathcal{X} ($\hat{\mathcal{X}}$). This maximum signalling threshold is encoded in the steady-state

probability distribution of this process, which can be computed from algorithm 2 (see appendix B), and which measures the potential of the system to reach any signalling threshold at sufficiently late times, for different ligand concentrations.

In figure 7, the distribution of the number of (phosphorylated and non-phosphorylated) bound dimers at steady state, for the IP and the DP models, is plotted for different ligand concentrations, $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}\}$. For low ligand concentrations, nearly all the n_L available ligands are forming phosphorylated bound dimers in steady state. This is particularly the case in the IP model, where no non-phosphorylated bound dimers exist. In the DP model, a small number of non-phosphorylated bound dimers can be found in steady state. These non-phosphorylated bound dimers in steady state explain why the distribution of the number of phosphorylated bound dimers in steady state is displaced to the left when phosphorylation is considered as a separate reaction in the DP model, in comparison with the same distribution in the IP model.

3.3.3. Dynamics of the receptor–ligand system

We now complement our previous results by carrying out a number of Gillespie simulations of the models, so that the time course of the different random variables in our processes ($M(t)$, $P(t)$, $\hat{M}(t)$, $\hat{D}(t)$ and $\hat{P}(t)$) can be studied. In particular, we plot in figure 8 the mean plus and minus (shadowed area) the standard deviation of the variables of interest ($M(t)$ and $D(t)$ in the IP model, and $\hat{M}(t)$, $\hat{D}(t)$ and $\hat{P}(t)$ in the DP model). The time course has been generated by means of Gillespie simulations, where we have broadened the VEGF-A concentration range by considering $n_L \in \{0.1n_R, 0.25n_R, 0.5n_R, 10n_R, 50n_R, 100n_R, 250n_R, 625n_R, 1250n_R\}$, which approximately corresponds to concentrations $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}, 0.1 \text{ nM}, 0.5 \text{ nM}, 1 \text{ nM}, 2.5 \text{ nM}, 6.25 \text{ nM}, 12.5 \text{ nM}\}$. We note that for small ligand concentrations the number of bound dimers grows as the VEGF-A concentration increases. For concentrations $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}\}$ the steady state is not reached in the first 60 min of the numerical simulations (figures 7 and 8). However, higher concentrations result in a saturated scenario, where we obtain lower numbers of P complexes for ligand concentrations higher than $c_L \sim 2.5 \text{ nM}$. Thus, concentrations around $0.1 \text{ nM} - 2.5 \text{ nM}$ may be considered as optimum when only surface dynamics of phosphorylated bound dimers is of interest.

As mentioned above, for ligand concentrations around $c_L \in \{6.25 \text{ nM}, 12.5 \text{ nM}\}$, the system exhibits a reduction in the number of bound dimers, which is caused by the fast and early formation of monomeric bound complexes (figure 8). In fact, for both IP and DP models and when focusing on the formation of bound monomers as a function of time, we observe, under optimum ligand concentrations, a peak of monomeric complexes in the first 5 min, which is followed by a decrease to the steady-state values. The same early peak can be observed under these ligand concentrations for non-phosphorylated bound dimers in the DP model, which is followed by an increase in the number of phosphorylated bound dimers. For high ligand concentrations, the steady-state value for monomeric complexes increases, so that formation of bound dimers is effectively blocked. The inhibition of bound dimer formation at high ligand concentrations is intrinsically related to the ligand-induced-dimerization assumption, where the formation of free receptor pre-dimers is not allowed. However, if free receptor pre-dimers were to be considered, their effect

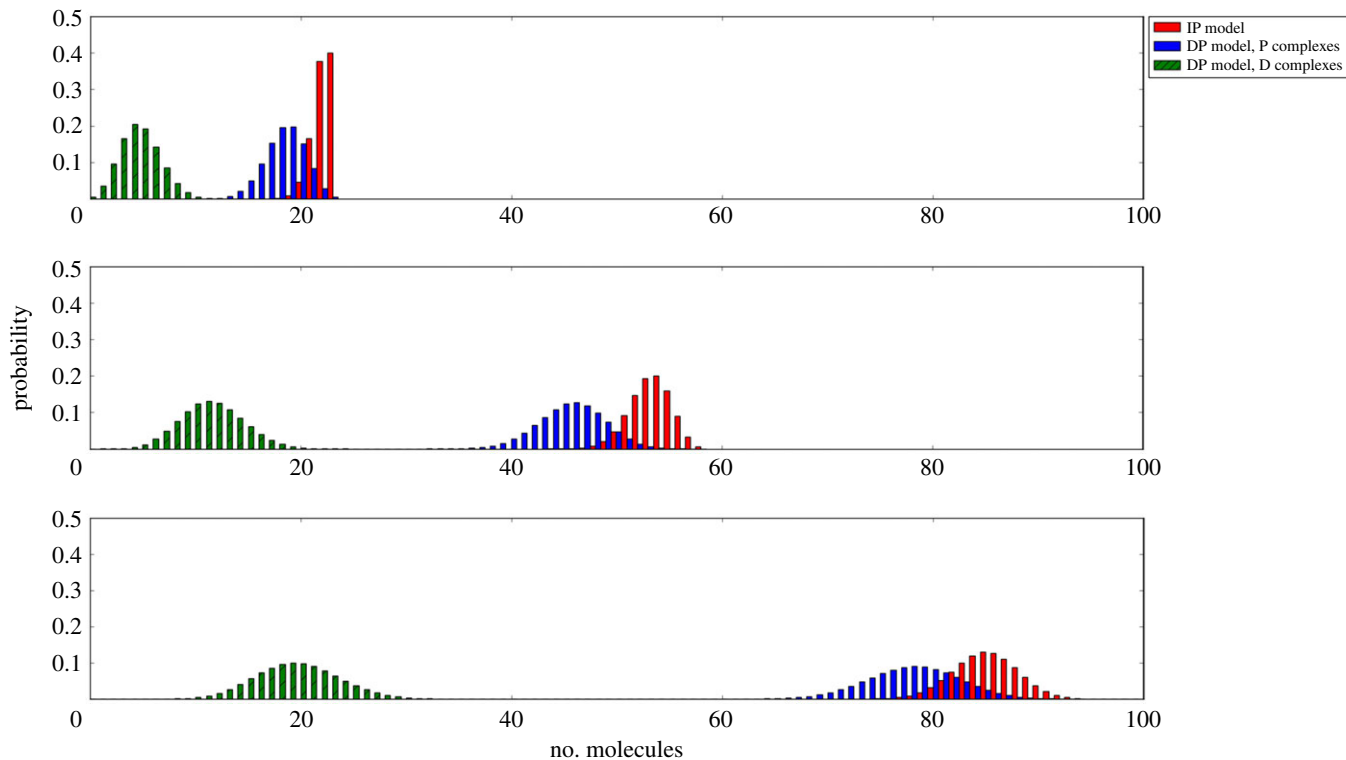


Figure 7. Probability distribution of the number of bound dimers in steady state for the processes \mathcal{X} (IP model, P bound dimers, red) and \mathcal{X} (DP model, D and P bound dimers, green and blue) for (from top to bottom) ligand concentrations, $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}\}$.

would be negligible for ligand concentrations below 1 nM [15], as our results in figure 8 also suggest.

3.3.4. Local sensitivity analysis

We study in this section the effect of the association, dissociation, phosphorylation and de-phosphorylation rates on the descriptors introduced, which can be estimated by means of the sensitivity analysis proposed in appendix D. In table 3, we present the elasticities (i.e. normalized derivatives) of the descriptors $E[T_{(0,0)}(N)]$, $E[T_{(0,0,0)}(N)]$, π_P and $\hat{\pi}_P$ (see appendices B and C), when N is chosen to be 25% of the total number of ligands n_L and for different ligand concentrations, c_L . As expected, the effect of each rate on any descriptor increases with increasing values of ligand concentration. It is also worth noting that the elasticities of the mean number of phosphorylated complexes in steady state are equal, with opposite sign, with respect to the association and dissociation rates (e.g. $(\partial\pi_P/\partial\alpha_+)/(\pi_P/\alpha_+) = -(\partial\pi_P/\partial\alpha_-)/(\pi_P/\alpha_-)$), which means that this variable only depends on the ratio of parameters: α_+/α_- , β_+/β_- and γ_+/γ_- . This can be easily understood since, from a deterministic perspective, the steady state corresponding to the DP model can be obtained as the solution of the following system of equations:

$$2 \frac{\alpha_+}{\alpha_-} R^* L^* - M^* = 0,$$

$$\frac{\beta_+}{\beta_-} M^* R^* - 2D^* = 0$$

$$\text{and } \frac{\gamma_+}{\gamma_-} D^* - P^* = 0,$$

which only depends on these parameter ratios. We also note that, according to the results of table 3, the rate α_+ plays an important role in all the descriptors. This can be explained as follows: once a ligand is 'destined' to form a bound

monomer complex, its most likely fate is to lead to a phosphorylation event before dissociation of the corresponding dimer occurs (see discussion in §4).

3.4. A study of receptor internalization and synthesis

It is well known that rapid internalization occurs for VEGFR2 following ligand binding and phosphorylation [39]. We briefly explore in this section how receptor synthesis and internalization events can have an impact on the molecular dynamics of the cell surface. In figure 9, we represent the IP and the DP models under the assumption that synthesis of new receptors, as well as internalization of free receptors, monomers and dimers, can also take place. We note that since modelling endosomal compartments is out of the scope of this paper, recycling events have not been explicitly considered in what follows: this would require tracking down the number of molecules in the different intra-cellular compartments, and thus, additional variables in the stochastic models. However, one can interpret the synthesis rate k_{syn} in figure 9 as an *insertion rate* [15], which implies a net contribution of new receptors on the cell surface, without having to specify whether these receptors have been truly synthesized and transported to the surface from the Golgi apparatus, or have been recycled to the surface from endosomal compartments. Since the parameter n_R is the basal (i.e. under no ligand stimulation) number of receptors on the cell surface, internalization and synthesis rates need to satisfy the condition $k_{syn} = n_R k_{int}$. Moreover, we set $k_{int} = 2.8 \times 10^{-4} \text{ s}^{-1}$ as previously determined [44], and consider that phosphorylated dimers can be internalized faster than non-phosphorylated ones [15,45], by setting $k_{int}^P = q k_{int}$ with $q \in \{1.0, 2.0, 5.0, 10.0\}$ (figure 9).

In figure 10, we plot analogous results to those of figure 8 for the models considered in figure 9 and values $q \in \{1, 2, 5, 10\}$. We focus here on the dynamics of phosphorylated ($P(t)$) in the

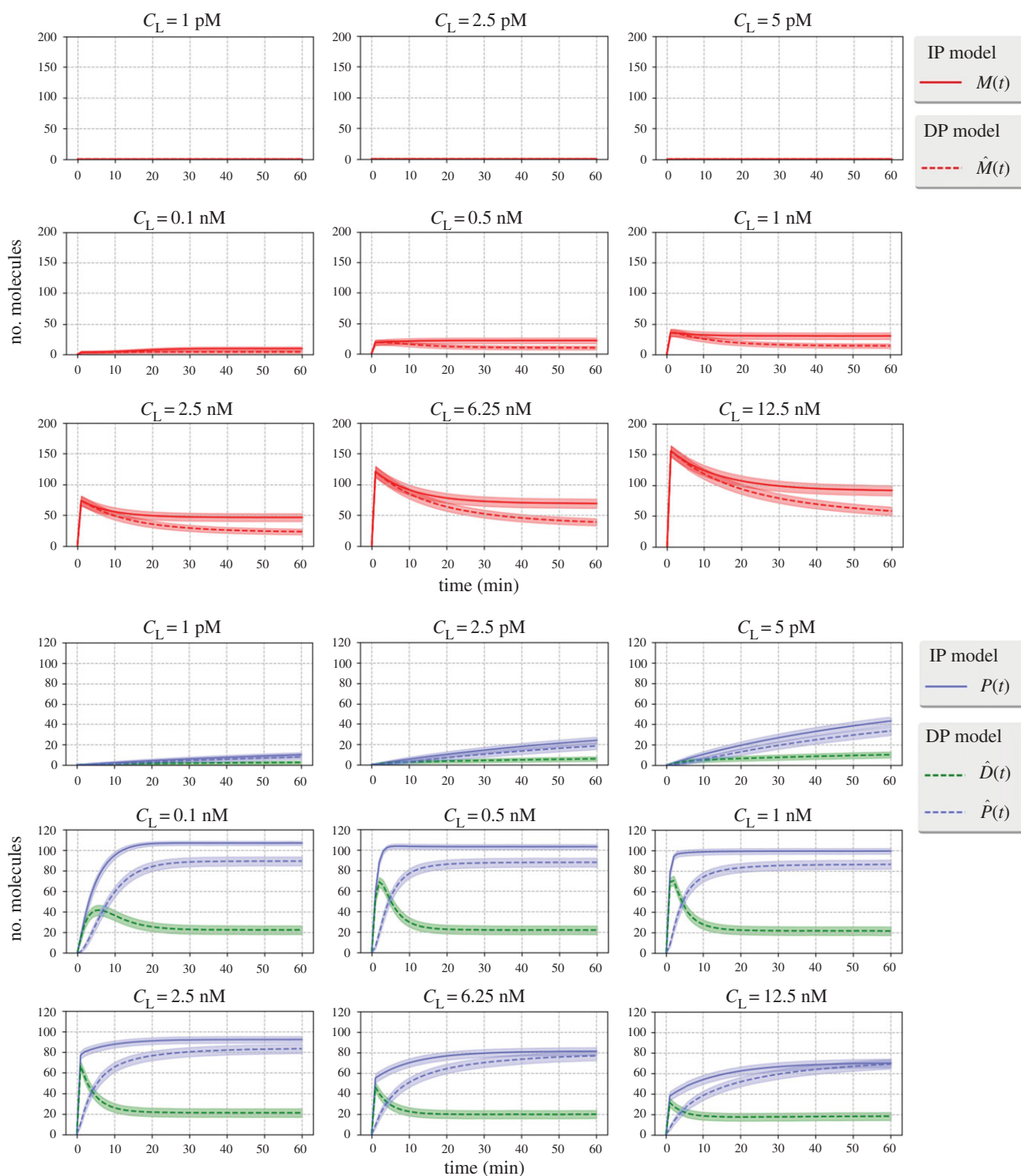


Figure 8. Gillespie simulations of the processes \mathcal{X} and $\hat{\mathcal{X}}$ for different initial ligand concentrations, $c_L \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}, 0.1 \text{ nM}, 0.5 \text{ nM}, 1 \text{ nM}, 2.5 \text{ nM}, 6.25 \text{ nM}, 12.5 \text{ nM}\}$. Dashed lines correspond to the IP model and solid lines correspond to the DP model. Time course for bound monomers (top) and dimers (bottom).

IP model, $\hat{P}(t)$ in the DP model) and non-phosphorylated ($\hat{D}(t)$ in the DP model) dimers, and consider concentrations $c_L \in \{0.1 \text{ nM}, 0.5 \text{ nM}, 1 \text{ nM}, 2.5 \text{ nM}, 6.25 \text{ nM}, 12.5 \text{ nM}\}$. If internalization of phosphorylated dimers does not occur fast enough (e.g. values $q \in \{1.0, 2.0\}$ in figure 10), a steady-state pool of phosphorylated dimers is maintained at late times on the cell surface. Under faster internalization ($q \in \{5.0, 10.0\}$), and for optimum ligand concentrations, a peak of phosphorylated dimers is observed after ligand stimulation (at time $t \sim 5$ min for the IP model and at time $t \sim 10$ min for the DP model). It is interesting to observe that the peak of non-phosphorylated dimers is well captured in figure 8 (i.e. when internalization and synthesis are not considered), and the same is true for the

time course of monomers (not reported in figure 10). It is only the peak of phosphorylated dimers which is significantly affected by internalization dynamics. Equally, optimum ligand concentrations are well characterized by the original IP and DP models; that is, similar optimum ligand concentrations are found, of the order of approximately 1 nM, in figures 8 and 10 (i.e. with and without receptor synthesis and internalization).

4. Discussion

In this paper, our aim was to quantify the signalling time-scales (or phosphorylation) for two different stochastic

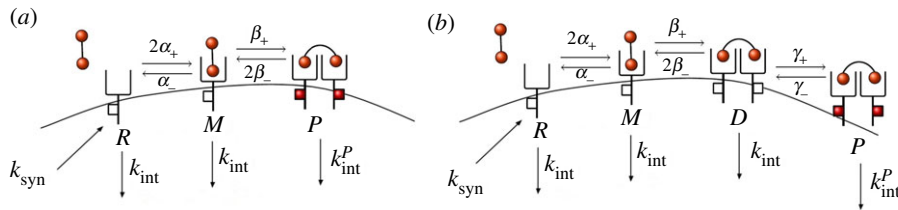


Figure 9. IP and DP models when synthesis of free receptors, as well as internalization of free receptors, monomers and dimers can take place. (a) The extended IP model and (b) the extended DP model.

Table 3. Elasticities for the stochastic descriptors $E[T_{(0,0)}(N)]$ and $E[T_{(0,0,0)}(N)]$ and mean values π_p and $\hat{\pi}_p$, with respect to each parameter, $\theta_i \in \{\alpha_+, \alpha_-, \beta_+, \beta_-, \gamma_+, \gamma_-\}$ for different ligand concentrations, $c_l \in \{1 \text{ pM}, 2.5 \text{ pM}, 5 \text{ pM}\}$.

elasticity	c_l	α_+	α_-	β_+	β_-	γ_+	γ_-
$\frac{\partial E[T_{(0,0)}(N)]/\partial \theta}{E[T_{(0,0)}(N)]/\theta}$	1 pM	-9.98×10^{-1}	1.61×10^{-2}	-2.17×10^{-2}	3.42×10^{-3}	—	—
	2.5 pM	-9.99×10^{-1}	1.78×10^{-2}	-2.36×10^{-2}	4.60×10^{-3}	—	—
	5 pM	-1.00	2.01×10^{-2}	-2.66×10^{-2}	6.02×10^{-3}	—	—
$\frac{\partial E[T_{(0,0,0)}(N)]/\partial \theta}{E[T_{(0,0,0)}(N)]/\theta}$	1 pM	-8.47×10^{-1}	1.22×10^{-2}	-1.73×10^{-2}	2.12×10^{-3}	-2.26×10^{-1}	8.82×10^{-2}
	2.5 pM	-8.60×10^{-1}	1.33×10^{-2}	-1.84×10^{-2}	2.59×10^{-3}	-2.68×10^{-1}	1.36×10^{-1}
	5 pM	-8.72×10^{-1}	1.51×10^{-2}	-2.07×10^{-2}	3.30×10^{-3}	-2.99×10^{-1}	1.76×10^{-1}
$\frac{\partial \pi_p/\partial \theta}{\pi_p/\theta}$	1 pM	3.45×10^{-2}	-3.45×10^{-2}	3.82×10^{-2}	-3.82×10^{-2}	—	—
	2.5 pM	6.67×10^{-2}	-6.67×10^{-2}	7.17×10^{-2}	-7.17×10^{-2}	—	—
	5 pM	1.03×10^{-1}	-1.03×10^{-1}	1.10×10^{-1}	-1.10×10^{-1}	—	—
$\frac{\partial \hat{\pi}_p/\partial \theta}{\hat{\pi}_p/\theta}$	1 pM	7.31×10^{-3}	-7.31×10^{-3}	8.08×10^{-3}	-8.08×10^{-3}	2.06×10^{-1}	-2.06×10^{-1}
	2.5 pM	1.73×10^{-2}	-1.73×10^{-2}	1.85×10^{-2}	-1.85×10^{-2}	2.15×10^{-1}	-2.15×10^{-1}
	5 pM	5.88×10^{-2}	-5.88×10^{-2}	6.12×10^{-2}	-6.12×10^{-2}	2.49×10^{-1}	-2.49×10^{-1}

models of receptor–ligand interaction (instantaneous phosphorylation, IP model, and delayed phosphorylation, DP model), and to analyse their late time behaviour, making use of new exact matrix-analytic techniques. Stochastic approaches are essential in order to explore the role of limited (and small) protein copy numbers in receptor–ligand signalling systems, since the stochastic nature of protein expression and quantitative differences in the abundance of proteins could dysregulate receptor-mediated signalling, as recently reported by Shi *et al.* [46].

We have assumed that bound dimers are instantaneously phosphorylated in the IP model, while in the DP model phosphorylation is considered a new and independent reaction. In these two models, matrix-analytic techniques have been applied (see appendices B and C, respectively) to study the time to reach a threshold number of phosphorylated bound dimers, P , on the cell membrane, and the steady-state probability distribution. We have identified these times as absorption times in conveniently defined auxiliary CTMCs, and their Laplace–Stieltjes transforms and different order moments have been computed algorithmically by means of a first-step analysis, while exploiting the quasi-birth-and-death structure of the infinitesimal generators associated with these processes. Moreover, the construction of the DP model as an extension of the IP model in §2 allows us not only to analyse the role played by phosphorylation events (see §3.3), but also to show how different reactions may be incorporated while adapting the matrix-analytic approach. A particular feature of this analytic

approach is that it allows one to study the role played by each kinetic rate, by computing the partial derivatives of the descriptors under consideration with respect to the corresponding model parameters.

Our numerical results in §3 have considered the interaction between receptor VEGFR2 and bivalent ligand VEGF-A in human vascular endothelial cells. Our results indicate that phosphorylation, as an additional reaction, only seems to quantitatively affect the timescales for signalling (or phosphorylation), but does not qualitatively change the dynamics of the process. Moreover, by sequentially incorporating receptor synthesis and internalization dynamics, we found that intra-cellular receptor trafficking plays an important role in shifting the original signal (in terms of phosphorylated dimers) found on the cell surface into endosomal compartments, but where the dynamics of free receptors, monomers and non-phosphorylated dimers are well characterized with mathematical models exclusively describing the cell surface. These cell surface models allowed us as well to identify optimum ligand concentrations, which were qualitatively unchanged if synthesis and internalization events are included (figure 10).

Our previous comments can be further illustrated by carrying out a single-molecule analysis; that is, by studying the fate of a bound monomer in the system. In particular, we consider a single ligand that has been captured by a receptor forming a bound monomer, and analyse the dynamics of this single complex, neglecting the effects due to other ligands or receptors in the system. Thus, we focus on the

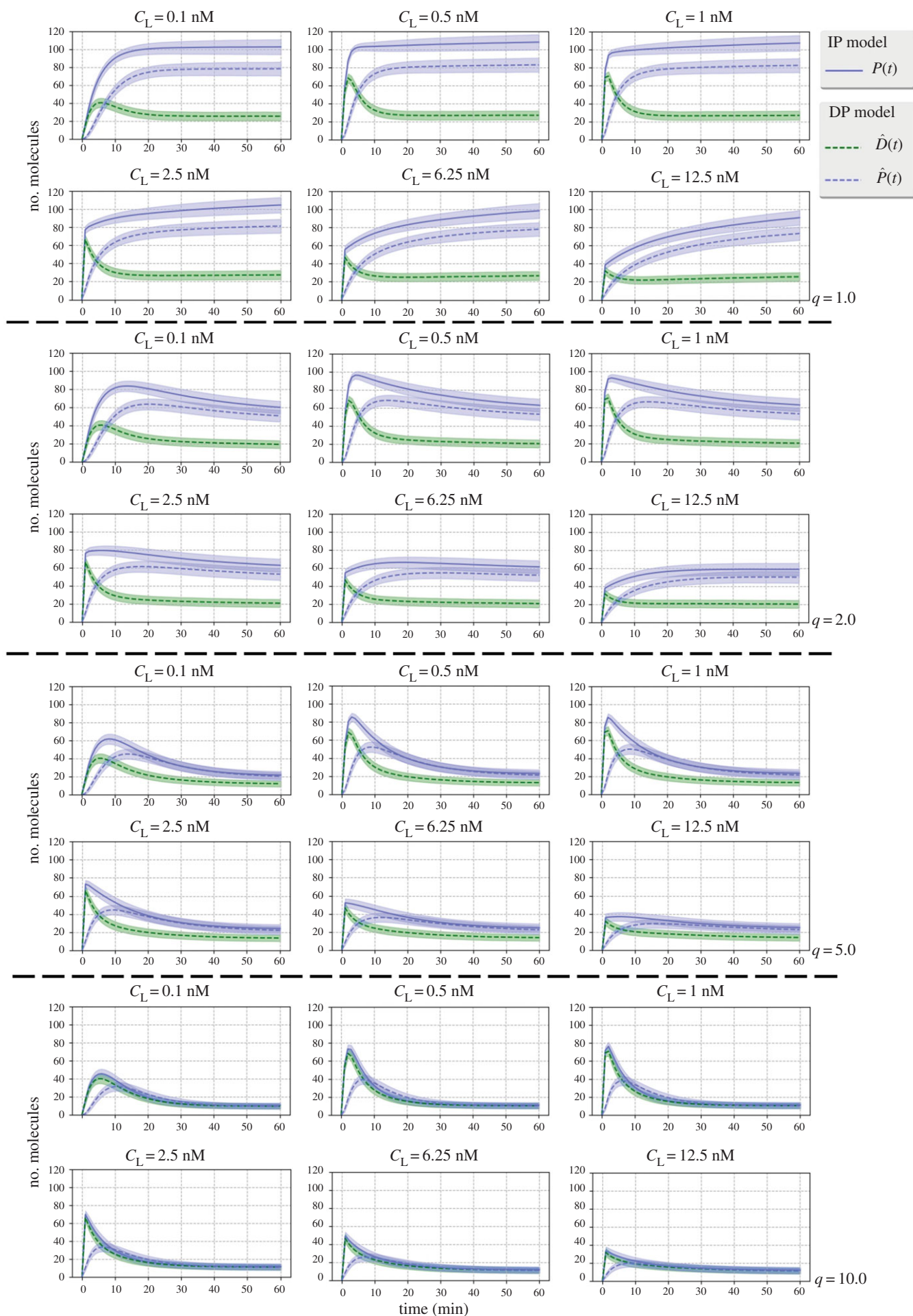


Figure 10. Gillespie simulations for the extended IP and DP models of figure 9, for different initial ligand concentrations, $c_L \in \{0.1 \text{ nM}, 0.5 \text{ nM}, 1 \text{ nM}, 2.5 \text{ nM}, 6.25 \text{ nM}, 12.5 \text{ nM}\}$ and different values of $q \in \{1, 2, 5, 10\}$. Dashed lines correspond to the IP model and solid lines correspond to the DP model. Time course for phosphorylated and non-phosphorylated dimers.

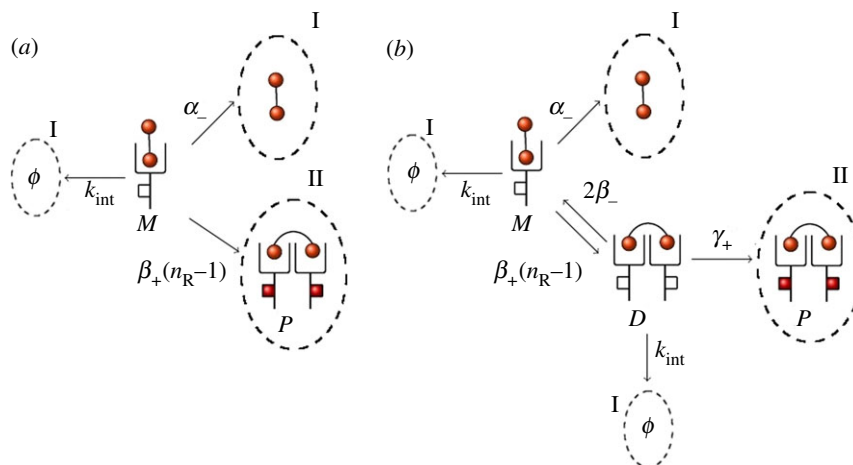


Figure 11. Individual bound monomer fate under (a) the extended IP model and (b) the extended DP model. Fate I: dissociation or internalization before signalling. Fate II: signalling before dissociation or internalization.

fate of this complex (phosphorylating or not before the bound monomer dissociates or internalizes), which depends on the kinetic rates, and is controlled by the stochastic processes illustrated in figure 11. We note that the original models (without internalization) can be obtained by setting $k_{\text{int}} = 0$ in figure 9, since in that case we also set $k_{\text{syn}} = n_R k_{\text{int}} = 0$.

If we define

p_{signal} = ‘probability of the complex signalling (or phosphorylating) before dissociation or internalization’;

that is, the probability of Fate II. This probability can be computed as follows:

— IP model (instantaneous phosphorylation):

$$p_{\text{signal}} = \frac{\beta_+(n_R - 1)}{k_{\text{int}} + \alpha_- + \beta_+(n_R - 1)}.$$

— DP model (delayed phosphorylation):

$$p_{\text{signal}} = \frac{\gamma_+ \beta_+(n_R - 1)}{(k_{\text{int}} + \alpha_-)(k_{\text{int}} + 2\beta_- + \gamma_+) + \beta_+(n_R - 1)(k_{\text{int}} + \gamma_+)}.$$

On the other hand, if we focus on the time to signalling and define

τ_{signal} = ‘mean time to complex signalling before dissociation or internalization, conditioned on this occurring’,

this conditioned mean time can be written as:

— IP model (instantaneous phosphorylation):

$$\tau_{\text{signal}} = \frac{1}{k_{\text{int}} + \alpha_- + \beta_+(n_R - 1)}.$$

— DP model (delayed phosphorylation):

$$\tau_{\text{signal}} = \frac{2\beta_- + \gamma_+}{\gamma_+} \left[\frac{1}{k_{\text{int}} + \alpha_- + \beta_+(n_R - 1)} + \frac{1}{k_{\text{int}} + 2\beta_- + \gamma_+} \right].$$

The values of p_{signal} and τ_{signal} are reported in table 4 for $k_{\text{int}} \in \{0, 2.8 \times 10^{-4} \text{ s}^{-1}\}$. From these results, it seems clear that once a ligand is bound to a monomeric receptor, the

Table 4. Probability of a single monomer signalling (i.e. dimerizing and becoming phosphorylated, p_{signal}) and conditioned time for this to occur (τ_{signal}).

k_{int}	model	p_{signal}	τ_{signal}
0	IP model	0.9874	9.5356 s
	DP model	0.9863	283.0799 s
$2.8 \times 10^{-4} \text{ s}^{-1}$	IP model	0.9847	9.5095 s
	DP model	0.9137	265.1743 s

probability to phosphorylate and, thus, to signal is almost one (for either model), when no internalization occurs. Internalization of complexes and delayed phosphorylation cannot decrease this probability on their own, and only when these two events are considered together, the single-molecule signalling probability of a monomer decreases approximately by 9%. However, the timescales to phosphorylate are mainly affected by the delayed phosphorylation. On the other hand, it might seem counterintuitive that the timescales for signal initiation are shorter when internalization takes place. We note here that these are conditioned times for signalling, that is, times conditioned on this signalling actually occurring. Thus, our results for τ_{signal} in table 4 should be interpreted as the fact that, if internalization can occur, only those monomers reaching dimerization and phosphorylation soon enough will initiate signalling before internalization takes place.

From a biological perspective, we note that the total number of VEGFR2s per cell varies according to other studies [30,39,47] and could be larger than the numbers used in our computations [42]. A larger number of VEGFR2 receptors on the cell surface would, however, only quantitatively change our results, and in particular a higher optimum ligand concentration threshold would be reported. The sensitivity analysis carried out for the descriptors enables us to show how the monomeric formation rate, α_+ , plays a crucial role in these models, with an effect which can be more than twice the effect of any other rate for some of the descriptors we have considered. Finally, the numerical results presented in §3 for the VEGF-A and VEGFR system have allowed us to quantify the effect of different ligand concentrations on the

timescales to signalling, the late time behaviour of the system and the time course dynamics of the individual molecular species. Increasing ligand concentration decreases the times to reach any signalling threshold and increases the maximum potential signalling thresholds to be reached. However, high ligand concentrations can result in saturated scenarios, where the phosphorylation of bound dimers is reduced and monomeric bound complexes are enhanced.

The approach presented here could be, in principle, applied to other RTKs, most notably the EGFR, which is driving cellular proliferation in a variety of epithelial tumours. This receptor is of special relevance in clinical oncology, since a series of promising anti-EGFR small-molecule RTK inhibitors have already been designed. Unfortunately, drug resistance usually emerges during the course of treatment and it is important to understand the molecular mechanisms that underlie the development of such drug resistance, which may involve both the wild-type and mutant receptors [48]. Other RTKs of interest, for example, are those of the fibroblast growth factor receptor family, insulin receptor family and the leucocyte RTK family.

Data accessibility. This article has no additional data.

Competing interests. We declare we have no competing interests.

Funding. This work has been supported by The Leverhulme Trust (RPG-2012-772) (M. López-García, G. Lythe, S. Ponnambalam and C. Molina-París). The research leading to these results has also received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7 under REA grant agreement PIRSES-GA-2008-230665 (2009–2012) (M. Nowicka, G. Lythe and C. Molina-París) and PIRSES-GA-2012-317893 (2013–2017) (M. López-García, G. Lythe and C. Molina-París). M. Nowicka is supported by EPSRC and AstraZeneca (CASE studentship).

Acknowledgements. We gratefully acknowledge the University of Leeds for the permission to use the High Performance Computing facilities ARC1 and ARC2.

Appendix A. Notation

In this appendix, we set some standard notation to be used in the paper. First, $\delta_{i,j}$ represents Kronecker's delta; that is,

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Given a set \mathcal{S} , $\#\mathcal{S}$ represents its cardinality. Matrices and vectors are always given in bold, where $\mathbf{0}_p$ (\mathbf{e}_q) represents a column vector of zeros (ones) with dimension p (q). The symbol T represents the transposition operator and, for a matrix $\mathbf{A}(\theta)$, we use the calculus notation

$$\mathbf{A}^{(l)}(0) = \left. \frac{d^l}{d\theta} \mathbf{A}(\theta) \right|_{\theta=0}.$$

Finally, when a matrix depends on different parameters, $\mathbf{A}(\alpha, \theta)$, its first-order partial derivatives with respect to each parameter are given by $\mathbf{A}^{(\alpha)}(\alpha, \theta)$ and $\mathbf{A}^{(\theta)}(\alpha, \theta)$, respectively.

Appendix B. Analysis of the IP model

The analysis carried out in this appendix requires the use of levels for the organization of the state space, Laplace–Stieltjes transforms, first-step arguments and auxiliary absorbing Markov chains. We first organize the state space \mathcal{S} , which

contains

$$\#\mathcal{S} = \frac{(n_L + 1)(n_L + 2)}{2},$$

states, by levels (groups of states) as

$$\mathcal{S} = \bigcup_{k=0}^{n_L} L(k),$$

where $L(k) = \{(n_1, n_2): n_2 = k\}$, $0 \leq k \leq n_L$, so that $J(k) = \#L(k) = n_L - k + 1$. That is, a level $L(k)$ comprises all the possible states (n_1, n_2) of the process with a total number of P complexes equal to k . Moreover, we order these levels as

$$L(0) \prec L(1) \prec \dots \prec L(n_L),$$

and states inside a level, $L(k) = \{(0, k), (1, k), \dots, (n_L - k, k)\}$, $0 \leq k \leq n_L$, are ordered as

$$(0, k) \prec (1, k) \prec \dots \prec (n_L - k, k).$$

Given the transitions of figure 2, it is clear that from a state (n_1, n_2) in level $L(n_2)$, the process can only move to states in the same level, $L(n_2)$, and to states in adjacent levels, $L(n_2 - 1)$ and $L(n_2 + 1)$. That is, if the state of the system is (n_1, n_2) (and then, the process is in level $L(n_2)$), the only possible transitions are to $(n_1 - 1, n_2)$ (if a bound monomer dissociates, in which case the process remains in level $L(n_2)$), to $(n_1 + 1, n_2)$ (if a bound monomer is formed, leaving the process in level $L(n_2)$), to $(n_1 + 1, n_2 - 1)$ (if a bound dimer dissociates, and the process then decreases to level $L(n_2 - 1)$) or to $(n_1 - 1, n_2 + 1)$ (if a bound dimer is created, increasing the level of the process to $L(n_2 + 1)$).

The organization of \mathcal{S} , previously proposed, becomes crucial to obtain a convenient structure for the infinitesimal generator \mathbf{Q} of \mathcal{X} , the matrix containing the transition rates in the Markov chain. In particular, the resulting \mathbf{Q} is of the quasi-birth-and-death type [18] (tridiagonal by blocks structure)

$$\mathbf{Q} = \begin{pmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} & \mathbf{0}_{J(0) \times J(2)} & \dots & \mathbf{0}_{J(0) \times J(n_L-1)} & \mathbf{0}_{J(0) \times J(n_L)} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \dots & \mathbf{0}_{J(1) \times J(n_L-1)} & \mathbf{0}_{J(1) \times J(n_L)} \\ \mathbf{0}_{J(2) \times J(0)} & \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \dots & \mathbf{0}_{J(2) \times J(n_L-1)} & \mathbf{0}_{J(2) \times J(n_L)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{J(n_L-1) \times J(0)} & \mathbf{0}_{J(n_L-1) \times J(1)} & \mathbf{0}_{J(n_L-1) \times J(2)} & \dots & \mathbf{A}_{n_L-1, n_L-1} & \mathbf{A}_{n_L-1, n_L} \\ \mathbf{0}_{J(n_L) \times J(0)} & \mathbf{0}_{J(n_L) \times J(1)} & \mathbf{0}_{J(n_L) \times J(2)} & \dots & \mathbf{A}_{n_L, n_L-1} & \mathbf{A}_{n_L, n_L} \end{pmatrix}, \quad (\text{B } 1)$$

where sub-matrices $\mathbf{A}_{k,k'}$ contain the infinitesimal transition rates of the transitions from states in level $L(k)$ to states in level $L(k')$, with $k' \in \{k - 1, k, k + 1\}$. In particular, matrices $\mathbf{A}_{k,k'}$ in (B 1) are obtained from (2.1) and are as follows:

— For $1 \leq k \leq n_L$,

$$(\mathbf{A}_{k,k-1})_{ij} = \begin{cases} 2\beta_- k, & \text{if } j = i + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k$, $0 \leq j \leq n_L - k + 1$.

— For $0 \leq k \leq n_L$,

$$(\mathbf{A}_{k,k})_{ij} = \begin{cases} 2\alpha_+(n_R - i - 2k)(n_L - i - k), & \text{if } j = i + 1, \\ \alpha_- i, & \text{if } j = i - 1, \\ -(2\alpha_+(n_R - i - 2k)(n_L - i - k) \\ + \alpha_- i + 2\beta_- k + \beta_+ i(n_R - i - 2k)), & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k$, $0 \leq j \leq n_L - k$.

— For $0 \leq k \leq n_L - 1$,

$$(\mathbf{A}_{k,k+1})_{ij} = \begin{cases} \beta_+ i (n_R - i - 2k), & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k$, $0 \leq j \leq n_L - k - 1$.

We consider the time to obtain a threshold number, $N > 0$, of P complexes. In particular, given an initial state of the process (n_1, n_2) , and a certain threshold $N > 0$, we consider the random variable

$$T_{(n_1, n_2)}(N) = \text{'time to reach a number } N \text{ of } P \text{ complexes in } \mathcal{X}, \\ \text{if the process starts at } (n_1, n_2) \in \mathcal{S}'.$$

We observe that this time is 0 for $N \leq n_2$. In order to study this descriptor for $N > n_2$, we make use of an auxiliary CTMC, $\mathcal{X}(N)$, which depends on the threshold value N . We define $\mathcal{X}(N)$ over the state space $\mathcal{S}(N)$ with

$$\mathcal{S}(N) = \mathcal{C}(N) \cup \{\bar{N}\},$$

where we denote $\mathcal{C}(N) = \cup_{k=0}^{N-1} L(k)$, and where \bar{N} is a macro-state that consists of all the states in the set $\cup_{k=N}^{n_L} L(k)$. Regarding the transitions of this auxiliary CTMC, we retain those transitions of \mathcal{X} between states in $\mathcal{C}(N)$, with \bar{N} an absorbing state, so that once $\mathcal{X}(N)$ enters \bar{N} , it does not leave this state. Transitions from states in level $L(N-1)$ to states in $L(N)$ of the original process \mathcal{X} become transitions from states in level $L(N-1)$ to the macro-state \bar{N} in $\mathcal{X}(N)$, where their infinitesimal transition rates are computed from the original ones as follows:

$$q_{(n_1, n_2), \bar{N}} = \sum_{(n'_1, n'_2) \in L(N)} q_{(n_1, n_2), (n'_1, n'_2)}, \quad \forall (n_1, n_2) \in L(N-1).$$

The process $\mathcal{X}(N)$ can be seen as the process \mathcal{X} until a number

N of P complexes are formed. Then, $\mathcal{X}(N)$ ends since \bar{N} is an absorbing state for this auxiliary process. With $\mathcal{X}(N)$ so defined, it is clear that the time taken to obtain a number N of P complexes in the original process \mathcal{X} is equal to the time until absorption at \bar{N} in the (absorbing) process $\mathcal{X}(N)$, which is known to follow a continuous phase-type (PH) distribution (e.g. [9,18]). The analysis of the exact distribution of a continuous phase-type random variable is, in general, a difficult problem. In our case, it would imply obtaining the exponential matrix $\exp(\mathbf{T}(N)) = \sum_{n=0}^{+\infty} (\mathbf{T}(N)^n / n!)$, where $\mathbf{T}(N)$ is a specific sub-matrix of the infinitesimal generator of $\mathcal{X}(N)$. Here, we instead make use of the Laplace-Stieltjes transform of $T_{(n_1, n_2)}(N)$, which completely determines its distribution, and which allows us to obtain any l th-order moment $E[T_{(n_1, n_2)}(N)^l]$. We can also efficiently calculate the l th-order moment by using the $(l-1)$ th-order moment, proceeding recursively, with the computational effort devoted to obtaining inverses of square blocks $\mathbf{A}_{k,kr}$ with dimension $J(k) = n_L - k + 1$. Again, the proposed organization of states is essential for the construction of an efficient algorithm. If we define the Laplace-Stieltjes transform of $T_{(n_1, n_2)}(N)$ as

$$\phi_{(n_1, n_2)}^N(z) = E[e^{-zT_{(n_1, n_2)}(N)}], \quad \Re(z) \geq 0,$$

then, the different l th-order moments of $T_{(n_1, n_2)}(N)$ can be obtained as

$$E[T_{(n_1, n_2)}(N)^l] = (-1)^l \left. \frac{d^l}{dz^l} \phi_{(n_1, n_2)}^N(z) \right|_{z=0}, \quad \forall l \geq 1.$$

We can apply a first-step argument in order to obtain a system of linear equations for the Laplace-Stieltjes transforms $\phi_{(n_1, n_2)}^N(z)$, given a state $(n_1, n_2) \in \mathcal{S}(N)$. We can write down the equation

$$\begin{aligned} \phi_{(n_1, n_2)}^N(z) &= (1 - \delta_{n_1+n_2, n_L}) \frac{2\alpha_+(n_R - n_1 - 2n_2)(n_L - n_1 - n_2)}{z + A_{(n_1, n_2)}} \phi_{(n_1+1, n_2)}^N(z) + (1 - \delta_{n_1, 0}) \\ &\times \frac{\alpha_- n_1}{z + A_{(n_1, n_2)}} \phi_{(n_1-1, n_2)}^N(z) + (1 - \delta_{n_1, 0}) \frac{\beta_+ n_1 (n_R - n_1 - 2n_2)}{z + A_{(n_1, n_2)}} (\delta_{n_2, N-1} \\ &+ (1 - \delta_{n_2, N-1}) \phi_{(n_1-1, n_2+1)}^N(z)) + (1 - \delta_{n_2, 0}) \frac{2\beta_- n_2}{z + A_{(n_1, n_2)}} \phi_{(n_1+1, n_2-1)}^N(z), \end{aligned} \quad (\text{B2})$$

where from now on $A_{(n_1, n_2)} = 2\alpha_+(n_R - n_1 - 2n_2)(n_L - n_1 - n_2) + \alpha_- n_1 + \beta_+ n_1 (n_R - n_1 - 2n_2) + 2\beta_- n_2$. Equation (B2) relates the Laplace-Stieltjes transforms of all the states in $\mathcal{S}(N)$, so that a system of linear equations is obtained. If we organize the Laplace-Stieltjes transforms in vectors by levels as follows:

$$\mathbf{g}^N(z) = (\mathbf{g}_0^N(z)^T, \mathbf{g}_1^N(z)^T, \mathbf{g}_2^N(z)^T, \dots, \mathbf{g}_{N-1}^N(z)^T)^T,$$

with $\mathbf{g}_k^N(z) = (\phi_{(0,k)}^N(z), \phi_{(1,k)}^N(z), \phi_{(2,k)}^N(z), \dots, \phi_{(n_L-k, k)}^N(z))^T$, for $0 \leq k \leq N-1$, then the system given in (B2) can be expressed in matrix form as

$$\mathbf{g}^N(z) = \mathbf{A}^N(z) \mathbf{g}^N(z) + \mathbf{a}^N(z), \quad (\text{B3})$$

with the matrix $\mathbf{A}^N(z)$ given by

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

and the vector $\mathbf{a}^N(z) = (\mathbf{0}_{J(0)}^T, \mathbf{0}_{J(1)}^T, \dots, \mathbf{0}_{J(N-2)}^T, \mathbf{a}_{N-1}^T(z))^T$. Submatrices $\mathbf{A}_{k,k'}(z)$ and sub-vector $\mathbf{a}_{N-1}(z)$ in (B 3) are given by

- $(\mathbf{a}_{N-1}(z))_i = \beta_+ i(n_R - i - 2(N - 1))/(z + A_{(i,N-1)})$, for $0 \leq i \leq n_L - N + 1$.
- For $1 \leq k \leq n_L$,

$$(\mathbf{A}_{k,k-1}(z))_{ij} = \begin{cases} \frac{2\beta_- k}{z + A_{(i,k)}}, & \text{if } j = i + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k, 0 \leq j \leq n_L - k + 1$.

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

$$(\mathbf{A}_{k,k}(z))_{ij} = \begin{cases} \frac{2\alpha_+(n_R - i - 2k)(n_L - i - k)}{z + A_{(i,k)}}, & \text{if } j = i + 1, \\ \frac{\alpha_- i}{z + A_{(i,k)}}, & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k, 0 \leq j \leq n_L - k$.

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

$$(\mathbf{A}_{k,k+1}(z))_{ij} = \begin{cases} \frac{\beta_- i(n_R - i - 2k)}{z + A_{(i,k)}}, & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k, 0 \leq j \leq n_L - k - 1$.

Exploiting the special block structure of $\mathbf{A}^N(z)$ allows for an efficient solution of the system in (B 3), in a recursive manner through a specialized block-Gaussian elimination process, leading to algorithm 1 (Part 1). The calculation of the Laplace–Stieltjes transforms in algorithm 1 (Part 1) has its own merit, since it determines the distribution of the random variable under consideration. Moreover, the calculation of the distribution function of $T_{(n_1, n_2)}(N)$ by numerical inversion of the transform is possible, although computationally expensive, and is not developed here (e.g. [49]).

Once the Laplace–Stieltjes transforms are in hand, we can obtain the different l th-order moments by successive differentiation of the system in (B 3). In particular, we can write

$$\mathbf{m}^{N,(l)} = \sum_{p=0}^l (-1)^p \binom{l}{p} \frac{d^p}{dz^p} \mathbf{A}^N(z) \Big|_{z=0} \mathbf{m}^{N,(l-p)} + (-1)^l \frac{d^l}{dz^l} \mathbf{a}^N(z) \Big|_{z=0}, \quad (\text{B 4})$$

where $\mathbf{m}^{N,(l)}$ is the column vector containing the desired moments $E[T_{(n_1, n_2)}(N)^l]$, for $(n_1, n_2) \in \mathcal{C}(N)$. We organize these moments in sub-vectors by levels as follows:

$$\mathbf{m}^{N,(l)} = (\mathbf{m}_0^{N,(l)T}, \mathbf{m}_1^{N,(l)T}, \mathbf{m}_2^{N,(l)T}, \dots, \mathbf{m}_{N-1}^{N,(l)T})^T,$$

with $\mathbf{m}_k^{N,(l)} = (E[T_{(0,k)}(N)^l], E[T_{(1,k)}(N)^l], E[T_{(2,k)}(N)^l], \dots, E[T_{(n_L - k, k)}(N)^l])^T$, for $0 \leq k \leq N - 1$. Note that the notation $\mathbf{m}^{N,(0)} = \mathbf{g}^N(0) = \mathbf{e}_{\#\mathcal{C}(N)}$ is implicit in (B 4). That is, the moment of order $l = 0$ is the Laplace–Stieltjes transform for $z = 0$. Finally, the system in (B 4) is rewritten following the notation presented in appendix A:

$$\mathbf{m}^{N,(l)} = \mathbf{A}^N(0) \mathbf{m}^{N,(l)} + \sum_{p=1}^l \binom{l}{p} (-1)^p \mathbf{A}^{N,(p)}(0) \mathbf{m}^{N,(l-p)} + (-1)^l \mathbf{a}^{N,(l)}(0). \quad (\text{B 5})$$

It is clear that the direct calculation of the inverse $(\mathbf{I}_{\#\mathcal{C}(N)} - \mathbf{A}^N(0))^{-1}$ involved in the solution of (B 5) can be avoided if working by levels and solving (B 5) in a similar way to algorithm 1 (Part 1). By starting with the known moment of order $p = 0$, we proceed recursively by calculating $\mathbf{m}^{N,(p)}$ from $\mathbf{m}^{N,(p-1)}$, until the desired order $p = l$ is reached, leading to algorithm 1 (Part 2). Matrices $\mathbf{A}^{N,(p)}(0)$ and $\mathbf{a}^{N,(p)}(0)$ in (B 5) are given by

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

$$\mathbf{a}^{N,(p)}(0) = \begin{pmatrix} \mathbf{0}_{J(0)} \\ \mathbf{0}_{J(1)} \\ \vdots \\ \mathbf{0}_{J(N-2)} \\ \mathbf{a}_{N-1}^{(p)}(0) \end{pmatrix},$$

where expressions for $\mathbf{a}_{N-1}^{(p)}(0)$ and $\mathbf{A}_{k,k'}^{(p)}(0)$, for $p \geq 1$, are as follows:

- $(\mathbf{a}_{N-1}^{(p)}(0))_i = (-1)^p p! (\beta_+ i(n_R - i - 2(N - 1)) / A_{(i, N-1)}^{p+1})$, for $0 \leq i \leq n_L - N + 1$.
- For $1 \leq k \leq n_L, p \geq 1$,

$$(\mathbf{A}_{k,k-1}^{(p)}(0))_{ij} = \begin{cases} (-1)^p p! \frac{2\beta_- k}{A_{(i,k)}^{p+1}}, & \text{if } j = i + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k, 0 \leq j \leq n_L - k + 1$.

- For $0 \leq k \leq n_L, p \geq 1$,

$$(\mathbf{A}_{k,k}^{(p)}(0))_{ij} = \begin{cases} (-1)^p p! \frac{2\alpha_+(n_R - i - 2k)(n_L - i - k)}{A_{(i,k)}^{p+1}}, & \text{if } j = i + 1, \\ (-1)^p p! \frac{\alpha_- i}{A_{(i,k)}^{p+1}}, & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k, 0 \leq j \leq n_L - k$.

— For $0 \leq k \leq n_L - 1, p \geq 1$,

$$(\mathbf{A}_{k,k+1}^{(p)}(0))_{ij} = \begin{cases} (-1)^p p! \frac{\beta_{+i}(n_L-i-2k)}{A_{(i,k)}^{p+1}}, & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - k, 0 \leq j \leq n_L - k - 1$.

Finally, the late time behaviour of the process is given by the stationary probability distribution of the CTMC; that is, the probabilities

$$\pi_{(n_1, n_2)} = \lim_{t \rightarrow +\infty} \mathbb{P}((M(t), P(t)) = (n_1, n_2)), \quad \forall (n_1, n_2) \in \mathcal{S},$$

which do not depend on the initial state. We can store this probability distribution in a row vector $\pi = (\pi_0, \pi_2, \dots, \pi_{n_L})$, where the row sub-vector π_k contains the ordered probabilities $\pi_{(n_1, n_2)}$ for states in level $L(k)$. Solving the system

$$\pi \mathbf{Q} = \mathbf{0}_{\#S}^T \quad \text{and} \quad \pi \mathbf{e}_{\#S} = 1,$$

and adapting arguments from Latouche & Ramaswami [9, ch. 10], we obtain algorithm 2. With π in hand, the long-term mean number of M and P complexes can be obtained as

$\pi_M =$ ‘mean number of M complexes in steady state’

$$= \sum_{k=0}^{n_L} k \left(\sum_{j=0}^{n_L} (\pi_j)_k \right),$$

$\pi_P =$ ‘mean number of P complexes in steady state’

$$= \sum_{k=0}^{n_L} k (\pi_k \mathbf{e}_{j(k)}).$$

Algorithm 1 (to obtain the Laplace–Stieltjes transforms $\mathbf{g}^{N,(l)}$ and the l th-order moments $\mathbf{m}^{N,(l)}$)

PART 1

$$\mathbf{H}_0^N(z) = \mathbf{I}_{j(0)} - \mathbf{A}_{0,0}(z);$$

For $k = 1, \dots, N - 1$:

$$\mathbf{H}_k^N(z) = \mathbf{I}_{j(k)} - \mathbf{A}_{k,k}(z) - \mathbf{A}_{k,k-1}(z) \\ \times \mathbf{H}_{k-1}^N(z)^{-1} \mathbf{A}_{k-1,k}(z);$$

$$\mathbf{g}_{N-1}^N(z) = \mathbf{H}_{N-1}^N(z)^{-1} \mathbf{a}_{N-1}(z);$$

$$\mathbf{m}_{N-1}^{N,(0)} = \mathbf{g}_{N-1}^N(0);$$

For $k = N - 2, \dots, 1, 0$:

$$\mathbf{g}_k^N(z) = \mathbf{H}_k^N(z)^{-1} \mathbf{A}_{k,k+1}(z) \mathbf{g}_{k+1}^N(z);$$

$$\mathbf{m}_k^{N,(0)} = \mathbf{g}_k^N(0);$$

PART 2

$$\mathbf{m}_{N-1}^{N,(0)} = \mathbf{g}_{N-1}^N(0);$$

For $k = N - 2, \dots, 1, 0$:

$$\mathbf{m}_k^{N,(0)} = \mathbf{g}_k^N(0);$$

For $p = 1, \dots, l$:

$$\mathbf{P}_0^{N,(p)} = \sum_{k=1}^p \binom{p}{k} (-1)^k (\mathbf{A}_{0,0}^{(k)}(0) \mathbf{m}_0^{N,(p-k)} \\ + \mathbf{A}_{0,1}^{(k)}(0) \mathbf{m}_1^{N,(p-k)});$$

For $j = 1, \dots, N - 1$:

$$\mathbf{P}_j^{N,(p)} = \mathbf{A}_{j,j-1}(0) \mathbf{H}_{j-1}^N(0)^{-1} \mathbf{P}_{j-1}^{N,(p)} \\ + \sum_{k=1}^p \binom{p}{k} (-1)^k (\mathbf{A}_{j,j-1}^{(k)}(0) \mathbf{m}_{j-1}^{N,(p-k)} \\ + \mathbf{A}_{j,j}^{(k)}(0) \mathbf{m}_j^{N,(p-k)} + (1 - \delta_{j,N-1}) \\ \times \mathbf{A}_{j,j+1}^{(k)}(0) \mathbf{m}_{j+1}^{N,(p-k)});$$

$$\mathbf{m}_{N-1}^{N,(p)} = \mathbf{H}_{N-1}^N(0)^{-1} (\mathbf{P}_{N-1}^{N,(p)} + (-1)^p \mathbf{a}_{N-1}^{(p)}(0));$$

For $j = N - 2, \dots, 1, 0$:

$$\mathbf{m}_j^{N,(p)} = \mathbf{H}_j^N(0)^{-1} (\mathbf{P}_j^{N,(p)} + \mathbf{A}_{j,j+1}(0) \mathbf{m}_{j+1}^{N,(p)});$$

Algorithm 2 (to obtain the stationary distribution π)

$$\mathbf{H}_0 = \mathbf{A}_{0,0};$$

For $k = 1, \dots, n_L - 1$:

$$\mathbf{H}_k = \mathbf{A}_{k,k} - \mathbf{A}_{k,k-1} \mathbf{H}_{k-1}^{-1} \mathbf{A}_{k-1,k};$$

$$\pi_{n_L}^* = \mathbf{1};$$

For $k = n_L - 1, \dots, 0$:

$$\pi_k^* = -\pi_{k+1}^* \mathbf{A}_{k+1,k} \mathbf{H}_k^{-1};$$

For $k = 0, \dots, n_L$:

$$\pi_k = \frac{1}{\sum_{j=0}^{n_L} \pi_j^* \mathbf{e}_{j(0)}} \pi_k^*;$$

Appendix C. Analysis of the DP model

To study the descriptors described in §3.3, we again define levels in the state space, and arrange $\hat{\mathcal{S}}$ in levels as follows:

$$\hat{\mathcal{S}} = \bigcup_{k=0}^{n_L} \hat{L}(k),$$

where $\hat{L}(k) = \{(n_1, n_2, n_3) \in \hat{\mathcal{S}} : n_3 = k\}$, for $0 \leq k \leq n_L$, so that

$$\hat{j}(k) = \#\hat{L}(k) = \frac{(n_L - k + 1)(n_L - k + 2)}{2}.$$

The three-dimensionality of our process implies that each level $\hat{L}(k)$ may be split into sub-levels, as follows:

$$\hat{L}(k) = \bigcup_{r=0}^{n_L-k} l(k; r),$$

with $l(k; r) = \{(n_1, n_2, n_3) \in \hat{\mathcal{S}} : n_2 = r, n_3 = k\}$, for $0 \leq r \leq n_L - k$, $0 \leq k \leq n_L$, and $J(k; r) = \#l(k; r) = n_L - r - k + 1$. That is,

$$l(k; r) = \{(0, r, k), (1, r, k), \dots, (n_L - r - k, r, k)\}, \\ 0 \leq r \leq n_L - k, \quad 0 \leq k \leq n_L,$$

and states in $l(k; r)$ are ordered as indicated above.

The given order of states and the organization by levels and sub-levels yield an infinitesimal generator similar to (B1), where quantities $J(k)$ and matrices $\mathbf{A}_{k,k'}$ are replaced by $\hat{J}(k)$ and $\hat{\mathbf{A}}_{k,k'}$, respectively. Matrix $\hat{\mathbf{A}}_{k,k'}$ contains the ordered infinitesimal transition rates corresponding to transitions from states in level $\hat{L}(k)$ to states in level $\hat{L}(k')$. Each matrix $\hat{\mathbf{A}}_{k,k'}$ is formed by sub-blocks $\mathbf{B}_{r,r'}^{k,k'}$ which contain the infinitesimal transition rates corresponding to transitions from states in sub-level $l(k; r) \subset \hat{L}(k)$ to states in sub-level $l(k'; r') \subset \hat{L}(k')$. We observe that the dimension of the matrix $\hat{\mathbf{A}}_{k,k'}$ is $\hat{J}(k) \times \hat{J}(k') = (n_L - k + 1)(n_L - k + 2)/2 \times (n_L - k' + 1)(n_L - k' + 2)/2$, while the dimension of the sub-block $\mathbf{B}_{r,r'}^{k,k'}$ inside $\hat{\mathbf{A}}_{k,k'}$ is $J(k; r) \times J(k'; r') = (n_L - r - k + 1) \times (n_L - r' - k' + 1)$. Expressions for these matrices are as follows:

— for $0 \leq k \leq n_L$

$$\hat{\mathbf{A}}_{k,k} = \begin{pmatrix} \mathbf{B}_{0,0}^{k,k} & \mathbf{B}_{0,1}^{k,k} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{1,0}^{k,k} & \mathbf{B}_{1,1}^{k,k} & \mathbf{B}_{1,2}^{k,k} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{2,1}^{k,k} & \mathbf{B}_{2,2}^{k,k} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{n_L-k-1, n_L-k-1}^{k,k} & \mathbf{B}_{n_L-k-1, n_L-k}^{k,k} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{n_L-k, n_L-k-1}^{k,k} & \mathbf{B}_{n_L-k, n_L-k}^{k,k} \end{pmatrix},$$

— for $0 \leq k \leq n_L - 1$,

$$\hat{\mathbf{A}}_{k,k+1} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{1,0}^{k,k+1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{2,1}^{k,k+1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{n_L-k,n_L-k-1}^{k,k+1} & \mathbf{0} \end{pmatrix},$$

— for $1 \leq k \leq n_L$,

$$\hat{\mathbf{A}}_{k,k-1} = \begin{pmatrix} \mathbf{0} & \mathbf{B}_{0,1}^{k,k-1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{1,2}^{k,k-1} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{n_L-k-1,n_L-k}^{k,k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{n_L-k,n_L-k+1}^{k,k-1} \end{pmatrix}.$$

We note that, although we are omitting the dimensions of the matrices $\mathbf{0}$ for the ease of notation, the dimension of each matrix $\mathbf{0}$, representing transitions from states in sub-level $l(k; r)$ to states in sub-level $l(k'; r')$, is $J(k; r) \times J(k'; r')$. The expressions for the matrices $\mathbf{B}_{r,r'}^{k,k'}$ are given as follows:

— For $0 \leq r \leq n_L - k$, $0 \leq k \leq n_L$,

$$(\mathbf{B}_{r,r}^{k,k})_{ij} = \begin{cases} \alpha_- i, & \text{if } j = i - 1, \\ -A_{(i,r,k)}, & \text{if } j = i, \\ 2\alpha_+ (n_R - i - 2r - 2k) & \\ \quad (n_L - i - r - k), & \text{if } j = i + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - r - k$, $0 \leq j \leq n_L - r - k$, and, from now on, $A_{(i,r,k)} = 2\alpha_+ (n_R - i - 2r - 2k)(n_L - i - r - k) + \alpha_- i + \beta_+ i (n_R - i - 2r - 2k) + 2\beta_- r + \gamma_+ r + \gamma_- k$.

— For $0 \leq r \leq n_L - k - 1$, $0 \leq k \leq n_L$,

$$(\mathbf{B}_{r,r+1}^{k,k})_{ij} = \begin{cases} \beta_+ i (n_R - i - 2r - 2k), & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - r - k$, $0 \leq j \leq n_L - r - k - 1$.

— For $1 \leq r \leq n_L - k$, $0 \leq k \leq n_L$,

$$(\mathbf{B}_{r,r-1}^{k,k})_{ij} = \begin{cases} 2\beta_- r, & \text{if } j = i + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - r - k$, $0 \leq j \leq n_L - r - k + 1$.

— For $1 \leq r \leq n_L - k$, $0 \leq k \leq n_L - 1$,

$$(\mathbf{B}_{r,r-1}^{k,k+1})_{ij} = \begin{cases} \gamma_+ r, & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - r - k$, $0 \leq j \leq n_L - r - k$.

— For $0 \leq r \leq n_L - k$, $1 \leq k \leq n_L$,

$$(\mathbf{B}_{r,r+1}^{k,k-1})_{ij} = \begin{cases} \gamma_- k, & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \leq i \leq n_L - r - k$, $0 \leq j \leq n_L - r - k$.

For an initial state $(n_1, n_2, n_3) \in \hat{\mathcal{S}}$ and a number $N > 0$, we are also interested in the random variable

$T_{(n_1, n_2, n_3)}(N)$ = 'time to reach a number N of P complexes, if the process starts at (n_1, n_2, n_3) '.

We omit N in the notation for convenience, and denote the random variable under study $T_{(n_1, n_2, n_3)}$. Again, this time is 0 for $N \leq n_3$. For $N > n_3$, we follow an argument similar to that of appendix B, so that $T_{(n_1, n_2, n_3)}$ is studied as an absorption time in a suitable auxiliary process.

In order to obtain the different l th-order moments in an efficient way, we define the Laplace–Stieltjes transform of $T_{(n_1, n_2, n_3)}$ as

$$\xi_{(n_1, n_2, n_3)}(z) = E[e^{-zT_{(n_1, n_2, n_3)}}], \quad \Re(z) \geq 0,$$

and the different l th-order moments of $T_{(n_1, n_2, n_3)}$ can be obtained as

$$E[T_{(n_1, n_2, n_3)}^l] = (-1)^l \frac{d^l}{dz^l} \xi_{(n_1, n_2, n_3)}(z) \Big|_{z=0}, \quad \forall l \geq 1.$$

By a first-step argument (omitted here since it is analogous to (B 2)), we obtain the system

$$\hat{\mathbf{g}}(z) = \hat{\mathbf{A}}(z) \hat{\mathbf{g}}(z) + \hat{\mathbf{a}}(z), \quad (\text{C } 1)$$

where the Laplace–Stieltjes transforms are stored in vectors $\hat{\mathbf{g}}(z)$, following the order given by the levels and sub-levels, and where the expressions for matrices $\hat{\mathbf{A}}(z)$ and $\hat{\mathbf{a}}(z)$ are omitted for brevity. By successive differentiation of the system in (C 1), we obtain the different l th-order moments $E[T_{(n_1, n_2, n_3)}^l]$ through an adapted version of algorithm 1, with the l th-order moments stored in the vectors $\hat{\mathbf{m}}^{(l)}$. We note that in the adapted version of algorithm 1 to solve (C 1), which is omitted, we need to deal with inverses of matrices with dimension $\hat{J}(k) = \#\hat{L}(k)$. The complexity of transitions between states does not allow us to gain further efficiency in our algorithms by working with inverses of matrices with the dimensions of the given sub-levels. However, in the special case $\gamma_- = 0$, that is, when de-phosphorylation is neglected, it is possible to improve the procedures so that the highest computational effort is placed on inverting matrices with the dimensions of sub-levels instead of levels, which would yield an algorithm 3, not described here.

Finally, we focus on the stationary distribution of the process, that is, the probabilities

$$\hat{\pi}_{(n_1, n_2, n_3)} = \lim_{t \rightarrow +\infty} \mathbb{P}((\hat{M}(t), \hat{D}(t), \hat{P}(t)) = (n_1, n_2, n_3)), \\ \forall (n_1, n_2, n_3) \in \hat{\mathcal{S}},$$

which do not depend on the initial state. Similar arguments to those considered in appendix B allow us to obtain the stationary distribution in a row vector $\hat{\pi} = (\hat{\pi}_0, \hat{\pi}_2, \dots, \hat{\pi}_{n_L})$, where $\hat{\pi}_k = (\hat{\pi}_0^k, \hat{\pi}_2^k, \dots, \hat{\pi}_{n_L-k}^k)$, and where row sub-vectors $\hat{\pi}_r^k$ contain, in an ordered manner, steady-state probabilities of states in sub-levels $l(k; r)$. An adapted version of algorithm 2 can be obtained, where the matrices $\mathbf{A}_{i,j'}$ in (B 1), would be now replaced by the matrices $\hat{\mathbf{A}}_{k,k'}$ previously defined. Once

these vectors are in hand, it is clear that

$\hat{\pi}_M$ = 'mean number of M complexes in steady state'

$$= \sum_{i=0}^{n_L} i \left(\sum_{k=0}^{n_L-i} \sum_{r=0}^{n_L-i-k} (\hat{\pi}_r^k)_i \right),$$

$\hat{\pi}_D$ = 'mean number of D complexes in steady state'

$$= \sum_{r=0}^{n_L} r \left(\sum_{k=0}^{n_L-r} \sum_{i=0}^{n_L-r-k} (\hat{\pi}_r^k)_i \right),$$

$\hat{\pi}_P$ = 'mean number of P complexes in steady state'

$$= \sum_{k=0}^{n_L} k \left(\sum_{r=0}^{n_L-k} \sum_{i=0}^{n_L-r-k} (\hat{\pi}_r^k)_i \right).$$

Appendix D. Local sensitivity analysis for the stochastic descriptors

The objective of this appendix is to develop a local sensitivity analysis to understand the effect that each of the (association, dissociation, phosphorylation or de-phosphorylation) rates (α_+ , α_- , β_+ , β_- , γ_+ and γ_-) has on the stochastic descriptors introduced (appendices B and C) for the DP and the IP models, in a given neighbourhood of parameter space. This selected neighbourhood of parameter space may be obtained from a parameter estimation of *in vitro* and *in silico* experiments, as shown in §3.1. Our aim then is to obtain the partial derivatives of our descriptors with respect to each parameter, so that these derivatives provide a measure of the effect of a perturbation of the parameters on the descriptors.

Sensitivity analysis for CTMC with absorbing states has been recently developed [22]. Although the Markov chains considered in our models are, in general, non-absorbing, the arguments in Caswell [22] can be clearly generalized to the CTMCs considered here; see Gómez-Corral & López-García [23] for how to adapt these arguments to structured Markov processes such as the ones considered in this study. We briefly explain how to adapt

them in what follows, while keeping the spirit of the matrix-analytic approach applied in previous sections.

We note that the descriptors of §3.3 are analysed in appendices B and C by following a matrix-oriented methodology, where algorithms 1 and 2 depend on the use of matrices $\mathbf{A}_{k,k}(z)$, $\mathbf{H}_k(z)$ and inverses $\mathbf{H}_k^{-1}(z)$, which are matrices that clearly depend on parameters (α_+ , α_- , β_+ , β_- , γ_+ , γ_-). Thus, when carrying out a local sensitivity analysis (in terms of partial derivatives) by adapting arguments of Caswell [22], one would need to compute the element-by-element partial derivatives of these matrices with respect to the parameters of interest. We note that given any matrix $\mathbf{B}_{m \times n}(\boldsymbol{\theta})$, that depends on some parameters vector $\boldsymbol{\theta}$, we denote by $\mathbf{B}^{(i)}(\boldsymbol{\theta})$ its element-by-element derivative with respect to $\theta_i \in \boldsymbol{\theta}$. It is possible to calculate the derivative of its inverse, $\mathbf{B}^{-1}(\boldsymbol{\theta})$, with respect to θ_i from $\mathbf{B}^{(i)}(\boldsymbol{\theta})$ as [50,51]

$$(\mathbf{B}^{-1})^{(i)}(\boldsymbol{\theta}) = -\mathbf{B}^{-1}(\boldsymbol{\theta})\mathbf{B}^{(i)}(\boldsymbol{\theta})\mathbf{B}^{-1}(\boldsymbol{\theta}).$$

We make use of this and other basic matrix calculus properties, as discussed in Caswell [22], to obtain algorithms 1S and 2S, which are given below, and can be obtained by sequentially differentiating all matrices in algorithms 1 and 2, respectively. Finally, the explicit expressions for matrices in these algorithms, consisting of the element-by-element partial derivative of the matrices defined in appendices B and C, with respect to any parameter, $\theta_i \in \{\alpha_+, \alpha_-, \beta_+, \beta_-, \gamma_+, \gamma_-\}$, are not reported here.

It is clear that, since our descriptors are stored in the vectors $\mathbf{m}^{N,(l)}$, $\hat{\mathbf{m}}^{(l)}$ (time to reach a threshold number of P complexes in the IP and DP models, respectively) and the quantities π_j and $\hat{\pi}_j$ (mean number of j complexes in steady state in the IP model ($j \in \{M, P\}$) and the DP model ($j \in \{M, D, P\}$), respectively), the objective in algorithms 1S and 2S is to obtain the derivative vectors $\mathbf{m}^{N,(l,\theta)}$, $\hat{\mathbf{m}}^{(l,\theta)}$, $\boldsymbol{\pi}^{(\theta)}$ and $\hat{\boldsymbol{\pi}}^{(\theta)}$. The first two vectors contain the derivatives of the l th-order moments of the time to reach a given threshold number of P complexes, and the last two yield the derivatives of the quantities π_j and $\hat{\pi}_j$, with respect to each rate $\theta_i \in \{\alpha_+, \alpha_-, \beta_+, \beta_-, \gamma_+, \gamma_-\}$.

Algorithm 1S (to obtain the derivative of the r th-order moments $E[T_{(n_1, n_2)}(N)^r]$ with respect $\theta_i \in \{\alpha_+, \alpha_-, \beta_+, \beta_-\}$)

$$\mathbf{H}_0^{N,(\theta)}(0) = -\mathbf{A}_{0,0}^{(\theta)}(0);$$

For $k = 1, \dots, N-1$:

$$\mathbf{H}_k^{N,(\theta)}(0) = -\mathbf{A}_{k,k}^{(\theta)}(0) - (\mathbf{A}_{k,k-1}^{(\theta)}(0)\mathbf{H}_{k-1}^N(0)^{-1}\mathbf{A}_{k-1,k}(0) - \mathbf{A}_{k,k-1}(0)\mathbf{H}_{k-1}^N(0)^{-1} \times \mathbf{H}_{k-1}^{N,(\theta)}(0)\mathbf{H}_{k-1}^N(0)^{-1}\mathbf{A}_{k-1,k}(0) + \mathbf{A}_{k,k-1}(0)\mathbf{H}_{k-1}^N(0)^{-1}\mathbf{A}_{k-1,k}^{(\theta)}(0));$$

$$\mathbf{m}_{N-1}^{N,(0,\theta)} = -\mathbf{H}_{N-1}^N(0)^{-1}\mathbf{H}_{N-1}^{N,(\theta)}(0)\mathbf{H}_{N-1}^N(0)^{-1}\mathbf{a}_{N-1}(0) + \mathbf{H}_{N-1}^N(z)^{-1}\mathbf{a}_{N-1}^{(\theta)}(z);$$

For $k = N-2, \dots, 1, 0$:

$$\mathbf{m}_k^{N,(0,\theta)} = -\mathbf{H}_k^N(0)^{-1}\mathbf{H}_k^{N,(\theta)}(0)\mathbf{H}_k^N(0)^{-1}\mathbf{A}_{k,k+1}(0)\mathbf{m}_{k+1}^{N,(0)} + \mathbf{H}_k^N(0)^{-1}\mathbf{A}_{k,k+1}^{(\theta)}(0)\mathbf{m}_{k+1}^{N,(0)} + \mathbf{H}_k^N(0)^{-1}\mathbf{A}_{k,k+1}(0)\mathbf{m}_{k+1}^{N,(0,\theta)};$$

For $j = 1, \dots, r$:

$$\mathbf{P}_0^{N,(j,\theta)} = \sum_{k=1}^j \binom{j}{k} (-1)^k (\mathbf{A}_{0,0}^{(k,\theta)}(0)\mathbf{m}_0^{N,(j-k)} + \mathbf{A}_{0,0}^{(k)}(0)\mathbf{m}_0^{N,(j-k,\theta)} + \mathbf{A}_{0,1}^{(k,\theta)}(0)\mathbf{m}_1^{N,(j-k)} + \mathbf{A}_{0,1}^{(k)}(0)\mathbf{m}_1^{N,(j-k,\theta)});$$

For $p = 1, \dots, N-1$:

$$\begin{aligned} \mathbf{P}_p^{N,(j,\theta)} &= \mathbf{A}_{p,p-1}^{(\theta)}(0)\mathbf{H}_{p-1}^N(0)^{-1}\mathbf{P}_{p-1}^{N,(j)} - \mathbf{A}_{p,p-1}(0)\mathbf{H}_{p-1}^N(0)^{-1}\mathbf{H}_{p-1}^{N,(\theta)}(0) \\ &\quad \times \mathbf{H}_{p-1}^N(0)^{-1}\mathbf{P}_{p-1}^{N,(j)} + \mathbf{A}_{p,p-1}(0)\mathbf{H}_{p-1}^N(0)^{-1}\mathbf{P}_{p-1}^{N,(j,\theta)} + \sum_{k=1}^j \binom{j}{k} (-1)^k \\ &\quad \times (\mathbf{A}_{p,p-1}^{(k,\theta)}(0)\mathbf{m}_{p-1}^{N,(j-k)} + \mathbf{A}_{p,p-1}^{(k)}(0)\mathbf{m}_{p-1}^{N,(j-k,\theta)} + \mathbf{A}_{p,p}^{(k,\theta)}(0)\mathbf{m}_p^{N,(j-k)} \\ &\quad + \mathbf{A}_{p,p}^{(k)}(0)\mathbf{m}_p^{N,(j-k,\theta)} + (1 - \delta_{p,N-1})(\mathbf{A}_{p,p+1}^{(k,\theta)}(0)\mathbf{m}_{p+1}^{N,(j-k)} + \mathbf{A}_{p,p+1}^{(k)}(0)\mathbf{m}_{p+1}^{N,(j-k,\theta)})); \end{aligned}$$

$$\mathbf{m}_{N-1}^{N,(j,\theta)} = -\mathbf{H}_{N-1}^N(0)^{-1}\mathbf{H}_{N-1}^{N,(\theta)}(0)\mathbf{H}_{N-1}^N(0)^{-1}(\mathbf{P}_{N-1}^{N,(j)} + (-1)^j \mathbf{a}_{N-1}^{(j)}(0)) + \mathbf{H}_{N-1}^N(0)^{-1}(\mathbf{P}_{N-1}^{N,(j,\theta)} + (-1)^j \mathbf{a}_{N-1}^{(j,\theta)}(0));$$

For $p = N-2, \dots, 1, 0$:

$$\begin{aligned} \mathbf{m}_p^{N,(j,\theta)} &= -\mathbf{H}_p^N(0)^{-1}\mathbf{H}_p^{N,(\theta)}(0)\mathbf{H}_p^N(0)^{-1}(\mathbf{P}_p^{N,(j)} + \mathbf{A}_{p,p+1}(0)\mathbf{m}_{p+1}^{N,(j)}) + \mathbf{H}_p^N(0)^{-1} \\ &\quad \times (\mathbf{P}_p^{N,(j,\theta)} + \mathbf{A}_{p,p+1}^{(\theta)}(0)\mathbf{m}_{p+1}^{N,(j)} + \mathbf{A}_{p,p+1}(0)\mathbf{m}_{p+1}^{N,(j,\theta)}); \end{aligned}$$

We point out that $\mathbf{m}^{N,(r,\theta)_k}$ and $\mathbf{A}^{(r,\theta)_k}(0)$ in algorithm 1S, which corresponds to the model with instantaneous phosphorylation, represent the derivatives of $\mathbf{m}_k^{N,(r)}$ and $\mathbf{A}_k^{(r)}(0)$, respectively, with respect θ_i , for $\theta_i \in \{\alpha_+, \alpha_-, \beta_+, \beta_-\}$.

Algorithm 2S (to obtain the derivative of π_M and π_P with respect $\theta_i \in \{\alpha_+, \alpha_-, \beta_+, \beta_-\}$)

$$\mathbf{H}_0^{(\theta)} = \mathbf{A}_{0,0}^{(\theta)};$$

For $k = 1, \dots, n_L - 1$:

$$\mathbf{H}_k^{(\theta)} = \mathbf{A}_{k,k}^{(\theta)} - (\mathbf{A}_{k,k-1}^{(\theta)} \mathbf{H}_{k-1}^{-1} \mathbf{A}_{k-1,k} - \mathbf{A}_{k,k-1} \times \mathbf{H}_{k-1}^{-1} \mathbf{H}_{k-1}^{(\theta)} \mathbf{H}_{k-1}^{-1} \mathbf{A}_{k-1,k} + \mathbf{A}_{k,k-1} \mathbf{H}_{k-1}^{-1} \mathbf{A}_{k-1,k}^{(\theta)});$$

$$\pi_{n_L}^{*(\theta)} = 0;$$

For $k = n_L - 1, \dots, 1, 0$:

$$\pi_k^{*(\theta)} = -(\pi_{k+1}^{*(\theta)} \mathbf{A}_{k+1,k} \mathbf{H}_k^{-1} + \pi_{k+1}^{*(\theta)} \mathbf{A}_{k+1,k}^{(\theta)} \mathbf{H}_k^{-1} - \pi_{k+1}^{*(\theta)} \mathbf{A}_{k+1,k} \mathbf{H}_k^{-1} \mathbf{H}_k^{(\theta)} \mathbf{H}_k^{-1});$$

For $k = 0, \dots, n_L$:

$$\pi_k^{(\theta)} = \frac{1}{\pi \mathbf{e}_{\#S}} (\pi_k^{*(\theta)} - \pi_k \pi^{*(\theta)} \mathbf{e}_{\#S});$$

$$\pi_M^{(\theta)} = \sum_{k=0}^{n_L} k \left(\sum_{j=0}^{n_L} (\pi_j^{(\theta)})_k \right);$$

$$\pi_P^{(\theta)} = \sum_{k=0}^{n_L} k (\pi_k^{(\theta)} \mathbf{e}_{J(k)});$$

References

- Lauffenburger DA, Linderman JJ. 1993 *Receptors: models for binding, trafficking, and signaling*, vol. 365. New York, NY: Oxford University Press.
- Starbuck C, Wiley HS, Lauffenburger DA. 1990 Epidermal growth factor binding and trafficking dynamics in fibroblasts: relationship to cell proliferation. *Chem. Eng. Sci.* **45**, 2367–2373. (doi:10.1016/0009-2509(90)80117-w)
- Tan WH, Popel AS, MacGabhann F. 2013 Computational model of VEGFR2 pathway to ERK activation and modulation through receptor trafficking. *Cell. Signal.* **25**, 2496–2510. (doi:10.1016/j.cellsig.2013.08.015)
- Olsson AK, Dimberg A, Kreuger J, Claesson-Welsh L. 2006 VEGF receptor signalling? in control of vascular function. *Nat. Rev. Mol. Cell Biol.* **7**, 359–371. (doi:10.1038/nrm1911)
- Alarcón T, Page KM. 2006 Stochastic models of receptor oligomerization by bivalent ligand. *J. R. Soc. Interface* **3**, 545–559. (doi:10.1098/rsif.2006.0116)
- MacGabhann F, Yang MT, Popel AS. 2005 Monte Carlo simulations of VEGF binding to cell surface receptors *in vitro*. *Biochim. Biophys. Acta Mol. Cell Res.* **1746**, 95–107. (doi:10.1016/j.bbamcr.2005.09.004)
- Currie J, Castro M, Lythe G, Palmer E, Molina-París C. 2012 A stochastic T cell response criterion. *J. R. Soc. Interface* **9**, 2856–2870. (doi:10.1098/rsif.2012.0205)
- Neuts MF. 1994 *Matrix-geometric solutions in stochastic models: an algorithmic approach*, 2nd edn. New York, NY: Courier Dover Publications.
- Latouche G, Ramaswami V. 1999 *Introduction to matrix analytic methods in stochastic modeling*. Philadelphia, PA: ASA-SIAM.
- Gómez-Corral A, López-García M. 2012 Extinction times and size of the surviving species in a two-species competition process. *J. Math. Biol.* **64**, 255–289. (doi:10.1007/s00285-011-0414-8)
- Gómez-Corral A, López-García M. 2012 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.* **64**, 236–259. (doi:10.1016/j.camwa.2012.01.058)
- Artalejo JR, Gomez-Corral A, López-García M, Molina-París C. 2016 Stochastic descriptors to study the fate and potential of naive T cell clonotypes in the periphery. *J. Math. Biol.* **74**, 673–708. (doi:10.1007/s00285-016-1020-6)
- Alarcón T, Page KM. 2007 Mathematical models of the VEGF receptor and its role in cancer therapy. *J. R. Soc. Interface* **4**, 283–304. (doi:10.1098/rsif.2006.0170)
- Cross MJ, Dixelius J, Matsumoto T, Claesson-Welsh L. 2003 VEGF-receptor signal transduction. *Trends. Biochem. Sci.* **28**, 488–494. (doi:10.1016/s0968-0004(03)00193-2)
- MacGabhann F, Popel AS. 2007 Dimerization of VEGF receptors and implications for signal transduction: a computational study. *Biophys. Chem.* **128**, 125–139. (doi:10.1016/j.bpc.2007.03.010)
- Ruch C, Skiniotis G, Steinmetz MO, Walz T, Ballmer-Hofer K. 2007 Structure of a VEGF–VEGF receptor complex determined by electron microscopy. *Nat. Struct. Mol. Biol.* **14**, 249–250. (doi:10.1038/nsmb1202)
- Bel G, Munsy B, Nemenman I. 2009 The simplicity of completion time distributions for common complex biochemical processes. *Phys. Biol.* **7**, 016003. (doi:10.1088/1478-3975/7/1/016003)
- Kulkarni VG. 1996 *Modeling and analysis of stochastic systems*. Texts in Statistical Science Series. London, UK: Chapman & Hall.
- Gillespie DT. 1977 Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361. (doi:10.1021/j100540a008)
- Gillespie CS. 2009 Moment-closure approximations for mass-action models. *IET. Syst. Biol.* **3**, 52–58. (doi:10.1049/iet-syb:20070031)
- Hespanha J. 2008 Moment closure for biochemical networks. In *3rd Int. Symp. on Communications, Control and Signal Processing, 2008, ISCCSP 2008, St Julian's, Malta, 12–14 March 2008*, pp. 142–147. IEEE.
- Caswell H. 2011 Perturbation analysis of continuous-time absorbing Markov chains. *Numer. Linear Algebra Appl.* **18**, 901–917. (doi:10.1002/nla.791)
- Gómez-Corral A, López-García M. 2018 Perturbation analysis in finite LD-QBD processes and applications to epidemic models. *Numer. Linear Algebra Appl.* **e2160**. (doi:10.1002/nla.2160)
- Weisz PB. 1973 Diffusion and chemical transformation: an interdisciplinary excursion. *Science* **179**, 433–440. (doi:10.1126/science.179.4072.433)
- DeLisi C. 1980 The biophysics of ligand–receptor interactions. *Q. Rev. Biophys.* **13**, 201–230. (doi:10.1017/S0033583500001657)
- Shoup D, Szabo A. 1982 Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys. J.* **40**, 33–39. (doi:10.1016/s0006-3495(82)84455-x)
- Keizer J. 1985 Theory of rapid bimolecular reactions in solution and membranes. *Acc. Chem. Res.* **18**, 235–241. (doi:10.1021/ar00116a002)
- Berezhkovskii AM, Szabo A. 2013 Effect of ligand diffusion on occupancy fluctuations of cell-surface receptors. *J. Chem. Phys.* **139**, 121910. (doi:10.1063/1.4816105)
- Faro J, Castro M, Molina-París C. 2017 A unifying mathematical framework for experimental TCR–pMHC kinetic constants. *Sci. Rep.* **7**, 46741. (doi:10.1038/srep46741)
- Ewan LC, Jopling HM, Jia H, Mittar S, Bagherzadeh A, Howell GJ, Walker JH, Zachary IC, Ponnambalam S. 2006 Intrinsic tyrosine kinase activity is required for vascular endothelial growth factor receptor 2 ubiquitination, sorting and degradation in endothelial cells. *Traffic* **7**, 1270–1282. (doi:10.1111/j.1600-0854.2006.00462.x)
- Mittar S, Ulyatt C, Howell GJ, Bruns AF, Zachary I, Walker JH, Ponnambalam S. 2009 VEGFR1 receptor tyrosine kinase localization to the Golgi apparatus is calcium-dependent. *Exp. Cell. Res.* **315**, 877–889. (doi:10.1016/j.yexcr.2008.12.020)
- Almqvist N, Bhatia R, Primbs G, Desai N, Banerjee S, Lal R. 2004 Elasticity and adhesion force mapping reveals real-time clustering of growth factor receptors and associated changes in local cellular rheological properties. *Biophys. J.* **86**, 1753–1762. (doi:10.1016/s0006-3495(04)74243-5)
- Erickson J, Goldstein B, Holowka D, Baird B. 1987 The effect of receptor density on the forward rate constant for binding of ligands to cell surface receptors. *Biophys. J.* **52**, 657–662. (doi:10.1016/s0006-3495(87)83258-7)
- MacGabhann F, Popel AS. 2005 Differential binding of VEGF isoforms to VEGF receptor 2 in the presence of neuropilin-1: a computational model. *Am. J. Physiol. Heart Circul. Physiol.*

- 288**, H2851–H2860. (doi:10.1152/ajpheart.01218.2004)
35. Linderman JJ, Lauffenburger DA. 2013 *Receptor/ligand sorting along the endocytic pathway*, vol. 78. Berlin, Germany: Springer Science & Business Media.
 36. Aird WC. 2007 Phenotypic heterogeneity of the endothelium I. structure, function, and mechanisms. *Circ. Res.* **100**, 158–173. (doi:10.1161/01.RES.0000255691.76142.4a)
 37. MacGabhann F, Popel AS. 2004 Model of competitive binding of vascular endothelial growth factor and placental growth factor to VEGF receptors on endothelial cells. *Am. J. Physiol. Heart Circul. Physiol.* **286**, H153–H164. (doi:10.1152/ajpheart.00254.2003)
 38. Bikfalvi A, Sauzeau C, Moukadir H, Maclouf J, Busso N, Bryckaert M, Plouet J, Tobelem G. 1991 Interaction of vasculotropin/vascular endothelial cell growth factor with human umbilical vein endothelial cells: binding, internalization, degradation, and biological effects. *J. Cell. Physiol.* **149**, 50–59. (doi:10.1002/jcp.1041490108)
 39. Ewan LC, Jopling HM, Jia H, Mittar S, Bagherzadeh A, Howell GJ, Walker JH, Zachary IC, Ponnambalam S. 2006 Intrinsic tyrosine kinase activity is required for vascular endothelial growth factor receptor 2 ubiquitination, sorting and degradation in endothelial cells. *Traffic* **7**, 1270–1282. (doi:10.1111/j.1600-0854.2006.00462.x)
 40. Huang X, Gottstein C, Brekken RA, Thorpe PE. 1998 Expression of soluble VEGF receptor 2 and characterization of its binding by surface plasmon resonance. *Biochem. Biophys. Res. Commun.* **252**, 643–648. (doi:10.1006/bbrc.1998.9717)
 41. Waltenberger J, Claesson-Welsh L, Siegbahn A, Shibuya M, Heldin CH. 1994 Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor. *J. Biol. Chem.* **269**, 26 988–26 995.
 42. Imoukhuede PI, Popel AS. 2012 Expression of VEGF receptors on endothelial cells in mouse skeletal muscle. *PLoS ONE* **7**, e44791. (doi:10.1371/journal.pone.0044791)
 43. Kut C, MacGabhann F, Popel AS. 2007 Where is VEGF in the body? A meta-analysis of VEGF distribution in cancer. *Br. J. Cancer* **97**, 978–985. (doi:10.1038/sj.bjc.6603923)
 44. MacGabhann F, Popel AS. 2007 Interactions of VEGF isoforms with VEGFR-1, VEGFR-2, and neuropilin *in vivo*: a computational model of human skeletal muscle. *Am. J. Physiol. Heart Circul. Physiol.* **292**, H459–H474. (doi:10.1152/ajpheart.00637.2006)
 45. Wendel Clegg L, MacGabhann F. 2015 Site-specific phosphorylation of VEGFR2 is mediated by receptor trafficking: insights from a computational model. *PLoS. Comput. Biol.* **11**, e1004158. (doi:10.1371/journal.pcbi.1004158)
 46. Shi T *et al.* 2016 Conservation of protein abundance patterns reveals the regulatory architecture of the EGFR-MAPK pathway. *Sci. Signal.* **9**, rs6. (doi:10.1126/scisignal.aaf0891)
 47. Napione L *et al.* 2012 Unraveling the influence of endothelial cell density on VEGF-A signaling. *Blood* **119**, 5599–5607. (doi:10.1182/blood-2011-11-390666)
 48. Wheeler DL, Dunn EF, Harari PM. 2010 Understanding resistance to EGFR inhibitors: impact on future treatment strategies. *Nat. Rev. Clin. Oncol.* **7**, 493–507. (doi:10.1038/nrclinonc.2010.97)
 49. Abate J, Whitt W. 1992 Numerical inversion of probability generating functions. *Operat. Res. Lett.* **12**, 245–251. (doi:10.1016/0167-6377(92)90050-D)
 50. Magnus JR, Neudecker H. 1985 Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. *J. Math. Psychol.* **29**, 474–492. (doi:10.1016/0022-2496(85)90006-9)
 51. Neudecker H, Magnus JR. 1988 *Matrix differential calculus with applications in statistics and econometrics*. New York, NY: John Wiley & Sons.