LiNA – A Graphical Matlab Tool for Analyzing Intrinsic Noise in Biochemical Reaction Networks

Tutorials and Theoretical Background

Ronny Straube and Axel von Kamp

Analysis and Redesign of Biological Networks Group
Max Planck Institute for Dynamics of Complex Technical Systems
Sandtorstrasse 1, 39106 Magdeburg, Germany

August 19, 2013
# Contents

1. **Installation and License Information** 4  
   1.1. Availability .................................................. 4  
   1.2. Installation .................................................. 4  
   1.3. License Information ........................................ 4  
   1.4. Contact Information ........................................ 4  

2. **Tutorials** 5  
   2.1. Noise Reduction Through Receptor Dimerization ............... 5  
      2.1.1. Setting up the Model ................................... 6  
      2.1.2. Taking A First Look ................................... 8  
      2.1.3. Volume Dependence of the Fluctuations ................. 8  
      2.1.4. Fano Factor ............................................. 9  
      2.1.5. Parameter Sets, Sessions and Figure Export ............ 10  
   2.2. Ultrasensitivity in the Goldbeter-Koshland Model ............ 11  
      2.2.1. Analysis of Systems with Conservation Relations ....... 12  
      2.2.2. Correlation Analysis ................................... 13  
   2.3. Bistability in Covalent Modification Cycles .................. 15  
      2.3.1. Model Setup ............................................. 16  
      2.3.2. Strategies for Dealing with Multiple Steady States .... 16  

3. **Technical Remarks** 19  
   3.1. Solving Eqs. (3.1) and (3.2) ................................ 19  
   3.2. Rescan vs. All Steady States ................................ 20  
   3.3. Data Management ............................................. 20  
   3.4. Output at MATLAB Console and \textit{vanKampen.out} ........... 20  

4. **Theoretical Background** 21  
   4.1. Reaction Networks and the Master Equation .................. 21  
   4.2. The Linear Noise Approximation ............................. 22  
   4.3. Calculation of Stochastic Quantities ....................... 24  
   4.4. Dealing with Mass Conservation ............................. 25  
   4.5. Examples ..................................................... 27  
      4.5.1. Noise Reduction Through Receptor Dimerization ....... 28  
      4.5.2. Receptor-Ligand Binding with Mass Conservation ....... 32  

A. **Derivation of the Linear Noise Approximation** 36  
   A.1. Transformation of the Master Equation ....................... 36
1. Installation and License Information

1.1. Availability

LiNA can be downloaded from

http://www.mpi-magdeburg.mpg.de/projects/LiNA

1.2. Installation

Unzip the LiNA.zip archive to your HOME directory or to any other directory of your choice. This will create a directory called LiNA containing the program files. Example networks described in the Tutorial Chapter (→ Chapter 2) can be found in the Examples folder which is located in the LiNA directory.

1.3. License Information

LiNA is free software. It is published under the terms of the GNU General Public License v3. A full copy of the license terms and conditions can be found in the file named COPYING which is located in the LiNA directory.

1.4. Contact Information

Please send your comments, suggestions or bug reports to

lina@mpi-magdeburg.mpg.de
2. Tutorials

If you are familiar with the linear noise approximation you may want to have a look at the following Tutorials which cover all of the functionality implemented in LiNA. If you don’t know the linear noise approximation you may find it useful to first read the background information in Chapter 4 and Appendix A before coming back.

Each tutorial highlights different aspects of LiNA’s functionality. For example, in the first tutorial one learns how to setup a model, how to define compounded parameters and how to use the graphical user interface for analyzing basic system properties. This tutorial also covers data management tasks such as loading/saving sessions and exporting figures. The second tutorial shows how LiNA handles systems with conservation relations and how to analyze correlations. The last tutorial discusses the functionality provided to analyze systems with multiple steady states. This tutorial refers to LiNA v1.

2.1. Noise Reduction Through Receptor Dimerization

In the first tutorial we shall analyze how dimerization affects the steady state fluctuations of receptor molecules. To this end, we consider the reaction system

\[ \emptyset \xrightleftharpoons{k_s} R \]

\[ R + R \xrightleftharpoons{k^+} R_2 \]

\[ R_2 \xrightarrow{k_{d2}} \emptyset \]

which represents an extension of the model considered by Hayot and Jayaprakash [5]. In the absence of dimerization, synthesis and degradation of receptor monomers, as described by Eq. (2.1), lead to Poissonian fluctuations where the average number of monomers is equal to the variance so that the Fano factor is 1. Hayot and Jayaprakash showed that taking dimerization into account (Eq. 2.2), but neglecting dimer degradation (Eq. 2.3), the steady state fluctuations for both, monomers and dimers, remain Poissonian within the linear noise approximation [5]. In Section 4.5.1 we re-analyze the system considered by Hayot and Jayaprakash, but for a finite lifetime of receptor dimers \( (k_{d2} \neq 0) \), which is shown to lead to sub-Poissonian fluctuations for monomers and dimers. Interestingly, these fluctuations can be minimized either for monomers or for dimers depending on the parameter regime. In the following, we will use LiNA to visualize some of the theoretical results obtained in Section 4.5.1.
2. Tutorials

2.1.1. Setting up the Model

To start LiNA, startup MATLAB first. If necessary change to the LiNA directory (see Section 1.2). Then, type `startlina` at the MATLAB prompt and press ENTER. This will bring up the Equations window (see Fig. 2.1) which allows you to enter new reactions, specify initial parameter values and to define new parameters through parameter substitutions. You can also load an existing model using the Load... button. In the opening file selection window change to the Examples folder and select the file `receptor_dimerization.mat` after which the Equations window will look as in Fig. 2.2.

The Equation Editor

In the first and in the third reaction the digit ‘0’ is used to denote unspecified sources and sinks. Irreversible reactions are entered using a single arrow (->), reversible reactions using a double arrow (<->). New reactions can be added with the Add Reaction button. Further information on this topic can be obtained by pressing the ? button in the upper right corner of the Equations window. Reactions can be deleted using the delete button (red circle with the white cross) which is located to the left of each reaction field.

For each reaction at least one parameter has to be specified; two for reversible reactions. The parameter in the left column refers to the forward reaction (→) whereas that in the right column refers to the backward reaction (←). Parameter names are predefined, but can be changed (only alphanumeric characters).

A Note on Units

The unit of a parameter depends on the order of the reaction and on the chosen unit for the species concentration. The default unit for the species concentration is $\mu$M = $10^{-6}$mol/liter. In that case, the value of the parameter $k_3$, which defines the rate with which monomers are synthesized ($0 \rightarrow R$), is expected to be entered in units of $\mu$M/time where ‘time’ could be seconds, minutes or hours. The particular time unit is unimportant.
since, in its current form, LiNA only calculates noise characteristics under steady state conditions which do not depend on the specific time scale. However, it is important that all parameter values are given with respect to the same time scale. The first order rate constants $k_{d1}$ ($kd1$), $k_{d2}$ ($kd2$) and $k^-$ ($km$), which describe the degradation of monomers, dimers and the dissociation of the dimer, respectively, have the unit $1/time$ where the time unit has to be the same as that used for $k_s$. Finally, the second-order rate constant $k^+$ ($kp$), which describes the association of two monomers into one dimer, is expected to be entered in units of $1/(\mu M \cdot time)$. If the species concentration is changed, e.g. from $\mu M$ to $nM$, the values for zeroth and second order rate constants ($k_s$ and $k^+$), are re-interpreted accordingly, i.e. as $nM/time$ for $k_s$ and $1/(nM \cdot time)$ for $k^+$. The values of first order rate constants are not affected by such a change.

### Why do I need to define the reaction volume?

Calculation of the standard deviation and, hence, calculation of the coefficient of variation depends on the reaction volume (cf. Eq. 4.17). The default volume is set to $V = 10^{-15} l = 1 fl = 1 \mu m^3$ which corresponds to the typical volume of a bacterial cell. In such a small volume a concentration of $x = 10nM$ means that there are approximately $n = V \cdot N_A \cdot x \approx 6$ particles in the system ($N_A = 6.022 \cdot 10^{23}$ particles/mol). This is the regime where one can expect stochastic effects to become important.

### Parameter Substitutions

In Subsection 4.5.1 we show that the steady states as well as the steady state fluctuations of the reaction system in Eqs. (2.1) - (2.3) only depend on certain parameter combinations.
which can be chosen as (cf. Eq. 4.49)

\[
R_s = \frac{k_s}{k_{d1}}, \quad \alpha = \frac{k_{d1}}{k_{d2}}, \quad \beta = \frac{k_{d2}}{k^-}, \quad K_D = \frac{k^-}{k^+}. \tag{2.4}
\]

In order to compare the results, generated by LiNA, with the theoretical results from Subsection 4.5.1 it is, thus, helpful to be able to define the same parameters in LiNA. This is the purpose of the Parameter Substitutions field shown in Figure 2.2. Using the expressions in Eq. (2.4) the parameters \(k_{d2}, k_{d1}, k_s\) and \(k^+\) have been replaced by \(R_s (Rs), \alpha (a), \beta (b)\) and \(K_D (Kd)\) according to

\[
k_{d2} = \beta k^-, \quad k_{d1} = \alpha \beta k^-, \quad k_s = R_s \alpha \beta k^-, \quad k^+ = k^-/K_D.
\]

The biological meaning of these parameters is immediately evident: \(R_s\) represents the steady state level of monomers in the absence of dimerization, \(\alpha\) differs from 1 if monomers and dimers are degraded at different rates, \(\beta\) compares the time scales for the dimer degradation and dimer dissociation, and \(K_D\) represents the dimer dissociation constant.

### 2.1.2. Taking A First Look

If the model setup is finished the analysis can be started by pressing the **Start** button (Fig. 2.2). This will bring up two new windows, denoted as Control and Plots (Fig. 2.3), while the Equations windows becomes ‘frozen’. The first quantity to be plotted is always the **Mean Values** of all species as a function of the first entry in the parameter list of the Parameters field (A), \(Kd\) in this case. The parameter list contains the newly defined parameters (cf. Eq. 2.4) as well as \(k^- (km)\) which has not been substituted. Note, however, that neither the steady states nor any of the stochastic quantities depends explicitly on \(k^-\). You can readily check this by choosing \(k^- (km)\) from the Parameter pull-down list in the X-Axis:Variable Parameter field (B). Then you should see 2 straight lines parallel to the x-axis.

### 2.1.3. Volume Dependence of the Fluctuations

To see how the reaction volume affects the magnitude of steady state fluctuations due to intrinsic reaction noise choose again \(Kd\) from the Parameter pull-down list in the X-Axis:Variable Parameter field (B) and select Mean with Errorbars from the pull-down list in the Y-Axis field (C). This should add some errorbars to the steady state curves. The errorbars represent the standard deviation as defined in Section 4.3. Now, if the reaction volume is decreased the average number of particles also decreases and particle fluctuations become larger. Conversely, if the reaction volume is increased particle fluctuations become reduced. You can test this idea by changing the value of the reaction volume: Select Vol. in the parameter list of the Parameters field (A) and change the value from \(1e-15\) to \(1e-14\) and, subsequently, to \(1e-16\). Don’t forget to press Apply Changes for parameter changes to take effect.
2.1.4. Fano Factor

If monomers and dimers are degraded at the same rate ($\alpha = 1$) and if dimer degradation occurs at a much slower rate than dimer dissociation ($\beta \ll 1$) the Fano factors of monomers and dimers can be approximated by (cf. Eqs. 4.55 and 4.56)

$$F_1 = \frac{\sigma^2_1}{x_1^2} \approx 1 - \frac{x_1^2/K_D}{(1 + 4x_1^2/K_D)^2}$$

$$F_2 = \frac{\sigma^2_2}{x_2^2} \approx 1 - \frac{1}{4} \frac{(4x_1^2/K_D)^2}{(1 + 4x_1^2/K_D)^2}$$

where $x_1^*$ and $x_2^*$ are the steady state values of monomers (Eq. 4.47) and dimers (Eq. 4.48), respectively. Hence, fluctuations in both, monomers and dimers, are sub-Poissonian. However, whereas monomer fluctuations, again, become Poissonian ($F_1 \approx 1$) as $x_1^* \gg K_D/4$ dimer fluctuations are reduced and the Fano factor approaches $F_2 \approx 3/4$.

To reproduce these results with LiNA choose $R_s$ (Rs) from the Parameter pull-down list in the X-Axis:Variable Parameter field (B), change the Range from 0.01 to 1000, change to log scale and choose Fano Factor from the pull-down list in the Y-Axis field (C). This should produce two curves as in Figure 2.4. Convince yourself that changing the binding affinity $K_D$ (Kd) merely shifts these curves to the left (if Kd is decreased) or to the right (if Kd is increased). Also, observe that increasing $\beta$ while keeping $\alpha = 1$ fixed successively shifts the Fano factor curve for monomers below that for dimers, so that monomer fluctuations are reduced. Now, keep $\beta = 0.01$ fixed and reduce $\alpha$ so that $\alpha \ll 1$. This reduces the Fano factor for monomers at intermediate values of $R_s$ where dimer fluctuations are still approximately Poissonian (region A in Fig. 2.5). At large
2. Tutorials

Figure 2.4.: Fano factors for monomers (R) and dimers (R2) as a function of $R_s = k_s/k_{d1}$ for $K_D = 1\mu M$, $\alpha = 1$ and $\beta = 0.01$.

values of $R_s$ the situation is reversed (region B in Fig. 2.5). Hence, if dimers are degraded at a much higher rate than monomers it depends on the value of $R_s$ whether fluctuations in monomers or dimers are reduced.

2.1.5. Parameter Sets, Sessions and Figure Export

As you keep playing around with different parameter combinations LiNA will add a corresponding Parameter set for each distinct parameter combination. The number of available parameter sets is shown in round brackets in the Parameters field. To choose a particular parameter set, simply change the number in the Parameter Set: field and press Enter. When you change the variables along the x-axis or the y-axis or if you change parameters LiNA adds a new figure to the plot list which can be accessed from the pull-down menu in the Plots window (D in Fig. 2.3).

The plots are labeled according to the format

```
parameter on the x-axis: quantity on the y-axis {parameter set}.
```

For example, in Fig. 2.5 the plot shows the Fano factor as a function of $R_s$ using parameter set 3. Note that parameter sets cannot be deleted. You can, however, delete a plot by pressing the Delete button next to the pull-down menu in the Plots window. If you also want to delete the calculations associated with a plot you have to use the delete button (red circle with a white cross) next to the Parameter pull-down list in the X-Axis: Variable Parameter field (B in Fig. 2.3).
2. Tutorials

If you wish to continue the analysis at a later time you can save the current session including parameter sets and plots using the *save* button (disk symbol right above the Y-Axis field) in the Control window. Saved sessions can be loaded using the Resume Session button from the startup window of LiNA (Fig. 2.1).

If you want to save a figure for further processing in MATLAB you can either directly save the plot as a MATLAB figure file (using the disk symbol in the Plots window) or you copy the contents of the current plot (using the copy symbol labeled by E in Fig. 2.3) which creates a new MATLAB figure so that you can add comments, change fonts or convert the figure to another format before saving. Note that the plots in a copied figure will not receive any further updates as you continue with your analysis.

2.2. Ultrasensitivity in the Goldbeter-Koshland Model

The Goldbeter-Koshland model is a classical model for the description of covalent modification systems [3] where a substrate molecule is interconverted between an unmodified (S) and a modified (S*) form by a pair of opposing converter enzymes. In the case of phosphorylation / dephosphorylation cycles the converter enzymes are called kinase (K) and phosphatase (P). In the simplest case the enzyme-catalyzed reactions are modeled as irreversible Michaelis-Menten type reactions of the form

\[
S + K \overset{k_1}{\underset{k_{-1}}{\rightleftharpoons}} S-K \overset{k_1}{\rightarrow} S^* + K, \quad S^* + P \overset{k_2}{\underset{k_{-2}}{\rightarrow}} S^*-P \overset{k_3}{\rightarrow} S + P. \tag{2.5}
\]
Here, we have assumed that the reactions occur under \textit{in vitro} conditions so that synthesis and degradation of substrate and converter enzymes can be neglected. Hence, there are three conservation relations

\[
[S] + [S^*] + [S-K] + [S^*-P] = S_T
\]
\[
[K] + [S-K] = K_T
\]
\[
[P] + [S^*-P] = P_T
\]

corresponding to the conservation of the total amounts of substrate ($S_T$), kinase ($K_T$) and phosphatase ($P_T$).

If the total substrate concentration is much larger than that of either converter enzyme ($S_T \gg \max(K_T, P_T)$) one can describe the dynamics of the modified form of the substrate by the effective equation [3]

\[
\frac{d[S^*]}{dt} \approx k_1 \frac{S_T - [S^*]}{K_1 + S_T - [S^*]} - k_2 \frac{[S^*]}{K_2 + [S^*]}
\]

where $K_1 = (k_1 + k_1^-)/k_1^+$ and $K_2 = (k_2 + k_2^-)/k_2^+$ denote the Michaelis-Menten constants of the kinase and the phosphatase, respectively. A hallmark of the Goldbeter-Koshland model is that, under steady state conditions $d[S^*]/dt = 0$, it can generate highly sigmoidal response curves if both converter enzymes operate in saturation (Fig. 2.7), i.e. if $\max(K_1, K_2) \ll S_T$. This phenomenon is called zero-order ultrasensitivity.

### 2.2.1. Analysis of Systems with Conservation Relations

To setup the model start LiNA from the MATLAB prompt with \texttt{startlina} and enter reactions and parameters as shown in Fig. 2.6. Alternatively, you may load the file Gold-
betet_Koshland_model.mat from the Examples folder. Note that the 'on' rate constants $k_i^+$ and $k_i^-$ have been replaced by the associated Michaelis-Menten constants according to $k_i^+ = (k_i + k_i^-)/K_i$ ($i = 1, 2$) where $K_i \leftrightarrow K_{mi}$. After starting the calculation (press Start) you will see three new parameters ($C_K$, $C_{Sp}$ and $C_P$) in the parameter list of the Parameters field (Fig. 2.7). LiNA automatically adds a new parameter for each conservation relation it detects in the network. Hence, $C_K$, $C_{Sp}$ and $C_P$ correspond to the total concentrations of kinase, substrate and phosphatase. LiNA displays the corresponding conservation relation in the X-Axis: Variable Parameter field of the Control window once you select one of these parameters from the Parameter list.

As networks become larger the number of species also increases. By default, LiNA plots the mean values of all species in the Plots window after startup, which can be confusing. Species can be selected or deselected by pressing the corresponding species button in the Variables field of the Control window. For large systems it can be faster to deselect all species first and then select only those species that are of interest: Use the right mouse button while holding the mouse pointer over the Variables field name.

To generate the ultrasensitive response curves shown in Fig. 2.7, choose $C_K$ from the Parameter list in the X-Axis: Variable Parameter field and change the value of $C_{Sp}$ from 1 to 10 so that $S_T \gg P_T$. Note that the condition $\max(K_1, K_2) \ll S_T$ is already fulfilled. Now, adjust the upper boundary of the Range to 2 and increase the number of Intervals from 10 to 40. Observe that both, $S (S)$ and $Sp (S^*)$, exhibit a sharp transition near $C_K=1$ ($K_T = 1$). This suggests that the steady state fluctuations of these two species are large near $C_K=1$. Convince yourself that this is true: Specifically, observe that the Fano factors of $S$ and $Sp$ change by a factor of more than 20 near $C_K=1$ (Choose Fano Factor from the pull-down list in the Y-Axis field of the Control window). In contrast, the coefficient of variation, which measures ratio between the standard deviation and the mean value, changes much less dramatically. (Choose Coefficient of Variation from the pull-down list in the Y-Axis field of the Control window)

2.2.2. Correlation Analysis

As systems become larger it can be interesting to study correlations among the species. As Figure 2.8 shows the strength of correlations changes as parameters are varied. There are two ways to start the correlation analysis. First, you can select Correlation from the pull-down list of the Y-Axis field in the Control window. By default, this will only plot the correlation coefficient among the first two species of the Variables field. Alternatively, you can press the Update Plot button in the Correlations field of the Control window which has the same effect.

To add new species pairs simply choose the corresponding species from the pull-down lists in the Correlations field and press Add. You may also select All species or all Selected species at once and press Add. For example, the combination Species/Selected will add all distinct pairs between Species and the selected species in the Variables field of the Control window. Similarly, the combination All/All would add all $n(n - 1)/2$ distinct species pairs (where $n$ is the total number of species in the system) to the list in the Correlations field. Species pairs in the Correlations field can be highlighted by clicking
2. Tutorials

Figure 2.7.: Ultrasensitivity in the Goldbeter-Koshland model for $K_1 = K_2 = 0.1\mu M$, $k_1 = k_2 = 1/s$, $S_T(C_{\text{Sp}}) = 10\mu M$ and $P_T(C_{\text{P}}) = 1\mu M$. Note that the curves for SK and SpP overlap.

on them (press the Shift or Control key for multiple selection). Highlighted species are removed via a context menu that opens upon a right-click with the mouse button within the Correlations field. This method is especially useful if you want to remove a small number of species pairs. However, imagine you have all $n(n-1)/2$ species pairs in the list and you want to remove only those that contain a particular species, say X. In that case it might be faster to choose the combination X/All from the pull-down list and press the Remove button.

Despite the fact that the existence of a correlation between species A and B, in general, does not imply a sense of interaction between them, the analysis of correlations can still provide some hints for the functionality of a network. To this end, delete the $S/K$ pair from the Correlations list in the Control window and add the pairs $S/Sp$, $K/Sp$ and $P/Sp$ and press Update Plot. As you can see the modified and the unmodified forms of the substrate are always anti-correlated (Fig. 2.8) which is a direct consequence of the approximate conservation relation for the substrate $S_T \approx [S] + [S^*]$, which holds under substrate excess. As the total substrate concentration is reduced this anti-correlation becomes weaker. (Try reducing $C_{\text{Sp}}$!) The almost perfect correlation between $S^*$ (Sp) and $K$ beyond the transition point $C_K=1$ means that essentially all of the modified substrate is generated through the reaction $S-K \rightarrow S^* + K$ in the parameter region $C_K>1$ and not through the dissociation of the $S^*-P$ complex - in agreement with the fact that the correlation coefficient for $S^*$ (Sp) and $P$ is essentially zero for $C_K>1$.  

14
2. Tutorials

Figure 2.8.: Correlation coefficients for selected species pairs in the Goldbeter-Koshland model (Eq. 2.5). Parameter values are the same as those used in Fig. 2.7.

2.3. Bistability in Covalent Modification Cycles

Here, we consider an extension of the Goldbeter-Koshland model which has been recently analyzed in the context of bistability and hysteresis [8]. The elementary reactions read

\[
S + KL \xrightleftharpoons{\kappa_1}{\kappa_1^-} SKL \xrightarrow{\kappa_1^+} S^* + KL
\]

\[
S + P \xrightleftharpoons{\kappa_2}{\kappa_2^-} S^*P \xrightarrow{\kappa_3} S + P
\]

\[
K + L \xrightleftharpoons{\kappa_3^+}{\kappa_3^-} KL
\]

\[
P + L \xrightleftharpoons{\kappa_4^+}{\kappa_4^-} PL
\]

where $L$ represents an allosteric effector that reciprocally affects the activities of the converter enzymes. Specifically, binding of $L$ is assumed to activate the kinase and to inactivate the phosphatase.

It can be shown that the steady states of this system depend on the Michaelis-Menten constants of the kinase ($K_1 = (k_1 + k_1^-)/k_1^+$) and the phosphatase ($K_2 = (k_2 + k_2^-)/k_2^+$), on the dissociation constants of the enzyme-effector complexes $KL$ ($K_{D3} = k_3^+ / k_3^-$) and $PL$ ($K_{D4} = k_4^- / k_4^+$), on the catalytic rate constants of the kinase ($k_1$) and the
Abbildung 2.9.: Steady state response curve for the extended Goldbeter-Koshland model in Eq. (2.6). In the region between the two saddle-node bifurcations (SN) two stable steady states (solid lines) coexist with one unstable steady state (dashed line). Parameter values: $K_1 = 1 \mu M$, $K_2 = 0.01 \mu M$, $K_{D3} = 10 \mu M$, $K_{D4} = 0.01 \mu M$, $k_1 = 10/s$, $k_2 = 0.1/s$, $S_T = 3 \mu M$, $K_T = P_T = 1 \mu M$.

phosphatase ($k_2$) as well as on the total concentrations of substrate ($S_T$) and converter enzymes ($K_T$, $P_T$). If certain conditions among these parameters are met this system can exhibit bistability for a range of effector concentrations (Fig. 2.9).

2.3.1. Model Setup

Load the `bistable_Goldbeter_Koshland_model.mat` file from the `Examples` folder. Note that, in the `Parameter Substitutions` field, the ‘on’ rate constants $k_i^+$ ($i = 1, \ldots, 4$) have been replaced by Michaelis-Menten ($K_1$, $K_2$) and dissociation constants ($K_{D3}$, $K_{D4}$) according to

$$k_i^+ = \frac{k_i + k_i^-}{K_i}, \quad \text{for } i = 1, 2 \quad \text{and} \quad k_i^+ = \frac{k_i^-}{K_{Di}}, \quad \text{for } i = 3, 4.$$ 

After starting the calculations you will find 4 new parameters in the `Parameters` list of the `Control` window which correspond, respectively, to the total concentration of substrate ($C_{Sp}$), converter enzymes ($C_K$ and $C_P$) and allosteric effector ($C_L$). Convince yourself that the steady states do not depend on the backward rate constants $k_i^-$ (km).

2.3.2. Strategies for Dealing with Multiple Steady States

LiNA cannot generate bifurcation diagrams, as in Fig. 2.9, directly which would require continuation methods [6]. Instead, LiNA uses a simple strategy to find the stable branches of the steady state curve, either by using different starting points in parameter space or by trying to numerically find all solutions of the steady state equations.
2. Tutorials

First Strategy

To generate the stable branches of the steady state curve shown in Fig. 2.9, choose C_L from the Parameter list in the X-Axis:Variable Parameter field, change the values of k1 and C_Sp from 1 to 10 and 3, respectively, adjust the upper boundary of the Range to 0.5, increase the number of Intervals from 10 to 40 and deselect all species except for Sp ($S^*$). This should result in a plot as in Fig. 2.10. Apparently, we have obtained the upper stable branch of the steady state curve in Fig. 2.9 as well as the part of the lower branch that is to the left of the saddle-node bifurcation (SN). The first strategy to obtain the lower stable branch is to generate the steady state curve in the opposite direction. This can be done using the Rescan buttons in the Multiple Steady States field of the Control window (A, Fig. 2.10). By repeatedly using the left and right arrow you alternatively obtain either the lower or the upper part of the steady state branch. (Try it!) If you wish to keep both branches press the copy button right next to Solution branch: 1 in the Multiple Steady States field. This will create a copy of the current solution branch. Then, use again the Rescan buttons to generate the complementary branch. Now you can access both branches independently either through the Solution branch pull-down list in the Multiple Steady States field or through the pull-down list in the Plots window (B, Fig. 2.10). In the latter case, the branch number is indicated in square brackets.

Second Strategy

If a system has more than two stable steady states it might be difficult, if not impossible, to find all stable branches with the first strategy. In that case you may press the All Steady States button in the Multiple Steady States field of the Control window. This function tries to find all steady state curves numerically. It checks for local stability and returns the stable branches which should then appear in the Solution branch pull-down list. For the current example, All Steady States yields the same result as the strategy using the Rescan buttons (Try it!). In general, it seems reasonable to combine both strategies. The method using the Rescan buttons is fast and gives a first indication whether a system admits multiple stable steady states in a certain parameter range. All Steady States can be used to test whether there are further stable branches which may exist, for example, inside the region between two saddle-node bifurcations.

Having found the stable branches of the steady state set one may plot the Fano factor or any of the other quantities from the pull-down list of the Y-axis field along these branches similar as in the case of a single steady state curve. Note, however, that the predictions, based on the LNA, may become worse close to the saddle-node bifurcations as the LNA fails at such bifurcation points.
Figure 2.10.: Stable branches of the steady state curve in Fig. 2.9. Parameter values are the same as in the caption of Fig. 2.9.
3. Technical Remarks

To generate parameter scans for graphical visualization LiNA solves the steady state equations that result from the linear noise approximation of the master equation (cf. Eqs. 4.8 and 4.14)

\[
\frac{d\bar{x}}{dt} = \mathbf{N} \cdot \tilde{f}_0 (\bar{x}, \bar{p}) = 0 \quad (3.1)
\]

\[
\frac{d\sigma}{dt} = \mathbf{K} (\bar{x}) \cdot \sigma + \sigma \cdot \mathbf{K}^T (\bar{x}) + \mathbf{D} (\bar{x}) = 0. \quad (3.2)
\]

For a given network topology (encoded by the stoichiometric matrix \( \mathbf{N} \)) and a given set of kinetic parameters \( \bar{p} \) Eq. (3.1) represents \( k \), in general nonlinear, algebraic equations for the unknown steady state concentrations of the \( k \) species \( \bar{x}^s = (x^s_1, \ldots, x^s_k) \). In contrast, Eq. (3.2) results in a set of linear algebraic equations for the \( k(k+1)/2 \) independent components of the variance-covariance matrix \( \sigma \). Hence, for a given set of parameters LiNA has to solve \( k + k(k+1)/2 \) algebraic equations. From the values of \( \bar{x}^s \) and \( \sigma \) all other quantities of interest can be derived (cf. Section 4.3).

Although the steady states are calculated numerically, at the beginning of the analysis LiNA sets up Eqs. (3.1) and (3.2) symbolically in order to apply parameter substitutions if desired. All symbolic calculations are performed with MuPAD which is available as symbolic toolbox for MATLAB since release 2007b. In order to make use of the symbolic results they are converted into character strings in a format which can be processed by the MATLAB parser. For actual computations, parameters are substituted by their values where necessary and the resulting strings are parsed into anonymous MATLAB functions which then can be used as inputs for the numerical methods. In this manner exact analytical results are used as far as possible and the numerical calculations can still be performed efficiently with the parsed functions.

3.1. Solving Eqs. (3.1) and (3.2)

The primary method for calculating steady states is the MATLAB root solver \texttt{fsolve}. Only non-negative steady states that fulfill the conservation relations (when present) and are locally stable are accepted as valid. In case the root solver returns an invalid steady state or fails to find a solution the ODE system in Eq. (3.1) is numerically integrated for a given amount of time and the resulting end point taken as a new starting point for the root solver. This procedure is iterated up to three times with the integration time increasing in each iteration. If an integration occurs a message is displayed at the MATLAB console. In order to improve the efficiency, a linear interpolation of already known steady states is kept from which start points for the calculation of further steady
states are derived. To solve Eq. (3.2) the matrices $K(\vec{x}^s)$ and $D(\vec{x}^s)$ are evaluated at the steady state solution $\vec{x}^s$ found by solving Eq. (3.1). The equation is then transformed into the standard form $A \cdot \vec{\sigma} = \vec{b}$ which can be readily solved using standard MATLAB functions. The vector $\vec{\sigma}$ contains the independent components of the variance-covariance matrix. For example, in the case of a reaction network with 2 species it would be given by $\vec{\sigma} = (\sigma_{11}, \sigma_{22}, \sigma_{12})^T$.

### 3.2. Rescan vs. All Steady States

In case Eq. (3.1) admits multiple stable steady states one can use the functions Rescan or All Steady States from the Control window to find them. The Rescan function uses the previous steady state as a starting point for calculating the next steady state along the parameter curve (in scan direction). In contrast, the function All Steady States tries to find all stable steady states in given parameter range by calling the MuPAD function numeric::solve. Although this is convenient in case of bifurcations it is also a quite time-consuming procedure. Note that neither rescan nor calculation of all steady states will introduce any new support points along the parameter curve but rather operates on those points that are visible in the current plot window.

### 3.3. Data Management

The primary results of steady states computations, i.e. the mean values and variances, are permanently kept by LiNA unless they are explicitly cleared. They are stored separately for each combination of parameter set and variable parameter that has been selected at any time. Covariances, which are needed for calculating correlations, are calculated for all pairs that have been displayed in a correlation plot for a given parameter set/variable parameter combination at any time and will be kept from that point onwards.

### 3.4. Output at MATLAB Console and vanKampen.out

After starting the analysis by pressing the Start button in the Equation editor (cf. Fig. 2.2) a message is displayed at the console indicating whether or not parameter substitution (if intended) has been successful and whether conservation relations have been detected. The console also displays the result of parameter scans each time a parameter or a parameter range is changed. In case something goes wrong during the computation an error message is displayed at the console which might help to trace possible problems in the computation.

For the record, the results of the symbolic computations are written to the file vanKampen.out each time a new analysis is started. Specifically, these are the stoichiometric matrix $N$, the Jacobian $K$, the diffusion matrix $D$ as well as the matrix $A$ and the vector $\vec{b}$ which are used to solve Eq. (3.2). The file also displays the right-hand side of the ODE system in Eq. (3.1) in symbolic form, lists the independent species together with the reduced stoichiometric matrix $N_R$ and conservation relations (if present).
4. Theoretical Background

In this Chapter we provide the necessary theoretical background to appreciate the functionality implemented in LiNA. To make the presentation as self-contained as possible we give a derivation of the linear noise approximation of the master equation for biochemical reaction networks in Appendix A. We also describe how to deal with mass-conservation relations. Specific examples are provided to illustrate the general method.

4.1. Reaction Networks and the Master Equation

We consider a volume \( V \) where \( r \) chemical reactions take place between \( k \) species. Such a stoichiometric reaction network can be written in the form

\[
\nu_{-j}^i X_1 + \ldots + \nu_{k_j}^i X_k \xrightarrow{W_j} \nu_{+j}^i X_1 + \ldots + \nu_{k_j}^+ X_k, \quad j = 1, \ldots, r
\]  

(4.1)

where the index \( j \) numbers the reactions. From the stoichiometric coefficients \( \nu_{ij}^\pm \in \{0, 1, 2, \ldots\} \) one can construct the stoichiometric matrix \( N \) which has dimension \( k \times r \).

Its components, which are defined by

\[
N_{ij} := \nu_{ij}^+ - \nu_{ij}^-, \quad i = 1, \ldots, k, \quad j = 1, \ldots, r,
\]

(4.2)

indicate how many molecules of species \( X_i \) are produced (\( N_{ij} > 0 \)) or consumed (\( N_{ij} < 0 \)) in the \( j \)th reaction.

Let \( n_i \) denote the number of molecules of species \( X_i \). Then the state of the reaction system is specified by the probability \( P(n_i, t) \) to have \( n_i \) molecules of species \( X_i \) at time \( t \) in the system. In a Markovian description of the reaction network (Eq. 4.1) the dynamics of \( P(n_i, t) \) is determined by transition rates \( W_j (n_i + N_{ij}|n_i) \) which denote the probability per unit time for a change of the number of molecules of species \( X_i \) by an amount \( N_{ij} \) as a result of reaction \( j \). For given transition rates the temporal evolution of \( P(n_i, t) \) is determined by the chemical master equation [9]

\[
\frac{dP(\vec{n}, t)}{dt} = \sum_{j=1}^{r} \left[ W_j \left( \vec{n}|\vec{n} - \vec{N}_j \right) P \left( \vec{n} - \vec{N}_j, t \right) - W_j \left( \vec{n} + \vec{N}_j|\vec{n} \right) P(\vec{n}, t) \right]
\]

(4.3)

where \( \vec{n} = (n_1, \ldots, n_k)^T \) is a column vector whose components denote the molecule numbers \( n_i \) of species \( X_i \) and \( \vec{N}_j = (N_{1j}, \ldots, N_{kj})^T \) represent a set of vectors that are constructed from the \( j \)th column of the stoichiometric matrix as defined in Eq. (4.2).

Despite the fact that Eq. (4.3) is linear in \( P(\vec{n}, t) \) the transition rates \( W_j \) often exhibit a nonlinear dependence on \( \vec{n} \) so that exact solutions of this equation are rare. As a result, one either relies on numerical simulations [2] or on approximation methods.
4. Theoretical Background

One such approximation method is the linear noise approximation (LNA) which is based on van Kampen’s system size expansion of the master equation [9]. To this end, one assumes that the molecule number vector can be decomposed as

\[ n_i = V_m x_i + V_m^{1/2} \xi_i, \quad i = 1, \ldots, k \]  

where

\[ x_i := \frac{\langle n_i \rangle}{V_m} \]

represents the average concentration of species \( X_i \) and

\[ V_m \equiv V \cdot N_A \quad (N_A \approx 6 \cdot 10^{23} \text{ particles/mol}) \]

represents the molar volume. Here, we have followed standard biochemical conventions according to which species concentrations are measured in units of \( \text{mol/liter} \).

The \( k \) quantities \( \xi_i \) in Eq. (4.4) represent the new stochastic variables describing the deviations from the deterministic behavior. Hence, the idea of the decomposition in Eq. (4.4) is to separate the average dynamics, described by \( x_i \), from the fluctuations \( \xi_i \) which are described by a new probability density \( \Pi(\xi_i, t) \).

4.2. The Linear Noise Approximation

The LNA is obtained by inserting the decomposition in Eq. (4.4) into the master equation (Eq. 4.3) and expanding the resulting equation in powers of \( V_m^{-1/2} \). The rationale behind this procedure is that one expects that, in the limit of large system size \( V \to \infty \), fluctuations become negligible and one obtains an effective equation for the average concentration \( \bar{x} \). To perform the expansion one has to make an assumption about how the transition rates \( W_j(\bar{n} + \bar{N}_j|\bar{n}) \equiv W_j(\bar{n}) \) in Eq. (4.3) depend on the system size. By looking at specific examples (see Section 4.5) it turns out that most cases of practical interest are covered by the Ansatz (cf. Ref. [9])

\[ W_j(\bar{n}) = V_m f_{0,j} \left( \frac{\bar{n}}{V_m} \right) + f_{1,j} \left( \frac{\bar{n}}{V_m} \right) + \frac{1}{V_m} f_{2,j} \left( \frac{\bar{n}}{V_m} \right) + \ldots \]

\[ = \sum_{l=0}^{\infty} V_m^{1-l} f_{l,j} \left( \bar{x} + V_m^{-\frac{1}{2}} \xi \right) \]  

which ensures that, in the limit \( V \to \infty \), the functions \( f_{l,j} \) only depend on the average concentration \( \bar{x} \) and contributions from terms with \( l \geq 2 \) become negligible for \( V \to \infty \). Also, the assumption that the leading order term is proportional to \( V \) reflects the fact that the probability for a particular transition increases with the reaction volume if the average concentration is kept constant. Later, we shall see that, in the LNA, the dynamics of \( \bar{x} \) and \( \xi \) in Eq. (4.4) is completely determined by the leading order term in Eq. (4.7).
4. Theoretical Background

For the average concentration \( \bar{x}(t) \) the expansion of the master equation yields a system of nonlinear ordinary differential equations (ODEs) given by (cf. Section A.2)

\[
\frac{dx_i}{dt} = \sum_{j=1}^{r} N_{ij} f_{0,j}(\bar{x}), \quad i = 1, \ldots, k \tag{4.8}
\]

whereas the density of the fluctuations around the average is described by the linear partial differential equation

\[
\frac{\partial \Pi(\vec{\xi},t)}{\partial t} = \sum_{i,i'=1}^{k} \left( -K_{ii'}(\bar{x}) \frac{\partial}{\partial \xi_i} \left[ \frac{\partial \Pi(\vec{\xi},t)}{\partial \xi_i} \right] + \frac{1}{2} D_{ii'}(\bar{x}) \frac{\partial^2 \Pi(\vec{\xi},t)}{\partial \xi_i \partial \xi_i'} \right), \tag{4.9}
\]

which is called Fokker-Planck equation. The drift matrix \( K \) and the diffusion matrix \( D \) in Eq. (4.9) are defined through

\[
K_{ii'}(\bar{x}) := \sum_{j=1}^{r} N_{ij} \frac{\partial f_{0,j}(\bar{x})}{\partial \xi_i'} \quad \text{and} \quad D_{ii'}(\bar{x}) := \sum_{j=1}^{r} f_{0,j}(\bar{x}) N_{ij} N_{i'j}. \tag{4.10}
\]

The nomenclature ‘linear noise approximation’, which is used to denote Eq. (4.9), results from the fact that the coefficients in front of the derivative terms in the Fokker-Planck equation are, at most, linear in \( \vec{\xi} \). From the definitions, given in Eq. (4.10), it is clear that \( K \) represents the Jacobian matrix associated with the reaction system in Eq. (4.8). In general, the coefficient matrices \( K(\bar{x}) \) and \( D(\bar{x}) \) depend implicitly on time through their dependence on the average concentration \( \bar{x}(t) \). Hence, to perform the LNA one first has to solve the ODE system in Eq. (4.8). Of course, in most practical cases this can only be done numerically. In a second step, one solves the Fokker-Planck equation (4.9) where the mean-field solution enters through the coefficient matrices \( K(\bar{x}) \) and \( D(\bar{x}) \). However, since the Fokker-Planck equation is linear with respect to \( \vec{\xi} \) its solution is always a multivariate Gaussian distribution given by

\[
\Pi(\vec{\xi},t) = \frac{1}{\sqrt{(2\pi)^k \det \sigma(t)}} \exp \left( -\sum_{i,i'=1}^{k} \frac{(\xi_i - \langle \xi_i(t) \rangle) \sigma^{-1}_{ii'}(t) (\xi_{i'} - \langle \xi_{i'}(t) \rangle)}{2} \right) \tag{4.11}
\]

where \( \sigma^{-1}_{ii'} \) are the components of the inverse variance-covariance matrix

\[
\sigma_{ii'} = \langle \xi_i \xi_{i'} \rangle, \quad i, i' = 1, \ldots, k \tag{4.12}
\]

and \( \det \sigma \neq 0 \) denotes its determinant. Using the Fokker-Planck equation in Eq. (4.9) it is straightforward to show that \( \langle \vec{\xi} \rangle \) and \( \sigma \) are determined by the linear ODE systems (cf. Section A.3)

\[
\frac{d\langle \vec{\xi} \rangle}{dt} = K \cdot \langle \vec{\xi} \rangle \tag{4.13}
\]

\[
\frac{d\sigma}{dt} = K \cdot \sigma + \sigma \cdot K^T + D \tag{4.14}
\]
4. Theoretical Background

i.e. the average fluctuations $\langle \xi \rangle$ obey the same equation as small deviations from a steady state of the mean-field equations (Eq. 4.8). This means that the LNA is only applicable near an asymptotically stable steady state where all eigenvalues of $K$ have a negative real part and the average fluctuations decay to zero in time. If the mean-field equations (Eq. 4.8) exhibit multiple stable steady states one can perform the LNA near each of the stable branches away from the bifurcation points.

The variance-covariance matrix, as defined in Eq. (4.12), has the same units as the average concentration $\bar{x}$ (cf. Eq. 4.4), i.e. mol/liter in our case. However, if one is interested in particle number fluctuations rather than concentration fluctuations one has to transform the Gaussian distribution in Eq. (4.11) back to original variables. Specifically, from Eq. (4.4) the following relations are apparent

$$\xi_i = \frac{n_i - V_m x_i}{V_m^2}$$

$$\xi_i - \langle \xi \rangle = \frac{n_i - V_m x_i - V_m^2 \langle \xi \rangle}{V_m^2} \equiv \frac{n_i - \langle n_i \rangle}{V_m^2}$$

$$d\xi = V_m^{-\frac{1}{2}} d\bar{n}$$

so that the Gaussian distribution in Eq. (4.11) can be rewritten as

$$P(\bar{n}, t) = \frac{1}{\sqrt{(2\pi)^k \det C(t)}} \exp \left( - \sum_{i,i'=1}^{k} \left( \frac{(n_i - \langle n_i \rangle)}{V_m^2} \right) C^{-1}_{ii'}(t) (n_{i'} - \langle n_{i'} \rangle(t)) \right)$$

where

$$C_{ii'} = V_m \sigma_{ii'}$$

(4.15)

denotes the variance-covariance matrix with respect to particle numbers. The dynamics of $C$ follows from that of $\sigma$ by multiplying Eq. (4.14) by $V_m$ which results in

$$\frac{dC}{dt} = K \cdot C + C \cdot K^T + V_m D.$$

4.3. Calculation of Stochastic Quantities

From the variance-covariance matrix $\sigma$ (Eq. 4.14) and the mean-field solution $\bar{x}$ (4.8) one can readily calculate several stochastic quantities of interest such as the Fano factor $F_i$ and the coefficient of variation $CV_i$ of the species $X_i$ as well as the correlation coefficient $r_{ii'}$ between species $X_i$ and $X_{i'}$. Using Eqs. (4.4) and (4.15) these quantities can be expressed as

$$F_i = \frac{C_{ii}}{\langle n_i \rangle} = \frac{\sigma_{ii}}{x_i}, \quad i = 1, \ldots, k$$

(4.16)

$$CV_i = \frac{\sqrt{C_{ii}}}{\langle n_i \rangle} = \frac{\sqrt{\sigma_{ii}}}{\sqrt{V_m x_i}}, \quad i = 1, \ldots, k$$

(4.17)

$$r_{ii'} = \frac{\sigma_{ii'}}{\sqrt{\sigma_{ii'}^2}}, \quad i, i' = 1, \ldots, k.$$

(4.18)
4. Theoretical Background

This shows that, within the LNA, the Fano factor and the correlation coefficients are independent of the reaction volume $V$. In particular, since $\xi_i \sim V_m^{-1/2}$ the components of the variance-covariance matrix $\sigma_{ii'} = \langle \xi_i \xi_{i'} \rangle$ have the same unit as the mean concentrations $x_i$. Also, from Eq. (4.17) we see that the standard deviation with respect to the mean concentration is given by $\sqrt{\sigma_{ii}/V_m}$ which has the same unit as $x_i$. In contrast, the coefficient of variation depends inversely on the molar volume and, hence, the $CV$ vanishes as the reaction volume $V$ increases.

4.4. Dealing with Mass Conservation

Species that are not actively synthesized or degraded obey mass conservation relations which are indicated by linear dependencies between some rows of the stoichiometric matrix $N$. This may happen, for example, if a transcription factor binds to specific target promoters on the DNA in which case the promoters can exist in two states: free or occupied by a transcription factor, but the total amount of promoters is constant in time, at least over the time scale of the cell cycle. Another important example, where mass conservation must be taken into account, is for the modeling of enzyme-catalyzed reactions under in vitro conditions. Here, substrates, enzymes and cofactors are typically supplied in fixed amounts so that their total concentrations remain constant in time.

In general, there are as many linearly dependent rows of $N$ as there are conserved species in the system. Mathematically, mass conservation relations are associated with the left null space of $N$ which is defined by

$$\mathbf{S} \cdot \mathbf{N} = 0. \quad (4.19)$$

The dimension of $\mathbf{S}$ is $(k - k') \times k$ where $k'$ is the number of linearly independent rows of $N$, i.e. $\mathbf{S}$ has a many rows as there are conserved species in the system, and the number of columns equals the number of species. In terms of $\mathbf{S}$ one can define the vector of conserved species $\mathbf{s}$ through

$$\mathbf{s} = \mathbf{S} \cdot \mathbf{x}. \quad (4.20)$$

Using the mean-field equations in Eq. (4.8) it is straightforward to see that $\mathbf{s}$ is conserved in time, i.e.

$$\frac{d\mathbf{s}}{dt} = \mathbf{S} \cdot \frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{N} \cdot \mathbf{f}_0(\mathbf{x}) = 0, \quad (4.21)$$

which means that

$$\mathbf{s}(t) = \mathbf{S} \cdot \mathbf{x}(t) \equiv \mathbf{s}_0 \quad \forall t \quad (4.22)$$

where $\mathbf{s}_0$ is a constant vector whose components denote the total (molar) concentrations of the conserved species. Writing Eq. (4.22) in terms of particle numbers, rather than concentrations, yields

$$\mathbf{S} \cdot \mathbf{n}(t) = \mathbf{n}_0 \quad (4.23)$$

where $\mathbf{n}_0 = V_m \mathbf{s}_0$ now denotes total particle numbers of the conserved species.
4. Theoretical Background

Remark

For \( \vec{s}_0 \) (or \( \vec{n}_0 \)) to be interpretable as a vector of total concentrations (or particle numbers) it is necessary that all elements of \( \mathbf{S} \) are positive. However, solving Eq. (4.19) generally yields a matrix \( \mathbf{S} \) which contains both positive and negative entries. Hence, before defining \( \vec{s} \) through Eq. (4.20), one first has to find a positive basis of the left null space of \( \mathbf{N} \), e.g. through appropriate linear combinations of the rows of \( \mathbf{S} \).

The conservation relations in Eq. (4.23) are \( k - k' \) linear equations for the \( k \) particle numbers of species \( X_i \) \((i = 1, \ldots, k)\), which can be used to partition the \( X_i \) into \( k' \) independent species \( X_{R,i} \) \((i = 1, \ldots, k')\) and \( k - k' \) dependent species \( Y_\alpha \) \((\alpha = 1, \ldots, k - k')\). Let’s denote by \( n_{R,i} \) \((x_{R,i})\) and \( m_\alpha \) \((y_\alpha)\) the particle numbers (concentrations) associated with species \( X_{R,i} \) and \( Y_\alpha \), respectively. Then, the repartitioned quantities can be obtained from the original quantities according to

\[
\vec{n}' \equiv \left( \begin{array}{c} \vec{n}_R \\ \vec{m} \end{array} \right) = \mathbf{P} \cdot \vec{n} \quad \text{and} \quad \vec{x}' \equiv \left( \begin{array}{c} \vec{x}_R \\ \vec{y} \end{array} \right) = \mathbf{P} \cdot \vec{x} \quad (4.24)
\]

where \( \mathbf{P} \) is an appropriate permutation matrix. Using these relations in conjunction with Eq. (4.4) one can also decompose the fluctuations into those associated with the independent species \( (\vec{\xi}_R) \) and those associated with the dependent species \( (\vec{\eta}) \) as

\[
\vec{\xi} = \left( \begin{array}{c} \vec{\xi}_R \\ \vec{\eta} \end{array} \right) = \mathbf{P} \cdot \vec{\xi} \quad (4.25)
\]

so that particle number vectors of the dependent and the independent species can be written in the form

\[
\vec{n}_R = V_m \vec{x}_R + V_m^\frac{1}{2} \vec{\xi}_R \\
\vec{m} = V_m \vec{y} + V_m^\frac{1}{2} \vec{\eta}
\]

(4.26)

In the next step, we solve Eqs. (4.22) and (4.23) for \( \vec{y} \) and \( \vec{m} \), respectively, which allows expressing the dependent quantities as linear combinations of the independent quantities \( (\vec{x}_R \text{ and } \vec{n}_R) \) through the so-called link matrix \( \mathbf{L} \) \([7]\) defined by

\[
\vec{y} = \vec{s}_0 + \mathbf{L} \cdot \vec{x}_R \\
\vec{m} = n_0 + \mathbf{L} \cdot \vec{n}_R
\]

(4.27)

(4.28)

Inserting the decompositions from Eq. (4.26) into Eq. (4.28) and comparing with Eq. (4.27) shows that dependent and independent fluctuations are related as

\[
\vec{\eta} = \mathbf{L} \cdot \vec{\xi}_R
\]

(4.29)

Finally, one obtains the dynamics of the reduced system in the form

\[
\frac{d\vec{x}_R}{dt} = \mathbf{N}_R \cdot f_0 (\vec{x}_R; \vec{s}_0)
\]

(4.30)
4. Theoretical Background

where the reduced stoichiometric matrix \( N_R \) is obtained by taking the first \( k' \) rows of the matrix \( P \cdot N \) and the reaction rate vector \( f_0^T \) is evaluated at the conservation relations using Eqs. (4.24) and (4.27). Similarly, we define the reduced drift matrix \( K_R \) and the reduced diffusion matrix \( D_R \) by

\[
K_{R,ii'}(\vec{x}_R; \vec{s}_0) := \sum_{j=1}^r N_{R,ij} \frac{\partial f_{0,j}(\vec{x}_R; \vec{s}_0)}{\partial x_{R,i'}} \quad \text{and} \quad \tag{4.31}
\]

\[
D_{R,ii'}(\vec{x}_R; \vec{s}_0) := \sum_{j=1}^r f_{0,j}(\vec{x}_R; \vec{s}_0) N_{R,ij} N_{R,i'j} \]

where the indices range between \( i, i' = 1, \ldots, k' \). Using \( K_R \) and \( D_R \) the dynamics of the variance-covariance matrix of the reduced system is determined by

\[
\frac{d\sigma_R}{dt} = K_R \cdot \sigma_R + \sigma_R \cdot K_R^T + D_R \quad \tag{4.32}
\]

where the components of \( \sigma_R \) are given by \( \sigma_{R,ii'} = \langle \xi_{R,i} \xi_{R,i'} \rangle \) for \( i, i' = 1, \ldots, k' \). Hence, one can apply the linear noise approximation to the reduced system, defined by Eq. (4.30), in a similar way as described in Section 4.2.

The components of the variance-covariance matrix associated with the conserved species can be calculated from the reduced variance-covariance matrix \( \sigma_R \) and Eq. (4.29). Specifically, one obtains

\[
\sigma_{(k' + \alpha)(k' + \beta)}' = \langle \eta_\alpha \eta_\beta \rangle = \left( \sum_{i=1}^{k'} L_{\alpha i} \xi_{R,i} \sum_{i'=1}^{k'} L_{\beta i'} \xi_{R,i'} \right) \tag{4.33}
\]

\[
= \sum_{i,i'=1}^{k'} L_{\alpha i} L_{\beta i'} \sigma_{R,ii'}, \quad \alpha, \beta = 1, \ldots, k - k' \]

where \( \sigma' \) denotes the variance-covariance matrix with respect to the transformed fluctuation vector \( \vec{\xi}' \) defined in Eq. (4.25). Similarly, the ‘mixed’ components are obtained from

\[
\sigma_{(k' + \alpha)(i)} = \langle \eta_\alpha \xi_{R,i} \rangle = \left( \sum_{i'=1}^{k'} L_{\alpha i'} \xi_{R,i'} \right) \xi_{R,i} \tag{4.34}
\]

\[
= \sum_{i'=1}^{k'} L_{\alpha i'} \sigma_{R,i'i}, \quad i = 1, \ldots, k', \quad \alpha = 1, \ldots, k - k'.
\]

4.5. Examples

In the next two subsections we demonstrate how to apply the linear noise approximation to specific reaction systems. In both cases it is possible to derive explicit expressions for the steady state fluctuations. In addition, the example in subsection 4.5.2 shows how to deal with mass-conservation relations.
4. Theoretical Background

4.5.1. Noise Reduction Through Receptor Dimerization

As a first example, we consider the reaction scheme

\[ \emptyset \xrightleftharpoons[k_{d1}]{k_s} R(X_1) \] (4.35)

\[ R(X_1) + R(X_1) \xrightarrow{k^+}{k^-} R_2(X_2) \] (4.36)

\[ R_2(X_2) \xrightleftharpoons[k_{d2}]{k_d} \emptyset \] (4.37)

which describes synthesis and degradation of a receptor \( R \) (Eq. 4.35) and its dimerization (Eq. 4.36). In addition, we assume that the dimerized receptor \( R_2 \) can also be degraded with a specific rate that may differ from that for the monomer (4.37). First, we note that in the absence of dimerization independent synthesis and degradation of receptor molecules, as described by Eq. (4.35), leads to Poissonian steady state fluctuations where the average concentration \( \langle R \rangle = k_s/k_{d1} \) is equal to its variance \( \sigma_R^2 \), so that the Fano factor \( F_1 = \sigma_R^2 / \langle R \rangle \) becomes 1. Moreover, Hayot and Jayaprakash observed [5] that, within the LNA, this holds also true for receptor dimers, i.e. \( F_2 = \sigma_{R_2}^2 / \langle R_2 \rangle = 1 \). Here, we show that a non-vanishing dimer degradation rate \( (k_{d2} \neq 0) \) leads to sub-Poissonian fluctuations for both, monomers and dimers. In addition, we find that the steady state fluctuations of dimerized receptors can become lower than those for monomers \( (F_2 < F_1) \) under certain conditions.

The stoichiometric matrix for the system in Eqs. (4.35) - (4.37) is given by

\[ \mathbf{N} = \frac{R}{R_2} \begin{pmatrix} 1 & -1 & -2 & 2 & 0 \\ 0 & 0 & 1 & -1 & -1 \end{pmatrix} \] (4.38)

Hence, there are five stoichiometric vectors

\[ \vec{N}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{N}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \vec{N}_3 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \vec{N}_4 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad \vec{N}_5 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \]

but no conservation relations.

Let’s denote the number of receptor monomers and dimers by \( n_1 \) and \( n_2 \), respectively. Then the five transition rates are given by

\[ W_1 \left( \vec{n} | \vec{n} - \vec{N}_1 \right) = \bar{k}_s, \quad W_2 \left( \vec{n} | \vec{n} - \vec{N}_2 \right) = k_{d1} n_1 \]

\[ W_3 \left( \vec{n} | \vec{n} - \vec{N}_3 \right) = \bar{k}^+ n_1 (n_1 - 1), \quad W_4 \left( \vec{n} | \vec{n} - \vec{N}_4 \right) = k^- n_2 \]

\[ W_5 \left( \vec{n} | \vec{n} - \vec{N}_5 \right) = k_{d2} n_2 \]

where \( \vec{n} = (n_1, n_2)^T \). The units of \( \bar{k}_s \) and \( \bar{k}^+ \) are

\[ [\bar{k}_s] = \text{particles} / s \quad \text{and} \quad [\bar{k}^+] = 1 \text{ particles} \cdot s \]
4. Theoretical Background

whereas the unit of the remaining three constants is 

\[ [k_{d1}] = [k^+] = [k_{d2}] = \frac{1}{s}. \]

Next, we expand the transition rates in the form (cf. Eq. 4.7) 

\[ W_j (\vec{n}) = V_m f_{0,j} (\vec{x}) + f_{1,j} (\vec{x}) + \ldots , \quad \vec{x} = \frac{\vec{n}}{V_m}. \]

For \( W_2 \), \( W_4 \) and \( W_5 \) this yields 

\[ W_2 = V_m k_{d1} x_1, \quad W_4 = V_m k^- x_2 \quad \text{and} \quad W_5 = V_m k_{d2} x_2 \]

which shows that 

\[ f_{0,2} (\vec{x}) = k_{d1} x_1, \quad f_{0,4} (\vec{x}) = k^- x_2 \quad \text{and} \quad f_{0,5} (\vec{x}) = k_{d2} x_2 \quad (4.39) \]

and \( f_{l,j} (\vec{x}) \equiv 0 \) for \( l \geq 1 \) and \( j = 2, 4, 5 \). Hence, first-order rate constants are unaffected when going from a mesoscopic to a macroscopic description.

However, expansion of the remaining transition rates 

\[ W_1 = V_m k_s \quad \text{and} \quad W_3 = \bar{k}^+ V_m x_1 (V_m x_1 - 1) \equiv V_m k^+ x_1^2 - k^+ x_1 \quad (4.40) \]

shows that zero-order rate constants become inversely proportional to the reaction volume \( (k_s = \bar{k}_s/V_m) \) whereas second-order rate constants become proportional to the reaction volume \( (k^+ = V_m \bar{k}^+) \) so that \( k_s \) and \( k^+ \), as defined in Eqs. (4.35) and (4.36), now have typical ‘macroscopic’ units

\[ [k_s] = \frac{\text{mol}}{\text{liter} \cdot s} \equiv \frac{M}{s} \quad \text{and} \quad [k^+] = \frac{\text{liter}}{\text{mol} \cdot s} \equiv \frac{1}{M \cdot s}. \]

From Eq. (4.40) we infer that 

\[ f_{0,1} (\vec{x}) = k_s \quad \text{and} \quad f_{0,3} (\vec{x}) = k^+ x_1^2 \quad (4.41) \]

with \( f_{l,1} (\vec{x}) \equiv 0 \) for \( l \geq 1 \) and \( f_{l,3} (\vec{x}) \equiv 0 \) for \( l \geq 2 \). We also see that, as a result of the dimerization reaction, there appears an \( \mathcal{O}(1) \)-term \( (f_{1,3} (\vec{x}) = -\bar{k}^+ x_1) \) in the expansion for \( W_3 \) which is, however, unimportant for the linear noise approximation.

The functions \( f_j^{(0)} \), defined in Eqs. (4.39) and (4.41), are used to construct the reaction rate vector 

\[ \vec{f}_0 = (k_s, k_{d1} x_1, k^+ x_1^2, k^- x_2, k_{d2} x_2)^T. \quad (4.42) \]

Then, the dynamics of the average concentration \( \bar{x} = (x_1, x_2)^T \) is determined by the ODE system (cf. Eq. 4.8) 

\[ \frac{d\bar{x}}{dt} = \mathbf{N} \cdot \vec{f}_0 (\bar{x}) \]
4. Theoretical Background

or, in components,

\[
\frac{dx_1}{dt} = k_s - k_{d1} x_1 - 2k^+ x_1^2 + 2k^- x_2 \quad (4.43)
\]

\[
\frac{dx_2}{dt} = k^+ x_1^2 - (k^- + k_{d2}) x_2.
\]

For the matrices \( K \) and \( D \), defined in Eq. (4.10), we obtain

\[
K = \begin{pmatrix}
-k_{d1} - 4k^+ x_1 & 2k^- \\
2k^+ x_1 & -(k^- + k_{d2})
\end{pmatrix}
\]

and

\[
D = N \cdot \text{diag} \begin{pmatrix} f_0 \end{pmatrix} \cdot N^T
\]

\[
= \begin{pmatrix}
k_s + k_{d1} x_1 + 4k^- x_2 + 4k^+ x_1^2 & -2k^- x_2 - 2k^+ x_2^2 \\
-2k^- x_2 - 2k^+ x_2^2 & k^+ x_1^2 + k^- x_1 + k_{d2} x_2
\end{pmatrix}
\]

where \( \text{diag} \begin{pmatrix} f_0 \end{pmatrix} \) is a diagonal matrix which has the components \( f_{0j} \) on its diagonal.

The solution of the steady state equations

\[
k_s - k_{d1} x_1 - 2k^+ x_1^2 + 2k^- x_2 = 0
\]

\[
k^+ x_1^2 - (k^- + k_{d2}) x_2 = 0,
\]

which result from Eq. (4.43) by setting \( dx_1/dt = 0 = dx_2/dt \), is given by

\[
x_2^s = \frac{1}{2} \alpha (R_s - x_1^s) \quad \text{and} \quad (4.47)
\]

\[
x_1^s = \frac{\gamma}{2} \left( \sqrt{1 + \frac{4R_s}{\gamma}} - 1 \right) \quad (4.48)
\]

where

\[
R_s = \frac{k_s}{k_{d1}}, \quad \alpha = \frac{k_{d1}}{k_{d2}}, \quad \beta = \frac{k_{d2}}{k^-}, \quad K_D = \frac{k^-}{k^+} \quad (4.49)
\]

and

\[
\gamma \equiv K_D (1 + \beta) \frac{\alpha}{2}.
\]

Note that in the limit \( 4R_s \ll \gamma \) expansion of the square root in Eq. (4.48) yields to leading order

\[
x_1^s \approx R_s \quad \text{and} \quad x_2^s \approx \frac{R_s^2}{K_D (1 + \beta)}. \quad (4.50)
\]

Hence, by increasing the dissociation constant \( K_D \) (weak association of monomers) or increasing \( \beta \) (fast dimer degradation) one recovers the case of simple receptor synthesis and degradation described by Eq. (4.35) where \( x_1^s \approx R_s \) and \( x_2^s \approx 0 \). However, in general one expects that dimer dissociation is much faster than dimer degradation so that \( \beta \ll 1. \)
4. Theoretical Background

To find the three independent components of the variance-covariance matrix

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

we have to solve the system of linear equations defined by

$$K(\bar{x}_s) \cdot \sigma + \sigma \cdot K(\bar{x}_s)^T + D(\bar{x}_s) = 0 \quad (4.51)$$

where the matrices $K$ and $D$, defined in Eqs. (4.44) and (4.45), are now evaluated at the steady state (cf. Eqs. 4.47 and 4.48). The solution of Eq. (4.51) is straightforward although it is typically not easy to interpret. However, the resulting expressions can be substantially simplified if one applies algebraic manipulations while repeatedly making use of the steady state relations in Eq. (4.46). The result is

$$\sigma_{11}^s = x_1^s \left(1 - \frac{(1 + \beta(1 + \alpha)) x_1^s/K_D + 4\beta x_2^s/K_D}{(1 + \beta(1 + \alpha) + 4x_1^s/K_D)(\alpha(1 + \beta) + 4x_1^s/K_D)}\right) \quad (4.52)$$

$$\sigma_{22}^s = x_2^s \left(1 - \frac{(2x_1^s/K_D)^2}{(1 + \beta(1 + \alpha) + 4x_1^s/K_D)(\alpha(1 + \beta) + 4x_1^s/K_D)}\right) \quad (4.53)$$

$$\sigma_{12}^s = -\frac{2(1 + \beta) x_2^s x_1^s/K_D}{(1 + \beta(1 + \alpha) + 4x_1^s/K_D)(\alpha(1 + \beta) + 4x_1^s/K_D)}. \quad (4.54)$$

From Eqs. (4.52) and (4.53) we see that receptor dimerization leads to sub-Poissonian fluctuations for both receptor monomers ($F_1 = \sigma_{11}^s/x_1^s < 1$) and receptor dimers ($F_2 = \sigma_{22}^s/x_2^s < 1$). Similarly, Eq. (4.54) tells us that fluctuations in receptor monomers are always anti-correlated with fluctuations in receptor dimers since $\sigma_{12}^s < 0$. This is, of course, not surprising since an increase of the dimer concentration must be accompanied by a decrease in the concentration of monomers. However, since a decrease in dimers does not necessarily result from an increase in monomers, but could also result from a degradation of dimers, the strength of the anti-correlation will vary with parameters.

To estimate how much the intrinsic noise can be reduced by receptor dimerization we consider the special case where $\alpha = 1$ (monomers and dimers are degraded at the same rate) and $\beta \ll 1$. In that case the Fano factors can be approximated by

$$F_1 \approx 1 - \frac{1}{4} \frac{4x_1^s/K_D}{(1 + 4x_1^s/K_D)^2} \quad (4.55)$$

$$F_2 \approx 1 - \frac{1}{4} \frac{(4x_1^s/K_D)^2}{(1 + 4x_1^s/K_D)^2} \quad (4.56)$$

which implies that

$$F_2 < F_1, \quad \text{for} \quad x_1^s > \frac{K_D}{4}. \quad (4.57)$$

In particular, one can show that, in the limit $x_1^s \gg K_D/4$, $F_1 \to 1$ whereas $F_2 \to 3/4$. Hence, dimerization can reduce the steady state fluctuations of dimerized receptors if $R_s$ is sufficiently large. Indeed, $x_1^s \sim \sqrt{\gamma R_s}$ for $R_s \gg \gamma/4$ (cf. Eq. 4.48).
Finally, we note that the case \( k_{d2} = 0 \), studied in Ref. [5], is recovered by letting
in Eqs. (4.52) - (4.54) \( \alpha \to \infty \) and \( \beta \to 0 \) while keeping the product \( \alpha \cdot \beta = k_{d1}/k^- \) constant (cf. Eq. 4.49) which leads to \( x_1^s \to R_s \) and \( x_2^s \to R_s^2/K_D \) (cf. Eq. 4.50) so that fluctuations in monomers and dimers become Poissonian (\( \sigma_{11}^s \to x_1^s \), \( \sigma_{22}^s \to x_2^s \)) and correlations between monomers and dimers vanish (\( \sigma_{12}^s \to 0 \)).

4.5.2. Receptor-Ligand Binding with Mass Conservation

As a second example, we consider the simple reaction scheme

\[
L(X_1) + R(X_2) \xrightleftharpoons{\kappa_1^-}{\kappa_1^+} R-L(X_3)
\]

which describes the binding of a ligand (\( L \)) to a receptor (\( R \)) under *in vitro* conditions. The transition rates are given by

\[
W_1 = k_1^+ n_1 n_2 \quad \text{and} \quad W_2 = k_1^- n_3
\]

where \( \kappa_1^+ \) and \( \kappa_1^- \) have the units \( 1/(\text{particle} \cdot \text{s}) \) and \( 1/\text{s} \), respectively. The stoichiometric matrix of the system in Eq. (4.57) reads

\[
N = \begin{pmatrix}
R & -1 & 1 \\
L & -1 & 1 \\
R-L & 1 & -1
\end{pmatrix}.
\]

In this example, we neglect reactions for synthesis and degradation of the ligand and receptor molecules so that their total numbers remain constant. Consequently, two of the three rows of \( N \) are linearly dependent. A positive basis of the left null space of \( N \) is given by

\[
S = \begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 1
\end{pmatrix}
\]

and the corresponding conservation relations read

\[
S \cdot \vec{x}(t) = \begin{pmatrix}
x_1(t) + x_3(t) \\
x_2(t) + x_3(t)
\end{pmatrix} = \vec{s}_0 \equiv \begin{pmatrix}
L_T \\
R_T
\end{pmatrix}.
\]

We choose \( X_3 \) as the independent, and \( X_1 \) and \( X_2 \) as the dependent species, i.e. we set (cf. Eqs. 4.26)

\[
\begin{pmatrix}
x_R \\
y_1 \\
y_2
\end{pmatrix} = P \cdot \begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix} = \begin{pmatrix}
x_3 \\
x_1 \\
x_2
\end{pmatrix}
\]

where \( P \) denotes the permutation matrix (cf. Eq. 4.24)

\[
P = \begin{pmatrix}
0 & 0 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{pmatrix}.
\]
4. Theoretical Background

The fluctuations are rearranged as

\[
\begin{pmatrix}
\xi_R \\
\eta_1 \\
\eta_2
\end{pmatrix}
= \mathbf{P} \cdot \begin{pmatrix}
\xi_1 \\
\xi_2 \\
\xi_3
\end{pmatrix}
= \begin{pmatrix}
\xi_3 \\
\xi_1 \\
\xi_2
\end{pmatrix}.
\] (4.61)

Using Eq. (4.59) we can express the dependent species concentrations as a function of the independent species concentration through

\[
y_1 = L_T - x_R
\]
\[
y_2 = R_T - x_R
\]

or

\[
\bar{y} = \bar{s}_0 + \mathbf{L} \cdot \bar{x}_R
\]

where the \((2 \times 1)\) link matrix \(\mathbf{L}\), as defined in Eq. (4.27), is given by

\[
\mathbf{L} = \begin{pmatrix}
-1 \\
-1
\end{pmatrix}.
\] (4.63)

Hence, fluctuations of the dependent species can be obtained from that of the independent species through (cf. Eq. 4.29)

\[
\eta_1 = -\xi_R
\]
\[
\eta_2 = -\xi_R.
\] (4.64)

The dynamics of the reduced system is then determined by

\[
\frac{dx_R}{dt} = \mathbf{N}_R \cdot \bar{f}_0 (x_R, \bar{s}_0) = k_1^+ (L_T - x_R)(R_T - x_R) - k_1^- x_R
\] (4.65)

where the reduced stoichiometric matrix is obtained from the first row of \(\mathbf{P} \cdot \mathbf{N}\) as

\[
\mathbf{N}_R = \begin{pmatrix}
1 \\
-1
\end{pmatrix}.
\]

The reaction rate vector is obtained by first expanding the transition rates, defined in Eq. (4.58), as

\[
W_1 = V_m k_1^+ x_1 x_2 = V_m k_1^+ y_1 y_2
\]
\[
W_2 = V_m k_1^- x_3 = V_m k_1^- x_R
\]

and then using the conservation relations in Eqs. (4.62) to replace \(y_1\) and \(y_2\) by \(x_R\) and \(\bar{s}_0 = (L_T, R_T)^T\) leading to

\[
f_0^T (x_R, \bar{s}_0) = \begin{pmatrix}
k_1^+ (L_T - x_R)(R_T - x_R) \\
k_1^- x_R
\end{pmatrix}.
\]

In Eq. (4.65) the second-order rate constant \(k_1^+\) has the unit liter/(mol \cdot s) \equiv 1/(M \cdot s).
4. Theoretical Background

Since the reduced system is 1-dimensional there is only 1 independent component of the variance-covariance matrix which is determined by

$$\sigma_R = -\frac{D_R}{2K_R}$$

where $\sigma_R = \langle \xi_R \xi_R \rangle = \langle \xi_3 \xi_3 \rangle \equiv \sigma_{33}$ (cf. Eq. 4.61). The reduced Jacobian and diffusion matrices ($K_R$ and $D_R$) are given by

$$K_R = -k^+_1 (L_T + R_T - 2x_R) - k^-_1$$
$$D_R = k^+_1 (L_T - x_R) (R_T - x_R) + k^-_1 x_R.$$  

Hence, under steady state conditions we have

$$k^+_1 (L_T - x_R) (R_T - x_R) \frac{1}{2} k^-_1 x_R$$

and the variance of the receptor-ligand complex can be written as

$$\sigma_R = \frac{1}{2} \frac{k^+_1 (L_T - x_R) (R_T - x_R) + k^-_1 x_R}{x_R}$$
$$= \frac{x_R}{L_T + R_T - 2x_R + K_D}$$

where $x_R$ is determined by the quadratic equation

$$x_R^2 - x_R (L_T + R_T + K_D) + L_TR_T = 0$$  

and $K_D = k^-_1 / k^+_1$ denotes the dissociation constant of the receptor-ligand complex. The biologically feasible solution of Eq. (4.67) is given by

$$x_R = \frac{L_T + R_T + K_D}{2} \left( 1 - \sqrt{1 - 4 \frac{L_T R_T}{(L_T + R_T + K_D)^2}} \right).$$  

The remaining components of the variance-covariance matrix can be calculated using the expressions in Eqs. (4.61) and (4.64) which follow from the general expressions in Eqs. (4.33) and (4.34). For example, the variances of ligand and free receptors are

$$\sigma_{11} = \langle \xi_1 \xi_1 \rangle = \langle \eta_1 \eta_1 \rangle = \langle (-\xi_R) (-\xi_R) \rangle = \sigma_R$$
$$\sigma_{22} = \langle \xi_2 \xi_2 \rangle = \langle \eta_2 \eta_2 \rangle = \langle (-\xi_R) (-\xi_R) \rangle = \sigma_R.$$  

Hence, the variances of all three species are the same ($\sigma_{11} = \sigma_{22} = \sigma_{33} \equiv \sigma_R$). Similarly, the calculation of the covariances yields

$$\sigma_{12} = \langle \xi_1 \xi_2 \rangle = \langle \eta_1 \eta_2 \rangle = \langle (-\xi_R) (-\xi_R) \rangle = \sigma_R$$
$$\sigma_{13} = \langle \xi_1 \xi_3 \rangle = \langle \eta_1 \xi_R \rangle = \langle (-\xi_R) \xi_R \rangle = -\sigma_R$$
$$\sigma_{23} = \langle \xi_2 \xi_3 \rangle = \langle \eta_2 \xi_R \rangle = \langle (-\xi_R) \xi_R \rangle = -\sigma_R.$$
4. Theoretical Background

This shows that the correlation coefficients, as defined in Eq. (4.18), are given by

$$r_{13} = r_{23} = -1 \quad \text{and} \quad r_{12} = 1$$

which expresses the obvious fact that the disappearance of a ligand or receptor molecule is always accompanied by the appearance of a receptor-ligand complex.

To obtain explicit expressions for other stochastic quantities such as the Fano factor one may use the expressions for $x_R$ and $\sigma_R$ in Eqs. (4.68) and (4.66). Alternatively, one can try to obtain simpler expressions for $x_R$ from suitable approximate solutions of Eq. (4.67). Specifically, if $K_D \ll R_T$ or $K_D \ll L_T$ the solution of Eq. (4.67) can be approximated by the discontinuous function

$$x_R \approx \begin{cases} L_T \left(1 - \frac{K_D}{R_T - L_T}\right) & L_T < R_T \\ R_T \left(1 - \frac{K_D}{L_T - R_T}\right) & R_T < L_T \end{cases}.$$  \hspace{1cm} (4.69)

In the opposite case, when $K_D \gg R_T$ or $K_D \gg L_T$ one obtains

$$x_R \approx \frac{L_T R_T}{L_T + R_T + K_D}. \hspace{1cm} (4.70)$$

Hence, in the first case (Eq. 4.69) the Fano factors are given to leading order by

$$F_1 = \frac{\sigma_R}{L_T - x_R} \approx \begin{cases} \frac{1}{L_T - x_R} & L_T < R_T \\ \frac{K_D}{(L_T - R_T) R_T} & R_T < L_T \end{cases},$$

$$F_2 = \frac{\sigma_R}{R_T - x_R} \approx \begin{cases} \frac{L_T K_D}{(R_T - L_T) R_T} & L_T < R_T \\ \frac{1}{R_T - L_T} & R_T < L_T \end{cases},$$

$$F_3 = \frac{\sigma_R}{x_R} \approx \begin{cases} \frac{K_D}{L_T - R_T} & L_T < R_T \\ \frac{K_D}{R_T - L_T} & R_T < L_T \end{cases},$$

whereas in the second case (Eq. 4.70) they are given to leading order by

$$F_1 \approx \frac{K_D R_T}{(L_T + K_D)(L_T + R_T + K_D)}, \hspace{1cm} (4.71)$$

$$F_2 \approx \frac{K_D L_T}{(R_T + K_D)(L_T + R_T + K_D)},$$

$$F_3 \approx \frac{K_D}{L_T + R_T + K_D}.$$  \hspace{1cm} (4.72)

From the expressions in Eqs. (4.71) we see that if receptor-ligand binding is tight, so that $K_D \ll \min(R_T, L_T)$, $F_1$ exhibits a switch-like transition from $F_1 \approx 1$ for $L_T < R_T$ (Poissonian noise) to $F_1 \approx 0$ for $L_T > R_T$ (noise suppression) and vice versa for $F_2$.

Problem

Check the validity of the approximate expressions in Eqs. (4.69) - (4.72) with LiNA!
A. Derivation of the Linear Noise Approximation

In the following, we give a derivation of the dynamic equations for the mean-field $\bar{x}$ (Eq. 4.8) and the probability density of the fluctuations $\Pi \left( \bar{\xi}, t \right)$ (Eq. 4.9), which are derived from the master equation (Eq. 4.3) using the decomposition of the particle number vector in Eq. (4.4). The derivation is a multivariate extension of that given in Ref. [9] for the special case of biochemical reaction networks where the jump sizes in the transition rates are determined by the entries of the stoichiometric matrix (cf. Eq. 4.3). An alternative derivation of the multivariate linear noise approximation, based on the step operator formalism, can be found in Ref. [1]. A derivation that goes beyond the linear noise approximation has been given by Grima in Ref. [4].

A.1. Transformation of the Master Equation

The probability densities $P (\bar{n}, t)$ and $\Pi \left( \bar{\xi}, t \right)$ are related through

$$ P (\bar{n}, t) = P \left( V_m \bar{x} + V_m^2 \bar{\xi}, t \right) = V_m^{-\frac{1}{2}} \Pi \left( \bar{\xi}, t \right) $$

(A.1)

which follows from the invariance of the probability measure

$$ P (\bar{n}, t) \, d\bar{n} = P (\bar{n}, t) V_m^\frac{1}{2} d\bar{\xi} = \Pi \left( \bar{\xi}, t \right) \, d\bar{\xi} $$

and Eq. (4.4). The probability density at the shifted argument can be obtained from

$$ P \left( \bar{n} - \bar{N}_j, t \right) = P \left( V_m \bar{x} + V_m^\frac{1}{2} \left( \bar{\xi} - V_m^{-\frac{1}{2}} \bar{N}_j \right), t \right) $$

$$ = V_m^{-\frac{1}{2}} \Pi \left( \bar{\xi} - V_m^{-\frac{1}{2}} \bar{N}_j, t \right). $$

(A.2)
A. Derivation of the Linear Noise Approximation

Under the transformation in Eq. (4.4) the time-derivative in Eq. (4.3) becomes

$$\frac{dP (\vec{n}, t)}{dt} = V_m^{-\frac{1}{2}} \left( \frac{\partial \Pi (\vec{\xi}, t)}{\partial t} \right) \vec{n}$$

$$= V_m^{-\frac{1}{2}} \left( \frac{\partial \Pi (\vec{\xi}, t)}{\partial t} \right) + \sum_{i=1}^{k} \left( \frac{\partial \Pi (\vec{\xi}, t)}{\partial \xi_i} \right)_t \left( \frac{d\xi_i}{dt} \right) \vec{n}$$

$$= V_m^{-\frac{1}{2}} \left[ \left( \frac{\partial \Pi (\vec{\xi}, t)}{\partial t} \right) \vec{\xi} + V_m^\frac{1}{2} \sum_{i=1}^{k} \left( \frac{\partial \Pi (\vec{\xi}, t)}{\partial \xi_i} \right) dx_i \right] \tag{A.3}$$

where the last step follows from Eq. (4.4) if $\vec{n}$ is kept constant.

Next, we have to transform the transition rates in Eq. (4.3). To this end we recall the expansion from Eq. (4.7)

$$W_j (\vec{n}) = \sum_{l=0}^{\infty} V_m^{1-l} f_{l,j} \left( \vec{x} + V_m^{\frac{1}{2}} \vec{\xi} \right)$$

which can be used to obtain the transition rates at the shifted argument as

$$W_j (\vec{n} - \vec{N}_j) = \sum_{l=0}^{\infty} V_m^{1-l} f_{l,j} \left( \vec{x} + V_m^{\frac{1}{2}} \left( \vec{\xi} - V_m^{\frac{1}{2}} \vec{N}_j \right) \right) \tag{A.4}$$

Inserting the expressions in Eqs. (A.1) - (A.4) into the master equation (Eq. 4.3) yields (after multiplication by $V_m^\frac{1}{2}$)

$$\frac{\partial \Pi (\vec{\xi}, t)}{\partial t} - V_m^\frac{1}{2} \sum_{i=1}^{k} \left( \frac{\partial \Pi (\vec{\xi}, t)}{\partial \xi_i} \right) dx_i = \sum_{l=0}^{\infty} V_m^{1-l} \sum_{j=1}^{r} \left[ f_{l,j} \left( \vec{x} + V_m^{\frac{1}{2}} \left( \vec{\xi} - V_m^{\frac{1}{2}} \vec{N}_j \right) \right) \Pi \left( \vec{\xi} - V_m^{\frac{1}{2}} \vec{N}_j, t \right) \right] \tag{A.5}$$

The idea of van Kampen’s system size expansion is to expand the right-hand side of this master equation in powers of $V_m^{\frac{1}{2}}$. Note that the left-hand side of the master equation (Eq. A.5) has already the form of an expansion in powers of $V_m^{\frac{1}{2}} \equiv \varepsilon$ which consists of terms of $O (\varepsilon^0)$ and $O (\varepsilon^{-1})$. In order to match these terms by corresponding terms on the right-hand side of the master equation (Eqs. A.6 and A.7) we expand the functions

$$f_{l,j} \left( \vec{x} + V_m^{\frac{1}{2}} \left( \vec{\xi} - V_m^{\frac{1}{2}} \vec{N}_j \right) \right) \Pi \left( \vec{\xi} - V_m^{\frac{1}{2}} \vec{N}_j, t \right)$$
A. Derivation of the Linear Noise Approximation

with respect to $V_m^{-\frac{1}{2}} \hat{N}_j$ in the argument $\hat{\xi} - V_m^{-\frac{1}{2}} \hat{N}_j$ which leads to

$$f_{l,j} \left( \vec{x} + V_m^{-\frac{1}{2}} \left( \vec{\xi} - V_m^{-\frac{1}{2}} \hat{N}_j \right) \right) \Pi \left( \vec{\xi} - V_m^{-\frac{1}{2}} \hat{N}_j, t \right)$$

$$= f_{l,j} \left( \vec{x} + V_m^{-\frac{1}{2}} \hat{\xi} \right) \Pi \left( \hat{\xi}, t \right)$$

(A.8)

$$- V_m^{-\frac{1}{2}} \sum_{i=1}^{k} \partial \left[ f_{l,j} \left( \vec{x} + V_m^{-\frac{1}{2}} \xi \right) \Pi \left( \xi, t \right) \right] \xi_i N_{ij}$$

(A.9)

$$+ \frac{1}{2} V_m^{-1} \sum_{i,i'=1}^{k} \partial^2 \left[ f_{l,j} \left( \vec{x} + V_m^{-\frac{1}{2}} \xi \right) \Pi \left( \xi, t \right) \right] \xi_i \xi_{i'} N_{ij} N_{i'j} + O \left( V_m^{-\frac{3}{2}} \right)$$

(A.10)

Note that the lowest order term in Eq. (A.8) cancels with the term in Eq. (A.7) of the master equation. The expressions in Eqs. (A.9) and (A.10) are of $O \left( V_m^{-\frac{1}{2}} \right)$ and $O \left( V_m^{-1} \right)$, respectively, so that they contribute terms of $O \left( V_m^{-\frac{l}{2}} \right)$ and $O \left( V_m^{-l} \right)$ to the master equation. However, since in Eq. (A.5) the lowest order term is of $O \left( 1 \right)$ the terms coming from Eqs. (A.9) and (A.10) can only be matched if $l = 0$.

The arguments of the functions $f_{l,j}$ in Eqs. (A.9) and (A.10) still depend on $V_m^{-\frac{1}{2}} \hat{\xi}$ which, after expansion with respect to $V_m^{-\frac{1}{2}} \hat{\xi}$

$$f_{l,j} \left( \vec{x} + V_m^{-\frac{1}{2}} \hat{\xi} \right) = f_{l,j} \left( \vec{x} \right) + V_m^{-\frac{1}{2}} \sum_{i=1}^{k} \partial f_{l,j} \left( \vec{x} \right) \xi_i + O \left( V_m^{-1} \right) ,$$

(A.11)

can potentially contribute to the matching with Eq. (A.5). However, for the second order term in Eq. (A.10) the higher order terms in Eq. (A.11) would yield terms of $O \left( V_m^{-\frac{3}{2}} \right)$ and higher which, thus, cannot be matched with the terms in Eq. (A.5). Hence, we may write Eq. (A.10) in the form

$$\frac{1}{2} V_m^{-1} \sum_{i,i'=1}^{k} \partial^2 \left[ f_{l,j} \left( \vec{x} + V_m^{-\frac{1}{2}} \xi \right) \Pi \left( \xi, t \right) \right] \xi_i \xi_{i'} N_{ij} N_{i'j}$$

$$= \frac{1}{2} V_m^{-1} \sum_{i,i'=1}^{k} f_{l,j} \left( \vec{x} \right) N_{ij} N_{i'j} \partial^2 \Pi \left( \xi, t \right) \xi_i \xi_{i'} + O \left( V_m^{-\frac{3}{2}} \right) .$$

(A.12)

In contrast, for the first order term in Eq. (A.9), the first two terms in Eq. (A.11) give a
A. Derivation of the Linear Noise Approximation

contribution so that we write this term in the form

\[ V_m^{-\frac{1}{2}} \sum_{i=1}^{k} \frac{\partial}{\partial \xi_i} \left[ f_{l,j}(\vec{x} + V_m^{-\frac{1}{2}} \vec{\xi}) \Pi(\vec{\xi}, t) \right] N_{ij} \]

\[ = V_m^{-\frac{1}{2}} \sum_{i=1}^{k} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} N_{ij} f_{l,j}(\vec{x}) \]

\[ + V_m^{-1} \sum_{i,i'=1}^{k} \frac{\partial f_{l,j}(\vec{x})}{\partial x_i} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} N_{ij} N_{i'j} + O\left(V_m^{-\frac{3}{2}}\right). \quad \text{(A.13)} \]

Finally, using the expressions in Eqs. (A.12) and (A.13) in Eqs. (A.10) and (A.9) and, the resulting expression in Eq. (A.6) gives

\[ \frac{\partial \Pi(\vec{\xi}, t)}{\partial t} - V_m^{\frac{1}{2}} \sum_{i=1}^{k} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} \frac{dx_i}{dt} = -V_m^{\frac{1}{2}} \sum_{j=1}^{r} \sum_{i=1}^{k} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} N_{ij} f_{0,j}(\vec{x}) \]

\[ - \sum_{j=1}^{r} \sum_{i,i'=1}^{k} N_{ij} \frac{\partial f_{0,j}(\vec{x})}{\partial x_i} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} + O\left(V_m^{-\frac{1}{2}}\right) \]

where we have retained only the \( l = 0 \) term.

A.2. Effective Equations for \( \vec{x}(t) \) and \( \Pi(\vec{\xi}, t) \)

Comparing first the terms of highest order in Eq. (A.14) gives

\[ O\left(V_m^{\frac{1}{2}}\right) : \sum_{i=1}^{k} \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} \left( \frac{dx_i}{dt} - \sum_{j=1}^{r} N_{ij} f_{0,j}(\vec{x}) \right) = 0. \]

This implies that \( \vec{x} \) has to fulfill the mean-field equations (cf. Eq. 4.8)

\[ \frac{dx_i}{dt} = \sum_{j=1}^{r} N_{ij} f_{0,j}(\vec{x}), \quad i = 1, \ldots, k \quad \text{(A.15)} \]

provided that

\[ \frac{\partial \Pi(\vec{\xi}, t)}{\partial \xi_i} \neq 0, \quad i = 1, \ldots, k \]
A. Derivation of the Linear Noise Approximation

which means that $\Pi$ must not be constant with respect to $\xi$ or, in other words, the mean-field equation (Eq. A.15) is only valid if there are fluctuations in the system.

Comparing the terms at next order yields a Fokker-Planck equation for the density of the fluctuations given by

$$\mathcal{O}(1): \quad \frac{\partial \Pi(\vec{\xi},t)}{\partial t} = -\sum_{j=1}^{r} \sum_{i,i'=1}^{k} N_{ij} \frac{\partial f_{0,j}(\vec{x})}{\partial x_{i'}} \frac{\partial^2 \Pi(\vec{\xi},t)}{\partial \xi_i} + \frac{1}{2} \sum_{j=1}^{r} f_{0,j}(\vec{x}) \sum_{i,i'=1}^{k} N_{ij} N_{i'j} \frac{\partial^2 \Pi(\vec{\xi},t)}{\partial \xi_i \partial \xi_{i'}}. \quad (A.16)$$

Introducing the drift matrix $K$ and the diffusion matrix $D$ by

$$K_{i'i'}(\vec{x}) := \sum_{j=1}^{r} N_{ij} \frac{\partial f_{0,j}(\vec{x})}{\partial x_{i'}} \quad \text{and} \quad D_{i'i'}(\vec{x}) := \sum_{j=1}^{r} f_{0,j}(\vec{x}) N_{ij} N_{i'j} \quad (A.17)$$

the Fokker-Planck equation can be written in compact notation

$$\frac{\partial \Pi(\vec{\xi},t)}{\partial t} = \sum_{i,i'=1}^{k} \left( -K_{i'i'}(\vec{x}) \frac{\partial [\xi_{i'} \Pi(\vec{\xi},t)]}{\partial \xi_i} + \frac{1}{2} D_{i'i'}(\vec{x}) \frac{\partial^2 \Pi(\vec{\xi},t)}{\partial \xi_i \partial \xi_{i'}} \right) \quad (A.18)$$

which agrees with Eq. (4.9).

A.3. Equations for $\langle \xi_i(t) \rangle$ and $\sigma_{ii'}(t) = \langle \xi_i \xi_{i'} \rangle$

With the help of the Fokker-Planck equation (A.18) one can derive equations for the temporal evolution of the average of a function $f(\vec{\xi})$ according to the scheme

$$\frac{d \langle f(\vec{\xi}) \rangle}{dt} = \frac{d}{dt} \int_{-\infty}^{\infty} d\xi_1 \cdots \int_{-\infty}^{\infty} d\xi_k f(\vec{\xi}) \Pi(\vec{\xi},t) \quad (A.19)$$

$$= \Pi_{i=1}^{k} \int_{-\infty}^{\infty} d\xi_i f(\vec{\xi}) \frac{\partial \Pi(\vec{\xi},t)}{\partial t}$$

$$= \sum_{j,j'=1}^{k} \int d\xi f(\vec{\xi}) \left( -K_{jj'} \frac{\partial [\xi_{j'} \Pi(\vec{\xi},t)]}{\partial \xi_j} + \frac{1}{2} D_{jj'} \frac{\partial^2 \Pi(\vec{\xi},t)}{\partial \xi_j \partial \xi_{j'}} \right) \quad (A.20)$$

where we have introduced the notation

$$\int d\vec{\xi} = \int_{-\infty}^{\infty} d\xi_1 \cdots \int_{-\infty}^{\infty} d\xi_k = \Pi_{i=1}^{k} \int_{-\infty}^{\infty} d\xi_i.$$
A. Derivation of the Linear Noise Approximation

Also, in order to rewrite the expressions in Eq. (A.20) in terms of average quantities we will have to do partial integrations of the form

\[ \int d\tilde{\xi} f(\tilde{\xi}) \frac{\partial}{\partial \tilde{\xi}_i} \left[ g(\tilde{\xi}) \Pi(\tilde{\xi}, t) \right] = -\int d\tilde{\xi} g(\tilde{\xi}) \Pi(\tilde{\xi}, t) \frac{\partial}{\partial \tilde{\xi}_i} f(\tilde{\xi}) \]

where the boundary term vanishes since \( \Pi(\tilde{\xi}, t) \), being a multivariate Gaussian (Eq. 4.11), decays exponentially as \( |\tilde{\xi}| \to \infty \).

For the mean value we obtain

\[
\frac{d\langle \xi_i \rangle}{dt} = \sum_{j,j'} k \int d\tilde{\xi} \tilde{\xi}_i \left( -K_{jj'} \frac{\partial}{\partial \tilde{\xi}_j} \frac{\partial}{\partial \tilde{\xi}_j} \Pi(\tilde{\xi}, t) + \frac{1}{2} D_{jj'} \frac{\partial^2 \Pi(\tilde{\xi}, t)}{\partial \tilde{\xi}_j \partial \tilde{\xi}_{j'}} \right) \\
= \sum_{j,j'} k \int d\tilde{\xi} K_{jj'} \xi_{j'} \Pi(\tilde{\xi}, t) \frac{\partial \xi_i}{\partial \tilde{\xi}_j} + \frac{1}{2} D_{jj'} \Pi(\tilde{\xi}, t) \frac{\partial^2 \xi_i}{\partial \tilde{\xi}_j \partial \tilde{\xi}_{j'}} \\
= \sum_{j,j'} k K_{jj'} \langle \xi_{j'} \rangle \delta_{ij} \\
= \sum_{j=1}^k K_{ij} \langle \xi_j \rangle
\]

or, in vector notation,

\[
\frac{d\langle \tilde{\xi} \rangle}{dt} = \mathbf{K} \cdot \langle \tilde{\xi} \rangle,
\]

i.e. the fluctuations \( \tilde{\xi} \) obey the same equation as small deviations from a steady state of the mean-field equations (Eq. A.15). Hence, if \( \mathbf{K}(\bar{x}) \) is evaluated at an asymptotically stable steady state all eigenvalues of \( \mathbf{K} \) have a negative real part so that the fluctuations decay to zero in time.
A. Derivation of the Linear Noise Approximation

The calculation for the variance-covariance matrix \( \sigma_{ij} = \langle \xi_i \xi_{i'} \rangle \) yields

\[
\frac{d \langle \xi_i \xi_{i'} \rangle}{dt} = \sum_{j,j'=1}^{k} \int d\vec{\xi} \xi_i \xi_{i'} \left( -K_{jj'} \frac{\partial}{\partial \xi_j} \left[ \xi_{i'} \Pi \left( \vec{\xi}, t \right) \right] + \frac{1}{2} D_{jj'} \frac{\partial^2 \Pi \left( \vec{\xi}, t \right)}{\partial \xi_j \partial \xi_{j'}} \right)
\]

\[
= \sum_{j,j'=1}^{k} \int d\vec{\xi} \left( K_{jj'} \xi_{i'} \Pi \left( \vec{\xi}, t \right) \frac{\partial}{\partial \xi_j} \left( \xi_i \xi_{i'} \right) + \frac{1}{2} D_{jj'} \Pi \left( \vec{\xi}, t \right) \frac{\partial^2 \left( \xi_i \xi_{i'} \right)}{\partial \xi_j \partial \xi_{j'}} \right)
\]

\[
+ \frac{1}{2} \sum_{j,j'=1}^{k} \int d\vec{\xi} D_{jj'} \Pi \left( \vec{\xi}, t \right) \left( \delta_{ij} \delta_{i'j'} \xi_{i'} + \delta_{i'j} \delta_{ij} \xi_i \right)
\]

\[
= \sum_{j=1}^{k} K_{ij} \langle \xi_j \xi_{i'} \rangle + \sum_{j'=1}^{k} K_{i,j'} \langle \xi_{j'} \xi_i \rangle + \frac{1}{2} \left( D_{ii'} + D_{i'i} \right)
\]

or, in matrix notation,

\[
\frac{d\sigma}{dt} = K \cdot \sigma + \sigma \cdot K^T + D \tag{A.22}
\]

where we have used the fact that \( D \) is symmetric \((D_{ii'} = D_{i'i})\).
Bibliography


