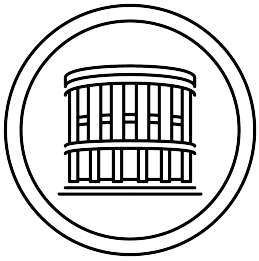
UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



Vnútorné a vonkajšie riešenia náhodných procesov

DIPLOMOVÁ PRÁCA

2025 Bc. Ivana Oravcová

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Vnútorné a vonkajšie riešenia náhodných procesov

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomicko-finančná matematika a modelovanie
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. Mgr. Pavol Bokes, PhD.

Bratislava 2025 **Bc. Ivana Oravcová**



62327950

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Ivana Oravcová
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Vnútorné a vonkajšie riešenia náhodných procesov
Inner and outer solutions to random processes

Anotácia: V aplikovanej matematike sa často zaujímame o modelovanie dynamiky určitého systému. V prípade deterministických systémov (pohyb planét, reaktor obsahujúci veľa molekúl, početný ekosystém) sú vhodným modelovacím nástrojom diferenciálne rovnice; v prípade stochastických systémov (bunka s malým počtom molekúl, ekosystém s malým počtom jedincov) sa používajú náhodné procesy. V oboch prípadoch sa často stretávame s problémom viacerých časových škál, ktorý nastáva keď systém kombinuje rýchle javy (trvajúce povedzme sekundy) s javmi ktoré sú pomalé (povedzme hodiny). Takéto systémy vieme zjednodušiť metódou vnútorných a vonkajších riešení. Práca sa zameria na použitie tejto metódy na stochastické modely motivované stochastickou expresiou génov v bunkách.

Vedúci: doc. Mgr. Pavol Bokes, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: doc. Mgr. Igor Melicherčík, PhD.

Dátum zadania: 11.01.2024

Dátum schválenia: 25.01.2024 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Poďakovanie Touto cestou by som sa chcela poďakovat môjmu vedúcemu diplomovej práce doc. Mgr. Pavlovi Bokesovi, PhD. za ochotu, pomoc, odborné rady, podnetné pripomienky a rýchle reakcie na moje otázky, ktoré mi pomohli pri písaní tejto práce.

Abstrakt v štátnom jazyku

ORAVCOVÁ, IVANA: Vnútorné a vonkajšie riešenia náhodných procesov [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. Mgr. Pavol Bokes, PhD., Bratislava, 2025, 79s.

V tejto práci budeme analyzovať model stochastickej génovej expresie pomocou aproximácie na kvázi stacionárne riešenie, rozdelením modelu na vnútorné a vonkajšie riešenia. Najprv však v prvej kapitole uvedieme potrebnú teóriu na pochopenie metódy modelovania základných biologických procesov. Ďalej sa pozrieme na vznik vnútorného a vonkajšieho riešenia modelu, ktoré vzniknú rozdelením procesov na tie, ktoré sa dejú rýchlo a tie, ktoré sa dejú pomaly. Túto teóriu si vysvetlíme aj na príklade Michaelis-Mentenovej kinetiky. Následne predstavíme ako sa modelovanie zmení, keď je v procese prítomná stochasticita. Predstavíme teóriu markovovských procesov a Chapman-Kolmogorovej rovnice a využijeme ju na odvodenie hlavnej rovnice, pomocou ktorej budeme modelovať stochastickú génovú expresiu. V druhej kapitole predstavíme model stochastickej génovej expresie, ktorý budeme analyzovať pomocou hlavnej rovnice. Vyjadríme si rozdelenie koncentrácie voľných bielkovín v popísanom modeli a budeme skúmať ako sa toto rozdelenie mení pre rozdielne voľby hodnoty parametrov.

Kľúčové slová: Vnútorné a vonkajšie riešenia, Zákon pôsobenia hmotností, Chapman-Kolmogorova rovnica, stochastická expresia génu, multimér

Abstract

ORAVCOVÁ, IVANA: Inner and outer solutions to random processes [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. Mgr. Pavol Bokes, PhD., Bratislava, 2025, 79p.

We will analyze the stochastic gene expression model using an approximation to a quasi-stationary solution by dividing the solution into an inner and outer solutions. In the first chapter, we will present the necessary theory to understand the method of modeling basic biological processes. We will look at the formulation of the inner and outer solutions of the model, which arise by dividing processes into those that occur quickly and those that occur slowly. We will also explain this theory using example of Michaelis-Menten kinetics. Then, we will present how modeling changes when stochasticity is present in the process. We will introduce the theory of Markov processes and the Chapman-Kolmogorov equation and use it to derive the master equation with which we will model stochastic gene expression. In the second chapter, we will present the stochastic gene expression model, which we will analyze using the master equation. We will derive the distribution of free protein concentration in the described model and examine how this distribution changes for different choices of parameters.

Keywords: Inner and outer solution, law of mass action, Chapman-Kolmogorov equation, stochastic expression of gene, multimer

Obsah

Úvod **8**

1 Modelovanie biologických procesov **10**

1.1 Základné biologické procesy bunky 10

1.2 Zákon pôsobenia hmotností a chemická kinetika 12

1.3 Teória singulárnych perturbácií - oddelenie časových mierok 15

1.3.1 Michaelis-Mentenová kinetika 18

1.4 Stochastická génová expresia 24

1.4.1 Markovovský proces a Chapman-Kolmogorova rovnica 24

1.4.2 Modelovanie stochastickej génovej expresie pomocou hlavnej rovnice 30

2 Model stochastickej génovej expresie a deterministického viazania voľných bielkovín do multimérov **33**

2.1 Formulácia a analýza modelu 33

2.2 Skúmanie priestoru parametrov 46

Záver **60**

Zoznam použitej literatúry **63**

Príloha A **66**

A.0.1 66

A.0.2 68

A.0.3 71

A.0.4 74

A.0.5 78

Úvod

Molekulárna biológia je rýchlo sa vyvíjajúce odvetvie vedy [2], ktoré úzko súvisí so štúdiom buniek a procesov dejúcich sa v nich. Na opísanie a štúdium procesov dejúcich sa v bunkách sa často používajú matematické modely. Napríklad regulačné procesy sú dôležité pre molekulárnu biológiu a často sú modelované matematickými procesmi [13].

Génová expresia je proces, počas ktorého sa syntetizuje genetická informácia a vytvárajú sa bielkoviny [17]. Bielkoviny plnia dôležitú úlohu pri základných biologických procesoch v rámci bunky. Produkcia bielkovín sa deje v malých počtoch, a preto množstvo bielkovín v rámci bunky podlieha náhodným fluktuáciám. Tieto náhodné fluktuácie v bielkovinových počtoch sú dôvodom rozdielnosti medzi jednotlivými bunkami [3]. Príkladom je, keď bunky reagujú rozdielne na rovnakú liečbu [3]. Fluktuácie v počte bielkovín sú škodlivé pre fungovanie bunky, keďže počet bielkovín musí byť v rámci určitých hodnôt pre ich optimálnu funkciu [19]. Bunky si preto vytvorili regulačné procesy, ktoré udržujú počet bielkovín v optimálnom rozhraní.

Články [3, 4, 19] analyzujú modely s alternatívnymi možnosťami regulácie náhodných fluktuácií v bielkovinovom počte. V článkoch [4, 19] do modelu génovej expresie zakomponovali voľné väzobné miesta, na ktoré sa voľné bielkoviny v prípade vysokého počtu vedia uviazať. Pričom v článku [4] viazané bielkoviny podliehajú zániku a v článku [19] sú viazané bielkoviny chránené pred zánikom. V článku [3] náhodné fluktuácie v počte bielkovín regulujú tak, že voľné bielkoviny majú možnosť zviazať sa do multimérov a pri úbytku voľných bielkovín sa vedia znovu rozviazať a doplniť tak počet voľných bielkovín.

V tejto práci sa v prvej kapitole oboznámime s potrebnými vedomosťami na zostavenie modelu génovej expresie. Pozrieme sa na zákon pôsobenia hmotností. Taktiež na to, ako vznikajú vonkajšie a vnútorné riešenia pri modelovaní biologických procesov, ktoré sa dejú v rôznych časových horizontoch. A ako sa modeluje stochastická génová expresia pomocou hlavnej rovnice (angl. master equation), popísaním stochastickej expresie ako markovoského procesu.

V druhej kapitole zostavíme nový model stochastickej génovej expresie. Inšpirujeme sa pri tom článkami [3, 4, 19]. Z článkov [4, 19] sa inšpirujeme stochastickým vznikom bielkovín pomocou náhodných vzplanutí (angl. burst) a z článku [3] sa inšpirujeme reguláciou náhodných fluktuácií pomocou viazania voľných bielkovín do multimérov. Po zostavení

modelu sa ho pokúsime analyzovať a pozrieť sa na rozdelenie koncentrácie voľných bielkovín.

1 Modelovanie biologických procesov

Biológia je veda o živote a procesoch spojených so životom [9]. Život je veľmi komplexný, preto ako sa uvádza v zdroji [9] je naša vedomosť o mechanizmoch, vďaka ktorým živé organizmy prežívajú veľmi čerstvá. Mechanizmy zabezpečujúce život a správanie buniek sa začali študovať v roku 1950 a odvtedy sa ich štúdium vyvíja. V poslednom polstoročí sa objavilo viacero detailov ohľadom bunkových a molekulárnych procesoch [9]. Detailnejšie skúmanie týchto procesov umožnilo matematické modelovanie procesov ako aj pokrok v bioinformatike.

Ako je vysvetlené v knihe [23], matematické modelovanie je pokus o popísanie stavov elementov systému a popísanie správania systému, konkrétne ako sa elementy navzájom ovplyvňujú. Model by mal byť dostatočne detailný, aby dobre zachytával správanie systému. Avšak pri zostavení modelu je dôležité premyslieť si, ktoré procesy chceme zahrnúť a aký stupeň detailu chceme použiť [23]. Napríklad zahrnúť do jedného modelu fungovanie celého organizmu by nebolo prospešné, keďže model by bol potom veľmi zložitý a ťažko by sa skúmal. Preto na opísanie fungovania jedného organizmu je rozumnejšie použiť viac modelov na rozdielnom stupni detailnosti procesov. Základom zostavenia dobrého modelu je zachytiť podstatné charakteristiky skúmaného procesu bez toho, aby sme model zaťažovali zbytočnými podrobnosťami [23]. Každý model je do istej miery zjednodušením reality, avšak aj napriek tomu sú užitočné, lebo nám pomáhajú pochopiť kvantitatívne správanie systému.

1.1 Základné biologické procesy bunky

V kontexte molekulárnej a bunkovej biológie, model môže opisovať procesy v rámci bunky a to podľa zdroja [23] napríklad transkripciu, transláciu, génovú reguláciu. Na niektoré z týchto procesov sa pozrieme bližšie.

V druhej kapitole budeme zostavovať model génovej expresie, preto sa teraz pozrieme, čo vlastne génová expresia znamená v biologickom kontexte. Vysvetlíme si to však zjednodušene, takéto vysvetlenie postačí na porozumenie modelu v kapitole 2 tejto práce. Podľa knihy [1] má génová expresia dve štádiá a to transkripciu a transláciu.

- **Transkripcia:** V prvom štádiu génovej expresie sa syntetizuje posol RNA (mRNA),

ktorý obsahuje akúsi šablónu pre syntézu bielkovín. Teda zabezpečuje, aby mali bielkoviny potrebné charakteristiky. mRNA je skopírovaná z DNA. Proces transkripcie opisujeme na základe knihy [1]. Označme gén g a mRNA r. Potom proces transkripcie vieme modelovať nasledovne: $g \rightarrow g + r$, [23].

- **Translácia:** Druhou fázou génovej expresie je zjednodušene povedané syntéza bielkoviny z mRNA. Proces translácie opisujeme na základe knihy [1]. Označme bielkovinu p, potom proces translácie môžeme modelovať nasledovne: $r \rightarrow r + p$, [23].

Génová expresia sa teda skladá z dvoch štádií a výsledkom je vytvorenie bielkoviny génom. Ako sme už vyššie spomenuli, pri modelovaní je potrebné si zvoliť stupeň detailnosti, preto sa niekedy procesy v modeli zjednodušujú. Preto tvorbu bielkovín vieme modelovať aj nasledovne $\rightarrow p$, aj keď v realite nič nevznikne z ničoho [23].

Ďalšia reakcia, ktorej bielkoviny môžu podliehať je dimerizácia, alebo viazanie [23]. V dimerizácii sa dve molekuly bielkovín spoja do jednej. Túto reakciu vieme zapísať nasledovne: $2p \rightarrow p_2$, kde p_2 predstavuje jednu molekulu, ktorá vznikla spojením dvoch molekúl bielkovín a nazývame ju dimér.

Opačná reakcia k dimerizácii je disociácia, alebo rozväzovanie [23]. V disociácii sa dimér rozviaže na dve molekuly bielkovín. Túto reakciu vieme zapísať nasledovne $p_2 \rightarrow 2p$.

Reakcia sa volá obojstranná, ak platí v oboch smeroch [23]. Obojstranné reakcie sú v biológii bežné a vieme ich napísať buď pomocou dvoch rovníc, alebo pomocou obojstrannej šípky: $2p \longleftrightarrow p_2$.

Posledný proces molekulárnej biológie, ktorý si na úvod predstavíme je zánik. Konkrétne najjednoduchšiu formu zániku, kedy sa molekula zmení na nič [23]. Proces zániku vieme zapísať nasledovne: $p \rightarrow \emptyset$, kde $\emptyset$ predstavuje prázdnu množinu.

Predstavili sme si niektoré základné procesy génovej expresie a regulácie, avšak keď

ich spojíme dokopy, získame jednoduchý autoregulačný model [23]:

$$\begin{aligned} g &\to g + r, \\ r &\to r + p, \\ 2p &\to p_2, \\ p_2 &\to 2p, \\ p &\to \emptyset, \\ r &\to \emptyset. \end{aligned}$$

Môžeme si uvedomiť, že tento model neobsahuje veľa detailov a je zjednodušením, avšak neskôr uvidíme, že na zostavenie modelu génovej expresie, ktorý bude mať výpovednú hodnotu, nám budú stačiť tieto základné reakcie.

1.2 Zákon pôsobenia hmotností a chemická kinetika

Keď sa na chemické rovnice pozrieme z diaľky, môžeme povedať, že pozostávajú z individuálnych sekvenčných krokov [18]. Každý krok sa zrealizuje, ak molekuly jedného alebo viacerých chemických druhov (závisí od konkrétnej reakcie) spolu reagujú prostredníctvom zrážok a premenia svoju chemickú podstatu, čím vzniknú molekuly iného chemického druhu [18].

Reakcie podliehajú takzvanému zákonu pôsobenia hmotností (angl. law of mass action), čo zjednodušene znamená, že rýchlosť reakcie je úmerná produktu koncentrácie reaktantov [14]. Teraz si to poďme však bližšie vysvetliť. Uvažujme jednoduchú reakciu, ktorá z dvoch reaktantov A, B vytvorí produkt C [13]. Pre jednoduchosť zatiaľ uvažujme, že táto rovnica nie je obojstranná, teda neplatí v oboch smeroch. Píšeme:

$$A + B \xrightarrow{k} C,$$

kde k je rýchlosť reakcie.

Predpokladáme, že zmena koncentrácie produktu v priebehu času je úmerná počtu zrážok reaktantov, teda molekúl A a B vynásobená pravdepodobnosťou, že v prípade zrážky molekúl, reakcia naozaj nastane [13]. Pravdepodobnosť zrážky molekúl nezávisí od počtu prítomných molekúl, ale od koncentrácie molekúl v celkovom objeme priestoru, kde sa nachádzajú. Predpokladáme, že faktory okolitého prostredia (napríklad teplota)

sú konštantné [13]. Koncentrácie môžu byť vyjadrené hmotnosťou látky na objem alebo počtom molekúl na objem. Uvedieme definíciu koncentrácie podľa zdroja [13]:

Definícia 1.1 (Hmotnosť látky na objem a počet molekúl na objem, Mol)**.** *Jeden mol látky obsahuje* 6.022045×10^{23} *molekúl tejto látky, referenciou je počet atómov izotopu* C^{12} *v* 12 *gramoch uhlíka, ktorý pozostáva výlučne z izotopu* C^{12}*. Avogadrova konštanta je počet molekúl v jednom mole,* $N_A = 6.022045 \times 10^{23} mol^{-1}$*. Koncentrácia sa vyjadruje v mol na liter* $1\ Mol = 1\ mol/liter$*. Molekulová hmotnosť látky je hmotnosť jedného molu. Na označenie molovej hmotnosti sa používa jednotka dalton (Da):* $1Da = 1\ g/mol$*.*

Koncentrácia látky môže byť teda udaná v počte molekúl na liter mol alebo ako hmotnosť na objem, pričom molová hmotnosť molekuly je súčtom hmotností atómov, z ktorých sa molekula skladá [13].

Vráťme sa naspäť k reakcii $A + B \xrightarrow{k} C$. Označme $X_A(t), X_B(t), X_C(t)$ počet molekúl jednotlivých látok v danom objeme V, potom koncentrácie jednotlivých látok A, B, C v moloch sa môžu vyjadriť nasledovne [13]:

$$x_A = [A] = \frac{X_A}{N_A V},$$
$$x_B = [B] = \frac{X_B}{N_A V},$$
$$x_C = [C] = \frac{X_C}{N_A V}.$$

Pričom hranaté zátvorky [.] reprezentujú koncentráciu danej látky [12]. Potom podľa zdroja [13] platí:

$$x_A + x_C = x_B + x_C = \text{konštanta v čase.}$$

Pre vhodnú konštantu r_1 potom výraz $r_1 x_A x_B \Delta t$ predstavuje očakávaný počet zrážok molekúl v čase Δt [13]. Avšak nie všetky tieto zrážky vedú k reakcii, len časť z nich r_2 vedie k reakcii. Keďže nie je isté, že pri zrážke molekúl v časovom intervale Δt dôjde k reakcii, molekulárne počty sú definované stochasticky [13]:

$$P\left(\begin{pmatrix} X_A(t+\Delta t) \\ X_B(t+\Delta t) \\ X_C(t+\Delta t) \end{pmatrix} = \begin{pmatrix} l-1 \\ m-1 \\ n+1 \end{pmatrix} \middle| \begin{pmatrix} X_A(t) \\ X_B(t) \\ X_C(t) \end{pmatrix} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}\right) = \tilde{k} x_A x_B \Delta t + O(\Delta t^2),$$

kde $\tilde{k} = r_1 r_2$. Čas medzi dvoma reakciami má exponenciálne rozdelenie s predpokladom $\frac{1}{\tilde{k} x_A x_B} = \frac{N_A^2 V^2}{k X_A X_B}$ [13]. Poznamenajme, že jednotka $\tilde{k}$ je $1/(\mathrm{Mol}^2\mathrm{s})$.

Ak sú molekulové počty veľké, rozptyl procesu sa zmenšuje: $E(X_C(t+\Delta t) - X_C) = = \tilde{k} x_A x_B \Delta t + O(\Delta t^2)$. Koncentrácia látky C je vyjadrená nasledovne: $x_C = \frac{X_C}{N_A V}$. Potom na základe zdroja [13] dostaneme:

$$\frac{\Delta x_C}{\Delta t} = \frac{\tilde{k} x_A x_B}{N_A V} + O(\Delta t).$$

Definovaním $k = \frac{\tilde{k}}{N_A V}$ dostávame $\Delta x_C = x_A x_B k$. Použitím $\Delta x_C = x_C(t+\Delta t) - x_C(t)$ a následným prechodom k limitnému prípadu $\Delta t \to 0$, podľa zdroja [13] získame:

$$\frac{\mathrm{d}x_C}{\mathrm{d}t} = k x_A x_B.$$

Takto sme na základe zdroja [13] sformulovali zákon pôsobenia hmotností a teda rýchlosť danej reakcie. Jednotka k má teraz rozmery $\frac{1}{\mathrm{mol}\times\mathrm{s}}$. Rovnakým postupom sa sformulujú aj rovnice pre koncentrácie x_A, x_B látok A, B a dostaneme sústavu rovníc opisujúcu danú reakciu:

$$\begin{aligned}
\frac{\mathrm{d}x_A}{\mathrm{d}t} &= -k x_A x_B, \\
\frac{\mathrm{d}x_B}{\mathrm{d}t} &= -k x_A x_B, \\
\frac{\mathrm{d}x_C}{\mathrm{d}t} &= k x_A x_B.
\end{aligned}$$

Teraz sa pozrieme na zložitejší prípad obojstrannej reakcie na základe zdroja [1]. Rovnicu ponecháme vo všeobecnom tvare a definujeme nasledovne:

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C,$$

kde k_-, k_+ sú rýchlosti reakcií. Potom, keď na túto obojstrannú reakciu použijeme zákon pôsobenia hmotností, dostaneme sústavu diferenciálnych rovníc reakcie [1]:

$$\begin{aligned}
\frac{\mathrm{d}x_A}{\mathrm{d}t} &= k_- x_C - k_+ x_A x_B, \\
\frac{\mathrm{d}x_B}{\mathrm{d}t} &= k_- x_C - k_+ x_A x_B, \\
\frac{\mathrm{d}x_C}{\mathrm{d}t} &= k_+ x_A x_B - k_- x_C.
\end{aligned}$$

Ekvilibrium je podľa knihy [23] také riešenie, kedy sa koncentrácie nemenia v čase. Teda ekvilibrium je riešenie rovníc pre zmenu koncentrácie v čase rovným nule, teda podľa

knihy [12] máme:

$$\frac{\mathrm{d}x_A}{\mathrm{d}t} = k_- x_C - k_+ x_A x_B = 0,$$
$$k_- x_C = k_+ x_A x_B,$$
$$\frac{x_C}{x_A x_B} = \frac{k_+}{k_-} = K,$$

kde K je ekvilibriová konštanta.

Nakoniec sa ešte pozrieme, ako by sme zákon pôsobenia hmotností použili na reakciu dimerizácie, uvedenú v podkapitole 1.1. Uvedieme však zovšeobecnenú verziu tejto reakcie na základe zdroja [22]:

$$aA + bB \xrightarrow{k} C.$$

Použitím zákona pôsobenia hmotností získame nasledovnú rovnicu zmeny koncentrácie v čase [22]:

$$\frac{\mathrm{d}x_A}{\mathrm{d}t} = -k x_A^a x_B^b,$$
$$\frac{\mathrm{d}x_B}{\mathrm{d}t} = -k x_A^a x_B^b,$$
$$\frac{\mathrm{d}x_C}{\mathrm{d}t} = k x_A^a x_B^b.$$

Pomocou zákona pôsobenia hmotností sa dajú vyjadriť reakcie pomocou sústavy diferenciálnych rovníc, ktoré opisujú zmenu koncentrácie látok v priebehu času. Vysvetlili sme si aj pojem ekvilibria.

1.3 Teória singulárnych perturbácií - oddelenie časových mierok

Reakcie sa často dejú na rozdielnych časových škálach. Niektoré sa dejú rýchlejšie a niektoré pomalšie. Na priblíženie myšlienky si uvedieme príklad o slnečnici z knihy [13]. Slnečnica denne otáča hlavu za slnkom. Zároveň aj rastie do výšky. Rast však trvá mesiace, zatiaľ čo hlavu otáča denne. Slnečnica sa hýbe teda dvomi spôsobmi: pomaly rastom do výšky a rýchlo otáčaním hlavy za slnkom. Predstavme si, že chceme zdokumentovať pohyb slnečnice. Ak urobíme fotku každú hodinu v priebehu niekoľkých dní, uvidíme len otáčanie hlavy slnečnice za slnkom a rast si nevšimneme. Ak však urobíme fotku raz denne v priebehu dvoch mesiacov, uvidíme rast kvetu do výšky, avšak otáčanie hlavy za

slnkom nebude pozorovateľné. Preto buď použijeme mierku vhodnú pre rýchly proces, alebo mierku vhodnú pre pomalý proces, pričom vždy uvidíme len jeden [13].

Rozdelenie reakcie na dve časové škály si ukážeme na všeobecných diferenciálnych rovniciach na základe zdroja [13]:

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= f(x, y), \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= \epsilon\, g(x, y), \end{aligned}$$

kde ϵ je malé, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Rýchlosť procesu opísaného premennou y je rádu ϵ. Tento proces je teda pomalý. Zatiaľ čo proces opísaný premennou x je rádu 1, ide teda o rýchly proces [13].

Na zachytenie pomalého procesu nám pomôže preškálovanie času podľa nasledovnej rovnice [13]:

$$\tau = \epsilon\, t.$$

Potom dostaneme [13]:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}\,\frac{\mathrm{d}\tau}{\mathrm{d}t} = \epsilon\,\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau},$$

keďže platí:

$$\begin{aligned} \tau &= \epsilon\, t, \\ \mathrm{d}\tau &= \epsilon\, \mathrm{d}t, \\ \frac{\mathrm{d}\tau}{\mathrm{d}t} &= \epsilon. \end{aligned}$$

Potom pre rýchly systém s časovou škálou t dostaneme podľa knihy [13] systém diferenciálnych rovníc:

$$\begin{aligned} \frac{\mathrm{d}x(t)}{\mathrm{d}t} &= f(x(t), y(t)), \\ \frac{\mathrm{d}y(t)}{\mathrm{d}t} &= \epsilon\, g(x(t), y(t)). \end{aligned}$$

Zatiaľ čo pre pomalý proces s časovou škálou τ dostaneme systém diferenciálnych rovníc [13]:

$$\epsilon \frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau} = f(x(\tau), y(\tau)),$$
$$\frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau} = g(x(\tau), y(\tau)).$$

Pre systém v rýchlej časovej škále t máme $\frac{\mathrm{d}x(t)}{\mathrm{d}t} = O(\epsilon^0)$ a $\frac{\mathrm{d}y(t)}{\mathrm{d}t} = O(\epsilon^1)$, z čoho vyplýva, že táto časová škála je vhodná pre rýchly proces [13].

Pre systém v časovej škále τ máme $\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau} = O(\epsilon^{-1})$ a $\frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau} = O(\epsilon^0)$, z čoho vyplýva, že táto časová škála je vhodná pre pomalý proces [13].

V takejto formulácii sa vieme zamerať na pomalú alebo rýchlu premennú podľa potreby. Pre $\epsilon > 0$ sú uvedené systémy ekvivalentné [13]. Toto neplatí pre limitné prípady $\epsilon \to 0$, vtedy dostávame rozdielne sústavy diferenciálnych rovníc [13]. Najprv sa pozrieme na rýchly proces pre rýchlu časovú škálu t a v limitnom prípade $\epsilon \to 0$ dostaneme sústavu diferenciálnych rovníc [13]:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, y),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = 0.$$

V tomto limitnom prípade rýchleho procesu dostávame konštantnosť pomalého procesu, teda $y(t)$ =konštanta. V čase sa mení len rýchla premenná x. Môžeme si teda uvedomiť, že takto sa nám podarilo zredukovať dimenziu vďaka oddeleniu časových škál jednotlivých procesov [13].

Teraz sa pozrieme na limitnú verziu $\epsilon \to 0$ systému diferenciálnych rovníc pre pomalý proces [13]:

$$0 = f(x, y),$$
$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = g(x, y).$$

V tomto prípade si môžeme uvedomiť, že predpokladáme, že rýchly proces je v ekvilibriu, teda jeho konštantnosť. Rovnako tak dosiahneme zredukovanie dimenzie [13].

Teória singulárnych perturbácií nám pomáha sledovať systém reakcií, kde sa reakcie dejú v rozdielnych časových škálach. Pomáha nám zredukovať dimenziu premenných a teda aj zjednodušiť sledovaný systém. V nasledujúcej podkapitole si rozdelenie procesu na rýchlu a pomalú časovú škálu ukážeme na konkrétnejšom príklade.

1.3.1 Michaelis-Mentenová kinetika

Uvažujme teraz reakciu, ktorá je katalyzovaná enzýmom, teda bielkovinami, ktoré sú kľúčové v biochémii. Sú dôležitými katalyzátormi, ktoré vyvolávajú biochemické reakcie tým, že znižujú aktivačnú energiu potrebnú na prebehnutie reakcie [2]. Veľa chemických procesov v rámci bunkovej biológie zahŕňa enzýmy [1]. Takúto reakciu s enzýmom, ktorý v priebehu reakcie nemení formu, opisuje Michaelis-Mentenová kinetika a my si na nej priblížime problematiku teórie singulárnych perturbácií. Michaelis-Mentenovú kinetiku si vysvetlíme hlavne na základe zdrojov [1, 2, 9, 12, 13, 14].

Reakcia opisovaná Michaelis-Mentenovou kinetikou zahŕňa substrát S reagujúci s enzýmom E, ktoré dokopy vytvoria komplex SE. Komplex sa nakoniec zmení na produkt P a enzým E [14]. Takáto reakcia sa reprezentuje nasledovne:

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE \xrightarrow{k_2} P + E, \tag{1}$$

kde k_1, k_{-1}, k_2 sú rýchlosti reakcií. Pričom k_2 sa označuje aj ako katalytická rýchlosť enzýmu [9]. Berúc do úvahy zákon pôsobenia hmotností, ktorý hovorí, že rýchlosť reakcie je úmerná koncentrácii reaktantov [14], si označíme koncentrácie nasledovne: $s = [S]$, $e = [E]$, $c = [SE]$, $p = [P]$ [1]. Teraz si môžeme sformulovať systém obyčajných diferenciálnych rovníc systému na základe zdroja [12]:

$$\begin{aligned}
\frac{\mathrm{d}s}{\mathrm{d}t} &= -k_1 se + k_{-1} c, \\
\frac{\mathrm{d}e}{\mathrm{d}t} &= -k_1 se + (k_{-1} + k_2) c, \\
\frac{\mathrm{d}c}{\mathrm{d}t} &= k_1 se - (k_{-1} + k_2) c, \\
\frac{\mathrm{d}p}{\mathrm{d}t} &= k_2 c.
\end{aligned} \tag{2}$$

Aby sme dokončili matematickú formuláciu, potrebujeme ešte počiatočné podmienky pre systém obyčajných diferenciálnych rovníc. Počiatočné podmienky majú nasledovný

tvar [14]:

$$s(0) = s_0, \quad e(0) = e_0, \quad c(0) = 0, \quad p(0) = 0. \tag{3}$$

Môžeme si uvedomiť, že sčítaním rovníc zmeny koncentrácie enzýmu a komplexu v čase dostaneme nulu, teda $\frac{\mathrm{d}e}{\mathrm{d}t} + \frac{\mathrm{d}c}{\mathrm{d}t} = 0$. Znamená to, že takýto súčet je konštantný v čase, teda dostávame, že $e + c = e_0$, kde e_0 je celková koncentrácie dostupných enzýmov [12]. Toto intuitívne dáva zmysel, keďže ako sme si mohli v schéme reakcie (1) všimnúť, enzým je len katalyzátor reakcie a vplyvom reakcie sa nemení [2].

Ďalej si môžeme uvedomiť, že súčet zmeny koncentrácie substrátu, komplexu a produktu v čase dáva tiež nulu, teda $\frac{\mathrm{d}}{\mathrm{d}t}(s + c + p) = 0$. Tento súčet je preto konštantný v čase a môžeme označiť $s + c + p = s_0$, kde s_0 je konštanta [2]. Táto rovnica vyjadruje, že substrát sa nachádza v reakcii (1) buď vo svojej originálnej podobe alebo viazaný na enzým alebo premenený na produkt [2].

Pomocou týchto dvoch pozorovaní vieme pôvodný systém obyčajných diferenciálnych rovníc (2) zjednodušiť na základe zdroja [2] nasledovne:

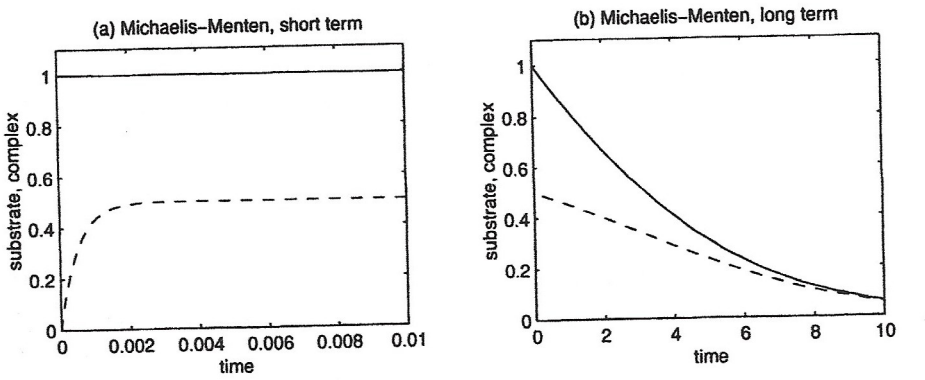
$$\begin{aligned} \frac{\mathrm{d}s}{\mathrm{d}t} &= -k_1 s(e_0 - c) + k_{-1} c, \\ \frac{\mathrm{d}c}{\mathrm{d}t} &= k_1 s(e_0 - c) - (k_{-1} + k_2) c, \\ e &= e_0 - c, \\ p &= s_0 - s - c. \end{aligned} \tag{4}$$

S počiatočnými podmienkami tvaru [2]:

$$s(0) = s_0, \quad e(0) = e_0, \quad c(0) = 0, \quad p(0) = 0. \tag{5}$$

Typické riešenie rovníc (4) si zobrazíme na obrázku 1.

Na obrázku 1 si môžeme naľavo všimnúť, veľmi rýchly rast koncentrácie komplexu prerušovanou čiarou a konštantnosť koncentrácie substrátu plnou čiarou. Na pravej strane obrázka 1 si môžeme všimnúť vývoj koncentrácií substrátu a komplexu v dlhšom časovom horizonte, pričom tuná náhly vzrast koncentrácie komplexu nevidno [2]. Vidíme teda, že najprv koncentrácia komplexu veľmi rýchlo vzrastie v krátkom časovom horizonte, pričom koncentrácia substrátu sa nemení. Následne v dlhšom časovom horizonte môžeme pozorovať vývoj koncentrácie komplexu a substrátu ako sa substrát mení na produkt za pomoci enzýmu [2].



Obr. 1: Koncentrácie substrátu (plná čiara) a komplexu (prerušovaná čiara) v Michaelis-Mentenovej reakcii ako funkcie závislé od času. Riešenie je počítané pre parametre: $k_1 = 4 \times 10^6 \mathrm{M}^{-1}\mathrm{s}^{-1}$, $k_{-1} = 25\mathrm{s}^{-1}$, $k_2 = 15\mathrm{s}^{-1}$, $s_0 = K_m = \frac{k_{-1}+k_2}{k_1} = 10^{-5}\mathrm{M}$ a $e_0 = 10^{-3}K_m = = 10^{-8}\mathrm{M}$ (Zdroj: [2])

Keďže koncentrácia komplexu dosahuje svoje ekvilibrium oveľa rýchlejšie ako koncentrácia substrátu, pre zjednodušenie riešenia budeme predpokladať, že koncentrácia komplexu je vo svojom ekvilibriu a nemení sa v čase [12]:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = k_1 s(e_0 - c) - (k_{-1} + k_2)c = 0. \tag{6}$$

Takýto prístup a predpoklad sa volá aproximácia kvázi stacionárneho stavu [12]. Pomocou tejto aproximácie dostávame:

$$k_1 s(e_0 - c) - (k_{-1} + k_2)c = 0,$$
$$k_1 s(e_0 - c) = (k_{-1} + k_2)c,$$

z toho podľa zdroja [12] dostávame:

$$c = \frac{k_1 s\, e_0}{k_{-1} + k_2 + k_1 s} = \frac{s\, e_0}{K_m + s}.$$

Navyše [2]:

$$V = \frac{\mathrm{d}p}{\mathrm{d}t} = k_2 c = \frac{k_2 s\, e_0}{K_m + s} = \frac{V_m s}{K_m + s}.$$

V je rýchlosť, ktorou sa formuje produkt. Táto rovnica sa nazýva aj Michaelis-Mentenová rovnica rýchlosti [2].

V_m sa nazýva maximálna reakčná rýchlosť a K_m je Michaelis-Mentenová konštanta [12], pričom:

$$V_m = k_2 e_0,$$
$$K_m = \frac{k_{-1} + k_2}{k_1}.$$

Takto sa dá reakcia zjednodušiť ešte viac za pomoci aproximácie kvázi stacionárneho riešenia. Avšak možnosť hľadať kvázi stacionárne riešenie vznikla na základe toho, že sme si všimli, že reakcie sa dejú v rozdielnych časových škálach. Ako však postupovať všeobecnejšie, keď túto informáciu nemáme? Myšlienka kvázi stacionárneho riešenia vznikla na základe existencie rýchleho procesu a pomalého procesu, čo nás navádza k použitiu asymptotickej aproximácie so zosúladením oblastí reakcie (angl. matched asymptotic expansions) [2]. Najprv si však musíme jasnejšie charakterizovať čo sa deje na rýchlej časovej škále a čo na pomalej a následne ich asymptoticky spojiť dokopy.

Začneme zbezrozmernením systému obyčajných diferenciálnych rovníc (4) na základe zdroja [13] nasledovne:

$$\tilde{s}(\tau) = \frac{s(t)}{s_0}, \quad \tilde{c}(\tau) = \frac{c(t)}{e_0}, \quad \tilde{e}(\tau) = \frac{e(t)}{e_0}, \quad \tilde{p}(\tau) = \frac{p(t)}{s_0}, \quad \tau = k_1 e_0 t.$$

Ako prvé sa pozrieme na vonkajšie riešenie systému, teda na pomalý proces s dlhším priebehom [2]. Najprv si však pomocou zbezrozmernených veličín upravíme systém obyčajných diferenciálnych rovníc (4) nasledovne [13]:

$$\frac{\mathrm{d}\tilde{s}(\tau)}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{s(t)}{s_0}\right) = \frac{\mathrm{d}s(t)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{1}{s_0} = \left(-k_1 s(e_0 - c) + k_{-1}c\right)\frac{1}{k_1 e_0 s_0},$$
$$\frac{\mathrm{d}\tilde{s}(\tau)}{\mathrm{d}\tau} = -\frac{s}{s_0} + \frac{sc}{s_0 e_0} + \frac{k_{-1}c}{k_1 s_0 e_0} = -\tilde{s} + \tilde{s}\tilde{c} + \kappa_e \tilde{c}.$$

Podobným spôsobom upravíme aj rovnicu pre koncentráciu komplexu a dostaneme nový zbezrozmernený systém obyčajných diferenciálnych rovníc [2]:

$$\frac{\mathrm{d}\tilde{s}}{\mathrm{d}\tau} = \kappa_e \tilde{c} - \tilde{s}(1 - \tilde{c}),$$
$$\epsilon \frac{\mathrm{d}\tilde{c}}{\mathrm{d}\tau} = \tilde{s}(1 - \tilde{c}) - \kappa_m \tilde{c}, \tag{7}$$

kde

$$\epsilon = \frac{e_0}{s_0}, \quad \kappa_e = \frac{k_{-1}}{k_1 s_0} = \frac{K_e}{s_0}, \quad \kappa_m = \frac{k_{-1} + k_2}{k_1 s_0} = \frac{K_m}{s_0}.$$

Konštanta $K_e = \frac{k_{-1}}{k_1}$ je ekvilibriová konštanta reakcie medzi substrátom a enzýmom, pričom κ_e je zbezrozmernená ekvilibriová konštanta a κ_m je zbezrozmernená verzia Michaelis-Mentenovej konštanty K_m [2].

Systém obyčajných diferenciálnych rovníc (7) budeme riešiť pre počiatočné podmienky tvaru [13]:

$$\tilde{s}(0) = 1, \quad \tilde{c}(0) = 0. \tag{8}$$

Za podmienky $\epsilon \ll 1$, čo sa často stáva, keďže v reakcii je väčšinou prítomného menej enzýmu oproti substrátu, budeme riešenie hľadať v tvare [2]:

$$\tilde{s}(t) = \sum_{n=0}^{\infty} \epsilon^n \tilde{s}_n(t), \quad \tilde{c}(t) = \sum_{n=0}^{\infty} \epsilon^n \tilde{c}_n(t).$$

Dosadením týchto tvarov do sústavy obyčajných diferenciálnych rovníc (7) a porovnaním mocnín parametra ϵ získame systém rovníc [2]:

$$\begin{aligned} \frac{\mathrm{d}\tilde{s}_0}{\mathrm{d}\tau} &= \kappa_e \tilde{c}_0 - \tilde{s}_0(1 - \tilde{c}_0), \\ 0 &= \tilde{s}_0(1 - \tilde{c}_0) - \kappa_m \tilde{c}_0. \end{aligned}$$

Po upravení máme:

$$\begin{aligned} \frac{\mathrm{d}\tilde{s}_0}{\mathrm{d}\tau} &= -\frac{\kappa \tilde{s}_0}{\kappa_m + \tilde{s}_0}, \\ \tilde{c}_0 &= \frac{\tilde{s}_0}{\kappa_m + \tilde{s}_0}, \end{aligned} \tag{9}$$

kde $\kappa = \kappa_m - \kappa_e$ [2]. Riešením rovnice pre počiatočné podmienky (8) integrovaním dostávame riešenie:

$$\kappa_m \log(\tilde{s}_0) = A - \kappa\tau, \tag{10}$$

kde A je integračná konštanta [2]. Avšak v tomto riešení je nezrovnalosť na základe zdroja [2]. Ak si zvolíme $A = 1$, aby sme splnili počiatočnú podmienku pre koncentráciu substrátu s, potom $\tilde{c}_0 = \frac{1}{\kappa_m + 1}$ a teda počiatočná podmienka pre koncentráciu komplexu c nie je splnená [2]. Problém spočíva v predpoklade, že c, s sú analytické funkcie premennej ϵ a teda môžu byť rozvinuté do mocninového radu v ϵ. Toto však nie je správny predpoklad [2]. Inak povedané, predpokladali sme, že problém P_ϵ pozostávajúci zo systému diferenciálnych rovníc (7) a počiatočného riešenia (8) má v limite $\epsilon \to 0$ rovnaké riešenie ako problém P_0

[2]. Avšak nezrovnalosť nastáva v tom, že problém P_0 pozostáva len z jednej diferenciálnej rovnice, preto jeho riešenie nemôže spĺňať dve počiatočné podmienky a problém P_0 teda vo všeobecnosti nemá riešenie [2].

Táto nezrovnalosť sa dá vyriešiť pomocou asymptotickej aproximácie so zosúladením oblasti reakcie. Systém rovníc (9) s riešením (10) správne opisuje riešenie problému P_ϵ, avšak nespĺňa počiatočné podmienky, teda riešenie nie je dobré v oblasti $\tau = 0$ a navyše riešenie obsahuje neznámu konštantu A na základe knihy [2]. Tieto nepresnosti sa dajú vyriešiť pomocou vnútorného riešenia, ktoré riešenie charakterizuje v oblasti $\tau = 0$ a teda v krátkom časovom horizonte, preto bude charakterizovať rýchlo sa dejúce reakcie [2]. Následne je potrebné požadovať, aby sa vonkajšie riešenie a vnútorné riešenie hladko prepojili.

Na to, aby sme našli vnútorné riešenie sa potrebujeme bližšie pozrieť na oblasť $\tau = 0$ a na to zadefinujeme novú nezávislú premennú T a závislé premenné $\bar{S}(T)$ a $\bar{C}(T)$ podľa knihy [2] nasledovne:

$$T = \frac{\tau}{\epsilon}, \quad \bar{S}(T) = \tilde{s}(\tau), \quad \bar{C}(T) = \tilde{c}(\tau).$$

Zavedieme si teda novú časovú škálu. Dosadením týchto zmenených premenných do systému obyčajných diferenciálnych rovníc (7) dostávame nasledovný systém obyčajných diferenciálnych rovníc [2]:

$$\begin{aligned} \frac{\mathrm{d}\bar{S}(T)}{\mathrm{d}T} &= \epsilon(\kappa_e \bar{C} - \bar{S}(1 - \bar{C})), \\ \frac{\mathrm{d}\bar{C}(T)}{\mathrm{d}T} &= \bar{S}(1 - \bar{C}) - \kappa_e \bar{C}. \end{aligned} \tag{11}$$

S počiatočnými podmienkami $\bar{S}(0) = 1$, $\bar{C}(0) = 0$. Riešenie budeme hľadať v tvare rozvoja $\bar{S}(T)$ a $\bar{C}(T)$ do mocninového radu ϵ. Následným dosadením do sústavy diferenciálnych rovníc (12) a následným porovnaním mocnín ϵ podľa zdroja [2], dostaneme riešenie:

$$\begin{aligned} \bar{S}_0(T) &= 1, \\ \bar{C}_0(T) &= \frac{1}{1 + \kappa_m} \left(1 - e^{-(1+\kappa_m)T}\right). \end{aligned} \tag{12}$$

Limitne pre $T \to \infty$ dostávame riešenie: $\left(\bar{S}_0(T), \bar{C}_0(T)\right) \to \left(1, \frac{1}{1+\kappa_m}\right)$. Nakoniec nám ešte ostáva nájsť hodnotu konštanty A tak, aby vonkajšie a vnútorné riešenie mali hladký prechod. Na základe knihy [2] toto nastáva pre hodnotu konštanty $A = 1$.

Podmienka $\epsilon = \frac{e_0}{s_0} \ll 1$ je dôležitá. Na základe zdroja [2] je ekvivalentná podmienke pre kvázi stacionárne riešenie, ktorú sme si vysvetlili vyššie. Podmienka býva často splnená vďaka vysokej efektívnosti enzýmov, preto stačí, aby boli prítomné v nízkej koncentrácii [2].

Ako sme si teda ukázali, existencia vnútorného a vonkajšieho riešenia, teda veľmi rýchlych reakcií a pomalých reakcií sa dá využiť pri hľadaní stacionárneho riešenia na redukciu dimenzie na základe toho, ktorý časový horizont chceme bližšie skúmať. Tento fakt využijeme aj neskôr pri skúmaní modelu génovej expresie.

1.4 Stochastická génová expresia

Geneticky identické bunky pod vplyvom rovnakých vonkajších podmienok môžu mať rozdielny molekulárny obsah a tým aj rozdielne reagovať na rovnaké vplyvy [1]. Táto variabilita je spôsobená hlavne tým, že v bunke sú jednotlivé molekuly prítomné v nízkych počtoch, teda aj v nízkych koncentráciách. Preto aj pridanie alebo odobranie jednej molekuly výrazne ovplyvní fungovanie bunky [21].

Šum v bunke sa delí na vnútorný šum a vonkajší šum [21]. Vonkajší šum opisuje rozdiely medzi bunkami, napríklad rozdiely medzi rýchlosťou viazania, odväzovania v rozdielnych bunkách [1]. Zatiaľ čo vnútorný šum vzniká vnútri bunky. Napríklad rozdiely v bielkovinových počtoch, ktoré vznikajú kvôli stochasticite ich vzniku a zániku [21].

Kvôli prítomnosti náhodného šumu, nevieme génovú expresiu modelovať deterministicky pomocou obyčajných diferenciálnych rovníc, ale potrebujeme ju modelovať stochasticky [10]. V tejto kapitole si ukážeme ako sa génová expresia modeluje stochasticky pomocou hlavnej rovnice. Najprv si však v nasledujúcej podkapitole uvedieme potrebnú teóriu spojitých markovovských procesov a Chapman-Kolmogorovej rovnice.

1.4.1 Markovovský proces a Chapman-Kolmogorova rovnica

Náhodné procesy sú voľne povedané procesy, ktoré sa náhodne, teda pravdepodobnostne vyvíjajú v čase [7]. Náhodné procesy sa volajú markovovské procesy za predpokladu, že spĺňajú markovovskú vlastnosť, ktorú si presne definujeme neskôr. Zjednodušene markovovská vlastnosť znamená, že budúci stav náhodného procesu nezávisí od predošlých stavov [8]. Markovovské procesy sa delia na základe charakteristiky času a stavu nasle-

dovne [8]:

- Markovovské reťazce s diskrétnym časom (množina stavov, do ktorých sa proces vie dostať je diskrétna),
- Markovovské reťazce so spojitým časom (množina stavov, do ktorých sa proces vie dostať je diskrétna),
- Markovovské procesy s diskrétnym časom (množina stavov, do ktorých sa proces vie dostať je spojitá),
- Markovovské procesy so spojitým časom (množina stavov, do ktorých sa proces vie dostať je spojitá).

My chceme sledovať génovú expresiu a teda produkciu molekúl bielkovín génom. Podľa zákona pôsobenia hmotností, ktorý sme uviedli v podkapitole 1.2, chceme sledovať koncentráciu molekúl bielkovín v bunke, preto transkripciu budeme modelovať pomocou markovovských procesov so spojitým časom a spojitou množinou možných stavov. Preto si teraz poriadne zadefinujeme markovovské procesy so spojitým časom a spojitou množinou možných stavov hlavne na základe zdrojov [5, 7, 15, 24].

Označme $(\Omega, \mathcal{F}, \mathbb{P})$ pravdepodobnostný priestor, kde Ω označuje množinu elementárnych javov, $\mathcal{F}$ označuje σ-algebru na množine elementárnych javov Ω, ktorá obsahuje všetky možné udalosti a $\mathbb{P}$ je pravdepodobnosť na množine $(\Omega, \mathcal{F})$ [24]. Priestor všetkých možných stavov označíme $(\mathbb{R}^d, \mathcal{B})$, kde $\mathcal{B}$ je borelovská σ-algebra. $X_t = (X_t)_{t\in\mathbb{R}_+}$ je trieda náhodných premenných $X_t : (\Omega, \mathcal{F}, P) \to (\mathbb{R}^d, \mathcal{B})$. Trieda náhodných premenných $(X_t)_{t\in\mathbb{R}_+}$ sa na základe zdroja [24] nazýva stochastický proces so spojitým časom. Ak $X_t = x \in (\mathbb{R}^d, \mathcal{B})$ vravíme, že proces sa nachádza v stave x v čase t. Pre každé $\omega \in \Omega$ sa množina $(X_t(\omega))_{t\in\mathbb{R}_+}$ nazýva realizácia alebo trajektória či dráha [24].

Ďalej definujeme filtráciu na priestore $(\Omega, \mathcal{F})$ ako neklesajúcu triedu $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ podmnožín σ-algebry $\mathcal{F}$ [15]:

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} \quad \text{for } s \leq t. \tag{13}$$

Zjednotenie takýchto σ-algebier označíme $\mathcal{F}_\infty = \cup_{t\in T}\mathcal{F}_t$. Filtráciu generovanú stochastickým procesom X_t označíme: $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$, kde $\sigma(X_s, s \leq t)$ označuje najmenšiu σ-algebru generovanú stochastickým procesom $(X_s)_{s\leq t}$ [15].

Teraz si zadefinujeme markovovský proces na základe zdroja [15]:

Definícia 1.2 (Markovovský proces)**.** *Nech X_t je stochastický proces definovaný na pravdepodobnostnom priestore $(\Omega, \mathcal{F}, P)$ s hodnotami z $\mathbb{R}^d$ a nech $\mathcal{F}_t^X$ je filtrácia generovaná stochastickým procesom $(X_t)_{t\in\mathbb{R}_+}$. Potom $(X_t)_{t\in\mathbb{R}_+}$ je markovovský proces ak spĺňa:*

$$\mathbb{P}(X_t \in \Gamma \mid \mathcal{F}_s^X) = \mathbb{P}(X_t \in \Gamma \mid X_s) \tag{14}$$

pre všetky s, $t \in T$ také, že $s \leq t$ a $\Gamma \in \mathcal{B}(\mathbb{R}^d)$.

Uvedomme si, že filtrácia $\mathcal{F}_t^X$ je generovaná množinou tvaru $\{\omega \mid X_{t_1} \in \Gamma_1,$ $X_{t_2} \in \Gamma_2, ..., X_{t_n} \in \Gamma_n\}$, kde $0 \leq t_1 \leq t_2 \leq ... \leq t_n \leq t$ a $\Gamma_i \in \mathcal{B}(\mathbb{R}^d)$. Potom je definícia markovovského procesu ekvivalentná rovnici:

$$\mathbb{P}(X_t \in \Gamma \mid X_{t_1}, X_{t_2}, ..., X_{t_n}) = \mathbb{P}(X_t \in \Gamma \mid X_{t_n}) \tag{15}$$

pre $n \geq 1$, $0 \leq t_1 \leq t_2 \leq ... \leq t_n \leq t$ a $\mathcal{B}(\mathbb{R}^d)$ [15]. Táto rovnica vyjadruje markovovskú vlastnosť, ktorá voľne hovorí, že markovovské procesy majú zabúdajúcu vlastnosť a teda to aký stav nadobudnú nezávisí od toho v akom stave boli v minulosti.

Teraz si uvedieme definíciu prechodovej pravdepodobnostnej funkcie markovovského procesu (angl. markov transition probability function) na základe zdroja [5]:

Definícia 1.3 (Prechodová pravdepodobnostná funkcia markovovského procesu)**.** *Nech $P(\Gamma, t \mid x, s)$ je nezáporná funkcia definovaná na $0 \leq s < t < \infty$, $x \in \mathbb{R}^d$, $\Gamma \in \mathcal{B}(\mathbb{R}^d)$. Potom P je prechodová pravdepodobnostná funkcia markovovského procesu ak spĺňa:*

1. *pre všetky $0 \leq s < t < \infty$ a pre všetky $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ je funkcia P $\mathcal{B}(\mathbb{R}^d)$-merateľná,*

2. *pre všetky $0 \leq s < t < \infty$ a pre všetky $x \in \mathbb{R}^d$ je $P(., t \mid x, s)$ prevdepodobnostná miera na $\mathcal{B}(\mathbb{R}^d)$,*

3. *Funkcia P spĺňa Chapman-Kolmogorovu rovnicu:*

$$P(\Gamma, t \mid x, s) = \int_{\mathbb{R}^d} P(\Gamma, t \mid y, u) P(\mathrm{d}y, u \mid x, s) \quad \forall x \in \mathbb{R}^d, s < u < t. \tag{16}$$

Ďalej si uvedieme definíciu homogenity markovovského procesu na základe zdroja [5]:

Definícia 1.4 (Homogénny markovovský proces)**.** *Hovoríme, že markovovský proces* $(X_t)_{t\in\mathbb{R}_+}$ *je časovo homogénny, ak jeho prechodová pravdepodobnostná funkcia* $P(\Gamma, t \mid x, s)$ *závisí od* t*,* s *len prostredníctvom ich rozdielu* $t-s$*. Teda pre všetky* $(s,t) \in \mathbb{R}_+^2$*,* $s < t$*, pre všetky* $u \in \mathbb{R}_+$*, pre všetky* $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ *a pre všetky* $x \in \mathbb{R}^d$ *platí:*

$$P(\Gamma, t \mid x, s) = P(\Gamma, t+u \mid x, s+u). \tag{17}$$

Znamená to teda, že prechodová pravdepodobnostná funkcia homogénneho markovovského procesu závisí len od rozdielu počiatočného a koncového času $t-s$ [15]. Prechodovú pravdepodobnostnú funkciu homogénneho markovovského procesu možno vyjadriť aj nasledovne podľa zdroja [15]:

$$P(\Gamma, t \mid x, s) = P(\Gamma, t-s \mid x, 0) = P(t-s, x, \Gamma) \tag{18}$$

pre všetky $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ a $x \in \mathbb{R}^d$. Ak počiatočný čas fixujeme na hodnotu 0. Potom Chapman-Kolmogorova rovnica pre časovo homogénny markovovský proces nadobudne nasledovný tvar [15]:

$$P(t+s, x, \Gamma) = \int_{\mathbb{R}^d} P(s, x, \mathrm{d}y) P(t, y, \Gamma). \tag{19}$$

Pokiaľ uvažujeme markovovský proces, ktorý má prechodovú pravdepodobnostnú funkciu vzhľadom na Lebesgueovu mieru [15], teda:

$$P(\Gamma, t \mid x, s) = \int_{\Gamma} p(y, t \mid x, s) \mathrm{d}y, \tag{20}$$

kde $p(y, t \mid x, s)$ nazveme prechodovú hustotu pravdepodobnosti. Je to funkcia so štyrmi argumentami a to počiatočným stavom x, počiatočným časom s, koncovým stavom y a koncovým časom t [15]. Chapman-Kolmogorova rovnica potom nadobudne tvar [15]:

$$\int_{\Gamma} p(y, t \mid x, s) \mathrm{d}y = \int_{\mathbb{R}^d} \int_{\Gamma} p(y, t \mid z, u) p(z, u \mid x, s) \mathrm{d}y \mathrm{d}z.$$

Keďže $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ je ľubovolná [15], dostaneme nasledovný tvar Chapman-Kolmogorovej rovnice:

$$p(y, t \mid x, s) = \int_{\mathbb{R}^d} p(y, t \mid z, u) p(z, u \mid x, s) \mathrm{d}z. \tag{21}$$

Za platnosti istých podmienok vieme Chapman-Kolmogorovu rovnicu zredukovať na diferenciálnu rovnicu [7]. Tieto predpoklady sú úzko spojené s vlastnosťami spojitosti daného procesu. Sú to konkrétne predpoklady pre všetky $\epsilon > 0$ [7]:

1. $\lim_{\Delta t \to 0} p(x, t + \Delta t \mid z, t)/\Delta t = \mathrm{W}(x \mid z, t)$ rovnomerne v x, z a t pre $\mid x - z \mid \geq \epsilon$,

2. $\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\epsilon} (x_i - z_i) p(x, t + \Delta t \mid z, t) \mathrm{d}x = A_i(z, t) + O(\epsilon)$ rovnomerne v z, ϵ a t,

3. $\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\epsilon} (x_i - z_i)(x_j - z_j) p(x, t + \Delta t \mid z, t) \mathrm{d}x = B_{ij}(z, t) + O(\epsilon)$ rovnomerne v z, ϵ a t.

Proces môže mať spojité trajektórie jedine ak sa funkcia $W(x \mid z, t)$ vynuluje pre všetky $x \neq z$ [7]. Preto musí táto funkcia vyjadrovať nespojitý pohyb, zatiaľ čo funkcie A_i, B_{ij} musia vyjadrovať spojitý pohyb [7].

Za predpokladu splnenia podmienok uvedených vyššie môžeme Chapman-Kolmogorovu rovnicu zderivovať podľa t. Proces derivovania nebudeme uvádzať celý, možno ho nájsť v knihe [7]. My si uvedieme len výsledok zderivovanej Chapman-Kolmogorovej rovnice [7]:

$$\begin{aligned} \frac{\partial p(z, t \mid y, t')}{\partial t} = &- \sum_i \frac{\partial}{\partial z_i} \left(A_i(z, t) p(z, t \mid y, t') \right) + \\ &+ \sum_{i,j} \frac{1}{2} \frac{\partial^2}{\partial z_i \partial z_j} \left(B_{ij}(z, t) p(z, t \mid y, t') \right) + \\ &+ \int \left(\mathrm{W}(z \mid x, t) p(x, t \mid y, t') - \mathrm{W}(x \mid z, t) p(z, t \mid y, t') \right) \mathrm{d}x. \end{aligned} \tag{22}$$

Pomocou tejto zderivovanej Chapman-Kolmogorovej rovnice vieme modelovať spojité markovovské procesy. Teda na základe nej vieme zostaviť aj model génovej expresie ako si ukážeme v nasledujúcej podkapitole. Najprv si však ešte zadefinujeme markovovský skokový proces (angl. markov jump process) a to na základe zdrojov [5, 24].

Uvažujme markovovský proces $(X_t)_{t \in \mathbb{R}_+}$, ktorého množina stavov je spočítateľná (napr. $\mathbb{N}$, $\mathbb{Z}$) [5]. Je to teda markovovský reťazec so spojitým časom. V tomto prípade bude prínosné, ak si zadefinujeme prechodovú pravdepodobnostnú funkciu v jednom bode [5]:

$$p_{ij}(s, t) = p(s, i, t, j) = P(X_t = j \mid X_s = i) \tag{23}$$

pre všetky $t_0 \leq s < t$, $i, j \in E$. Zo všeobecných vlastností markovovských procesov vychádzajú nasledovné vlastnosti prechodovej pravdepodobnostnej funkcie v jednom bode [5]:

a) $p_{ij}(s, t) \geq 0$,

b) $\sum_{j\in E} p_{i,j}(s,t) = 1$,

c) $p_{i,j}(s,t) = \sum_{k\in E} p_{i,k}(s,r)p_{k,j}(r,t)$, kde $t_0 \leq s \leq r \leq t$ v $\mathbb{R}_+$ a $i,j \in E$,

d)

$$\lim_{t\to s^+} p_{i,j}(s,t) = \delta_{ij} = \begin{cases} 1 & \text{pre } i = j, \\ 0 & \text{pre } i \neq j. \end{cases}$$

Pokiaľ je markovovský proces časovo homogénny dostaneme prechodové pravdepodobnostné funkcie $(\bar{p}_{ij}(t))_{t\in\mathbb{R}_+}$, pre ktoré podľa zdroja [5] platí:

$$p_{ij}(s,t) = \bar{p}_{ij}(t-s), \quad s \leq t.$$

Zároveň pre časovo homogénny markovovský proces platí podľa knihy [5] nasledujúca veta:

Veta 1.5. *Prechodové pravdepodobnostné funkcie $(p_{ij}(t))_{t\in\mathbb{R}_+}$ časovo homogénneho markovovského procesu so spočítateľnou množinou stavov E sú rovnomerne spojité v $t \in \mathbb{R}_+$ pre každé fixné $i,j \in E$.*

Túto vetu možno dokázať na základe zdroja [5].

Uvažujme triedu matíc $(P(t))_{t\in\mathbb{R}_+}$ so zložkami $(p_{ij}(t))_{t\in\mathbb{R}_+}$ pre $i,j \in E$. Potom môžeme podmienky c), d) prepísať do maticového tvaru nasledovne [5]:

c') $P(s+t) = P(s)P(t)$ pre hocijaké $s,t \geq 0$,

d') $\lim_{h\to 0^+} P(h) = P(0) = I$.

Trieda stochastických funkcií spĺňajúcich podmienky c'), d') je maticová prechodová funkcia [5]. Pokiaľ maticová prechodová funkcia spĺňa:

$$\sum_{j\in E,\, i\neq j} q_{ij} = -q_{ii} = q_i < \infty$$

pre hocijaké $i \in E$ nazýva sa konzervatívna [5]. Matica $Q = (q_{ij})_{i,j\in E}$ sa nazýva matica intenzít (angl. intensity matrix). Nasledujúcu vetu uvádzame z knihy [5].

Veta 1.6. *Pokiaľ E je konečná, potom maticová prechodová funkcia môže byť vyjadrená prostredníctvom matice intenzít:*

$$P(t) = e^{tQ}, \quad \text{pre } t \geq 0.$$

Túto vetu možno dokázať na základe zdroja [5].

Uvažujme konzervatívny markovovský skokový proces so stacionárnymi časovo homogénnymi prechodovými pravdepodobnosťami s maticou intenzít Q, potom máme:

$$P(X_u = i \ \ \forall u \in [s, s+t] \mid X_s = i) = e^{-q_i t}$$

pre každé $s, t \in \mathbb{R}_+$ a $i \in E$ [5]. Toto nám ukazuje, že čas zotrvania v stave i má exponenciálne rozdelenie s parametrom q_i [5].

Markovovský skokový proces je teda taký proces, ktorý ostane v jednom stave náhodnú časovú dĺžku, ktorá má exponenciálne rozdelenie a následne skočí do iného stavu [24]. Tento proces využijeme pri skladaní modelu stochastickej génovej expresie v nasledujúcej kapitole. Teraz na to už máme postačujúcu teóriu.

1.4.2 Modelovanie stochastickej génovej expresie pomocou hlavnej rovnice

Ako sme si v podkapitole 1.1 povedali, génová expresia sa skladá z dvoch krokov a to z transkripcie mRNA a z translácie, kedy sa z mRNA syntetizuje bielkovina. Ako sa však píše v zdroji [10], na základe experimentálnych štúdií bolo odpozorované, že tvorba proteínu sa deje vo vzplanutiach a teda vzniká viac bielkovín naraz a následne sa chvíľu nedeje nič. Toto vzniká v dôsledku toho, že život mRNA je podstatne kratší ako život bielkoviny a produkcia bielkovín sa deje počas života mRNA [10]. Preto keď sa na proces pozeráme v dlhšom časovom horizonte môžeme proces translácie zanedbať a predpokladať, že bielkoviny sa syntetizujú ihneď po vytvorení mRNA [10]. Za týchto predpokladov sa bielkoviny produkujú v nezávislých vzplanutiach s náhodnou veľkosťou [6] (teda množstvom vyprodukovaných bielkovín).

Koncentráciu bielkovín v molekule budeme modelovať ako spojitú premennú x, ktorá udáva koncentráciu molekúl bielkovín v bunke v čase t, $t \geq 0$ [1]. Bielkoviny vznikajú vzplanutiami (angl. burst), pričom množstvo vzniknutých bielkovín v každom vzplanutí má exponenciálne rozdelenie, čo bolo experimentálne pozorované na základe článku [6]. Pričom frekvenciu vzniku vzplanutí označíme A a priemernú veľkosť vzplanutia označíme b. Bielkoviny taktiež podliehajú degradácii a tú vyjadríme funkciou $c(x)$, teda degradácia môže byť závislá od aktuálnej hladiny bielkovín v bunke [4]. Takto zadefinovaný model

génovej expresie vieme modelovať pomocou hlavnej rovnice [4], ktorá má tvar:

$$\frac{\partial p(x,t)}{\partial t} = \frac{\partial}{\partial x}(c(x)p(x,t)) + \int_0^x \left(b^{-1}e^{-\frac{x-y}{b}} - \delta(x-y)\right) Ap(y,t)\,\mathrm{d}y, \tag{24}$$

kde $p(x,t)$ je hustota pravdepodobnosti koncentrácie bielkovín x v čase t. Príslušná počiatočná podmienka má tvar $p(x,t_0) = \delta(x-x_0)$ [19]. Môžeme si všimnúť, že hlavná rovnica je upravená Chapman-Kolmogorova rovnica, ktorú sme odvodili v podkapitole 1.4.1. Rovnica opisuje zmenu v čase funkcie hustoty koncentrácie bielkovín x. Pričom prvý člen na pravej strane rovnice opisuje zmenu bielkovinovej koncentrácie x na základe degradácie bielkovín a druhý člen predstavuje zmenu na základe vzniku bielkovín exponenciálne rozdelenými vzplanutiami [1].

Funkcia $\delta(x)$ v hlavnej rovnici (24) je Diracova delta funkcia, ktorú si zadefinujeme podľa zdroja [20]. Diracovu delta funkciu môžeme zadefinovať nasledovne:
$\delta(x) = lim_{n\to\infty}\frac{n}{\sqrt{\pi}}e^{-n^2x^2}$. Toto nám dáva funkciu, ktorej integrál cez hodnoty premennej x je jeden a je veľmi úzka a špicatá v oblasti $x = 0$. Diracova delta funkcia má vlastnosti [20]:

$$\delta(x) = 0 \quad \text{pre vetky } x \neq 0,$$
$$\int_{-\infty}^{\infty} \delta(x)\mathrm{d}x = 1.$$

Vďaka týmto vlastnostiam má Diracova delta funkcia dobré využitie, keďže v integráli násobku Diracovej delta funkcie s inou funkciou dostaneme nenulovú hodnotu druhej funkcie v bode $x = 0$ [20], teda:

$$\int_{-\infty}^{\infty} f(x)\delta(x)\mathrm{d}x = f(0).$$

A všeobecnejšie:

$$\int_{-\infty}^{\infty} f(x)\delta(x-y)\mathrm{d}x = f(y).$$

Teraz by sme chceli nájsť stacionárne riešenie rovnice (24). Toto sa dá pomocou Laplaceovej transformácie, ako je ukázané v zdrojoch [1, 6]. My však stacionárne riešenie nájdeme na základe zdroja [4], pričom budeme úplne sledovať postup tohto riešenia.

Rovnica (24) môže byť napísaná v tvare [4]:

$$\frac{\partial p}{\partial t} + \frac{\partial J}{\partial x},$$

pričom

$$J = -c(x)p(x,t) + \int_0^x e^{-\frac{x-y}{b}} Ap(y,t)\,\mathrm{d}y.$$

V stacionárnom riešení $p(x,t) = p(x)$ musí platiť $J = 0$ a taktiež $\frac{\mathrm{d}J}{\mathrm{d}x} = 0$ [4], teda:

$$c(x)p(x) = \int_0^x e^{-\frac{x-y}{b}} Ap(y)\,\mathrm{d}y, \tag{25}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}(c(x)p(x)) = Ap(x) - \frac{1}{b}\int_0^x e^{-\frac{x-y}{b}} Ap(y)\,\mathrm{d}y. \tag{26}$$

Predelením rovnice (25) konštantou b a následným sčítaním s rovnicou (26) dostávame:

$$\frac{\mathrm{d}}{\mathrm{d}x}(c(x)p(x)) + \frac{c(x)p(x)}{b} = Ap(x)\mathrm{d}y. \tag{27}$$

Integráciou rovnice (27) získame všeobecný predpis funkcie hustoty pre koncentráciu bielkovín v stacionárnom stave:

$$p(x) = \frac{\kappa}{c(x)}\exp\left(-\frac{x}{b} + \int \frac{A}{c(x)}\,\mathrm{d}x\right), \tag{28}$$

kde κ je normalizačná konštanta.

Teraz sme si ukázali ako sa vo všeobecnosti modeluje stochastická génová expresia pomocou hlavnej rovnice. A taktiež ako sa rozdiely v rýchlosti procesov dajú využiť na zjednodušenie modelu. Konkrétne keď sme zanedbali transláciu v procese tvorby bielkovín. V nasledujúcej kapitole využijeme uvedenú hlavnú rovnicu a odvodené stacionárne riešenie na analýzu konkrétneho modelu génovej expresie.

2 Model stochastickej génovej expresie a deterministického viazania voľných bielkovín do multimérov

Mnoho bielkovín vzniká v malom molekulárnom počte v rámci jednotlivých buniek, a preto podliehajú náhodným fluktuáciám v ich počte [3]. Vďaka týmto náhodným fluktuáciám dochádza k rozdielom medzi jednotlivými bunkami, keďže ako je uvedené v článku [19], počty bielkovín musia byť udržané v rámci optimálnych hodnôt pre správnu funkciu bunky. Preto bunky reagujú rôzne na mnohé vonkajšie faktory. Príkladom podľa zdroja [3] je, keď rakovinové bunky reagujú rôzne na liečbu, niektoré prežijú smrteľnú liečbu a iné nie. Toto sa deje kvôli náhodným fluktuáciám v rámci buniek. Preto si bunky vytvorili rôzne regulačné mechanizmy na udržanie bielkovín v správnych hodnotách [19]. My sa pozrieme na alternatívny spôsob regulácie náhodných fluktuácií počtu bielkovín v rámci bunky, vytvorením modelu stochastickej génovej expresie hlavne na základe zdrojov [3, 4, 19].

Ako sme už ukázali v podkapitole 1.4.2, bielkoviny budú vznikať v náhodných vzplanutiach [4]. Vzplanutiami myslíme, že gén bude krátko vo fáze intenzívnej tvorby bielkovín, aktívnej fáze. Aktívnu fázu bude striedať dlhšia fáza nečinnosti, kedy bielkoviny nebudú vznikať [16]. Počet vzniknutých bielkovín bude závisieť od toho, ako dlho bol gén v aktívnej fáze [4]. Tieto stochastické vzplanutia vzniku bielkovín budeme rovnako ako v článku [4], modelovať exponenciálnym rozdelením. Toto vyplýva aj z vlastností markovovského skokového procesu, uvedeného v podkapitole 1.4.1. Stochastický vznik bielkovín budeme vyvažovať deterministickým zánikom a deterministickým viazaním bielkovín do multimérov [3].

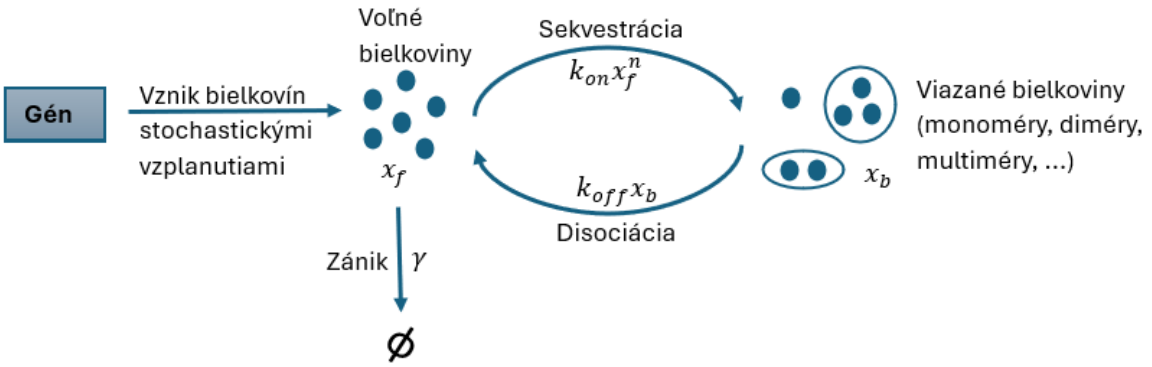
2.1 Formulácia a analýza modelu

Ako sme už spomenuli vyššie, vznik bielkovín budeme modelovať stochastickými vzplanutiami, ktoré sa budú riadiť exponenciálnym rozdelením [4]. Pričom frekvenciu vzniku vzplanutí označíme A. Priemernú veľkosť vzplanutia označíme b. Bielkoviny vzniknuté vo vzplanutí, ktoré sa nestihli zviazať, budeme označovať voľné bielkoviny. Voľné bielkoviny budú podliehať zániku, ktorý sa bude diať deterministicky rýchlosťou γ. Voľné bielkoviny sa budú vedieť viazať do takzvaných multimérov. Multimér je n bielkovín zviazaných do-

kopy. Takéto bielkoviny zviazané do multimérov, budeme volať viazané bielkoviny. Viazané bielkoviny nepodliehajú zániku. V závislosti od nastavenia parametra veľkosti multiméru n, vieme v modeli simulovať viazanie bielkovín do monomérov $n = 1$, dimérov $n = 2$ a multimérov vyššieho počtu. Pričom monomér vznikne zviazaním jednej voľnej bielkoviny do viazanej bielkoviny. Dimér vznikne zviazaním dvoch voľných bielkovín a takýmto princípom ďalej. Takéto viazanie bielkovín sme zvolili na základe článku [3]. Viazanie voľných bielkovín sa deje rýchlosťou k_{on} a je to deterministický proces. Viazané bielkoviny, multiméry sa vedia odviazať na n voľných bielkovín rýchlosťou k_{off}. Toto je tiež deterministický proces. Koncentráciu voľných bielkovín budeme označovať x_f, koncentráciu viazaných bielkovín budeme označovať x_b a celkovú koncentráciu bielkovín označíme x. Pričom medzi nimi platí nasledovný vzťah:

$$x = x_f + x_b. \tag{29}$$

Naformulovaný model s jednotlivými procesmi zobrazíme na obrázku 2.



Obr. 2: Model stochastickej génovej expresie a deterministického viazania voľných bielkovín do multimérov (Vlastný zdroj)

Mnoho biologických procesov sa na základe zdrojov [2, 13, 14] deje v rozličných časových dĺžkach, ako sme si už vysvetlili v podkapitole 1.3. Niektoré sa dejú v priebehu sekúnd a iné v priebehu hodín. Táto časová inkonzistencia vytvára vnútorné a vonkajšie riešenia. Pričom vnútorné riešenie vznikne na základe biologických procesov, ktoré sa dejú rýchlo a vonkajšie riešenie vznikne na základe procesov, ktoré majú dlhší priebeh. My budeme na základe zdrojov [4, 19] predpokladať, že stochastický vznik bielkovín a deterministický zánik bielkovín sa bude diať v dlhšom časovom horizonte. Viazanie bielkovín do multimérov a ich odväzovanie sa bude diať veľmi rýchlo. Tento predpoklad využijeme

na redukciu dimenzie modelu. Na základe zdroja [4] budeme predpokladať, že viazanie a odväzovanie bielkovín sa deje tak rýchlo, že v hocijakom časovom okamihu môžu byť v kvázi stacionárnom stave, teda bude platiť:

$$k_{on}x_f^n = k_{off}x_b. \tag{30}$$

Na základe rovníc (29) a (30) zredukujeme dimenziu modelu. Spravíme to tak, že z rovnice (30) si vyjadríme vzťah pre premennú koncentrácie voľných bielkovín x_b. Dostaneme vzťah:

$$x_b = \frac{k_{on}}{k_{off}}x_f^n. \tag{31}$$

Toto dosadíme do rovnice (29) a dostaneme:

$$x = \frac{k_{on}}{k_{off}}x_f^n + x_f. \tag{32}$$

Celkovú koncentráciu bielkovín x budeme chápať ako funkciu od koncentrácie voľných bielkovín x_f. Funkciu zániku bielkovín označíme $c(x)$. Vieme, že v našom modeli podliehajú zániku len voľné bielkoviny, preto máme $c(x) = \gamma x_f$.

Zlomok $\frac{k_{on}}{k_{off}}$ sa v našich vzťahoch a výpočtoch bude veľa opakovať, preto si ho označíme $\frac{k_{on}}{k_{off}} = k$, čím si zjednodušíme zápis.

Chceli by sme pozorovať ako sa mení koncentrácia voľných bielkovín v našom modeli, preto by sme chceli poznať rozdelenie koncentrácie voľných bielkovín. Ako sme si už ukázali v podkapitole 1.4.2, stochastická génová expresia sa modeluje pomocou hlavnej rovnice (24). My by sme sa chceli pozrieť na stacionárne riešenie $p(x,t) = p(x)$ tejto rovnice, ktoré sme si taktiež odvodili v podkapitole 1.4.2 na základe zdroja [4] a pre pripomenutie ju uvedieme znova:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(c(x)p(x)\right) + \frac{c(x)p(x)}{b} = Ap(x). \tag{33}$$

Ako sme si už ukázali v podkapitole 1.4.2, integráciou rovnice (33) získame všeobecný predpis funkcie hustoty pre celkovú koncentráciu bielkovín v stacionárnom stave:

$$p(x) = \frac{\kappa}{c(x)}\exp\left(-\frac{x}{b} + \int \frac{A}{c(x)}\,\mathrm{d}x\right). \tag{34}$$

Do vzťahu (34) dosadíme parametre nášho modelu. Najprv si vypočítame integrál $\int \frac{A}{c(x)}\,\mathrm{d}x$ pre parametre nášho modelu:

$$\int \frac{a(x)}{c(x)}\,\mathrm{d}x = \int \frac{A}{\gamma x_f}\,\mathrm{d}x = \int \frac{A}{\gamma x_f}\,\frac{\mathrm{d}x}{\mathrm{d}x_f}\,\mathrm{d}x_f. \tag{35}$$

Potrebujeme si vypočítať výraz $\frac{\mathrm{d}x}{\mathrm{d}x_f}$. Urobíme to nasledovne z rovnice (32):

$$\begin{aligned} x &= x_f + k\, x_f^n, \\ \mathrm{d}x &= \mathrm{d}x_f + k\, n\, x_f^{n-1} \mathrm{d}x_f, \\ \mathrm{d}x &= \mathrm{d}x_f \left(1 + k\, n\, x_f^{n-1}\right), \end{aligned}$$

$$\frac{\mathrm{d}x}{\mathrm{d}x_f} = 1 + k\, n\, x_f^{n-1}. \tag{36}$$

Tento výsledok dosadíme do vzťahu (35) a dopočítame integrál:

$$\begin{aligned} \int \frac{A}{\gamma x_f} \frac{\mathrm{d}x}{\mathrm{d}x_f} \mathrm{d}x_f &= \int \frac{A}{\gamma x_f} \left(1 + k\, n\, x_f^{n-1}\right) \mathrm{d}x_f, \\ &= \frac{A}{\gamma} \int \left(\frac{1}{x_f} + k\, n\, x_f^{n-2}\right) \mathrm{d}x_f, \\ &= \frac{A}{\gamma} \left(\ln(x_f) + k\, n \frac{x_f^{n-1}}{n-1}\right). \end{aligned}$$

Z tohto výrazu vidíme, že všeobecné riešenie neplatí pre $n = 1$, preto možnosť viazania do monomérov bude náš špeciálny prípad a pozrieme sa naň samostatne.

Najprv si vyjadríme funkciu hustoty koncentrácie bielkovín pre $n = 2, 3, 4, ...$ Dosadíme teda medzivýpočty do rovnice (34):

$$\begin{aligned} p(x) &= \frac{\kappa}{c(x)} \exp\left(-\frac{x}{b} + \int \frac{a(x)}{c(x)} \mathrm{d}x\right) = \frac{\kappa}{\gamma x_f} \exp\left(-\frac{x}{b} + \frac{A}{\gamma}\left(\ln(x_f) + k\, n \frac{x_f^{n-1}}{n-1}\right)\right), \\ &= \frac{\kappa}{\gamma x_f} e^{-\frac{x}{b}} \exp\left(\frac{A}{\gamma} \ln(x_f)\right) \exp\left(\frac{A}{\gamma} k\, n \frac{x_f^{n-1}}{n-1}\right) = \frac{\kappa}{\gamma} e^{-\frac{x}{b}} x_f^{(\frac{A}{\gamma}-1)} \exp\left(\frac{A}{\gamma} k\, n \frac{x_f^{n-1}}{n-1}\right). \end{aligned}$$

Nás však nezaujíma funkcia hustoty pre celkovú koncentráciu bielkovín, ale pre koncentráciu voľných bielkovín. Použijeme preto prevod hustoty rozdelenia celkovej koncentrácie bielkovín na hustotu rozdelenia koncentrácie voľných bielkovín [4]:

$$p(x(x_f)) \frac{\mathrm{d}x}{\mathrm{d}x_f} = \tilde{p}(x_f). \tag{37}$$

Tento prevod použijeme na funkciu hustoty koncentrácie bielkovín pre $n = 2, 3, 4, ...$

$$\tilde{p}(x_f) = \frac{\kappa}{\gamma} \exp\left(-\frac{x_f + k\, x_f^n}{b}\right) x_f^{\frac{A}{\gamma}-1} \exp\left(\frac{A}{\gamma} k\, n \frac{x_f^{n-1}}{n-1}\right) \left(1 + k\, n\, x_f^{n-1}\right). \tag{38}$$

Dostaneme tak funkciu hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$, kde κ je normalizačná konštanta, ktorou zaručíme, aby platila rovnosť $\int_0^\infty \tilde{p}(x_f)\,\mathrm{d}x_f = 1$, teda:

$$\kappa = \frac{1}{\int_0^\infty \frac{1}{\gamma}\exp\left(-\frac{x_f + k\,x_f^n}{b}\right) x_f^{\frac{A}{\gamma}-1}\exp\left(\frac{A}{\gamma}k\,n\frac{x_f^{n-1}}{n-1}\right)\left(1 + k\,n\,x_f^{n-1}\right)\,\mathrm{d}x_f}. \tag{39}$$

Takto sme dostali funkciu hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Ako sme už spomínali vyššie, tento vzorec neplatí pre voľbu $n = 1$. Túto možnosť, teda musíme vypočítať osobitne. Najprv si ale uvedomíme, že rovnica (36) vyzerá pre voľbu $n = 1$ nasledovne:

$$\frac{\mathrm{d}x}{\mathrm{d}x_f} = 1 + k\,. \tag{40}$$

Postupujeme vypočítaním integrálu (35) pre voľbu $n = 1$:

$$\begin{aligned}\int \frac{A}{\gamma x_f}\,\frac{\mathrm{d}x}{\mathrm{d}x_f}\,\mathrm{d}x_f &= \int \frac{A}{\gamma x_f}\,(1+k)\,\mathrm{d}x_f,\\ &= \frac{A}{\gamma}\,(1+k)\int \frac{1}{x_f}\,\mathrm{d}x_f,\\ &= \frac{A}{\gamma}\,(1+k)\ln(x_f).\end{aligned}$$

Tento výsledok dosadíme do rovnice (34):

$$p(x) = \frac{\kappa}{\gamma x_f}\exp\left(-\frac{x}{b} + \frac{A}{\gamma}\,(1+k)\ln(x_f)\right) = \frac{\kappa}{\gamma x_f}e^{-\frac{x}{b}}x_f^{\frac{A}{\gamma}(1+k)} = \frac{\kappa}{\gamma}e^{-\frac{x}{b}}x_f^{\frac{A}{\gamma}(1+k)-1}.$$

Vyjadrili sme si funkciu hustoty celkovej koncentrácie bielkovín pre prípad $n = 1$. Avšak rovnako ako v prípade pre $n = 2, 3, 4, ...$ nás zaujíma funkcia hustoty rozdelenia koncentrácie voľných bielkovín. Použijeme preto vzťah (37):

$$\tilde{p}(x_f) = \frac{\kappa}{\gamma}e^{-\frac{x_f(1+k)}{b}}x_f^{\frac{A}{\gamma}(1+k)-1}\,(1+k)\,. \tag{41}$$

Dostali sme tak funkciu hustoty koncentrácie voľných bielkovín pre prípad $n = 1$, kde κ je normalizačná konštanta, ktorou zaručíme, aby platila rovnosť $\int_0^\infty \tilde{p}(x_f)\,\mathrm{d}x_f = 1$, teda:

$$\kappa = \frac{1}{\int_0^\infty \frac{1}{\gamma}e^{-\frac{x_f(1+k)}{b}}x_f^{\frac{A}{\gamma}(1+k)-1}\,(1+k)\,\mathrm{d}x_f}. \tag{42}$$

Po dôkladnej analýze sme si všimli, že funkcia hustoty koncentrácie voľných bielkovín pre prípad $n = 1$ sa podobá na funkciu hustoty gama rozdelenia. Uvedieme preto definíciu gama rozdelenia zo zdroja [11].

Definícia 2.1 (Gama rozdelenie)**.** *Náhodná premenná má gama rozdelenie pravdepodobnosti ak má hustotu tvaru:*

$$f(x) = \frac{\lambda^l}{\Gamma(l)} e^{-\lambda x} x^{l-1} \qquad pre \quad x \geq 0,$$

kde $l \in \mathbb{R}^+$ *je parameter tvaru a* $\lambda \in \mathbb{R}^+$ *je parameter škály.* $\Gamma(l) = \int_0^\infty e^{-x} x^{l-1} \mathrm{d}x$ *je gama funkcia. Stredná hodnota má tvar* $\mu = \frac{l}{\lambda}$*, disperzia má tvar* $\sigma^2 = \frac{l}{\lambda^2}$ *a koeficient šikmosti má tvar* $m = \frac{2}{\sqrt{l}}$*.*

Napriek podobe funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ na funkciu hustoty gama rozdelenia, funkcia hustoty nie je v presnom tvare gama rozdelenia. Pokúsime sa preto funkciu hustoty koncentrácie voľných bielkovín pre $n = 1$ upraviť pomocou substitúcie na funkciu hustoty gama rozdelenia.

Začneme použitím substitúcie pre normalizačnú konštantu κ:

$$\begin{aligned}
\frac{1}{\kappa} &= \int_0^\infty \frac{1}{\gamma} e^{-\frac{x_f(1+k)}{b}} x_f^{\frac{A}{\gamma}(1+k)-1} (1+k)\, \mathrm{d}x_f, \\
&= \left| \frac{x_f\,(1+k)}{b} = y \right| = \left| x_f = y\left(\frac{b}{1+k}\right) \right|, \\
&= \int_0^\infty \frac{1}{\gamma} e^{-y} y^{\frac{A}{\gamma}(1+k)-1} \left(\frac{b}{1+k}\right)^{\frac{A}{\gamma}(1+k)-1} (1+k) \left(\frac{b}{1+k}\right) \mathrm{d}y, \\
&= \int_0^\infty \frac{b}{\gamma} e^{-y} y^{\frac{A}{\gamma}(1+k)-1} \left(\frac{b}{1+k}\right)^{\frac{A}{\gamma}(1+k)-1} \mathrm{d}y.
\end{aligned}$$

Normalizačnú konštantu upravenú pomocou substitúcie dosadíme do funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ (41) a dostaneme:

$$\begin{aligned}
\tilde{p}(x_f) &= \frac{\left(\frac{1+k}{b}\right)^{\frac{A}{\gamma}(1+k)-1}}{\gamma \int_0^\infty \frac{b}{\gamma} e^{-y} y^{\frac{A}{\gamma}(1+k)-1} \mathrm{d}y} e^{-\frac{x_f(1+k)}{b}} x_f^{\frac{A}{\gamma}(1+k)-1} (1+k)\,, \\
&= \frac{\left(\frac{1+k}{b}\right)^{\frac{A}{\gamma}(1+k)-1}}{b \int_0^\infty e^{-y} y^{\frac{A}{\gamma}(1+k)-1} \mathrm{d}y} e^{-\frac{x_f(1+k)}{b}} x_f^{\frac{A}{\gamma}(1+k)-1} (1+k)\,,
\end{aligned}$$

$$\tilde{p}(x_f) = \frac{\left(\frac{1+k}{b}\right)^{\frac{A}{\gamma}(1+k)}}{\int_0^\infty e^{-y} y^{\frac{A}{\gamma}(1+k)-1} \mathrm{d}y} e^{-\frac{x_f(1+k)}{b}} x_f^{\frac{A}{\gamma}(1+k)-1}. \tag{43}$$

Vidíme, že táto funkcia hustoty koncentrácie voľných bielkovín pre $n = 1$ už je v tvare funkcie hustoty gama rozdelenia. Pre parameter tvaru platí vzťah:

$$l = \frac{A}{\gamma}(1 + k).$$

Parameter škály má tvar:

$$\lambda = \frac{1 + k}{b}.$$

Strednú hodnotu koncentrácie voľných bielkovín pre $n = 1$ vyjadríme nasledovne:

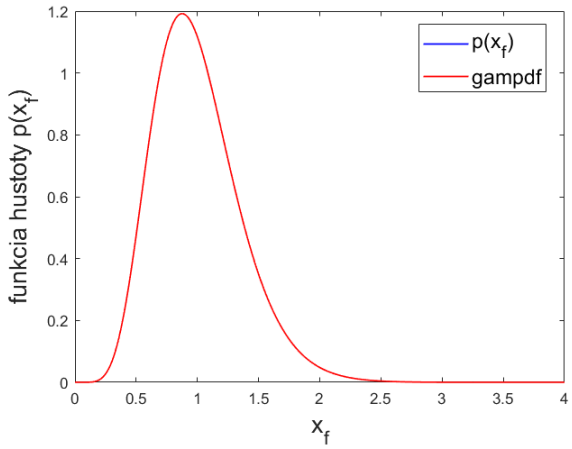
$$\mu = \frac{l}{\lambda} = \frac{\frac{A}{\gamma}(1 + k)}{\frac{1+k}{b}} = \frac{Ab}{\gamma}. \tag{44}$$

Disperziu koncentrácie voľných bielkovín pre $n = 1$ vyjadríme nasledovne:

$$\sigma^2 = \frac{l}{\lambda^2} = \frac{\frac{A}{\gamma}(1 + k)}{\left(\frac{1+k}{b}\right)^2} = \frac{Ab^2}{\gamma(1 + k)}. \tag{45}$$

Koeficient šikmosti koncentrácie voľných bielkovín pre $n = 1$ vyjadríme nasledovne:

$$m = \frac{2}{\sqrt{l}} = \frac{2}{\sqrt{\frac{A}{\gamma}(1 + k)}}. \tag{46}$$

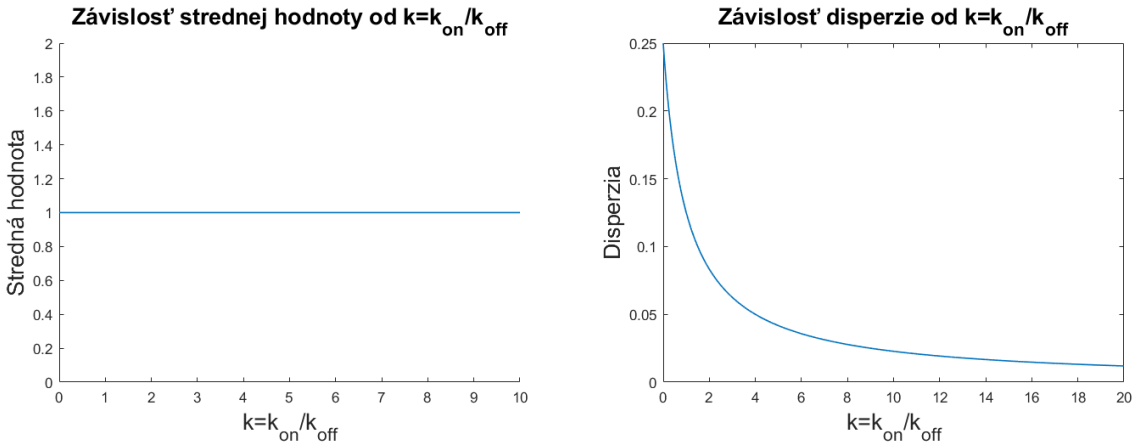


Obr. 3: Funkcia hustoty koncentrácie voľných bielkovín pre $n = 1$ a voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$ a $k_{on} = k_{off} = 0.5$ v porovnaní s funkciou hustoty gama rozdelenia (Vlastný zdroj)

Na obrázku 3 sme si vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 1$ modrou farbou a hustotu gama rozdelenia s rovnakými parametrami červenou farbou. Nie je žiadnym prekvapením, že tieto dve funkcie sa prekrývajú. Funkciu hustoty koncentrácie

voľných bielkovín pre $n = 1$ na obrázku 3 sme vykreslili pre nasledovnú voľbu parametrov: frekvencia vzniku vzplanutí $A = 4$, priemerná veľkosť vzplanutia $b = \frac{1}{4}$, rýchlosť zániku $\gamma = 1$ a rýchlosť viazania a odviazania $k_{on} = k_{off} = 0.5$ a teda $k = \frac{k_{on}}{k_{off}} = 1$.

Ďalej by sme sa chceli pozrieť na závislosť strednej hodnoty a disperzie koncentrácie voľných bielkovín pre $n = 1$ od parametra k. Túto závislosť si zobrazíme na obrázku 4. Strednú hodnotu a disperziu koncentrácie voľných bielkovín pre $n = 1$ vykreslíme pre voľbu parametrov: $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$.



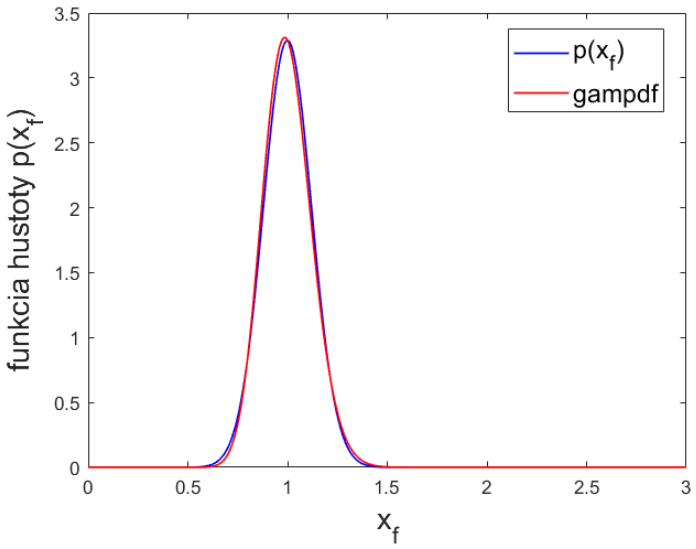
Obr. 4: Závislosť strednej hodnoty a disperzie koncentrácie voľných bielkovín pre $n = 1$ a voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$ od hodnoty parametra k (Vlastný zdroj)

Na obrázku 4 si môžeme všimnúť, že stredná hodnota koncentrácie voľných bielkovín pre $n = 1$ má v závislosti od hodnoty parametra k konštantný priebeh. Toto je konzistentné so vzorcom (44) pre strednú hodnotu koncentrácie voľných bielkovín pre $n = 1$. Stredná hodnota koncentrácie voľných bielkovín pre $n = 1$ nám vyšla konštantne 1, táto hodnota závisí od zvolených hodnôt parametrov modelu. Stredná hodnota koncentrácie voľných bielkovín pre $n = 1$ teda nie je závislá od rýchlosti viazania a odväzovania.

Na obrázku 4 si taktiež môžeme všimnúť, že disperzia koncentrácie voľných bielkovín pre $n = 1$ má v závislosti od hodnoty parametra k klesajúci priebeh. Tento fakt je konzistentný so vzorcom (45) pre disperziu koncentrácie voľných bielkovín pre $n = 1$.

Pre špeciálny prípad $n = 1$ sa nám teda podarilo nájsť pekné rozdelenie koncentrácie voľných bielkovín. Avšak všeobecnejšie pre $n = 2, 3, 4, ...$ sa nám nepodarilo funkciu hustoty koncentrácie voľných bielkovín (38) krajšie upraviť. Skúsime ju však porovnať s funkciou hustoty gama rozdelenia, ktorá bude mať rovnakú strednú hodnotu a disperziu

ako má funkcia hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Pokúsime sa tak zistiť, či aj pre prípad $n = 2, 3, 4, ...$ sa rozdelenie koncentrácie voľných bielkovín správa podobne ako gama rozdelenie.

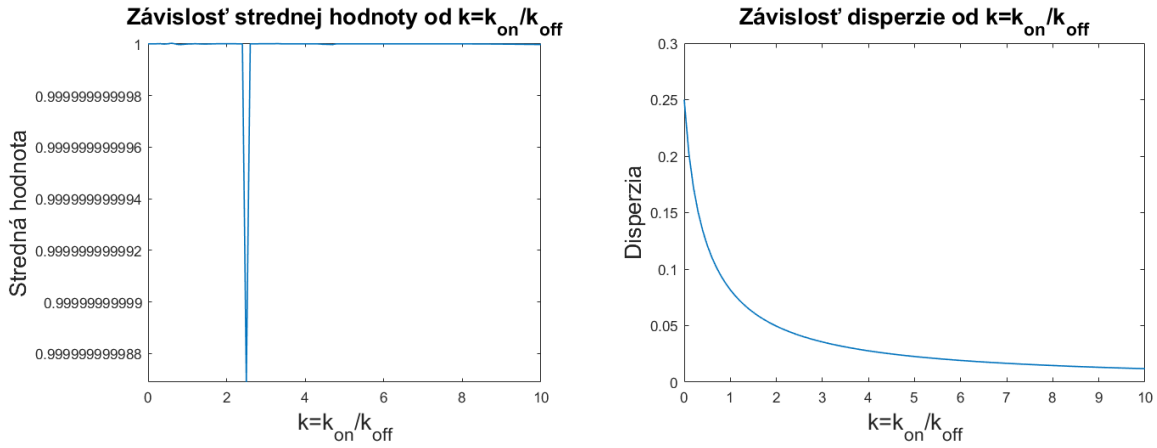


Obr. 5: Funkcia hustoty koncentrácie voľných bielkovín pre $n = 2$ a voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$, $k_{on} = 0.8$ a $k_{off} = 0.1$ a funkcia hustoty gama rozdelenia (Vlastný zdroj)

Na obrázku 5 sme modrou farbou vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 2$ (teda pre viazanie do dimérov) a voľbu parametrov: $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$, $k_{on} = 0.8$ a $k_{off} = 0.1$. Parameter $k = \frac{k_{on}}{k_{off}} = \frac{0.8}{0.1} = 8$. Červenou farbou sme vykreslili funkciu hustoty gama rozdelenia s rovnakou disperziou a rovnakou strednou hodnotou ako má funkcia hustoty koncentrácie voľných bielkovín pre $n = 2$. Na obrázku 5 môžeme vidieť, že funkcie sa veľmi podobajú, avšak nie sú úplne totožné.

Ďalej si na obrázku 6 vykreslíme strednú hodnotu a disperziu v závislosti od veľkosti parametra k. Strednú hodnotu a disperziu koncentrácie voľných bielkovín pre $n = 2$ vykreslíme pre voľbu parametrov: $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$.

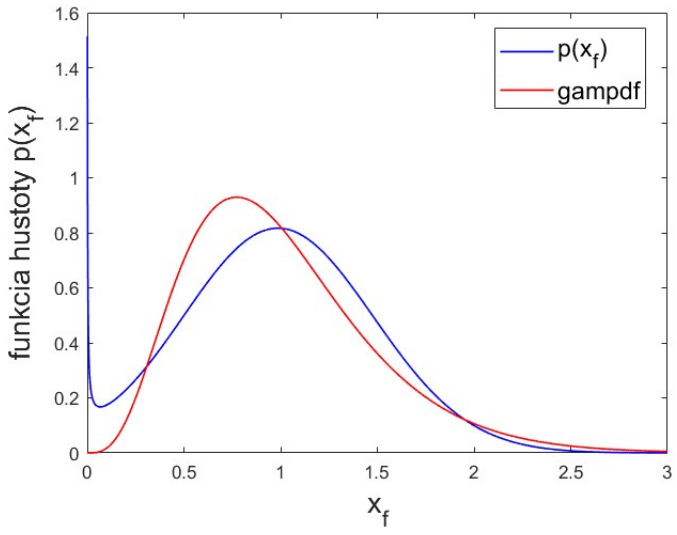
Ako si môžeme na obrázku 6 všimnúť, stredná hodnota koncentrácie voľných bielkovín pre $n = 2$ a uvedenú voľbu parametrov v závislosti od veľkosti parametra k je takmer konštantná hodnota 1, avšak vidíme tam jedno mierne vychýlenie od hodnoty 1. Keby sa koncentrácia voľných bielkovín pre $n = 2$ správala podľa gama rozdelenia, stredná hodnota by bola konštantná v závislosti od veľkosti parametra k. Nastáva teda otázka, či je tento malý výkyv spôsobený numerickou aproximáciou, alebo tam naozaj je.



Obr. 6: Závislosť strednej hodnoty a disperzie koncentrácie voľných bielkovín pre $n = 2$ a voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$, $k_{on} = 0.8$ a $k_{off} = 0.1$ od hodnoty parametra k (Vlastný zdroj)

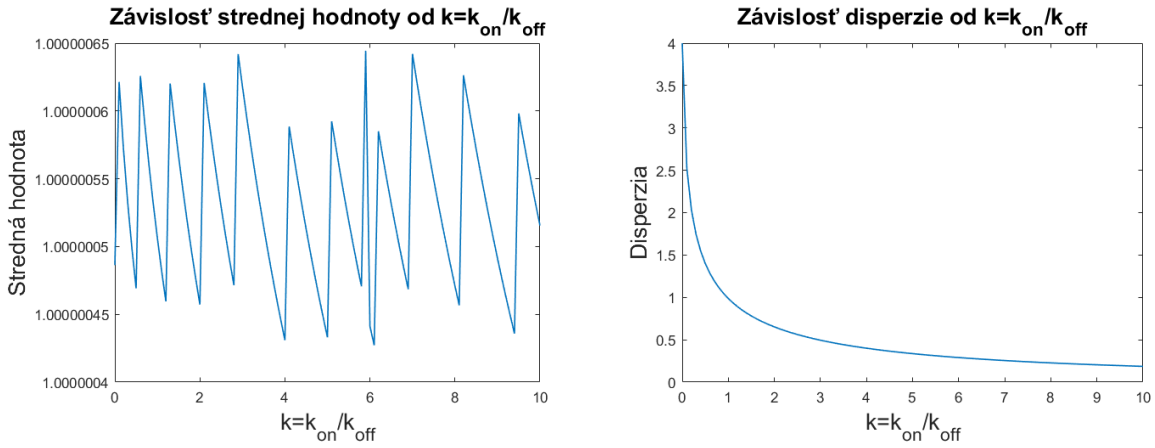
Na obrázku 6 si môžeme všimnúť aj disperziu koncentrácie voľných bielkovín pre $n = 2$ pre uvedenú voľbu parametrov v závislosti od veľkosti parametra k. Môžeme si všimnúť, že má podobný charakter ako disperzia koncentrácie voľných bielkovín pre $n = 1$ v závislosti od veľkosti parametra k.

Teraz si vykreslíme funkciu hustoty koncentrácie voľných bielkovín pre $n = 2$ pre trochu inú voľbu parametrov a porovnáme ju s funkciou hustoty gama rozdelenia s rovnakou strednou hodnotou a disperziou.



Obr. 7: Funkcia hustoty koncentrácie voľných bielkovín pre $n = 2$ a voľbu parametrov $A = \frac{1}{4}$, $b = 4$, $\gamma = 1$, $k_{on} = 0.8$ a $k_{off} = 0.1$ a funkcia hustoty gama rozdelenia (Vlastný zdroj)

Na obrázku 7 sme si modrou farbou vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 2$ a voľbu parametrov: $A = \frac{1}{4}$, $b = 4$, $\gamma = 1$, $k_{on} = 0.8$ a $k_{off} = 0.1$. Teda parameter $k = \frac{k_{on}}{k_{off}} = \frac{0.8}{0.1} = 8$. Môžeme si všimnúť, že sme zmenili len frekvenciu vzniku vzplanutí a priemernú veľkosť vzplanutia. Červenou farbou sme vykreslili funkciu hustoty gama rozdelenia s rovnakou strednou hodnotou a disperziou. Vidíme, že tieto funkcie sú výraznejšie odlišné. Pre funkciu hustoty koncentrácie voľných bielkovín pre $n = 2$ môžeme pozorovať bimodalitu, teda dva lokálne extrémy, ktoré dávajú dve rôzne stacionárne riešenia. Skúsime sa pozrieť na strednú hodnotu a disperziu v závislosti od veľkosti parametra k. Strednú hodnotu a disperziu koncentrácie voľných bielkovín pre $n = 2$ vykreslíme pre voľbu parametrov: $A = \frac{1}{4}$, $b = 4$, $\gamma = 1$ na obrázok 8.



Obr. 8: Závislosť strednej hodnoty a disperzie koncentrácie voľných bielkovín pre $n = 2$ a voľbu parametrov $A = \frac{1}{4}$, $b = 4$, $\gamma = 1$, $k_{on} = 0.8$ a $k_{off} = 0.1$ od hodnoty parametra k (Vlastný zdroj)

Ako si môžeme na obrázku 8 všimnúť, stredná hodnota koncentrácie voľných bielkovín pre $n = 2$ a uvedenú voľbu parametrov v závislosti od veľkosti parametra k nevyzerá byť konštantná. Hodnoty sú blízko 1, avšak vidíme tam skoky. Toto pozorovanie by sme chceli ďalej overiť a zistiť, či sú tieto odchýlky od hodnoty 1 spôsobené numerickými chybami, alebo tam naozaj sú.

Na obrázku 8 si môžeme všimnúť aj disperziu koncentrácie voľných bielkovín pre $n = 2$ a uvedenú voľbu parametrov v závislosti od veľkosti parametra k. Vidíme, že má podobný priebeh ako disperzia koncentrácie voľných bielkovín pre $n = 1$ uvedená na obrázku 4, avšak môžeme pozorovať, že klesá trochu rýchlejšie s rastom hodnoty parametra k.

Vyššie sme poukázali na fakt, že sme si nie istí konštantnosťou strednej hodnoty kon-

centrácie voľných bielkovín pre $n = 2, 3, 4, ...$ v závislosti od hodnoty parametra k. Pokúsime sa pomocou rovnice (24) ukázať, že stredná hodnota pre koncentráciu voľných bielkovín x_f je konštantná v závislosti od hodnoty parametra k, nezávisle od voľby parametra n. Budeme postupovať nasledovne. Rovnicu (24) prenásobíme premennou x a zintegrujeme:

$$\int_0^\infty x \left(\frac{\partial p(x,t)}{\partial t} - \frac{\partial}{\partial x}(c(x)p(x,t)) \right) \mathrm{d}x = \int_0^\infty x \int_0^x \left(b^{-1} e^{-\frac{x-y}{b}} - \delta(x-y) \right) Ap(y,t) \,\mathrm{d}y \,\mathrm{d}x. \tag{47}$$

Začneme úpravou ľavej strany rovnice:

$$\begin{aligned} &\int_0^\infty x \left(\frac{\partial p(x,t)}{\partial t} - \frac{\partial}{\partial x}(c(x)p(x,t)) \right) \mathrm{d}x = \\ =&\int_0^\infty x \left(\frac{\partial p(x,t)}{\partial t} \right) \mathrm{d}x - \int_0^\infty x \left(\frac{\partial}{\partial x}(c(x)p(x,t)) \right) \mathrm{d}x. \end{aligned}$$

Prvý integrál je pre stacionárne riešenie rovný 0, preto pokračujeme úpravou druhého integrálu. Pričom na úpravu tohto integrálu použijeme metódu per partes:

$$\begin{aligned} &- \int_0^\infty x \left(\frac{\partial}{\partial x}(c(x)p(x,t)) \right) \mathrm{d}x = \left| \begin{smallmatrix} u=x,\, u'=1 \\ v'=\left(\frac{\partial}{\partial x}(c(x)p(x,t))\right),\, v=(c(x)p(x,t)) \end{smallmatrix} \right| = \\ =&- \left[xc(x)p(x,t) \right]_0^\infty + \int_0^\infty c(x)p(x,t) \,\mathrm{d}x = \left| \begin{smallmatrix} c(x)=\gamma x_f \\ p(x,t)\frac{\mathrm{d}x}{\mathrm{d}x_f}=\tilde{p}(x_f,t) \end{smallmatrix} \right| = \\ =&- \left[xc(x)p(x,t) \right]_0^\infty + \gamma \int_0^\infty x_f \tilde{p}(x_f,t) \,\mathrm{d}x_f = \gamma \int_0^\infty x_f \tilde{p}(x_f,t) \,\mathrm{d}x_f. \end{aligned}$$

Môžeme predpokladať, že prvý člen je nulový, keďže funkcia hustoty celkovej koncentrácie bielkovín exponenciálne klesá a exponenciálny pokles je rýchlejší ako lineárny rast. Druhý člen, integrál je stredná hodnota koncentrácie voľných bielkovín x_f prenásobená konštantou γ.

Teraz sa pozrieme na pravú stranu rovnice (47):

$$\begin{aligned} &\int_0^\infty x \int_0^x \left(b^{-1} e^{-\frac{x-y}{b}} - \delta(x-y) \right) Ap(y,t) \,\mathrm{d}y \,\mathrm{d}x = \\ =&\int_0^\infty \int_0^x x \left(b^{-1} e^{-\frac{x-y}{b}} - \delta(x-y) \right) Ap(y,t) \,\mathrm{d}y \,\mathrm{d}x = \\ =&\int_0^\infty \int_0^x x \left(b^{-1} e^{-\frac{x-y}{b}} \right) Ap(y,t) \,\mathrm{d}y \,\mathrm{d}x - \int_0^\infty \int_0^x x\delta(x-y) Ap(y,t) \,\mathrm{d}y \,\mathrm{d}x. \end{aligned}$$

Najprv upravíme prvý integrál, kde najskôr zmeníme poradie integrácie a použijeme me-

tódu per partes na dopočítanie:

$$\int_0^\infty \int_0^x x\left(b^{-1}e^{-\frac{x-y}{b}}\right) Ap(y,t)\,\mathrm{d}y\,\mathrm{d}x = \int_0^\infty \int_y^\infty x\left(b^{-1}e^{-\frac{x-y}{b}}\right) Ap(y,t)\,\mathrm{d}x\,\mathrm{d}y =$$
$$= \int_0^\infty \frac{Ap(y,t)}{b} e^{\frac{y}{b}} \int_y^\infty xe^{-\frac{x}{b}}\,\mathrm{d}x\,\mathrm{d}y = \left|\begin{smallmatrix} u=x,\, u'=1 \\ v'=e^{-\frac{x}{b}},\, v=-be^{-\frac{x}{b}} \end{smallmatrix}\right| =$$
$$= \int_0^\infty \frac{Ap(y,t)}{b} e^{\frac{y}{b}} \left(\left[-xbe^{-\frac{x}{b}}\right]_y^\infty + \int_y^\infty e^{-\frac{x}{b}}\,\mathrm{d}x\right)\mathrm{d}y =$$
$$= \int_0^\infty \frac{Ap(y,t)}{b} e^{\frac{y}{b}} \left(\left[-xbe^{-\frac{x}{b}}\right]_y^\infty - \left[be^{-\frac{x}{b}}\right]_y^\infty\right)\mathrm{d}y =$$
$$= \int_0^\infty \frac{Ap(y,t)}{b} e^{\frac{y}{b}} \left(ybe^{-\frac{y}{b}} + b^2e^{-\frac{y}{b}}\right)\mathrm{d}y = \int_0^\infty \left(Ap(y,t)y + Ap(y,t)b\right)\mathrm{d}y =$$
$$= A\int_0^\infty p(y,t)y\,\mathrm{d}y + Ab\int_0^\infty p(y,t)\,\mathrm{d}y = A\int_0^\infty yp(y,t)\,\mathrm{d}y + Ab.$$

Vidíme, že posledný integrál, ktorý nám ostal, je stredná hodnota celkovej koncentrácie bielkovín.

Teraz sa pozrieme na druhý integrál:

$$\int_0^\infty \int_0^x x\delta(x-y)Ap(y,t)\,\mathrm{d}y\,\mathrm{d}x = \int_0^\infty xA\int_0^x \delta(x-y)p(y,t)\,\mathrm{d}y\,\mathrm{d}x =$$
$$= \int_0^\infty xAp(x,t)\,\mathrm{d}x = A\int_0^\infty xp(x,t)\,\mathrm{d}x.$$

Dostali sme strednú hodnotu celkovej koncentrácie bielkovín $E(x)$ prenásobenú konštantou A.

Teraz spojíme všetky medzivýpočty do jednej rovnice:

$$\gamma\int_0^\infty x_f\tilde{p}(x_f,t)\,\mathrm{d}x_f = A\int_0^\infty yp(y,t)\,\mathrm{d}y + Ab - A\int_0^\infty xp(x,t)\,\mathrm{d}x,$$
$$\gamma\int_0^\infty x_f\tilde{p}(x_f,t)\,\mathrm{d}x_f = Ab,$$

$$E(x_f) = \int_0^\infty x_f\tilde{p}(x_f,t)\,\mathrm{d}x_f = \frac{Ab}{\gamma}. \tag{48}$$

Odvodili sme si tak vzorec pre výpočet strednej hodnoty koncentrácie voľných bielkovín. Vidíme, že stredná hodnota koncentrácie voľných bielkovín je nezávislá od hodnoty parametra k a teda je v závislosti od hodnoty tohto parametra konštantná pre ľubovoľnú hodnotu parametra n. Výkyvy na obrázkoch 6 a 8 sú teda spôsobené numerickými aproximáciami pri výpočte strednej hodnoty.

2.2 Skúmanie priestoru parametrov

V tejto podkapitole by sme chceli porovnať správanie funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ a funkcie hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ v závislosti od zmeny hodnoty parametrov k, b. Preto by sme ich chceli mať v istom zmysle podobné. Spravíme to tak, že najprv fixujeme hodnoty parametrov $\gamma = 1$, $Ab = 1$. Budeme fixovať aj hodnotu parametra b. Tú fixujeme rôzne pre jednotlivé grafy. Keďže chceme funkcie hustoty v istom zmysle rovnaké, budeme požadovať, aby sa stredná hodnota koncentrácie voľných bielkovín pre $n = 1$ rovnala strednej hodnote koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Zároveň budeme požadovať, aby sa stredná hodnota koncentrácie viazaných bielkovín pre $n = 1$ rovnala strednej hodnote koncentrácie viazaných bielkovín pre $n = 2, 3, 4, ...$ Budeme voliť hodnotu parametra k rôzne pre prípady $n = 1$ a $n = 2, 3, 4, ...$ tak, aby sa nám stredné hodnoty rovnali.

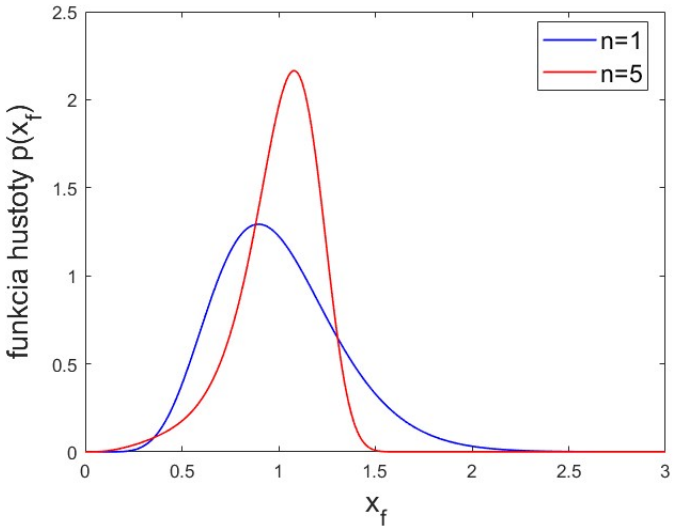
Najprv si vyjadríme vzťah pre výpočet strednej hodnoty pre koncentráciu viazaných bielkovín. Využijeme pri tom vzťah (31), ktorý udáva závislosť medzi koncentráciou voľných bielkovín a koncentráciou viazaných bielkovín. Pre strednú hodnotu koncentrácie viazaných bielkovín potom bude platiť:

$$E(x_b) = \int_0^\infty x_b \, \tilde{p}(x_f) \, \mathrm{d}x_f = \int_0^\infty k \, x_f^n \, \tilde{p}(x_f) \, \mathrm{d}x_f = kE(x_f^n).$$

Pre špeciálny prípad $n = 1$ bude platiť vzťah $E(x_b) = kE(x_f)$, pričom vieme, že funkcia hustoty koncentrácie voľných bielkovín pre $n = 1$ sa správa podľa gama rozdelenia a stredná hodnota sa dá vyjadriť vzťahom (44). Potom pre strednú hodnotu koncentrácie viazaných bielkovín pre $n = 1$ dostaneme vzťah $E(x_b) = kE(x_f) = k\frac{Ab}{\gamma}$. Vyššie sme si povedali, že hodnoty parametrov γ, Ab budeme fixovať na hodnotu 1. Dostaneme teda vzťah $E(x_b) = k$ pre špeciálny prípad $n = 1$.

Parametre na vykreslenie funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ a $n = 2, 3, 4, ...$ budeme voliť nasledovným postupom. Ako prvé si zvolíme hodnotu parametra k pre prípad $n = 2, 3, 4, ...$ a dopočítame strednú hodnotu koncentrácie viazaných bielkovín $E(x_b)$ pre $n = 2, 3, 4, ...$ Následne zvolíme hodnotu parametra k pre $n = 1$ tak, aby sa stredné hodnoty koncentrácií viazaných bielkovín pre oba prípady rovnali. Vieme, že pre $n = 1$ platí vzťah $E(x_b) = k$, preto hodnotu parametra k pre $n = 1$ zvolíme rovnú strednej hodnote koncentrácie viazaných bielkovín pre $n = 2, 3, 4, ...$ Takto zabezpečíme akúsi podobnosť parametrov pre prípady $n = 1$ a $n = 2, 3, 4, ...$

Teraz si vykreslíme funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov: $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$. Parameter k má rôznu hodnotu v závislosti od hodnoty parametra n a to konkrétne $k = 1$ pre $n = 5$ a $k = 1.3998$ pre voľbu parametra $n = 1$. Funkciu hustoty koncentrácie voľných bielkovín vykreslíme pre $n = 1$ a pre $n = 5$. Výsledný graf môžeme vidieť na obrázku 9.

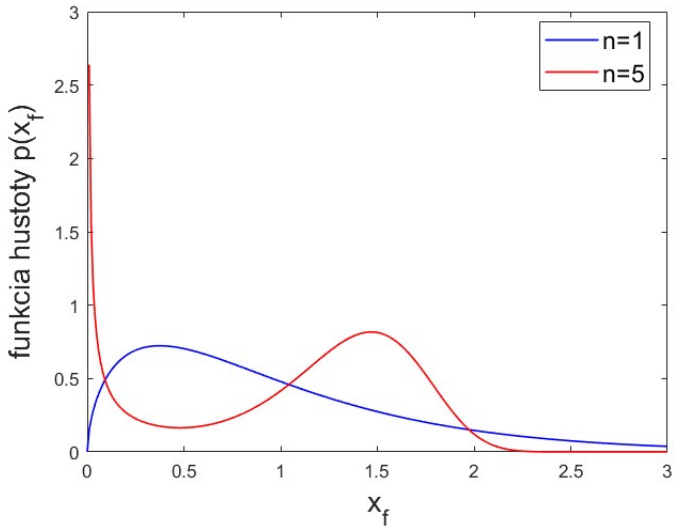


Obr. 9: Funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$. Parameter $k = 1$ pre $n = 5$ a $k = 1.3998$ pre $n = 1$. (Vlastný zdroj)

Na obrázku 9 sme vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 1$ modrou farbou a funkciu hustoty koncentrácie voľných bielkovín pre $n = 5$ červenou farbou. Vidíme, že rozdelenie koncentrácie voľných bielkovín pre rozdielnu voľbu hodnoty parametra n sa líši. Vidíme, že rozdelenie koncentrácie voľných bielkovín pre voľbu $n = 1$ má ťažšie chvosty oproti voľbe $n = 5$. Taktiež sa nám na základe obrázka 9 javí, že funkcia hustoty koncentrácie voľných bielkovín pre $n = 1$ je mierne vyšikmená doľava, zatiaľ čo funkcia hustoty koncentrácie voľných bielkovín pre $n = 5$ je mierne vyšikmená doprava. Toto pozorovanie si overíme výpočtom koeficienta šikmosti pre obe funkcie hustoty. Koeficient šikmosti funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ má hodnotu 0.6455. Hodnota koeficienta šikmosti funkcie hustoty koncentrácie voľných bielkovín pre $n = 5$ je -0.8327. A teda opačné vyšikmenie týchto funkcií hustoty sa nám potvrdilo.

Teraz si vykreslíme funkcie hustoty koncentrácie voľných bielkovín pre trochu inú voľbu parametrov: $A = \frac{1}{4}$, $b = 4$, $\gamma = 1$. Parameter k má rôznu hodnotu v závislosti od

hodnoty parametra n a to konkrétne $k = 1$ pre $n = 5$ a $k = 5.3995$ pre voľbu parametra $n = 1$. Môžeme si všimnúť, že oproti obrázku 9 sme upravili len hodnoty parametrov A, b. Funkciu hustoty koncentrácie voľných bielkovín vykreslíme pre $n = 1$ a pre $n = 5$. Výsledný graf môžeme vidieť na obrázku 10.

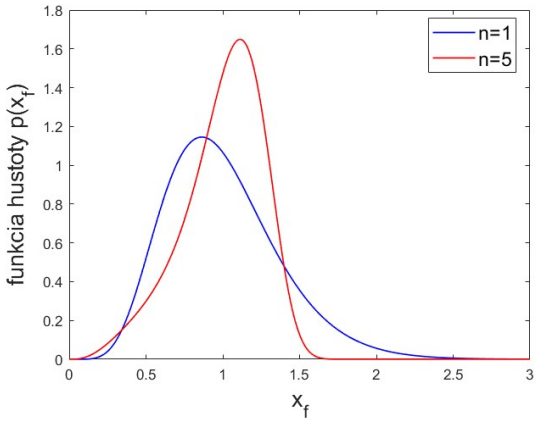


Obr. 10: Funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov $A = \frac{1}{4}$, $b = 4$, $\gamma = 1$. Parameter $k = 1$ pre $n = 5$ a $k = 5.3995$ pre $n = 1$. (Vlastný zdroj)

Na obrázku 10 sme vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 1$ modrou farbou a funkciu hustoty koncentrácie voľných bielkovín pre $n = 5$ červenou farbou. Vidíme, že rozdelenie koncentrácie voľných bielkovín pre rozdielnu voľbu hodnoty parametra n sa líši. Pre funkciu hustoty koncentrácie voľných bielkovín pre $n = 5$ vykreslenú červenou farbou pozorujeme bimodalitu. Teda máme dva lokálne extrémy, ktoré nám dávajú dve rôzne stacionárne riešenia. Taktiež môžeme pozorovať rozdielne vyšikmenie funkcií hustôt, vypočítame si preto koeficienty šikmosti. Koeficient šikmosti funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ má hodnotu 1.5812, zatiaľ čo koeficient šikmosti funkcie hustoty koncentrácie voľných bielkovín pre $n = 5$ má hodnotu -0.3929. Vidíme teda, že funkcie hustôt sú vyšikmené do opačných strán, toto vyšikmenie je konzistentné s vyšikmením na obrázku 9.

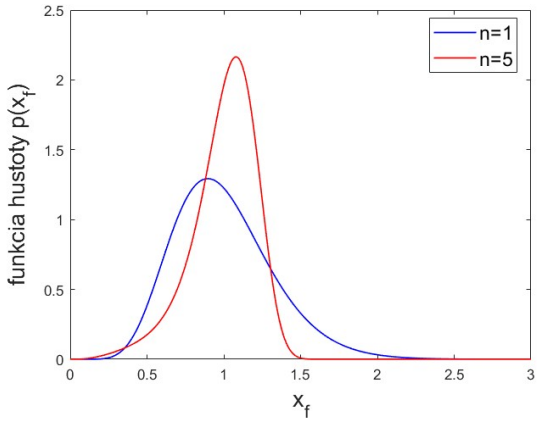
Teraz sa pozrieme ako sa funkcia hustoty koncentrácie voľných bielkovín mení v závislosti od zmeny hodnoty parametra k. Na obrázok 11 si vykreslíme funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ a $n = 5$ pre voľbu parametrov $A = 4$, $b = \frac{1}{4}$,

$\gamma = 1$ a tri rôzne hodnoty parametra k, ktoré závisia od voľby hodnoty parametra n. Tento rozdiel v hodnote parametra k vznikol preto, lebo hodnotu parametra k sme volili tak, aby sa stredné hodnoty koncentrácie viazaných bielkovín pre $n = 1$ a $n = 5$ rovnali. Konkrétne $k = 0.8261$, $k = 1.3998$ a $k = 2.4537$ pre $n = 1$ a $k = 0.5$, $k = 1$ a $k = 2$ pre $n = 5$. Budeme tak vedieť pozorovať ako sa funkcie hustoty koncentrácie voľných bielkovín menia v závislosti od zmeny hodnoty parametra $k = \frac{k_{on}}{k_{off}}$.



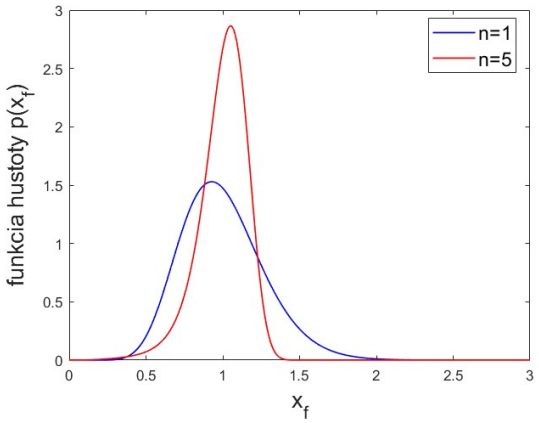
$k = 0.8261$ pre $n = 1$

$k = 0.5$ pre $n = 5$



$k = 1.3998$ pre $n = 1$

$k = 1$ pre $n = 5$



$k = 2.4537$ pre $n = 1$

$k = 2$ pre $n = 5$

Obr. 11: Funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$ (Vlastný zdroj)

Na obrázok 11 sme modrou farbou vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 1$ a červenou farbou funkciu hustoty koncentrácie voľných bielkovín pre

$n = 5$ a voľbu parametrov $A = 4$, $b = \frac{1}{4}$, $\gamma = 1$. Hodnotu parametra k sme na každom grafe volili rôzne. Na prvý pohľad sa všetky grafy na obrázku 11 zdajú podobné. Rozdiel si všimneme, keď sa pozrieme na zvislú os. Pre najnižšiu voľbu hodnoty parametra k, zobrazenú v ľavom hornom rohu, si môžeme všimnúť, že funkcie hustôt sú nižšie a majú trochu ťažšie chvosty. Pre voľbu hodnoty parametra $k = 1.3998$ pre $n = 1$ a $k = 1$ pre $n = 5$, zobrazenú v pravom hornom rohu, si môžeme všimnúť, že funkcie hustôt sú o trochu vyššie a majú trochu ľahšie chvosty. Nakoniec pre poslednú najväčšiu voľbu hodnoty parametra k, zobrazenú v ľavom dolnom rohu, si môžeme všimnúť, že funkcie hustôt sú ešte o trochu vyššie a majú ešte trochu ľahšie chvosty. Teda čím je hodnota parametra k väčšia, tým sú funkcie hustôt koncentrácie voľných bielkovín vyššie a majú ľahšie chvosty. Pričom ak je hodnota parametra k väčšia ako jeden, tak to znamená že $k_{on} > k_{off}$, teda voľné bielkoviny sa rýchlejšie viažu do multimérov ako sa viazané bielkoviny rozväzujú na voľné bielkoviny.

Ako sme sa vyššie presvedčili, náš model povoľuje pre určitú voľbu parametrov bimodalitu riešenia. My by sme sa chceli pozrieť, pre aké hodnoty parametrov bimodalita riešenia nastáva. Preto chceme preskúmať priestor parametrov. Budeme postupovať podobne ako v článku [4]. Bimodalita rozdelenia koncentrácie voľných bielkovín nastáva, keď funkcia hustoty koncentrácie voľných bielkovín $\tilde{p}(x_f)$ má viac lokálnych riešení [4]. Teda keď existuje viacej riešení rovnice $\frac{\mathrm{d}}{\mathrm{d}x_f}\tilde{p}(x_f) = 0$. Deriváciu funkcie hustoty koncentrácie voľných bielkovín si vyjadríme pomocou vzťahu (37) nasledovne:

$$\begin{aligned} \tilde{p}(x_f) &= p(x(x_f))\frac{\mathrm{d}x}{\mathrm{d}x_f}, \\ \frac{\mathrm{d}}{\mathrm{d}x_f}\tilde{p}(x_f) &= \frac{\mathrm{d}}{\mathrm{d}x}p(x(x_f))\left(\frac{\mathrm{d}x}{\mathrm{d}x_f}\right)^2 + p(x(x_f))\frac{\mathrm{d}^2x}{\mathrm{d}x_f^2}. \end{aligned} \tag{49}$$

Z rovnice (33) si vyjadríme vzťah pre $\frac{\mathrm{d}}{\mathrm{d}x}p(x)$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\left(c(x)p(x)\right) + \frac{c(x)p(x)}{b} &= Ap(x), \\ \frac{\mathrm{d}}{\mathrm{d}x}\left(c(x)\right)p(x) + \frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\right)c(x) &= Ap(x) - \frac{c(x)p(x)}{b}, \\ \frac{\mathrm{d}}{\mathrm{d}x}(p(x))c(x) &= Ap(x) - \frac{\mathrm{d}}{\mathrm{d}x}(c(x))p(x) - \frac{c(x)p(x)}{b}, \\ \frac{\mathrm{d}}{\mathrm{d}x}(p(x))c(x) &= \left(A - \frac{\mathrm{d}}{\mathrm{d}x}(c(x)) - \frac{c(x)}{b}\right)p(x). \end{aligned}$$

Do rovnice si vieme dosadiť funkciu zániku bielkovín, platnú pre náš model: $c(x) = \gamma x_f$:

$$\frac{\mathrm{d}}{\mathrm{d}x}(p(x))\gamma x_f = \left(A - \frac{\mathrm{d}}{\mathrm{d}x}(\gamma x_f) - \frac{\gamma x_f}{b}\right) p(x),$$
$$\frac{\mathrm{d}}{\mathrm{d}x}(p(x)) = \frac{1}{\gamma x_f}\left(A - \gamma\frac{\mathrm{dx_f}}{\mathrm{d}x} - \frac{\gamma x_f}{b}\right) p(x).$$

V rovnici (49) potrebujeme vyčísliť $\gamma\frac{\mathrm{d}x_f}{\mathrm{d}x}$, spravíme to použitím vzťahu (36) a dostaneme:

$$\frac{\mathrm{d}}{\mathrm{d}x}(p(x)) = \frac{1}{\gamma x_f}\left(A - \frac{\gamma}{1 + k\,n\,x_f^{n-1}} - \frac{\gamma x_f}{b}\right) p(x). \tag{50}$$

Pomocou vzťahu (36) si vyjadríme členy $\left(\frac{\mathrm{d}x}{\mathrm{d}x_f}\right)^2$, $\frac{\mathrm{d}^2x}{\mathrm{d}x_f^2}$ potrebné na vyjadrenie rovnice (49):

$$\left(\frac{\mathrm{d}x}{\mathrm{d}x_f}\right)^2 = \left(1 + k\,n\,x_f^{n-1}\right)^2,$$
$$\frac{\mathrm{d}^2x}{\mathrm{d}x_f^2} = \frac{\mathrm{d}}{\mathrm{d}x_f}\frac{\mathrm{d}x}{\mathrm{d}x_f} = \frac{\mathrm{d}}{\mathrm{d}x_f}\left(1 + k\,n\,x_f^{n-1}\right) = k\,n\,(n-1)\,x_f^{n-2}.$$

Teraz všetky medzivýpočty dosadíme do vzťahu (49) a dostaneme:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x_f}\tilde{p}(x_f) =& \frac{1}{\gamma x_f}\left(A - \frac{\gamma}{1 + k\,n\,x_f^{n-1}} - \frac{\gamma x_f}{b}\right) p(x(x_f))\left(1 + k\,n\,x_f^{n-1}\right)^2 + \\ &+ p(x(x_f))\left(k\,n\,(n-1)\,x_f^{n-2}\right), \\ \frac{\mathrm{d}}{\mathrm{d}x_f}\tilde{p}(x_f) =& p(x)\left(\frac{1}{\gamma x_f}\left(A - \frac{\gamma}{1 + k\,n\,x_f^{n-1}} - \frac{\gamma x_f}{b}\right)\left(1 + k\,n\,x_f^{n-1}\right)^2 + \left(k\,n\,(n-1)\,x_f^{n-2}\right)\right).\end{aligned} \tag{51}$$

Už vieme, že bimodalita rozdelenia koncentrácie voľných bielkovín nastáva, keď platí $\frac{\mathrm{d}}{\mathrm{d}x_f}\tilde{p}(x_f) = 0$. Potom z rovnice (51) vidíme, že bimodalitu rozdelenia koncentrácie voľných bielkovín získame, keď sa zátvorka na pravej strane rovnice rovná 0. Takto sme si vyjadrili prvú rovnicu pre výpočet bimodálnej oblasti v priestore parametrov pre náš model:

$$\left(\frac{1}{\gamma x_f}\left(A - \frac{\gamma}{1 + k\,n\,x_f^{n-1}} - \frac{\gamma x_f}{b}\right)\left(1 + k\,n\,x_f^{n-1}\right)^2 + \left(k\,n\,(n-1)\,x_f^{n-2}\right)\right) = 0. \tag{52}$$

Na základe článku [4] vieme, že počet lokálnych extrémov sa mení, ak sa derivácia ľavej strany rovnice (52) podľa x_f rovná 0. Teda ak platí:

$$\frac{\mathrm{d}}{\mathrm{d}x_f}\left(\frac{1}{\gamma x_f}\left(A - \frac{\gamma}{1 + k\,n\,x_f^{n-1}} - \frac{\gamma x_f}{b}\right)\left(1 + k\,n\,x_f^{n-1}\right)^2 + \left(k\,n\,(n-1)\,x_f^{n-2}\right)\right) = 0,$$
$$\frac{\mathrm{d}}{\mathrm{d}x_f}\left(\left(\frac{A}{\gamma x_f} - \frac{1}{x_f + k\,n\,x_f^n} - \frac{1}{b}\right)\left(1 + k\,n\,x_f^{n-1}\right)^2 + \left(k\,n\,(n-1)\,x_f^{n-2}\right)\right) = 0,$$

$$\left(\frac{-A}{\gamma x_f^2}+\frac{1+k\,n^2\,x_f^{n-1}}{(x_f+k\,n\,x_f^n)^2}\right)\left(1+k\,n\,x_f^{n-1}\right)^2+$$
$$+\left(\frac{A}{\gamma x_f}-\frac{1}{x_f+k\,n\,x_f^n}-\frac{1}{b}\right)2\left(1+k\,n\,x_f^{n-1}\right)\left(k\,n\,(n-1)\,x_f^{n-2}\right)+$$
$$+\left(k\,n\,(n-1)\,(n-2)\,x_f^{n-3}\right)=0. \tag{53}$$

Rovnice (52) a (53) určujú kombináciu parametrov modelu, pre ktorú dochádza k zmene z unimodality na bimodalitu pre funkciu hustoty $\tilde{p}(x_f)$ koncentrácie voľných bielkovín x_f. Pričom parametre modelu sú: A, b, γ, k, n. Teda naša množina premenných má dimenziu 5. Nie je praktické analyzovať premenné s takou veľkou dimenziou, preto budeme hodnotu niektorých parametrov fixovať. Presnejšie budeme fixovať hodnotu rýchlosti zániku $\gamma = 1$, veľkosti multiméru n a fixujeme aj hodnotu frekvencie vzniku vzplanutí A a to rovnicou $Ab = 1$. Budeme sa teda pozerať na hodnoty parametrov b, k ktoré určujú plochy bimodality a unimodality pre konkrétne hodnoty parametrov A, γ, n pre konkrétnu hladinu koncentrácie voľných bielkovín x_f.

Fixované parametre, konkrétne $\gamma = 1$ a $A = \frac{1}{b}$ si dosadíme do rovníc pre výpočet hranice medzi bimodalitou a unimodalitou (52), (53) a dostaneme:

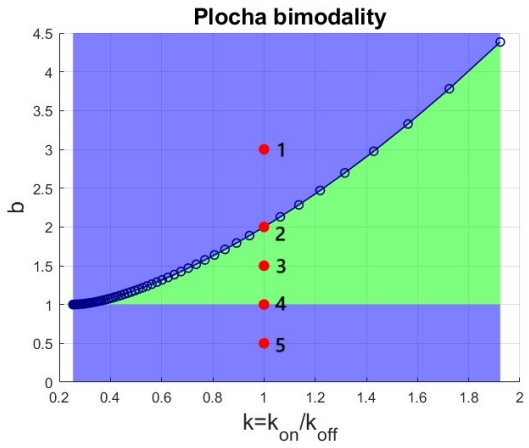
$$\left(\frac{1}{x_f}\left(\frac{1-x_f}{b}-\frac{1}{1+k\,n\,x_f^{n-1}}\right)\left(1+k\,n\,x_f^{n-1}\right)^2+\left(k\,n\,(n-1)\,x_f^{n-2}\right)\right)=0, \tag{54}$$

$$\left(\frac{-1}{bx_f^2}+\frac{1+k\,n^2\,x_f^{n-1}}{(x_f+k\,n\,x_f^n)^2}\right)\left(1+k\,n\,x_f^{n-1}\right)^2+$$
$$+2\left(\frac{1}{bx_f}-\frac{1}{x_f+k\,n\,x_f^n}-\frac{1}{b}\right)\left(1+k\,n\,x_f^{n-1}\right)\left(k\,n\,(n-1)\,x_f^{n-2}\right)+$$
$$+\left(k\,n\,(n-1)\,(n-2)\,x_f^{n-3}\right)=0. \tag{55}$$

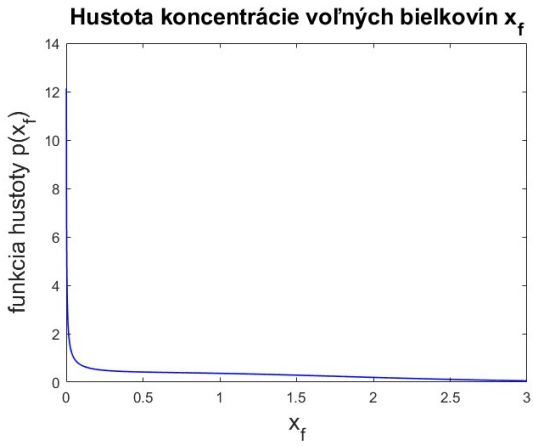
Teraz sa pozrieme na oblasti unimodality a bimodality v závislosti od hodnôt parametrov b, k pri fixovanej hodnote parametrov A, γ, n pre hladinu koncentrácie voľných bielkovín x_f, ktorú budeme voliť z intervalu hodnôt pre rozdielne hodnoty parametra n. Hranicu medzi unimodalitou a bimodalitou vypočítame na základe rovníc (54) a (55).

Na obrázku 12 v ľavom hornom rohu sme vykreslili rozdelenie plôch unimodality a bimodality rozdelenia koncentrácie voľných bielkovín pre parametre $Ab = 1$, $\gamma = 1$ a $n = 2$. Modrou farbou je vyznačené územie unimodality a zelenou farbou je vyznačené územie bimodality rozdelenia koncentrácie voľných bielkovín pre tvorbu dimérov ($n = 2$).

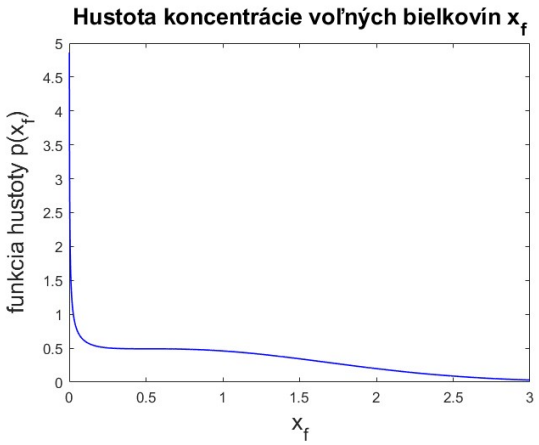
Tmavomodrá krivka znázorňuje riešenie rovníc (54) a (55), pričom tmavomodrými krúžkami pozdĺž krivky sú znázornené riešenia pre konkrétne hodnoty koncentrácie voľných bielkovín x_f. Priamka $b = 1$ rozdeľuje unimodalitu od bimodality rozdelenia koncentrácie voľných bielkovín pre tvorbu dimérov ($n = 2$) iným spôsobom. Toto rozhranie nevyšlo ako riešenie rovníc (54) a (55) ale všimli sme si ho podrobnejšou analýzou. Ide totiž o iný prechod medzi unimodalitou a bimodalitou ako v prípade riešenia rovníc (54) a (55). Červenými bodmi na obrázku 12 v ľavom hornom rohu sme označili voľbu parametrov k, b, pre ktoré sme vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 2$, aby sme zanalyzovali prechod medzi unimodalitou a bimodalitou na oboch rozhraniach. Funkciu hustoty koncentrácie voľných bielkovín pre prvú voľbu parametrov $n = 2$, $k = 1$, $b = 3$ sme vykreslili v pravom hornom rohu obrázka 12. Vidíme, že zvolené hodnoty parametrov ležia v oblasti unimodality, preto nás neprekvapí unimodalita funkcie hustoty koncentrácie voľných bielkovín pre túto voľbu parametrov. Funkciu hustoty koncentrácie voľných bielkovín pre druhú voľbu parametrov $n = 2$, $k = 1$, $b = 2$ sme vykreslili vľavo do stredu na obrázku 12. Tieto hodnoty parametrov sme volili z hranice medzi unimodalitou a bimodalitou rozdelenia koncentrácie voľných bielkovín pre $n = 2$ vypočítanou pomocou rovníc (54) a (55). Na grafe funkcie hustoty koncentrácie voľných bielkovín pre $n = 2$ môžeme pozorovať inflexný bod pre hodnotu koncentrácie voľných bielkovín $x_f = 0.5$, pre túto hodnotu koncentrácie voľných bielkovín boli totiž rovnice (54), (55) pre dané parametre počítané. Funkciu hustoty koncentrácie voľných bielkovín pre tretiu voľbu parametrov $n = 2$, $k = 1$, $b = 1.5$ sme vykreslili vpravo v strede na obrázku 12. Tieto hodnoty parametrov sme volili z oblasti bimodality rozdelenia koncentrácie voľných bielkovín pre $n = 2$. Bimodalita tejto funkcie hustoty koncentrácie voľných bielkovín pre $n = 2$ nás teda neprekvapuje. Funkciu hustoty koncentrácie voľných bielkovín pre štvrtú voľbu parametrov $n = 2$, $k = 1$, $b = 1$ sme vykreslili v ľavom dolnom rohu obrázka 12. Hodnoty parametrov sme volili z druhej hranice medzi unimodalitou a bimodalitou rozdelenia koncentrácie voľných bielkovín pre $n = 2$. Môžeme si uvedomiť, že funkcia hustoty koncentrácie voľných bielkovín pre túto voľbu parametrov je unimodálna, avšak zmena z unimodality na bimodalitu sa nestala prechodom cez inflexný bod, ale stočením ľavého chvosta funkcie hustoty. Preto sme túto hranicu zmeny bimodality na unomodalitu rozdelenia koncentrácie voľných bielkovín nezachytili pomocou riešenia rovníc (54), (55).



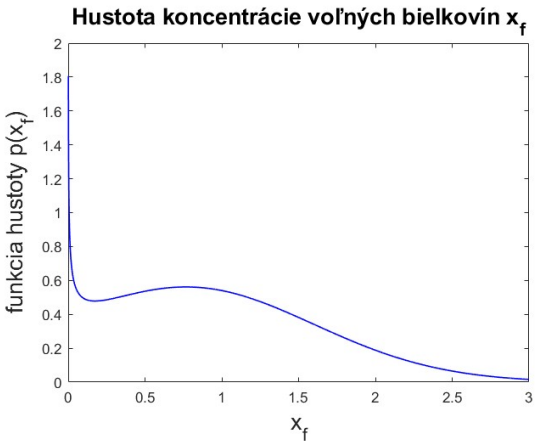
$n = 2$



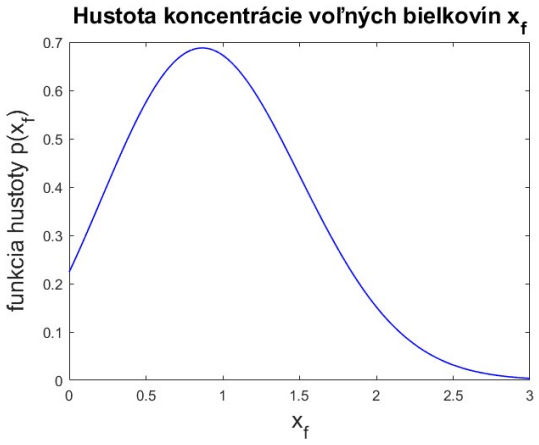
1) $n = 2$, $k = 1$, $b = 3$



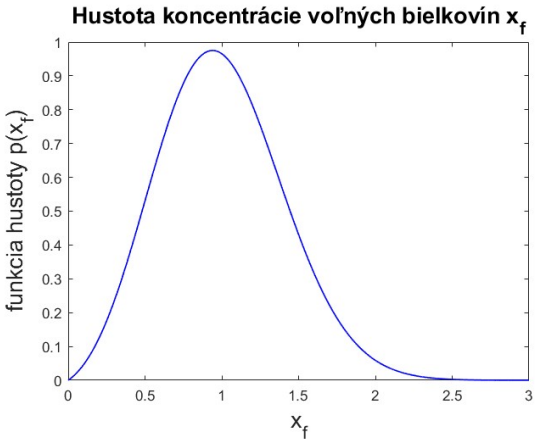
2) $n = 2$, $k = 1$, $b = 2$



3) $n = 2$, $k = 1$, $b = 1.5$



4) $n = 2$, $k = 1$, $b = 1$

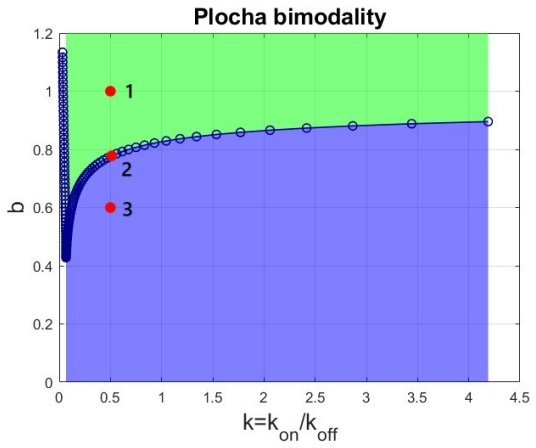


5) $n = 2$, $k = 1$, $b = 0.5$

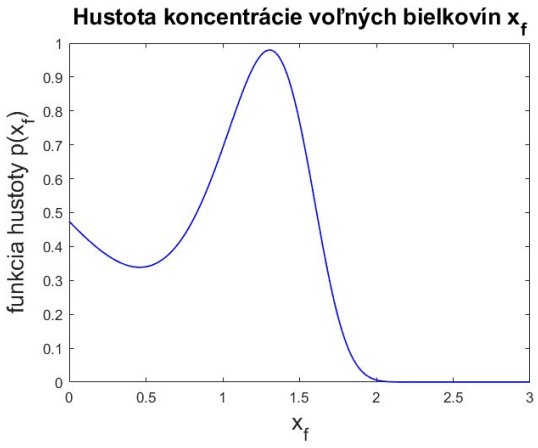
Obr. 12: Hranica medzi unimodalitou a bimodalitou funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov $Ab = 1$, $\gamma = 1$, $n = 2$ a vykreslenie niektorých funkcií hustoty koncentrácie voľných bielkovín pre rôzne voľby parametrov k, b (Vlastný zdroj)

Funkciu hustoty koncentrácie voľných bielkovín pre poslednú piatu voľbu parametrov $n = 2$, $k = 1$, $b = 0.5$ sme vykreslili v pravom dolnom rohu obrázka 12. Hodnoty parametrov sme volili z oblasti unimodality rozdelenia koncentrácie voľných bielkovín pre $n = 2$. Unimodalita tejto funkcie hustoty koncentrácie voľných bielkovín pre $n = 2$ nás teda neprekvapuje.

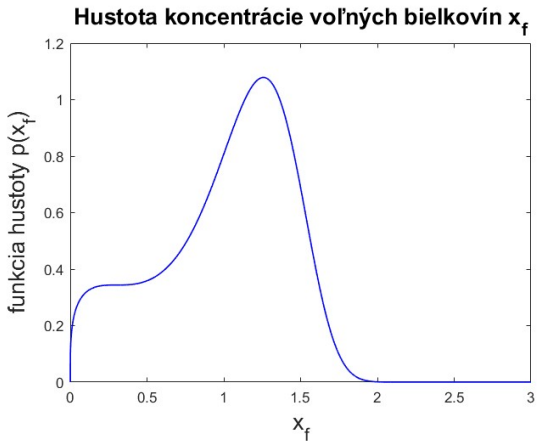
Teraz sa podobne pozrieme na rozhranie medzi unimodalitou a bimodalitou rozdelenia koncentrácie voľných bielkovín pre tvorbu multimérov $n = 5$. Na obrázok 13 v ľavom hornom rohu sme vykreslili rozdelenie plôch unimodality a bimodality rozdelenia koncentrácie voľných bielkovín pre parametre $Ab = 1$, $\gamma = 1$ a $n = 5$. Modrou farbou je vyznačené územie unimodality a zelenou farbou je vyznačené územie bimodality rozdelenia koncentrácie voľných bielkovín pre tvorbu multimérov ($n = 5$). Tmavomodrá krivka znázorňuje riešenie rovníc (54) a (55), pričom tmavomodrými krúžkami pozdĺž krivky sú znázornené riešenia pre konkrétne hodnoty koncentrácie voľných bielkovín x_f. Červenými bodmi sme označili voľbu parametrov k, b, pre ktoré sme vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 5$, aby sme zanalyzovali prechod medzi unimodalitou a bimodalitou. Funkciu hustoty koncentrácie voľných bielkovín pre prvú voľbu parametrov $n = 5$, $k = 0.5$, $b = 1$ sme vykreslili v pravom hornom rohu obrázka 13. Vidíme, že zvolené hodnoty parametrov ležia v oblasti bimodality, a preto nás neprekvapí bimodalita funkcie hustoty koncentrácie voľných bielkovín pre túto voľbu parametrov. Funkciu hustoty koncentrácie voľných bielkovín pre druhú voľbu parametrov $n = 5$, $k = 0.5128$, $b = 0.7778$ sme vykreslili v ľavom dolnom rohu na obrázku 13. Tieto hodnoty parametrov sme volili z hranice medzi unimodalitou a bimodalitou rozdelenia koncentrácie voľných bielkovín pre $n = 5$ vypočítanou pomocou rovníc (54) a (55). Na grafe funkcie hustoty koncentrácie voľných bielkovín pre $n = 5$ môžeme pozorovať inflexný bod pre hodnotu koncentrácie voľných bielkovín $x_f = 0.3$, pre túto hodnotu koncentrácie voľných bielkovín boli totiž rovnice (54), (55) pre dané hodnoty parametrov počítané. Funkciu hustoty koncentrácie voľných bielkovín pre tretiu voľbu parametrov $n = 5$, $k = 0.5$, $b = 0.6$ sme vykreslili v pravom dolnom rohu na obrázku 13. Tieto hodnoty parametrov sme volili z oblasti unimodality rozdelenia koncentrácie voľných bielkovín pre $n = 5$. Unimodalita tejto funkcie hustoty koncentrácie voľných bielkovín pre $n = 5$ nás teda neprekvapuje.



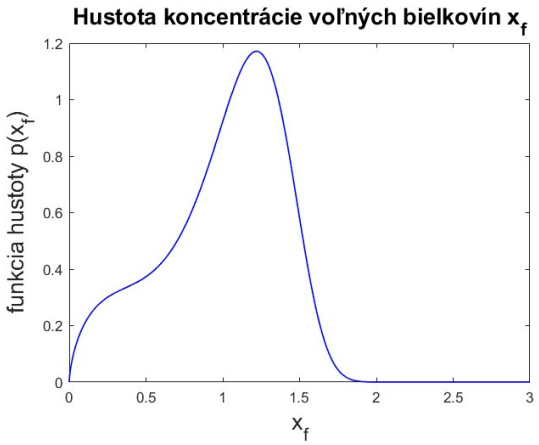
$n = 5$



1) $n = 5$, $k = 0.5$, $b = 1$



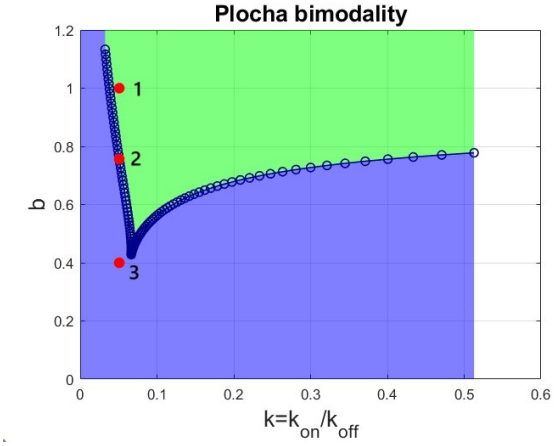
2) $n = 5$, $k = 0.5128$, $b = 0.7778$



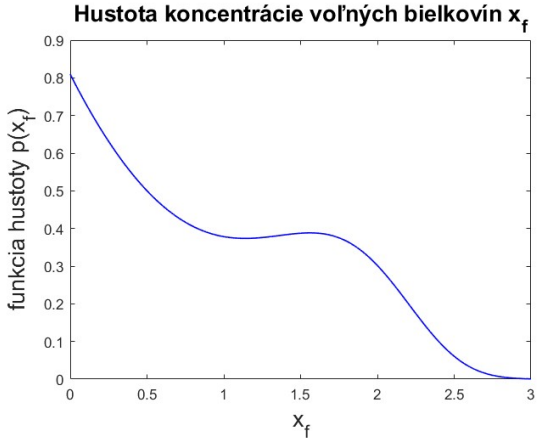
3) $n = 5$, $k = 0.5$, $b = 0.6$

Obr. 13: Hranica medzi unimodalitou a bimodalitou funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov $Ab = 1$, $\gamma = 1$, $n = 5$ a vykreslenie niektorých funkcií hustoty koncentrácie voľných bielkovín pre rôzne voľby parametrov k, b (Vlastný zdroj)

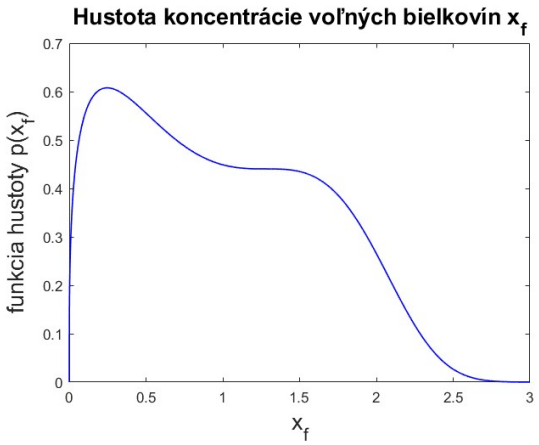
Graf rozhrania medzi unimodalitou a bimodalitou rozdelenia koncentrácie voľných bielkovín pre tvorbu multimérov $n = 5$ na obrázku 13 v ľavom hornom rohu je vykreslený pre väčšie hodnoty parametrov k, b, preto tam poriadne nevidíme jednu vetvu prechodu medzi unimodalitou a bimodalitou. Vykreslíme si preto nový obrázok 14, kde rozhranie medzi unimodalitou a bimodalitou vykreslíme pre menšie hodnoty parametrov k, b a pozrieme sa bližšie na druhú vetvu rozdelenia unimodality a bimodality rozdelenia koncentrácie voľných bielkovín pre $n = 5$.



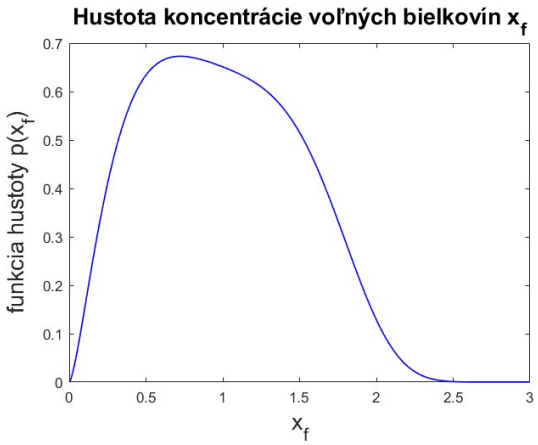
$n = 5$



1) $n = 5$, $k = 0.0509$, $b = 1$



2) $n = 5$, $k = 0.0509$, $b = 0.7565$



3) $n = 5$, $k = 0.0509$, $b = 0.4$

Obr. 14: Hranica medzi unimodalitou a bimodalitou funkcie hustoty koncentrácie voľných bielkovín pre voľbu parametrov $Ab = 1$, $\gamma = 1$, $n = 5$ a vykreslenie niektorých funkcií hustoty koncentrácie voľných bielkovín pre rôzne voľby parametrov k, b (Vlastný zdroj)

Na obrázok 14 v ľavom hornom rohu sme vykreslili rozdelenie plôch unimodality a bimodality rozdelenia koncentrácie voľných bielkovín pre parametre $Ab = 1$, $\gamma = 1$ a $n = 5$ pre menšie hodnoty parametrov k, b. Modrou farbou je vyznačené územie unimodality a zelenou farbou je vyznačené územie bimodality rozdelenia koncentrácie voľných bielkovín pre tvorbu multimérov ($n = 5$). Tmavomodrá krivka znázorňuje riešenie rovníc (54) a (55), pričom tmavomodrými krúžkami pozdĺž krivky sú znázornené riešenia pre konkrétne hodnoty koncentrácie voľných bielkovín x_f. Červenými bodmi sme označili voľbu parametrov k, b, pre ktoré sme vykreslili funkciu hustoty koncentrácie voľných bielkovín pre $n = 5$, aby sme zanalyzovali prechod medzi unimodalitou a bimodalitou. Funkciu

hustoty koncentrácie voľných bielkovín pre prvú voľbu parametrov $n = 5$, $k = 0.0509$, $b = 1$ sme vykreslili v pravom hornom rohu obrázka 14. Vidíme, že zvolené hodnoty parametrov ležia v oblasti bimodality, a preto nás neprekvapí bimodalita funkcie hustoty koncentrácie voľných bielkovín pre túto voľbu parametrov. Funkciu hustoty koncentrácie voľných bielkovín pre druhú voľbu parametrov $n = 5$, $k = 0.0509$, $b = 0.7565$ sme vykreslili v ľavom dolnom rohu na obrázku 14. Tieto hodnoty parametrov sme volili z hranice medzi unimodalitou a bimodalitou rozdelenia koncentrácie voľných bielkovín pre $n = 5$ vypočítanou pomocou rovníc (54) a (55). Na grafe funkcie hustoty koncentrácie voľných bielkovín pre $n = 5$ môžeme pozorovať inflexný bod pre hodnotu koncentrácie voľných bielkovín $x_f = 1.28$, pre túto hodnotu koncentrácie voľných bielkovín boli totiž rovnice (54), (55) pre dané parametre počítané. Funkciu hustoty koncentrácie voľných bielkovín pre tretiu voľbu parametrov $n = 5$, $k = 0.0509$, $b = 0.4$ sme vykreslili v pravom dolnom rohu na obrázku 14. Tieto hodnoty parametrov sme volili z oblasti unimodality rozdelenia koncentrácie voľných bielkovín pre $n = 5$. Unimodalita tejto funkcie hustoty koncentrácie voľných bielkovín pre $n = 5$ nás teda neprekvapuje.

Oblasti unimodality a bimodality rozdelenia koncentrácie voľných bielkovín pre $n = 3, 4, 5, ...$ sa správajú podobne ako pre prípad $n = 5$, preto ich nebudeme podrobnejšie analyzovať.

Nakoniec sa ešte pozrieme ako vyzerajú rovnice (54), (55) pre voľbu parametra $n = 1$, teda pre viazanie voľných bielkovín do monomérov. Ako sme si už vyššie odvodili, koncentrácia voľných bielkovín pre tvorbu monomérov ($n = 1$) sa spáva podľa gama rozdelenia hustoty. Z čoho vyplýva, že bimodalita pre rozdelenie koncentrácie voľných bielkovín pri viazaní do monomérov nenastáva. Overíme to ešte pomocou rovníc (54), (55). Najprv si vyjadríme rovnicu (54) pre voľbu $n = 1$:

$$\left(\frac{1}{x_f}\left(\frac{1-x_f}{b} - \frac{1}{1+k}\right)(1+k)^2\right) = 0. \tag{56}$$

Druhú rovnicu si odvodíme zderivovaním rovnice (56) nasledovne:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x_f}\left(\frac{1}{x_f}\left(\frac{1-x_f}{b}-\frac{1}{1+k}\right)(1+k)^2\right)=0,\\
&\frac{\mathrm{d}}{\mathrm{d}x_f}\left(\left(\frac{1-x_f}{bx_f}-\frac{1}{x_f(1+k)}\right)(1+k)^2\right)=0,\\
&\left(\frac{-1}{bx_f^2}+\frac{1}{x_f^2(1+k)}\right)(1+k)^2=0,\\
&\frac{(1+k)^2}{bx_f^2}=\frac{(1+k)}{x_f^2},\\
&\frac{(1+k)}{b}=1.
\end{aligned} \tag{57}$$

Teda rovnica (57) sa rovná nule pokiaľ platí $b = 1 + k$. Aby však nastala bimodalita, potrebujeme aby zároveň platila aj rovnica (56). Skúsime preto do rovnice (56) dosadiť $b = k + 1$:

$$\begin{aligned}
&\frac{1}{x_f}\left(\frac{1-x_f}{b}-\frac{1}{b}\right)b^2=0,\\
&\frac{1}{x_f}\left(\frac{-x_f}{b}\right)b^2=0,\\
&-b=0.
\end{aligned}$$

Teda aby sa rovnica (57) rovnala nule, musí platiť $b = k+1$ a aby sa zároveň rovnica (56) rovnala nule musí zároveň platiť $b = 0$. Avšak b sa nemôže rovnať nule, keďže v oboch rovniciach (56), (57) sa parameter b nachádza v menovateli. Preto bimodalita pre viazanie do monomérov ($n = 1$) nemôže nikdy nastať.

Záver

Hlavným cieľom našej diplomovej práce bolo zostavenie nového modelu stochastickej génovej expresie. Spravili sme tak skombinovaním modelov stochastickej génovej expresie zo zdrojov [3, 4, 19]. Následne sme chceli vytvorený model analyzovať.

V kapitole 1 sme uviedli potrebnú teóriu na zostavenie modelu stochastickej génovej expresie. Začali sme uvedením základných procesov génovej expresie v podkapitole 1.1 a to hlavne na základe zdrojov [1, 23]. V nasledujúcej podkapitole 1.2 sme si vysvetlili zákon pôsobenia hmotností, ktorý nám pomohol vyjasniť otázku rýchlosti reakcií. V tejto podkapitole sme čerpali hlavne zo zdrojov [1, 12, 13, 14, 18, 22, 23]. Následne sme sa v podkapitole 1.3 pozreli na vznik vonkajšieho a vnútorného riešenia v procesoch, na základe rozdelenia procesov na tie, ktoré sa dejú rýchlo, v krátkom časovom horizonte a na tie, ktoré sa dejú pomaly, v dlhom časovom horizonte. Pri tejto podkapitole sme čerpali hlavne zo zdroja [13]. Teóriu singulárnych perturbácií sme bližšie vysvetlili na príklade Michaelis-Mentenovej kinetiky, ktorú sme uviedli v podkapitole 1.3.1. V tejto podkapitole sme predstavili aj pojem kvázi stacionárneho riešenia, ktoré vzniká za predpokladu, že rýchle procesy, ktoré sa dejú v krátkom časovom horizonte, sú v stacionárnom stave. Michaelis-Mentenovú kinetiku sme si vysvetlili hlavne na základe zdrojov [1, 2, 9, 12, 13, 14]. Následne sme si v podkapitole 1.4 vysvetlili, prečo môžeme očakávať stochasticitu v modeli génovej expresie, že je to spôsobené nízkym počtom molekúl bielkovín prítomných v bunke. V ďalšej podkapitole 1.4.1 sme si preto predstavili teóriu markovovských procesov a Chapman-Kolmogorovej rovnice, aby sme vedeli zostaviť model aj pre stochastický proces. V tejto podkapitole sme čerpali hlavne zo zdrojov [5, 7, 8, 15, 24]. V poslednej podkapitole 1.4.2 sme si odvodili hlavnú rovnicu z Chapman-Kolmogorovej rovnice na základe zdrojov [1, 4, 6, 10, 21].

V kapitole 2 sme zostavili stochastický model génovej expresie s deterministickým zánikom, deterministickým viazaním do multimérov a deterministickým odväzovaním na voľné bielkoviny. Pri zostavovaní modelu sme sa inšpirovali hlavne zdrojmi [3, 4, 19]. Sformulovaný model sme zobrazili na obrázku 2. Model sme vyjadrili pomocou upravenej Chapman-Kolmogorovej rovnice, ktorej ďalšou úpravou a dosadením parametrov nášho modelu sme odvodili funkciu hustoty rozdelenia koncentrácie voľných bielkovín. Všimli sme si, že všeobecná funkcia hustoty rozdelenia koncentrácie voľných bielkovín (38) platí

len pre prípad $n = 2, 3, 4, ...$ Preto sme funkciu hustoty rozdelenia koncentrácie voľných bielkovín pre $n = 1$ odvodili osobitne a všimli sme si, že substitúciou ju vieme upraviť na funkciu hustoty gama rozdelenia (43). Funkciu hustoty rozdelenia koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ sa nám nepodarilo upraviť na funkciu hustoty gama rozdelenia, preto sme ju chceli porovnať s týmto rozdelením a zistiť, či sa správa podobne. Porovnanie pre rôzne hodnoty parametrov sme znázornili na obrázkoch 5, 7. Zistili sme, že pre istú voľbu parametrov sa funkcia hustoty rozdelenia koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ podobá na gama rozdelenie a pre inú nie. Taktiež sme po prvý krát spozorovali bimodalitu funkcie hustoty rozdelenia koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Pozreli sme sa aj na stredné hodnoty a disperzie koncentrácie voľných bielkovín. Stredná hodnota koncentrácie voľných bielkovín pre $n = 1$ a danú voľbu parametrov nám vyšla konštantne 1 a zobrazili sme ju na obrázku 4. Zatiaľ čo stredná hodnota koncentrácie voľných bielkovín pre $n = 2$, ktorú sme zobrazili na obrázkoch 6, 8 pre rozdielne voľby parametrov nám skákala veľmi blízko hodnoty 1, ale nebola konštantná, preto sme úpravou rovnice (24) prišli na vzťah (48), ktorý jasne ukazuje konštantnosť strednej hodnoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ A teda pozorované výkyvy boli len numerické chyby. Ďalej sme chceli porovnať funkciu hustoty rozdelenia koncentrácie voľných bielkovín pre $n = 1$, teda pre viazanie do monomérov, s funkciou hustoty rozdelenia koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$, teda viazanie do multimérov, avšak chceli sme mať zvolené hodnoty parametrov pre prípady $n = 1$ a $n = 2, 3, 4, ...$ v istom zmysle podobné. Dosiahli sme to tak, že sme hodnoty parametrov volili tak, aby sa stredné hodnoty koncentrácie voľných bielkovín pre $n = 1$ a pre $n = 2, 3, 4, ...$ rovnali a zároveň, aby aj stredné hodnoty pre koncentráciu viazaných bielkovín pre $n = 1$ a $n = 2, 3, 4, ...$ boli rovnaké. Výsledné porovnanie sme pre rozdielne voľby parametrov zobrazili na obrázkoch 9, 10. Nakoniec sme chceli ešte preskúmať pre aké hodnoty parametrov nastáva bimodalita rozdelenia koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Odvodili sme si preto vzťah pre deriváciu funkcie hustoty koncentrácie voľných bielkovín a z nej sme odvodili dve podmienky pre bimodalitu rozdelenia (52), (53). Následne sme fixovali hodnoty parametrov $\gamma = 1$, $Ab = 1$ a taktiež sme fixovali hodnotu parametra n. Pozerali sme sa na územie bimodality v závislosti od hodnoty parametrov k, b a pre $n = 2$ sme ho zobrazili na obrázku 12 a pre $n = 5$ na obrázkoch 13, 14. Nakoniec sme už len

dokázali, že bimodalita rozdelenia koncentrácie voľných bielkovín pre $n = 1$ nenastáva.

Za jeden z prínosov našej práce môžeme považovať zhrnutie teórie k zostaveniu modelu stochastickej génovej expresie ako aj vzniku vnútorného a vonkajšieho riešenia. Vytvorenie a analýzu nového modelu stochastickej génovej expresie a spísania výsledkov analýzy. Ďalším z prínosov tejto práce je sprístupnenie danej tematiky slovenským čitateľom, keďže väčšina literatúry je písaná v anglickom jazyku. Možným rozšírením tejto práce by mohlo byť pozmenenie nášho modelu tak, aby aj viazané bielkoviny x_b podliehali zániku. Funkcia zániku by sa teda zmenila do tvaru $c(x) = \gamma(x_f + x_b)$. Následne by sme mohli sledovať ako to ovplyvní rozdelenie koncentrácie voľných bielkovín. Iným rozšírením by mohlo byť, pozrieť sa pomocou simulácií na rozdelenie koncentrácie voľných bielkovín a porovnanie výsledkov s odvodenými výsledkami v tejto práci.

Zoznam použitej literatúry

[1] Bressloff, P. C.: *Stochastic Processes in Cell Biology Volume 1*, Springer, Berlin, 2014, 2. vydanie

[2] Britton, N. F.: *Essential Mathematical Biology*, Springer, London, 2005, 3. vydanie

[3] Biswas, A., Nieto, C., Bokes, P., Singh, A.: *Invariance of Poisson Statistics in Molecular Counts to Sequestration Mechanisms*, dostupné na internete (3.12.2024): https://osf.io/uq2vk/download

[4] Bokes, P., Singh, A.: *Protein Copy Number Distributions for a Self-Regulating Gene in the Presence of Decoy Binding Sites*, (2015), dostupné na internete (3.12.2024): https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120555#sec017

[5] Capasso, V., Bakstein, D.: *An Introduction to Continuous-Time Stochastic Processes*, Springer, Boston, 2005

[6] Friedman, N., Cai, L., Xie, X. S.:*Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression*, (2006), dostupné na internete (21.4.2025):

[7] Gardiner, C. W.: *Handbook of Stochastic Methods for Physics Chemistry and the Natural Science*, Springer, New York, 1985, 2. vydanie

[8] Ibe, O. C.: *Markov Processes for Stochastic Modeling*, Elsevier Academic Press, Burlington, 2008

[9] Ingalls, B. P.: *Mathematical Modeling in Systems Biology: An Introduction*, MIT press, 2013

[10] Jia, T.:*Stochastic Modeling of Gene Expression and Post-transcriptional Regulation*, (2011), dostupné na internete (20.4.2025): https://vtechworks.lib.vt.edu/items/74d547d2-ebe7-4741-b864-87f65433332a

[11] Joyce, D.:*Common Probability Distributions*, (2006), dostupné na internete (27.12.2024): https://qiangbo-workspace.oss-cn-shanghai.aliyuncs.com/2018-11-11-common-probability-distributions/distributions2.pdf

[12] Liu, W.: *Introduction to modeling biological cellular control systems*, Springer Science & Business Media, Miláno, 2012

[13] Müller, J., Kuttler, C.: *Methods and Models in Mathematical Biology. Lecture Notes on Mathematical Modelling in Life Sciences*, Springer, Berlin, 2015

[14] Murray, J., D.: *Mathematical Biology I: An Introduction*, Springer, Berlin, 2007, 3. vydanie

[15] Pavliotis, G. A.: *Stochastic Processes and Applications*, (2014), dostupné na internete (21.4.2025): https://link.springer.com/content/pdf/10.1007/978-1-4939-1323-7.pdf

[16] Peccound, J., Ycart, B.: *Markovian Modelling of Gene Product Synthesis*, (1995), dostupné na internete (27.12.2024):https://www.sciencedirect.com/science/article/abs/pii/S0040580985710271?via%3Dihub

[17] Robert, P.: *Mathematical models of gene expression*, (2019), dostupné na internete (2.1.2024): https://projecteuclid.org/journals/probability-surveys/volume-16/issue-none/Mathematical-models-of-gene-expression/10.1214/19-PS332.full

[18] Santillán, M.: *Chemical Kinetics, Stochastic Processe, and Irreversible Thermodynamics*, Springer, Heidelberg, 2014

[19] Soltani, M., Bokes, P., Fox, Z., Singh, A.: *Nonspecific Transcription Factor Binding Can Reduce Noise in the Expression of Downstream Proteins*, (2015), dostupné na internete (3.12.2024): https://iopscience.iop.org/article/10.1088/1478-3975/12/5/055002/ampdf

[20] Swain, P. S.: *Lecture Notes on Stochastic Models in Systems Biology*, (2016), dostupné na internete (20.4.2025): https://arxiv.org/pdf/1607.07806

[21] Swain, P. S., Elowitz, M. B., Siggia, E. D.: *Intrinsic and Extrinsic Contributions to Stochasticity in Gene Expression*, (2002), dostupné na internete (20.4.2025): https://www.pnas.org/doi/pdf/10.1073/pnas.162041399

[22] Upadhyay, S., K.: *Chemical Kinetics and Reaction Dynamics*, Springer, New York, 2006

[23] Wilkinson, D., J.: *Stochastic Modelling for Systems Biology*, CRC Press Taylor and Francis group, Boca Raton, 2006

[24] Winkelmann, S., Schütte Ch.: *Stochastic Dynamics in Computational Biology*, Springer, Cham, 2020

Príloha A

Priložené kódy v programovacom jazyku programu Matlab boli použité na výpočty a vykreslenie obrázkov v druhej kapitole tejto práce.

A.0.1

Nasledovný kód slúži na vykreslenie funkcie hustoty koncentrácie voľných bielkovín pre tvorbu monomérov $n = 1$ v porovnaní s funkciou hustoty gama rozdelenia a na vykreslenie strednej hodnoty a disperzie koncentrácie voľných bielkovín pre tvorbu monomérov $n = 1$ v závislosti od hodnoty parametra k. Pomocou tohto kódu boli vytvorené obrázky 3, 4.

```
n=1;  %veľkosť multiméru, špeciálny prípad monomér
k_off=0.5; %rýchlosť odviazania
k_on=0.5;   %rýchlosť viazania
k=1;
gamma=1;  %rýchlosť zániku
A=4;  %transkripčná rýchlosť
b=1/4;  %priemerná veľkosť vzplanutia

%výpočet funkcie hustoty pre koncentráciu voľných bielkovín pre n=1
p_1_1=@(x) (1/gamma).*x.^((A./gamma).*(1+k)-1).*exp((-x-(k).*x)/(b))
    .*(1+(k));
kappa=1/integral(p_1_1,0,Inf); %normalizačná konštanta pre n=1
p=@(x) (kappa/gamma).*x.^((A./gamma).*(1+k)-1).*exp((-x-(k).*x)/(b))
    .*(1+(k)); %funkcia hustoty pre koncentráciu voľných bielkovín pre
                %n=1
W=integral(p,0,Inf); %kontrola, W=1
W

%výpočet parametrov pre funkciu hustoty gama rozdelenia
shape=((A./gamma)+((A./gamma).*(k)));
rate=((1+(k))./(b));
scale=1/rate;

%vykreslenie funkcie hustoty pre koncentráciu voľných bielkovín pre n=1
figure
o=0:0.01:4;
```

```
y1 = gampdf(o,shape,scale);
plot(o,p(o),'b','LineWidth', 1)
xlabel('x_f', 'FontSize', 16)
ylabel('funkcia hustoty p(x_f)', 'FontSize', 16)
%title('Hustota koncentrácie voľných bielkovín x_f', 'FontSize', 16)
hold on
plot(o,y1,'r','LineWidth', 1)
legend('p(x_f)','gampdf', 'Location', 'best', 'FontSize', 14)
hold off

%výpočet strednej hodnoty koncentrácie voľných bielkovín a vykreslenie
%závislosti strednej hodnoty od hodnoty parametra k
mu=(A.*b)./(gamma);

figure
o=0:0.01:10;
mu_k=ones(length(o),1);
mu_k=mu*mu_k;
line(o,mu_k,'LineWidth', 1)
xlabel('k=k_{on}/k_{off}', 'FontSize', 16)
ylabel('Stredná hodnota', 'FontSize', 16)
title('Závislosť strednej hodnoty od k=k_{on}/k_{off}', 'FontSize', 16)

%výpočet disperzie koncentrácie voľných bielkovín a vykreslenie
%závislosti disperzie od hodnoty parametra k
var=(A.*b.^2)./(gamma.*(1+k));

o=0:0.01:20;
var_k=zeros([length(o) 0]);

for j=0:(length(o)-1)
var_k(j+1)= (A.*b.^2)/(gamma.*(1+o(j+1)));
end

figure
plot(o,var_k,'LineWidth', 1)
xlabel('k=k_{on}/k_{off}', 'FontSize', 16)
ylabel('Disperzia', 'FontSize', 16)
```

```
title('Závislosť disperzie od k=k_{on}/k_{off}', 'FontSize', 16)
```

A.0.2

Nasledovný kód slúži na vykreslenie funkcie hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ v porovnaní s funkciou hustoty gama rozdelenia a na vykreslenie strednej hodnoty a disperzie koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ v závislosti od hodnoty parametra k. Pomocou tohto kódu boli vytvorené obrázky 5, 6, 7, 8.

```
%Nastavenie parametrov
n=2;  %veľkosť multiméru, n=2,3,4,...
k_off=0.1; %rýchlosť odviazania (rozpadu multimérov)
k_on=0.8;   %rýchlosť viazania do multimérov
%k=k_on/k_off;
gamma=1;  %rýchlosť zániku
A=4;  %transkripčná rýchlosť
b=1/4;  %priemerná veľkosť vzplanutia

%výpočet funkcie hustoty pre koncentráciu voľných bielkovín
%pre n=2,3,4,...
p_1=@(x) (1/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(k_on/k_off).*n.*((
    x.^(n-1))/(n-1))-((x+(k_on/k_off).*x.^(n))/(b))).*(1+(k_on/k_off).*n
    .*x.^(n-1));
kappa=1/integral(p_1,0,Inf); %normalizačná konštanta
% funkcia hustoty koncentrácie voľných bielkovín
p=@(x) (kappa/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(k_on/k_off).*n
    .*((x.^(n-1))/(n-1))-((x+(k_on/k_off).*x.^(n))/(b))).*(1+(k_on/k_off)
    .*n.*x.^(n-1));
S=integral(p,0,Inf);
S %kontrola, S=1

%Vykreslenie funkcie hustoty koncentrácie voľných bielkovín
figure
o=0:0.01:3;
plot(o,p(o),'b','LineWidth', 1)
xlabel('x_f')
```

```
ylabel('funkcia hustoty p(x_f)')
%title('Hustota koncentrácie voľných bielkovín x_f')

%výpočet strednej hodnoty koncentrácie voľných bielkovín
p_mu=@(x) x.*p(x);
mu=integral(p_mu,0,Inf);
mu

%výpočet disperzie koncentrácie voľných bielkovín
p_var=@(x) ((x-mu).^(2)).*p(x);
var=integral(p_var,0,Inf);
var

%výpočet koeficienta šikmosti koncentrácie voľných bielkovín
p_skew=@(x) (((x-mu)./sqrt(var)).^3).*p(x);
skew=integral(p_skew,0,Inf);
skew

%Vykreslenie závislosti strednej hodnoty koncentrácie voľných bielkovín
   od hodnoty parametra k=k_on/k_off
o=0:0.1:10;
mu_k=zeros([(length(o)) 1]);

for k=0:(length(o)-1)
p_1_k=@(x) (1/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(o(k+1)).*n.*(x
   .^(n-1)/(n-1))-((x+(o(k+1)).*x.^(n))/(b))).*(1+(o(k+1)).*n.*x.^(n-1))
   ;
kappa_k=1/integral(p_1_k,0,Inf); %normalizačná konštanta
p_mu_k=@(x) x.*(kappa_k/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(o(k+1)
   ).*n.*((x.^(n-1))/(n-1))-((x+(o(k+1)).*x.^(n))/(b))).*(1+(o(k+1)).*n
   .*x.^(n-1));
mu_k(k+1)= integral(p_mu_k,0,Inf);
end

figure
k=0:0.1:10;
plot(k,mu_k,'LineWidth', 1)
xlabel('k=k_{on}/k_{off}', 'FontSize', 16)
```

```
ylabel('Stredná hodnota', 'FontSize', 16)
title('Závislosť strednej hodnoty od k=k_{on}/k_{off}', 'FontSize', 16)

%Vykreslenie závislosti disperzie koncentrácie voľných bielkovín od
    hodnoty parametra k=k_on/k_off
var_k=zeros([(length(o)) 1]);

for k=0:(length(o)-1)
p_1_var_k=@(x) (1/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(o(k+1)).*n
    .*(x.^(n-1)/(n-1))-((x+(o(k+1)).*x.^(n))/(b))).*(1+(o(k+1)).*n.*x.^(n
    -1));
kappa_var_k=1/integral(p_1_var_k,0,Inf); %normalizačná konštanta
p_var_k=@(x) ((x-mu_k(k+1)).^(2)).*(kappa_var_k/gamma).*x.^((A/gamma)-1)
    .*exp((A/gamma).*(o(k+1)).*n.*((x.^(n-1))/(n-1))-((x+(o(k+1)).*x.^(n)
    )/(b))).*(1+(o(k+1)).*n.*x.^(n-1));
var_k(k+1)= integral(p_var_k,0,Inf);
end

figure
k=0:0.1:10;
plot(k,var_k,'LineWidth', 1)
xlabel('k=k_{on}/k_{off}', 'FontSize', 16)
ylabel('Disperzia', 'FontSize', 16)
title('Závislosť disperzie od k=k_{on}/k_{off}', 'FontSize', 16)

%Výpočet parametrov tvaru a škály pre vykreslenie funkcie hustoty gama
    rozdelenia

z = optimvar('z',2);
eq1 = z(1)*z(2) == mu;
eq2 = z(1)*z(2)^2 == var;
prob = eqnproblem;
prob.Equations.eq1 = eq1;
prob.Equations.eq2 = eq2;
show(prob)
z0.z = [1 1];
[sol,fval,exitflag] = solve(prob,z0)
disp(sol.z)
```

```
%Porovnanie gamma funkcie hustoty a funkcie hustoty koncentrácie voľných
    bielkovín

figure
o=0:0.001:3;
plot(o,p(o),'b','LineWidth', 1)
xlabel('x_f', 'FontSize', 16)
ylabel('funkcia hustoty p(x_f)', 'FontSize', 16)
%title('Hustota koncentrácie voľných bielkovín x_f', 'FontSize', 16)
hold on
shape=sol.z(1);
scale=sol.z(2);
y1 = gampdf(o,shape,scale);
plot(o,y1,'r','LineWidth', 1)
legend('p(x_f)','gampdf', 'Location', 'best', 'FontSize', 14)
hold off
```

A.0.3

Nasledovný kód slúži na vykreslenie funkcie hustoty koncentrácie voľných bielkovín pre $n = 1$ v porovnaní s funkciou hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Pomocou tohto kódu boli vytvorené obrázky 9, 10, 11.

```
%Nastavenie parametrov
n=5;  %veľkosť multiméru, n=2,3,4,...
gamma=1;  %rýchlosť zániku
A=4;  %transkripčná rýchlosť
b=1/4;  %priemerná veľkosť vzplanutia
%k=k_on/k_off;

%výpočet parametrov pre prípady n=2,3,4,... a n=1 tak, aby mali rovnaké
   E(x_f) a E(x_b)
%Fixujeme parametre A*b=1, hodnotu A, hodnotu b, gamma=1
%Dopočítavame hodnotu parametra k=k_on/k_off
```

```
%KROK 1 - zvolíme hodnotu k pre n=2,3,4,... a dopočítame E(x_b)
kn=1; %Volím parameter
p_n_1=@(x) (1/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(kn).*n.*((x.^(n
    -1))/(n-1))-((x+(kn).*x.^(n))/(b))).*(1+(kn).*n.*x.^(n-1));%funkcia
    pre výpočet normalizačnej konštanty pre n=2,3,4,...
kappa_n=1/integral(p_n_1,0,Inf); %normalizačná konštanta pre n=2,3,4,...
Exbn1=@(x)  kn.*x.^n.*(kappa_n/gamma).*x.^((A/gamma)-1).*exp((A/gamma)
    .*(kn).*n.*((x.^(n-1))/(n-1))-((x+(kn).*x.^(n))/(b))).*(1+(kn).*n.*x
    .^(n-1));
Exbn=integral(Exbn1,0,Inf);%výpočet strednej hodnoty koncentrácie
                           %viazaných bielkovín pre n=2,3,4,...
Exbn

%KROK 2 - Dopočítame hodnotu k pre n=1 tak, aby Exbn=E(x_b)=Exb1, ibaže
%pri voľbe parametrov A*b=1, hodnotu A, hodnotu b, gamma=1 máme
%E(x_b)=kE(x_f)=k pre n=1, teda k1=(E(x_b),n=2,3,4,...)

k1=Exbn;
p_1_1=@(x) (1/gamma).*x.^((A./gamma).*(1+k1)-1).*exp((-x-(k1).*x)/(b))
    .*(1+(k1));%funkcia pre výpočet normalizačnej konštanty pre n=1
kappa_1=1/integral(p_1_1,0,Inf); %normalizačná konštanta pre n=1
Exb11=@(x)  ((k1).*x).*(kappa_1/gamma).*x.^((A./gamma).*(1+k1)-1).*exp
    ((-x-(k1).*x)/(b)).*(1+(k1));
Exb1=integral(Exb11,0,Inf);%výpočet strednej hodnoty koncentrácie
                           %viazaných bielkovín pre n=1
Exb1

%KROK 3 - Vykreslenie výsledných funkcií na graf

% funkcia hustoty koncentrácie voľných bielkovín pre n=2,3,4,...
p_n_1=@(x) (1/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(kn).*n.*((x.^(n
    -1))/(n-1))-((x+(kn).*x.^(n))/(b))).*(1+(kn).*n.*x.^(n-1));%funkcia
    pre výpočet normalizačnej konštanty pre n=2,3,4,...
kappa_n=1/integral(p_n_1,0,Inf); %normalizačná konštanta pre n=2,3,4,...
p_n=@(x) (kappa_n/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(kn).*n.*((x
    .^(n-1))/(n-1))-((x+(kn).*x.^(n))/(b))).*(1+(kn).*n.*x.^(n-1));%
    funkcia hustoty voľných bielkovín pre n=2,3,4,...
S=integral(p_n,0,Inf);
```

```
S
%výpočet strednej hodnoty koncentrácie voľných bielkovín pre n=2,3,4,...
p_mu_n=@(x) x.*p_n(x);
mu_n=integral(p_mu_n,0,Inf);
mu_n

%výpočet disperzie koncentrácie voľných bielkovín pre n=2,3,4,...
p_var_n=@(x) ((x-mu_n).^(2)).*p_n(x);
var_n=integral(p_var_n,0,Inf);
var_n

%výpočet koeficienta šikmosti koncentrácie voľných bielkovín
%pre n=2,3,4,...
p_skew_n=@(x) (((x-mu_n)./sqrt(var_n)).^3).*p_n(x);
skew_n=integral(p_skew_n,0,Inf);
skew_n

% funkcia hustoty koncentrácie voľných bielkovín pre n=1
p_1_1=@(x) (1/gamma).*x.^((A./gamma).*(1+k1)-1).*exp((-x-(k1).*x)/(b))
    .*(1+(k1));%funkcia pre výpočet normalizačnej konštanty pre n=1
kappa_1=1/integral(p_1_1,0,Inf); %normalizačná konštanta pre n=1
p_1=@(x) (kappa_1/gamma).*x.^((A./gamma).*(1+k1)-1).*exp((-x-(k1).*x)/(b
    )).*(1+(k1));% funkcia hustoty voľných bielkovín pre n=1

%výpočet strednej hodnoty koncentrácie voľných bielkovín pre n=1
p_mu_1=@(x) x.*p_1(x);
mu_1=integral(p_mu_1,0,Inf);
mu_1

%výpočet disperzie koncentrácie voľných bielkovín pre n=1
p_var_1=@(x) ((x-mu_1).^(2)).*p_1(x);
var_1=integral(p_var_1,0,Inf);
var_1

%výpočet šikmosti koncentrácie voľných bielkovín pre n=1
p_skew_1=@(x) (((x-mu_1)./sqrt(var_1)).^3).*p_1(x);
skew_1=integral(p_skew_1,0,Inf);
skew_1
```

```
%vykreslenie funkcií hustoty pre koncentráciu voľných bielkovín pre n=1
    a n=2,3,4,...
figure
o=0:0.01:3;
plot(o,p_1(o),'b','LineWidth', 1)
xlabel('x_{f}', 'FontSize', 16)
ylabel('funkcia hustoty p(x_f)', 'FontSize', 16)
%title('Hustota koncentrácie voľných bielkovín x_f', 'FontSize', 16)
hold on
plot(o,p_n(o),'r','LineWidth', 1)
legend('n=1','n=5', 'Location', 'best', 'FontSize', 14)
hold off
```

A.0.4

Nasledovný kód slúži na vykreslenie hraníc bimodality funkcie hustoty koncentrácie voľných bielkovín v závislosti od veľkosti parametrov k, b a pri konštantnej hodnote parametrov $Ab = 1$, $\gamma = 1$, $n = 2, 3, 4, ...$ Pomocou tohto kódu boli vytvorené obrázky 12, 13, 14.

```
%Nastavenie parametrov
n=2;  %veľkosť multiméru, n=2,3,4,...
gamma=1;  %rýchlosť zániku
%A=1/b;  transkripčná rýchlosť
%b  priemerná veľkosť vzplanutia
%Ab=1
%k_off rýchlosť odviazania (rozpadu multimérov)
%k_on rýchlosť viazania do multimérov
%k=k_on/k_off;

%Rovnice pre plochu bimodality
%eq1 = (1/x)*(((1-x)/(z(2)))-((1)/(1+z(1)*n*x^(n-1))))*(1+z(1)*n*x^(n-1)
    )^(2)+z(1)*n*(n-1)*x^(n-2)==0;
%eq2 = (((-1)/(z(2)*x^(2)))+(1+z(1)*n^(2)*x^(n-1))/((x+z(1)*n*x^(n))^2))
    *(1+z(1)*n*x^(n-1))^2+2*(((1)/(z(2)*x))-((1)/(x+z(1)*n*x^n))-((1)/(z
```

```
    (2))))*(1+z(1)*n*x^(n-1))*(z(1)*n*(n-1)*x^(n-2))+(z(1)*n*(n-1)*(n-2)*
    x^(n-3))==0;
%z(1)=k; z(2)=b
solution=zeros(150,2);
z_0=[1 1]; %počiatočný bod
for j=0:150
x=j*0.01+0.01; %hodnoty koncentrácie voľných bielkovín
z = optimvar('z',2);
eq1 = (1/x)*(((1-x)/(z(2)))-((1)/(1+z(1)*n*x^(n-1))))*(1+z(1)*n*x^(n-1))
    ^(2)+z(1)*n*(n-1)*x^(n-2)==0;
eq2 = (((-1)/(z(2)*x^(2)))+(1+z(1)*n^(2)*x^(n-1))/((x+z(1)*n*x^(n))^2))
    *(1+z(1)*n*x^(n-1))^2+2*(((1)/(z(2)*x))-((1)/(x+z(1)*n*x^n))-((1)/(z
    (2))))*(1+z(1)*n*x^(n-1))*(z(1)*n*(n-1)*x^(n-2))+(z(1)*n*(n-1)*(n-2)*
    x^(n-3))==0;
prob = eqnproblem;
prob.Equations.eq1 = eq1;
prob.Equations.eq2 = eq2;
%show(prob);
z0.z = z_0;
[sol,fval,exitflag] = solve(prob,z0);

solution(j+1,1)=sol.z(1);
solution(j+1,2)=sol.z(2);
z_0=sol.z;%nová voľba počiatočných hodnôt = nastavíme ako riešenie
            %predošlej iterácie
end
solution

%Vykreslenie riešení, pre rozdielne voľby parametra n sú relevantné
%rozdielne intervaly koncentrácie voľných bielkovín x_f, preto
%pre rozdielne hodnoty parametra n, vykresľujeme rozdielne obrázky

%n=2
figure
hold on
h = fill([solution(1:58,1)', fliplr(linspace(0.2538071, 1.9230769, 58))
    ],[solution(1:58,2)', 4.5*ones(58,1)'], 'b'); % Vyznačenie plochy
    unimodality modrou farbou
```

```
set(h, 'FaceAlpha', 0.5)
h.EdgeColor = 'none';
hold on
h1 = fill([solution(1:58,1)', fliplr(linspace(0.2538071, 1.9230769, 58))
    ],[solution(1:58,2)', ones(58,1)'], 'g'); % Vyznačenie plochy
    bimodality zelenou farbou
set(h1, 'FaceAlpha', 0.5)
h1.EdgeColor = 'none';
hold on
h2 = fill([linspace(0.2538071, 1.9230769, 58), fliplr(linspace
    (0.2538071, 1.9230769, 58))],[ones(58,1)', zeros(58,1)'], 'b'); %
    Vyznačenie plochy unimodality modrou farbou
set(h2, 'FaceAlpha', 0.5)
h2.EdgeColor = 'none';
hold on
plot(solution(1:58,1),solution(1:58,2),'-o', 'Color', '#00008B','
    LineWidth', 1)
hold on
plot(1, 3, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')
hold on
plot(1, 2, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')
hold on
plot(1, 1.5, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')
hold on
plot(1, 1, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')
hold on
plot(1, 0.5, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')


xlabel('k=k_{on}/k_{off}', 'FontSize', 16);
ylabel('b', 'FontSize', 16);
title('Plocha bimodality', 'FontSize', 16);
grid on
hold off

%n=5
figure
plot(solution(14:101,1),solution(14:101,2),'-o','LineWidth', 1)
```

```
hold on
a=area(solution(14:101,1), solution(14:101,2), 'FaceColor', 'b', '
    FaceAlpha', 0.5); % Vyznačenie plochy unimodality modrou farbou
a.EdgeColor = 'none';
hold on
h = fill([solution(14:101,1)', fliplr(solution(14:101,1)')],[solution
    (14:101,2)', 1.2*ones(88,1)'], 'g'); % Vyznačenie plochy bimodality
    zelenou farbou
set(h, 'FaceAlpha', 0.5)
h.EdgeColor = 'none';
hold on
plot(solution(14:151,1),solution(14:151,2),'-o', 'Color', '#00008B','
    LineWidth', 1)
hold on
plot(0.5, 1, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')
hold on
plot(0.51282818, 0.77784582255, 'ro', 'MarkerSize', 7, 'MarkerFaceColor'
    , 'r')
hold on
plot(0.5, 0.6, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')

xlabel('k=k_{on}/k_{off}', 'FontSize', 16);
ylabel('b', 'FontSize', 16);
title('Plocha bimodality', 'FontSize', 16);

grid on;
hold off

%n=5
figure
plot(solution(30:151,1),solution(30:151,2),'-o', 'Color', '#00008B','
    LineWidth', 1)
hold on
a=area(solution(30:151,1), solution(30:151,2), 'FaceColor', 'b', '
    FaceAlpha', 0.5); % Vyznačenie plochy unimodality modrou farbou
a.EdgeColor = 'none';
hold on
h1 = fill([solution(30:151,1)', fliplr(solution(30:151,1)')],[solution
```

```
    (30:151,2)', 1.2*ones(122,1)'], 'g'); % Vyznačenie plochy bimodality
    zelenou farbou
set(h1, 'FaceAlpha', 0.5)
h1.EdgeColor = 'none';
hold on
h1 = fill([linspace(0, 0.03241,7), fliplr(linspace(0, 0.03241,7))],[
    zeros(7,1)', 1.2*ones(7,1)'], 'b'); % Vyznačenie plochy unimodality
    modrou farbou
set(h1, 'FaceAlpha', 0.5)
h1.EdgeColor = 'none';
hold on
plot(solution(30:151,1),solution(30:151,2),'-o', 'Color', '#00008B','
    LineWidth', 1)
hold on
plot(0.0508611321, 0.75649759308, 'ro', 'MarkerSize', 7, '
    MarkerFaceColor', 'r')
hold on
plot(0.0508611321, 1, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')
hold on
plot(0.0508611321, 0.4, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r')

xlabel('k=k_{on}/k_{off}', 'FontSize', 16);
ylabel('b', 'FontSize', 16);
title('Plocha bimodality', 'FontSize', 16);

grid on;
hold off
```

A.0.5

Nasledovný kód slúži na vykreslenie funkcie hustoty koncentrácie voľných bielkovín pre $n = 2, 3, 4, ...$ Pomocou tohto kódu boli vytvorené obrázky 12, 13, 14.

```
%Nastavenie parametrov
n=5;  %veľkosť multiméru, n=2,3,4,...
gamma=1;  %rýchlosť zániku
A=1/b;  %transkripčná rýchlosť
```

```
k=0.0508611321; %k=k_on/k_off
b=0.4;  %priemerná veľkosť vzplanutia

p_1=@(x) (1/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(k).*n.*((x.^(n-1))
    /(n-1))-((x+(k).*x.^(n))/(b))).*(1+(k).*n.*x.^(n-1));%funkcia
          %pre výpočet normalizačnej konštanty
kappa=1/integral(p_1,0,Inf); %normalizačná konštanta
% funkcia hustoty koncentrácie voľných bielkovín pre n=2,3,4,...
p=@(x) (kappa/gamma).*x.^((A/gamma)-1).*exp((A/gamma).*(k).*n.*((x.^(n
    -1))/(n-1))-((x+(k).*x.^(n))/(b))).*(1+(k).*n.*x.^(n-1));
S=integral(p,0,Inf); % kontrola správnosti normalizačnej konštanty, S=1
S
%Vykreslenie funkcie hustoty koncentrácie voľných bielkovín
%pre n=2,3,4,...
figure
o=0:0.001:3;
plot(o,p(o),'b','LineWidth', 1)
xlabel('x_f', 'FontSize', 16)
ylabel('funkcia hustoty p(x_f)', 'FontSize', 16)
title('Hustota koncentrácie voľných bielkovín x_f', 'FontSize', 16)
```