

Univerza  
v Ljubljani

Fakulteta *za računalništvo  
in informatiko*



Miha Moškon

MODELI IN METRIKE DINAMIKE PREKLOPA  
V ENOSTAVNIH BIOLOŠKIH SISTEMIH  
ZA POTREBE RAČUNALNIŠKIH STRUKTUR  
PRIHODNOSTI

Doktorska disertacija

Mentor: izr. prof. dr. Miha Mraz

Ljubljana, 2012



Univerza v Ljubljani  
Fakulteta za računalništvo in informatiko

Miha Moškon

## **Modeli in metrike dinamike preklopa v enostavnih bioloških sistemih za potrebe računalniških struktur prihodnosti**

### **POVZETEK**

Sintezna biologija (angl. *synthetic biology*) je hitro razvijajoča se veda, ki se ukvarja s spreminjanjem obstoječih in gradnjo novih bioloških sistemov z vnaprej določenimi funkcionalnostmi. Pri tem združuje znanja raziskovalcev z različnih področij, saj inženirske pristope povezuje z biološkimi. Biološko računalništvo (angl. *biological computing*) je relativno novo podpodročje sintezne biologije, ki se ukvarja z izgradnjo gradnikov zmožnih procesiranja podatkov oziroma izgradnjo računalnika v bioloških sistemih. Z določeno stopnjo abstrakcije, tj. z vpeljavo modelov za simuliranje dinamike tovrstnih sistemov, postane področje obvladljivo tudi za strokovnjake s področja računalništva, ki v izgradnji tovrstnih bioloških sistemov vidijo predvsem hipotetično procesno platformo prihodnosti.

V zadnjem času je bilo na področju biološkega računalništva razvitih že veliko število bioloških sistemov, ki bi lahko služili kot osnovni gradniki pri gradnji biološkega računalnika (npr. kombinatorna logična vrata, oscilatorji in pomnilne celice). Po drugi strani razpolagamo s številnimi modeli, ki omogočajo simuliranje dinamike tovrstnih sistemov. S stališča snovanja kompleksnejših bioloških sistemov zmožnih procesiranja podatkov se poraja potreba po analizi modelirane dinamike v bioloških sistemih. V kontekstu raziskav s področja računalništva pod pojmom dinamike smatramo spremembe prisotnosti opazovanih kemijskih vrst (angl. *chemical species*), ki jih lahko interpretiramo kot informacijo. Kriterije za ocenjevanje primernosti in medsebojne povezljivosti bioloških sistemov lahko vzpostavimo na podlagi metrik, ki pa še niso bile vpeljane. Vpeljava metrik bi omogočila objektivnost pri izbiri ustreznih bioloških logičnih gradnikov kot osnov bioloških sistemov namenjenih procesiranju podatkov. Posledično bi bila možna oziroma veliko lažja izgradnja kompleksnejših bioloških sistemov, kot je naprimer biološki računalnik.

V delu pojasnimo princip gradnje novih bioloških sistemov na podlagi gensko regulatornih omrežij. Opišemo različne pristope k modeliranju tovrstnih sistemov in podamo

zglede modelov enostavnih bioloških sistemov zmožnih procesiranja podatkov. Na podlagi ugotovljenih lastnosti se lotimo vzpostavitve metrik, ki po eni strani izhajajo iz karakteristik elektronskih digitalnih vezij, po drugi strani pa iz matematičnega področja nelinearnih dinamičnih sistemov, saj lahko na tovrstne gradnike gledamo iz obeh perspektiv. Vzpostavljene metrike ovrednotimo na že prej podanih zgledih. Uporabnost metrik poleg tega demonstriramo z modularnim povezovanjem osnovnih gradnikov v kompleksnejšo celoto oziroma kompleksnejši biološki sistem zmožen procesiranja podatkov. Na podlagi dobljenih rezultatov opredelimo prednosti in slabosti tovrstnih sistemov, ki jih moramo upoštevati pri načrtovanju ciljnih aplikacij in podamo izhodišča za nadaljnje delo.

**Ključne besede:** biološko računalništvo, dinamika preklopa, modeliranje bioloških sistemov, nekonvencionalno procesiranje, sintezna biologija, genska regulatorna omrežja, metrike.

University of Ljubljana  
Faculty of Computer and Information Science

Miha Moškon

## **Computer structures perspective on switching dynamics of simple biological systems**

### **ABSTRACT**

Synthetic biology is a rapidly evolving discipline that copes with the modifications of existent and with the construction of new biological systems with novel functionalities. Its interdisciplinarity arises from combining of engineering and biological sciences. Biological computing is a relatively new research field that is analysing the possibilities of constructing a biological computer. Synthetic biology approaches can also be used in order to build biological computer. Certain levels of abstraction, i.e. with the introduction of models which can be used in order to simulate the dynamics of such biological systems, gives an opportunity to the scientists from different disciplines, such as computer science, to perform in-depth researches on these fields.

Large number of biological systems that are capable of data processing, such as combinatorial logical gates, oscillators and flip-flops, have already been realized with the synthetic biology approaches in the last years. Even more, many models that can be used for simulating the dynamics of such systems already exist. In order to construct more complex biological systems with the data processing capabilities analysis of their *switching dynamics* has to be made. A principal challenge in this field is the development of metrics that would estimate the information processing capabilities of basic primitives and possibilities of building more complex systems with their interconnectivity. With the introduction of such metrics the characterization of such systems could be made more straightforwardly and objectively. Construction of more complex biological systems with the data processing capabilities and consequently constructing a biological computer would become a possible task to perform.

We present the basics of construction of biological systems based on gene regulatory networks. We develop the basic approaches in modeling of such systems and demonstrate them with few example models of simple biological systems with data processing capabil-

ities. Based on the results of simulated switching dynamics metrics are established. Metrics are established regarding the characteristics which are used to describe electronic digital systems and regarding the mathematical field of nonlinear dynamical systems. Evaluation of metrics is demonstrated on simulation results of examples presented before. Metrics are also used in order to evaluate the interconnectivity of presented primitives and to modularly connect these primitives in a more complex biological system with data processing capabilities. Based on the results we discuss advantages and disadvantages of processing in such systems which have to be considered when choosing their target applications.

**Key words:** biological computing, switching dynamics, modeling of biological systems, unconventional computing, synthetic biology, gene regulatory networks, metrics.

## ZAHVALA

*Na prvem mestu se zahvaljujem mentorju, izr. prof. dr. Mihi Mrazu, in vodji Laboratorija za računalniške strukture in sisteme prof. dr. Nikolaju Zimicu za vodenje in usmerjanje v času podiplomskega študija. Hvala tudi laboratorijskim sodelavcem, zaradi katerih je bil marsikateri delovnik krajši. Hvala Miha, Primož, Iztok, Domen, Mattia, Miran in Vito. Še posebej bi se rad zahvalil svojim staršem za vso pomoč, varnost in razumevanje skozi celotno obdobje formalnega šolanja. Hvala Stane in Olga. Za pomoč pri premagovanju marsikaterih težav, za poslušanje in podporo v času pisanja tega dela pa se zahvaljujem Zali.*

— Miha Moškon, Ljubljana, 2012.





## KAZALO

<b>Povzetek</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Zahvala</b>	<b>v</b>
<b>1 Uvod</b>	<b>5</b>
1.1 Osnovne hipoteze . . . . .	7
1.2 Prispevki znanosti . . . . .	7
1.3 Metodologije dela . . . . .	8
1.4 Pregled naloge . . . . .	8
<b>2 Osnovni pojmi iz sintezne biologije</b>	<b>11</b>
2.1 Kemijska zvrst . . . . .	11
2.2 DNA in gen . . . . .	11
2.3 Promotor . . . . .	12
2.4 Kodirajoče zaporedje proteina . . . . .	12
2.5 Sinteza proteina . . . . .	12
2.5.1 Transkripcija . . . . .	12
2.5.2 Translacija . . . . .	12
2.6 Transkripcijski faktorji . . . . .	13
2.7 Gensko regulatorna omrežja . . . . .	14
2.8 Sinteza biologija in biološko računalništvo . . . . .	15
2.9 Gostitelj . . . . .	15
2.10 Razgradnja kemijske zvrsti . . . . .	16
2.11 Šum . . . . .	16

2.12	Koncentracija kemijske zvrsti . . . . .	18
2.12.1	Koncentracija pri deterministični kinetiki . . . . .	18
2.12.2	Koncentracija pri stohastični kinetiki . . . . .	18
<b>3</b>	<b>Procesiranje podatkov v bioloških sistemih</b>	<b>21</b>
3.1	Osnove procesiranja podatkov v digitalnih elektronskih vezjih . . . . .	21
3.2	Osnove procesiranja podatkov v bioloških sistemih . . . . .	22
3.3	Primeri enostavnih bioloških sistemov zmožnih procesiranja podatkov . .	23
3.3.1	Biološki gonilnik . . . . .	23
3.3.2	Biološki negator . . . . .	24
3.3.3	Biološka AND vrata . . . . .	25
3.3.4	Biološka NOR vrata . . . . .	25
3.3.5	Preklopno stikalo . . . . .	26
3.3.6	Represilator . . . . .	30
3.4	Zaključek poglavja . . . . .	31
<b>4</b>	<b>Matematični modeli za opisovanje gensko regulatornih omrežij</b>	<b>33</b>
4.1	Deterministični modeli . . . . .	34
4.1.1	Primerjava determinističnih in stohastičnih modelov . . . . .	35
4.1.2	Zakon o vplivu koncentracij . . . . .	35
4.1.3	Enačba hitrosti reakcije . . . . .	36
4.1.4	Michaelis-Menten kinetika . . . . .	38
4.1.5	Hillove enačbe . . . . .	39
4.2	Stohastični modeli . . . . .	43
4.2.1	Opis stohastičnosti kinetike kemijskih procesov . . . . .	44
4.2.2	Pristop z glavno enačbo . . . . .	46
4.2.3	Gillespiejev algoritem za stohastično simuliranje . . . . .	46
4.2.4	Metoda $\tau$ -skokov . . . . .	47
4.2.5	Langevinova enačba in enačba hitrosti reakcije . . . . .	48
4.2.6	Stohastično modeliranje s Poissonovo porazdelitvijo . . . . .	49
4.3	Semikvantitativni modeli . . . . .	50
4.4	Primeri modeliranja enostavnih bioloških sistemov zmožnih procesiranja podatkov . . . . .	50

4.4.1	Izhodišča pri vzpostavitvi modelov . . . . .	50
4.4.2	Modeliranje biološkega gonilnika . . . . .	55
4.4.3	Modeliranje biološkega negatorja . . . . .	59
4.4.4	Modeliranje bioloških AND vrat . . . . .	62
4.4.5	Modeliranje bioloških NOR vrat . . . . .	64
4.4.6	Modeliranje preklopnega stikala . . . . .	68
4.4.7	Modeliranje razširjenega preklopnega stikala . . . . .	73
4.4.8	Modeliranje represilatorja . . . . .	76
4.5	Zaključek poglavja . . . . .	79
<b>5</b>	<b>Analiza stabilnosti bioloških sistemov</b>	<b>83</b>
5.1	Analiza bioloških sistemov z ravnovesnimi stanji . . . . .	84
5.1.1	Izračun Jacobijeve matrike . . . . .	85
5.1.2	Določitev stabilnosti ravnovesne točke . . . . .	85
5.1.3	Prikazovanje stabilnosti . . . . .	86
5.2	Analiza oscilatornih sistemov . . . . .	86
5.3	Bifurkacijska analiza . . . . .	88
5.3.1	Tipi bifurkacij . . . . .	89
5.3.2	Zgledi analize bifurkacij . . . . .	92
5.4	Zaključek poglavja . . . . .	97
<b>6</b>	<b>Metrike</b>	<b>99</b>
6.1	Izhodišča pri izbiri metrik . . . . .	100
6.1.1	Metrike za ocenjevanje zmogljivosti digitalnih vezij . . . . .	100
6.1.2	Ocenjevanje stabilnosti nelinearnih dinamičnih sistemov . . . . .	107
6.2	Metrike za vrednotenje bioloških sistemov . . . . .	108
6.2.1	Izbira nosilca informacije . . . . .	108
6.2.2	Določitev šuma . . . . .	108
6.2.3	Določitev logičnih nivojev . . . . .	109
6.2.4	Določitev prepovedanega področja . . . . .	110
6.2.5	Določitev preklopnih časov . . . . .	111
6.2.6	Določitev porabe vhoda . . . . .	111
6.2.7	Določitev periode oscilacij . . . . .	111
6.2.8	Določitev časa veljavnosti izhodnega signala . . . . .	111

6.2.9	Določitev časa osveževanja . . . . .	112
6.2.10	Določitev robustnosti . . . . .	112
6.3	Analiza osnovnih bioloških sistemov, zmožnih procesiranja podatkov . . .	113
6.3.1	Analiza biološkega gonilnika . . . . .	114
6.3.2	Analiza biološkega negatorja . . . . .	120
6.3.3	Analiza bioloških AND vrat . . . . .	121
6.3.4	Analiza bioloških NOR vrat . . . . .	124
6.3.5	Analiza preklopnega stikala . . . . .	124
6.3.6	Analiza razširjenega preklopnega stikala . . . . .	129
6.3.7	Analiza represilatorja . . . . .	133
6.4	Zaključek poglavja . . . . .	138
<b>7</b>	<b>Izgradnja kompleksnejšega biološkega sistema</b>	<b>139</b>
7.1	Povezovanje osnovnih gradnikov . . . . .	139
7.1.1	Ortogonalnost sestavnih delov . . . . .	140
7.1.2	Kompatibilnost gradnikov v smislu logičnih nivojev . . . . .	140
7.1.3	Kompatibilnost gradnikov v smislu časovnih karakteristik . . . . .	140
7.1.4	Ohranjanje logičnih nivojev . . . . .	141
7.2	Zgled kompleksnejšega biološkega sistema . . . . .	141
7.2.1	Sinhronska <i>RS</i> pomnilna celica . . . . .	141
7.2.2	Model sinhronskega preklopnega stikala . . . . .	142
7.2.3	Analiza rezultatov modela sinhronskega preklopnega stikala . . . . .	144
7.3	Zaključek poglavja . . . . .	145
<b>8</b>	<b>Zaključek</b>	<b>149</b>
	<b>Literatura</b>	<b>151</b>
	<b>Izjava</b>	<b>163</b>

## SPISEK UPORABLJENIH KRATIC IN SIMBOLOV

V nalogi so uporabljeni sledeči simboli in kratice:

- $\mathbf{x}(t)$ : vektor nenegativnih realnih števil, ki predstavljajo koncentracije kemijskih zvrsti v determinističnem modelu ob času  $t$ ;
- $x_i(t)$ : nenegativno realno število, ki predstavlja koncentracijo kemijske zvrsti  $i$  v determinističnem modelu ob času  $t$ ;
- $\mathbf{y}(t)$ : vektor nenegativnih celih števil, ki predstavljajo koncentracije kemijskih zvrsti v stohastičnem modelu ob času  $t$ ;
- $y_i(t)$ : nenegativno celo število, ki predstavlja koncentracijo kemijske zvrsti  $i$  v stohastičnem modelu ob času  $t$ ;
- $k_f$ : hitrostna konstanta vodilne reakcije (angl. *rate constant for forward reaction*);
- $k_b$ : hitrostna konstanta povratne reakcije (angl. *rate constant for backward reaction*);
- $K$ : ravnotežna konstanta (angl. *equilibrium constant*);
- $K_d$ : Michaelisova oziroma disociacijska konstanta (angl. *dissociation constant*);
- $\alpha$ : maksimalna hitrost transkripcije mRNA proteina (angl. *maximal transcription rate*) - določa maksimalno aktivnost promotorja;
- $\alpha_0$ : osnovno izražanje gena (angl. *basal expression*) oziroma minimalna hitrost transkripcije mRNA proteina - določa minimalno aktivnost promotorja; imenovan tudi koeficient puščanja (angl. *leakage coefficient*);
- $\beta$ : hitrost translacije (angl. *translation rate*) mRNA proteina v protein;

- $\delta$ : koeficient razgradnje proteina (angl. *degradation coefficient*);
- $t_h$ : razpolovna doba kemijske zvrsti (angl. *half-life*);
- $t_l$ : življenjska doba kemijske zvrsti (angl. *life time*);
- $N_A$ : Avogardovo število, ki določa število elementarnih entitet (npr. atomov) v enem *molu*; en mol vsebuje  $N_A = 6,0221415 \cdot 10^{23}$  elementarnih entitet;
- $V$ : reakcijska prostornina (angl. *reaction volume*) je prostornina, znotraj katere se odvijajo opazovani kemijski procesi; ponavadi predpostavljamo, da je to prostornina celice, v kateri se nahaja opazovani sistem ( $V = 2 \cdot 10^{-15} L$  [1]);
- $\Omega$ : velikost sistema, ki je pri determinističnih modelih že upoštevana v samih koncentracijah, saj jih podajamo v  $\frac{mol}{L}$ , pri stohastičnih modelih pa je potrebno reakcijsko velikost sistema eksplicitno upoštevati; izračunamo jo kot  $\Omega = V \cdot N_A$ ;
- CME: kemijska glavna enačba (angl. *Chemical Master Equation*);
- $S_i$ : ena od opazovanih kemijskih zvrsti ( $i$ -ta kemijska zvrst), ki določa stanje sistema;
- $R_j$ : ena od opazovanih kemijskih reakcij ( $j$ -ta kemijska reakcija), ki vpliva na stanje sistema;
- $\nu_j$ : vektor stehiometričnih koeficientov reakcije  $j$ , ki določa spremembo opazovanih kemijskih zvrsti po izvedbi reakcije  $R_j$ ; določa namreč razmerje med reaktanti in produkti reakcije  $j$ ;
- $a_j(\mathbf{y}(t))$ : funkcija, ki določa nagnjenost k izvedbi reakcije  $j$  (angl. *propensity function*) in posredno določa verjetnost izvedbe funkcije  $j$  v določenem času;
- $J(s)$ : Jacobijeva matrika, ki predstavlja linearizacijo sistema v okolici točke  $s$ ; sestavljena je iz parcialnih odvodov diferencialnih enačb, ki predstavljajo spremembe koncentracij kemijskih zvrsti v opazovanem sistemu;
- $\mathbf{x}^*$ : ravnovesno stanje sistema;
- DNA: deoksiribonukleinska kislina (angl. *deoxyribonucleic acid*);
- RNA: ribonukleinska kislina (angl. *ribonucleic acid*);
- mRNA: sporočilna RNA (angl. *messenger RNA*);

- $\sigma_N$ : določa standardno deviacijo signala kot posledico šuma;
- $C_L$ : nizki logični nivo izhodne koncentracije nosilca informacije, določen na podlagi stabilnega stanja determinističnega modela sistema;
- $C_H$ : visoki logični nivo izhodne koncentracije nosilca informacije, določen na podlagi stabilnega stanja determinističnega modela sistema;
- $C_{OH}$ : visoki nivo izhodne koncentracije;
- $C_{OH(\min)}$ : spodnja meja visokega nivoja izhodne koncentracije;
- $C_{OL}$ : nizki nivo izhodne koncentracije;
- $C_{OL(\max)}$ : zgornja meja nizkega nivoja izhodne koncentracije;
- $C_{IH(\min)}$ : spodnja meja visokega nivoja vhodne koncentracije;
- $C_{IL(\max)}$ : zgornja meja nizkega nivoja vhodne koncentracije;
- $N_H$ : maksimalen pričakovani šum pri signalu v visokem logičnem stanju;
- $N_L$ : maksimalen pričakovani šum pri signalu v nizkem logičnem stanju;
- $N$ : maksimalen pričakovani šum v izhodnem signalu;
- $NM_O$ : izhodno prepovedano področje;
- $NM_I$ : vhodno prepovedano področje;
- $t_r$ : čas vzpona (angl. *rise time*) izhodnega signala;
- $t_f$ : čas padca (angl. *fall time*) izhodnega signala;
- $t_s$ : čas preklopa (angl. *switch time*) izhodnega signala;
- $C_I$ : količina izhodnega proteina, ki se porabi ob vezavi le-tega v vlogi vhoda v obravnavani gradnik (*poraba vhoda*);
- $t_{v_H}$ : čas veljavnosti izhodnega signala v visokem stanju;
- $t_{v_L}$ : čas veljavnosti izhodnega signala v nizkem stanju;
- $t_v$ : čas veljavnosti izhodnega signala;

- $t_{REF}$ : čas osveževanja;
- $R_d$ : robustnost sistema kot razdalja med stabilnimi stanji;
- $R_b$ : robustnost sistema kot razdalja od točke bifurkacije;
- $R_c$ : robustnost sistema kot velikost stabilnega limitnega cikla.



# 1 Uvod

Ideja o procesiranju podatkov v bioloških sistemih (na DNA temelječih medijih) je stara že približno 25 let. V praksi je bila prvič potrjena kot obetaven koncept z izvedbo Adlemanovega eksperimenta za reševanje kompleksnih procesnih problemov [2]. Tudi sama gostota zapisa dušikovih baz adenina, citozina, gvanina in timina na DNA verigi, ki jih lahko interpretiramo kot kodirane podatke, je za nekaj velikostnih redov večja, kot je danes izvedljiva na področju digitalnih pomnilnih medijev, kot so recimo trdi diski [3]. Poleg tega je s komplementarnostjo dvojne vijačnice na DNA verigi že dosežena redundanca zapisa.

V zadnjem času se zelo hitro razvija področje sintezne biologije (angl. *synthetic biology*) [4–6], ki se ukvarja z umetno sintezo bioloških sistemov z vnaprej določenimi funkcionalnostmi. Z vidika področja računalništva so tovrstni gradniki zanimivi, saj bi ob obstoju ustreznih bioloških sistemov, ki pokrivajo funkcionalnosti logičnih primitivov, lahko zgradili biološki računalnik.

Sinteza novih bioloških sistemov poteka večinoma na podlagi regulacije genetskih omrežij (angl. *gene regulatory networks*). Tako kot na področju računalniškega proce-

siranja tudi tu operiramo z vhodi, navodili za procesiranje in izhodi. V vlogi vhodov nastopajo proteini in fizikalno-kemijske količine z vplivom na funkcijo proteinov (npr. IPTG, temperatura), ki vršijo nalogo transkripcijskih faktorjev. Navodila za procesiranje so v genetskih omrežjih zapisana na DNA verigi, ki določa samo strukturo izhodnih proteinov in odvisnost njihove pojavnosti od prisotnosti vhodnih proteinov. Tudi v vlogi izhodov nastopajo proteini. Ti lahko spet vršijo vlogo vhodov v opazovanem sistemu (povratna zanka) ali v nekem drugem sistemu.

Sintezna biologija z vidika računalniškega procesiranja danes že omogoča realizacijo enostavnih logičnih primitivov, kot so kombinatorna logična vrata [7–9], pomnilni element [10, 11] in oscilator [12], kot tudi kompleksnejših logičnih gradnikov, kot sta naprimer števec [13] in sistem za formacijo vzorcev [14].

Kompleksnost bioloških sistemov s številom vhodov narašča in postane razmeroma hitro neobvladljiva. Količino eksperimentalnega dela v fazi načrtovanja *umetnega biološkega sistema* lahko zmanjšamo z uporabo matematičnih modelov za simuliranje kemijskih procesov [15, 16]. Z vidika področja računalništva bi namembnost tovrstnih modelov lahko opredelili z modeliranjem prisotnosti in koncentracij kemijskih zvrsti v opazovanem biološkem sistemu. Na ta način lahko brez eksperimentalnega dela pridemo do osnovnih zakonitosti dinamike v ciljnem biološkem sistemu, kar nam močno olajša njegovo realizacijo.

V grobem lahko matematične modele za simuliranje dinamike v bioloških sistemih razdelimo v tri skupine [15], in sicer na deterministične, stohastične in semikvantitativne. Deterministični modeli so relativno enostavni, omogočajo pa zgolj aproksimacijo bodisi povprečnega odziva celice v populaciji bodisi odziva posamezne celice v daljšem časovnem obdobju. Do modeliranja heterogenega odziva populacije enakih celic lahko pridemo s stohastičnimi modeli, ki upoštevajo stohastičnost kemijskih reakcij, pri čemer so ti modeli veliko kompleksnejši od determinističnih. Kompromis med obema pristopoma je semikvantitativno modeliranje, ki z določenimi poenostavitvami stohastičnega modeliranja kombinira oba pristopa. Več tipov modelov ločimo tudi znotraj posameznih skupin, večinoma pa deterministični modeli temeljijo na Hillovih enačbah [16], stohastični pa na reševanju kemijske glavne enačbe (angl. *chemical master equation*) [17].

S postopki sintezne biologije je bilo torej razvitih že veliko število umetnih bioloških sistemov, ki predstavljajo potencialne gradnike sistemov za procesiranje podatkov [7, 9, 13, 18–21]. Prav tako že razpolagamo s številnimi modeli za simuliranje dinamike v

tovrstnih bioloških sistemih. Osnovna pomanjkljivosti na področju modeliranja v sintezni biologiji pa je po našem mnenju v odsotnosti metrik. Z vidika procesiranja podatkov bi slednje potrebovali za ocenjevanje *ustreznosti dinamike* v modeliranih bioloških sistemih. Pod pojmom *ustreznosti dinamike* na tem mestu smatramo vse časovne dinamične značilnosti nosilca podatkov, ki ustrezajo značilnostim digitalne preklopne logike. Metrike ovrednotene na množici bioloških sistemov v vlogi potencialnih gradnikov za procesiranje podatkov bi tako lahko uporabili za identifikacijo gradnikov z najustreznejšimi dinamičnimi lastnostmi, za analizo kompatibilnosti med osnovnimi gradniki pri gradnji kompleksnejših bioloških sistemov ali za namene avtomatskega načrtovanja bioloških sistemov s sposobnostjo procesiranja podatkov.

## 1.1 Osnovne hipoteze

Glavni motivi pričujoče disertacije so potrditve naslednjih hipotez:

- obstoj fizikalne količine izbrane kemijske zvrsti v vlogi izhoda (npr. koncentracije), ki je sposobna vršiti funkcijo nosilca podatkov, na podoben način kot napetost v klasični elektroniki,
- koristnost vpeljave metrik za analizo modelirane dinamike preklopa v bioloških sistemih (slednje ne velja le za področje procesiranja podatkov, temveč tudi za področja medicine, farmacije itd.),
- obstoj nabora bioloških gradnikov oziroma možnost njihove realizacije s postopki sintezne biologije, ki imajo zelene metrične in logične karakteristike in predstavljajo funkcijsko poln nabor logičnih funkcij.

## 1.2 Prispevki znanosti

V doktorski disertaciji predstavimo naslednja prispevka k znanosti:

- *Vzpostavitev metrik za ocenjevanje ustreznosti dinamike preklopa posameznih bioloških sistemov kot procesnih elementov*: Vzpostavljene bodo metrike, ki bodo omogočale vzpostavitev objektivnega ocenjevanja dinamike opazovanih bioloških sistemov kot sistemov za procesiranje podatkov.
- *Vzpostavitev nabora bioloških gradnikov, ki predstavljajo funkcijsko poln nabor logičnih funkcij*: Vzpostavljena bodo gensko regulatorna omrežja gradnikov, katerih

modelirana dinamika odseva zelene dinamične lastnosti opredeljene s pomočjo vzpostavljenih metrik in katerih delovanje kot celota predstavlja funkcijsko poln nabor logičnih funkcij.

### 1.3 Metodologije dela

Pri izdelavi doktorskega dela so bile uporabljene naslednje metode:

- Pregled in analiza relevantne literature: Narejen je bil pregled in analiza literature s področja sintezne biologije, modeliranja bioloških sistemov, procesiranja podatkov z biološkimi sistemi in analize obstoječih bioloških struktur.
- Implementacija ustreznih modelov: V programskem okolju so bili implementirani različni modeli bioloških struktur. Prav tako so bili postavljeni modeli kompleksnejših bioloških struktur, ki temeljijo na osnovnih modelih.
- Analiza modelov: Opravljena je bila analiza različnih tipov modelov in izbira najustrežnejših modelov za simuliranje delovanja bioloških sistemov kot sistemov za procesiranje podatkov.
- Vzpostavitev metrik za analizo bioloških sistemov z vidika procesnih elementov: Vzpostavljene so bile metrike za ocenjevanje ustreznosti posameznih bioloških sistemov kot procesnih elementov.
- Eksperimentalno delo s simulatorjem: V simulacijskem okolju so bili opravljeni eksperimenti nad postavljenimi modeli posameznih bioloških sistemov. Opravljena je bila analiza rezultatov z vidika metrik.
- Primerjalna analiza: Na podlagi metrik je bila izvedena primerjalna analiza med postavljenimi modeli različnih bioloških sistemov. Prav tako je bila izvedena analiza kompleksnejših bioloških sistemov na podlagi najustrežnejših modelov.

### 1.4 Pregled naloge

V nalogi se ukvarjamo z vzpostavitvijo modelov različnih bioloških sistemov in njihovim vrednotenjem z vpeljavo metrik za ocenjevanje primernosti funkcionalnosti gradnikov za procesiranje podatkov. V poglavju 2 predstavimo osnovne pojme sintezne biologije, ki jih v nalogi uporabljamo in bralcu z računalniškim znanjem po vsej verjetnosti niso

poznani. V poglavju 3 razložimo pristope pri procesiranju v sodobnih računalniških sistemih in jih vpeljemo na področje bioloških sistemov, natančneje na gensko regulatorna omrežja. Pri tem navedemo zglede tovrstnih bioloških sistemov, zmožnih procesiranja podatkov. Modeliranje gensko regulatornih omrežij je razloženo v poglavju 4, kjer poleg tega opišemo vzpostavitev modelov bioloških sistemov, navedenih v poglavju 3. V poglavju 5 predstavimo analizo stabilnosti bioloških sistemov, ki izhaja iz matematičnega področja nelinearnih dinamičnih sistemov. Rezultati modelov predstavljenih bioloških sistemov so vrednoteni na podlagi metrik, vzpostavljenih v poglavju 6, ki izhajajo iz karakteristik elektronskih digitalnih vezij, in iz stabilnostne analize obravnavanih sistemov. Na podlagi vrednotenja modelov v poglavju 7 vzpostavimo model kompleksnejšega biološkega gradnika zmožnega procesiranja podatkov sestavljenega iz osnovnih gradnikov opisanih v prejšnjih poglavjih. Na koncu ovrednotimo ustreznost tovrstnih bioloških sistemov kot gradnikov za procesiranje podatkov in jih primerjamo s klasičnimi elektronskimi elementi, ki sestavljajo današnje računalnike. Zaključek z izhodišči za nadaljnje delo sledi v poglavju 8.



# 2 Osnovni pojmi iz sintezne biologije

V pričujočem poglavju so razloženi osnovni pojmi uporabljeni v nalogi, ki so bralcu z računalniškega področja verjetno nepoznani.

## 2.1 Kemijska zvrst

*Kemijska zvrst* (angl. *chemical specie*) predstavlja osnovno entiteto, ki jo opazujemo znotraj biološkega sistema. Primeri opazovanih kemijskih zvrsti so npr. *proteini*, *DNA molekule*, *mRNA molekule* itd. Njihovo prisotnost ponavadi merimo s koncentracijo molekul (glej razdelek 2.12) opazovane kemijske zvrsti v nekem okolju (npr. celici).

## 2.2 DNA in gen

*Deoksiribonukleinska kislina* oziroma DNA (angl. *deoxyribonucleic acid - DNA*) je kislina, ki vsebuje genetski zapis organizma. Sestavljena je iz štirih različnih *baz*, in sicer *adenina* (A), *timina* (T), *citozina* (C) in *guanina* (G). DNA vsebuje večje število genov, ki določajo pogoje za izražanje določenega proteina in zaporedje aminokislin, ki ta protein sestavljajo. V tem smislu lahko predpostavljamo, da so geni sestavljeni iz dveh

funkcionalnih delov, in sicer iz *promotorja* in *kodirajočega zaporedja proteina*.

### 2.3 Promotor

*Promotor* (angl. *promoter*) je del gena, na katerega se ob prisotnosti oziroma odsotnosti drugih *transkripcijskih faktorjev* veže encim *RNA polimeraza*. Vezava RNA polimeraze na promotor povzroči začetek procesa *transkripcije* (glej razdelek 2.5.1). V tem procesu se kodirajoče zaporedje proteina, ki sledi promotorju, prepíše v t. i. *mRNA* (angl. *messenger RNA*), imenovano tudi *sporočilna RNA*.

### 2.4 Kodirajoče zaporedje proteina

*Kodirajoče zaporedje proteina* je del gena, ki določa strukturo sintetiziranega izhodnega proteina. V njem je zapisana sekvenca aminokislin, ki sestavljajo ciljni protein. Pogoji za začetek izražanja oziroma transkripcije kodirajočega zaporedja proteina so določeni s soležnim promotorjem.

### 2.5 Sinteza proteina

Sinteza izhodnega proteina povzroči povečanje njegove koncentracije znotraj opazovanega okolja in je posledica *izražanja* določenega *gena* (angl. *gene expression*). Opišemo jo lahko s sekvenco dveh procesov, in sicer procesa *transkripcije* (angl. *transcription*) in procesa *translacije* (angl. *translation*) [22]. Oba procesa sta podrobneje predstavljena v nadaljevanju.

#### 2.5.1 Transkripcija

Prepisovanje tistega dela gena, ki določa strukturo izhodnega sintetiziranega proteina, tj. kodirajoče zaporedje proteina, v mRNA imenujemo *transkripcija* [22]. Intenzivnost procesa transkripcije je pogojena s prisotnostjo oziroma odsotnostjo določenih transkripcijskih faktorjev (glej razdelek 2.6).

#### 2.5.2 Translacija

Na podlagi mRNA *ribosomi* (angl. *ribosomes*) v procesu translacije [22] povežejo aminokislino v verigo, ki sestavlja ciljni protein. Predpostavljamo lahko, da proces translacije

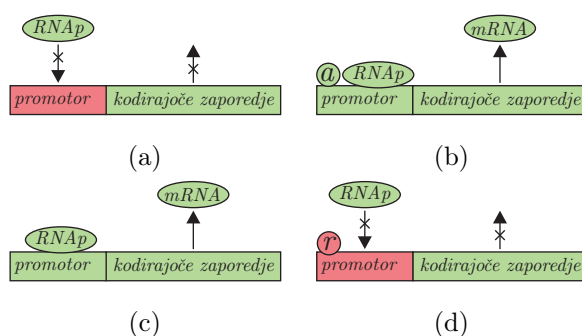


poteka brezpogojno - iz mRNA se sintetizirajo proteini z določeno časovno zakasnitvijo brez dodatnih zunanjih pogojev, vse dokler se mRNA ne razgradi.

## 2.6 Transkripcijski faktorji

Potek in hitrost transkripcije je odvisna od prisotnosti oziroma odsotnosti t. i. *transkripcijskih faktorjev* (angl. *transcription factors*) [21]. Proces transkripcije se sproži, ko se na promotor gena veže transkripcijski faktor RNA polimeraza (angl. *RNA polymerase*). Ostale transkripcijske faktorje lahko razdelimo v dve skupini, in sicer na *aktivatorje* (angl. *activators*) in *represorje* (angl. *repressors*). Aktivatorji proces vezave RNA polimeraze na promotor omogočijo oziroma pospešujejo (glej sliki 2.1 (a) in (b)), represorji pa ta proces zavirajo oziroma preprečujejo (glej sliki 2.1 (c) in (d)).

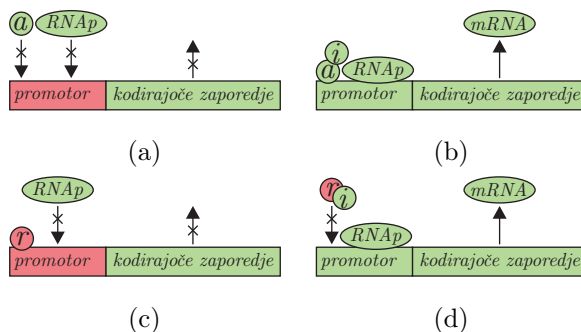
S svojo vezavo na promotor aktivatorji tako pospešujejo oziroma omogočijo vezavo na promotor RNA polimerazi. Po drugi strani se represorji vežejo na poseben del promotora (t. i. *operator*) in s tem zavirajo oziroma preprečijo vezavo RNA polimeraze na promotor. V naši obravnavi tovrstnih sistemov predpostavljamo, da je RNA polimeraza v celici vedno na razpolago, prisotnost ostalih transkripcijskih faktorjev pa se s časom spreminja.



**Slika 2.1** Regulacija transkripcije, pri čemer sliki (a) in (b) prikazujeta promotor, ki je aktiven ob prisotnosti aktivatorja, sliki (c) in (d) pa promotor, ki je aktiven ob odsotnosti represorja. RNA polimeraza je na sliki predstavljena z  $RNA_p$ , mRNA z  $mRNA$ , aktivator z  $a$ , represor pa z  $r$ .

Transkripcijski faktorji so lahko v določenih primerih v *aktivnih* stanjih, v katerih se lahko normalno vežejo na promotor, ali *neaktivnih* stanjih, v katerih se na promotor ne morejo vezati. V primeru, da opazujemo aktivator, je le-ta v aktivnem stanju, ko je nanj vezan t. i. *induktor* (angl. *inducer*), ki spremeni afiniteto aktivatorja in s tem njegovo zmožnost vezave na promotor (glej sliki 2.2 (a) in (b)). Brez induktorja je aktivator v neaktivnem stanju in se tako ne more vezati na promotor. V primeru, da opazujemo

repressor, je situacija ravno nasprotna, saj se lahko na promotor veže le prost represor, z induktorjem pa se na promotor ne more vezati (glej sliki 2.2 (c) in (d)).

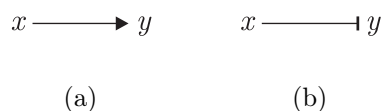


**Slika 2.2** Regulacija transkripcije, pri čemer sliki (a) in (b) prikazujeta promotor, ki je aktiven ob prisotnosti aktivatorja in induktorja, sliki (c) in (d) pa promotor, ki je aktiven ob odsotnosti represorja oziroma ob prisotnosti induktorja. RNA polimeraza je na sliki predstavljena z  $RNAP$ , mRNA z  $mRNA$ , aktivator z  $a$ , represor z  $r$ , induktor pa z  $i$ .

## 2.7 Gensko regulatorna omrežja

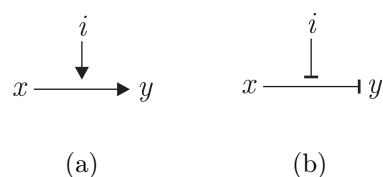
V vlogi transkripcijskih faktorjev, ki določajo hitrost vezave RNA polimeraze na promotor (glej razdelek 2.3) in s tem intenzivnost izražanja posameznega gena, lahko nastopajo proteini, ki so rezultat izražanja nekega drugega gena. Še več, izhodni protein, določen z opazovanim genom, lahko vpliva na transkripcijo svoje lastne mRNA. S spreminjanjem strukture promotorja lahko določimo, kateri proteini bodo nastopali kot transkripcijski faktorji in kakšna bo njihova vloga, tj. ali bodo transkripcijo zavirali (represorji) ali pa pospeševali (aktivatorji). Z regulacijo preko transkripcijskih faktorjev lahko na ta način posamezne dele DNA povezujemo med seboj v *gensko regulatorna omrežja* (angl. *gene regulatory networks*).

Strukturo posameznega gensko regulatornega omrežja lahko ponazorimo z *usmerjenim grafom* (angl. *directed graph*), pri katerem kot vozlišča nastopajo proteini, ki so lahko posledica izražanja posameznega gena v omrežju, transkripcijski faktor za nek del omrežja ali oboje. Relacije med proteini lahko ponazorimo z dvema tipoma povezav, pri čemer prvi tip povezav ponazarja aktivatorski vpliv (slika 2.3(a)), drugi tip pa represorski vpliv nekega proteina na izražanje drugega proteina (slika 2.3(b)). Z razširjeno notacijo lahko prikazujemo tudi vpliv proteinov, ki imajo vlogo induktorjev, in sicer tako, da vpeljemo povezave med vozlišči in drugimi povezavami. Spet razpolagamo z dvema tipoma



**Slika 2.3** Protein  $x$  lahko aktivira izražanje proteina  $y$  (a) ali pa njegovo izražanje zavira (b).

povezav, in sicer s povezavami, ki ponazarjajo vpliv induktorjev na vezavo aktivatorja (slika 2.4(a)) in s povezavami, ki ponazarjajo vpliv induktorjev na vezavo represorja (slika 2.4(b)).



**Slika 2.4** Vezava induktorja na aktivator omogoči njegovo vezavo na promotor (a), vezava induktorja na represor pa zavira njegovo vezavo na promotor (b).

## 2.8 Sintezna biologija in biološko računalništvo

Sintezna biologija (angl. *synthetic biology*) [4-6, 23] se ukvarja s sintezo novih bioloških sistemov s specifično uporabnostjo. Biološko računalništvo (angl. *biological computing*) je podpodročje sintezne biologije, ki se ukvarja z izgradnjo sistemov, zmožnih procesiranja podatkov oziroma izgradnjo računalnika na podlagi bioloških molekul [20]. Največ poskusov v smeri izgradnje biološkega računalnika je bilo v zadnjih letih posvečeno izgradnji bioloških sistemov, zmožnih procesiranja podatkov na podlagi gensko regulatornih omrežij s postopki sintezne biologije, zato se na to podpodročje biološkega računalništva naslanjamo tudi v tem delu.

## 2.9 Gostitelj

Končna realizacija biološkega sistema je izvedena z vnosom sintetiziranih DNA zaporedij, ki določajo funkcionalnosti le-tega, v celice ciljnega organizma oziroma v t. i. *gostitelja*. V ta namen se na področju sintezne biologije običajno uporabljajo celice bakterij *Escherichia coli*, ki spadajo v skupino *prokariotskih organizmov*, ali celice kvasovk *Sac-*

*charomyces cerevisiae*, ki spadajo v skupino *evkariontskih organizmov*. V zadnjem času kot gostitelji pogosto nastopajo tudi sesalske celice [24].

## 2.10 Razgradnja kemijske zvrsti

Razgradnja kemijske zvrsti povzroči zmanjšanje njene koncentracije in je proces, katerega rezultat je nasproten njeni sintezi. Hitrost razgradnje lahko do neke mere kontroliramo z zunanjimi vplivi (npr. z UV-svetlobo), do določene mere pa je razgradnja vselej prisotna.

Razgradnjo kemijske zvrsti  $A$  lahko predstavimo s kemijsko reakcijo



kjer je koeficient  $\delta$  hitrost razgradnje definiran kot

$$\delta = \frac{\ln(2)}{t_h}. \quad (2.2)$$

Pri tem  $t_h$  predstavlja razpolovno dobo kemijske zvrsti [12, 25]. Na podoben način lahko definiramo življenjsko dobo kemijske zvrsti z izrazom

$$t_l = \frac{t_h}{\ln(2)}. \quad (2.3)$$

## 2.11 Šum

Šum je neželen pojav, do katerega pride zaradi različnih faktorjev in povzroči heterogenost odziva populacije enakih celic z enakimi vhodi oziroma različne odzive v sicer enakih celicah znotraj kolonije. Ločimo več vrst šuma [26], in sicer:

- *Notranji šum* (angl. *internal noise*) izhaja iz sistema samega in ga lahko razdelimo na dva podtipa. Prvi je *intrinzičen šum* (angl. *intrinsic noise*) [25, 27], ki izhaja iz stohastičnosti kemijskih procesov, kot so transkripcija, translacija, aktivacija promotorja in razgradnja kemijskih zvrsti. Še posebej pride do izraza pri majhnih koncentracijah kemijskih zvrsti v sistemu [28]. Intrinzičen šum povzroči, da se enaki geni znotraj iste celice odzivajo na drugačen način. Drugi tip notranjega šuma je *ekstrinzičen šum* (angl. *extrinsic noise*), ki je predvsem posledica razlik med različnimi celicami (npr. velikost celice, starost celice) in vpliva na vse gene znotraj iste celice na enak način.
- *Zunanji šum* (angl. *external noise*) izhaja iz okolja, v katerem se celica nahaja in je torej posledica nihanj zunanjih faktorjev, ki vplivajo na sistem (npr. temperature

ali svetlobe), oziroma je v sistem vnesen preko šuma v vhodnih signalih. Zunanji šum z velikostjo sistema narašča.

Eksperimentalno je bilo dokazano, da je šum občutno manjši pri veliki količini mRNA (angl. *high mRNA abundance*) in manjših hitrostih prepisovanja mRNA v ciljni protein, tj. pri veliki količini mRNA in manjši *učinkovitosti translacij* (angl. *translation efficiency*) kot pri visoki učinkovitosti translacij in majhni količini mRNA (angl. *low mRNA abundance*) [15, 27]. Količina mRNA je posredno določena s hitrostjo transkripcij. Povprečna vrednost količine sintetiziranega proteina je v obeh primerih enaka. Po drugi strani se šum zmanjšuje z naraščajočo velikostjo celic (glej razdelek 4.1.1), kar je posredno povezano z učinkovitostjo translacij in količino mRNA.

Količino šuma oziroma razmerje med šumom in signalom lahko z različnimi metodami zmanjšamo, po drugi strani pa šum lahko izkoristimo. Primer sistema, ki izkorišča šum, je transkripcijsko omrežje, ki na podlagi šuma proži oscilacije (angl. *noise induced oscillations*) [29].

Osnovna mera za šum je običajno podana s *standardno deviacijo*  $\sigma_N$  neke stohastične spremenljivke  $y$ . Po drugi strani lahko s pomočjo standardne deviacije definiramo *Fanojev faktor* oziroma t. i. *moč šuma* kot

$$F = \frac{\sigma_y^2}{\langle y \rangle}, \quad (2.4)$$

pri čemer  $\langle y \rangle$  predstavlja srednjo vrednost stohastične spremenljivke  $y$ . Če je šum porazdeljen po Poissonovi porazdelitvi, velja, da je Fanojev faktor enak 1 [25].

*Relativna varianca*  $\eta$  je preprostejša mera in je definirana z razmerjem med standardno deviacijo in srednjo vrednostjo odziva

$$\eta = \frac{\sigma_y}{\langle y \rangle}. \quad (2.5)$$

Moč šuma je načeloma boljši pokazatelj dejanske prisotnosti šuma v sistemu, saj obstajajo situacije, ko se le-ta poveča, relativna varianca pa ne [27]. Relativna varianca se namreč z različnimi predznaki, a enako količino šuma lahko izniči.

Uporabna mera za oceno šuma je v našem primeru tudi maksimalni šum v določenem stanju signala. V tem primeru merimo maksimalna odstopanja od povprečne vrednosti signala. Maksimalni šum v visokem logičnem stanju signala označujemo z  $N_H$ , maksimalni šum v nizkem logičnem stanju signala pa z  $N_L$ .

## 2.12 Koncentracija kemijske zvrsti

Koncentracija določa prisotnost opazovane kemijske zvrsti znotraj nekega sistema. V okviru našega dela predpostavljamo, da je prisotnost določene kemijske zvrsti homogena, torej je koncentracija v vseh segmentih sistema enaka. Obravnavanje koncentracije kemijskih zvrsti se razlikuje glede na to, ali na sistem gledamo z determinističnega ali stohastičnega vidika [30].

### 2.12.1 Koncentracija pri deterministični kinetiki

Predpostavljajmo, da je prostornina, znotraj katere potekajo kemijske reakcije in znotraj katere se nahajajo kemijske zvrsti (*reakcijska prostornina*), fiksna. Okolje je homogeno, sistem pa je v termičnem ravnovesju (temperatura sistema se ne spreminja).

V tem primeru opazujemo prisotnost posamezne kemijske zvrsti v številu elementarnih entitet (npr. molekul) na enoto prostornine oziroma molih na liter ( $\frac{\text{mol}}{L} = M$ ), pri čemer je v enem molu  $N_A = 6,0221415 \cdot 10^{23}$  elementarnih entitet, ki določajo opazovano kemijsko zvrst [31].

### 2.12.2 Koncentracija pri stohastični kinetiki

Stohastičen način opazovanja kinetike kemijskih procesov je običajno natančnejši od determinističnega, saj kemijske procese upošteva natančneje. Stohastično opazovanje temelji na evoluciji diskretnega vektorja, ki predstavlja stanje sistema v danem času. Pri tem posamezna komponenta vektorja predstavlja količino posamezne kemijske zvrsti oziroma število elementarnih entitet, ki posamezno kemijsko zvrst predstavljajo. Mera za koncentracijo je v tem primeru število molekul. Prehod iz stohastične kinetike na deterministično določa enačba

$$x_i(t) = \langle y_i(t) \rangle \cdot \Omega, \quad (2.6)$$

kjer  $x_i(t)$  določa koncentracijo kemijske zvrsti  $S_i$  na določeno prostornino ob času  $t$  ("deterministično" koncentracijo kemijske zvrsti  $i$ ),  $y_i(t)$  število molekul ob času  $t$ ,  $\langle y_i(t) \rangle$  povprečje le-teh,  $\Omega$  pa velikost sistema definirana kot

$$\Omega = V \cdot N_A, \quad (2.7)$$

s pomočjo katere pridemo iz koncentracij, ki jih uporabljamo v determinističnih modelih ( $M = \frac{\text{mol}}{L}$ ), do koncentracij, ki jih uporabljamo v stohastičnih modelih (število

molekul). Pri tem je  $V$  prostornina reakcijskega prostora (angl. *reaction volume*),  $N_A$  pa Avogadrovo število.





# 3

## Procesiranje podatkov v bioloških sistemih

Pričujoče poglavje opiše osnove procesiranja podatkov na podlagi dvojiške logike v digitalnih elektronskih vezjih. Teorijo dvojiškega procesiranja uvede na področje bioloških sistemov. Biološki sistemi imajo v tem kontekstu vlogo dvostanjske procesne platforme in tako predpostavljajo alternativo procesiranju s klasičnimi digitalnimi elektronskimi vezji. Procesiranje s tovrstnimi sistemi je prikazano na osnovnih procesnih gradnikih, zasnovanih na podlagi gensko regulatornih omrežij.

### 3.1 Osnove procesiranja podatkov v digitalnih elektronskih vezjih

Sodobni računalniški sistemi in njihove strukture večinoma temeljijo na dvojiški logiki, ki izhaja iz predstavitve podatkov v dvojiškem sistemu. Vrednosti signalov imajo torej dve veljavni stanji, in sicer so lahko v *nizkem* (logična 0) ali v *visokem logičnem stanju* (logična 1). Pri tovrstnem procesiranju podatkov izhajamo iz *Booleove algebre* in njenih postulatov [32, 33].

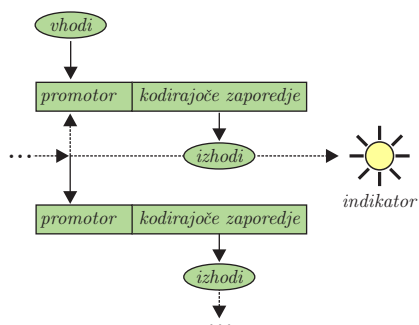
Procesiranje podatkov v dvojiški logiki temelji na t. i. *preklopnih funkcijah*, ki po-

samezni kombinaciji *vhodnih spremenljivk* (*neodvisnih spremenljivk*) priredijo eno od dveh možnih *izhodnih vrednosti odvisnih spremenljivk* na vsakem izhodu. Iz postulatov Booleove algebre sledi, da lahko s preklopnimi funkcijami, ki predstavljajo t. i. *funkcijsko poln sistem*, izrazimo poljubno preklavno funkcijo. Primer funkcijsko polnega sistema so npr. *konjunkcija*, *disjunkcija* in *negacija*. Z realizacijo osnovnih *logičnih operatorjev* oziroma *osnovnih logičnih vrat* in njihovim medsebojnim povezovanjem lahko torej pridemo do poljubnih gradnikov za procesiranje podatkov.

Do dvojiške interpretacije signala lahko v realnih sistemih pridemo z binarizacijo fizikalne količine, ki predstavlja *nosilca informacije*. Logično stanje, ki ga nivo binariziranega nosilca informacije predstavlja, je določeno s *konvencijo o pozitivni* (angl. *positive-logic convention*) oziroma *negativni logiki* (angl. *negative-logic convention*). V primeru CMOS in TTL logičnih vezij v vlogi nosilca informacije nastopa napetost, in sicer je pri pozitivni logiki logična 0 povezana z nizkim, logična 1 pa z visokim nivojem napetosti [33]. Realizacija osnovnih logičnih vrat v teh vezjih temelji na medsebojnem povezovanju elektronskih elementov, kot so npr. tranzistorji in diode. Z njihovo realizacijo, ustreznim medsebojnim povezovanjem in binarizacijo napetosti torej pridemo do sistema, ki je sposoben izračuna poljubne preklapne funkcije oziroma sistema sposobnega procesiranja podatkov - računalnika.

### 3.2 Osnove procesiranja podatkov v bioloških sistemih

Pri procesiranju podatkov v bioloških sistemih se bomo osredotočili na gensko regulatorna omrežja (glej razdelek 2.7), pri katerih lahko s spreminjanjem strukture promotorja določimo, kateri proteini bodo nastopali kot transkripcijski faktorji in kakšna bo njihova vloga, s strukturo kodirajočega zaporedja proteina pa kateri proteini bodo po aktivaciji promotorja sintetizirani. Realizacija osnovnih logičnih vrat tako temelji na medsebojnem povezovanju gensko regulatornih omrežij preko transkripcijskih faktorjev (glej sliko 3.1), pri čemer vlogo signala prevzema stopnja prisotnosti določene kemijske zvrsti (npr. proteina, molekul mRNA itd.). Stopnjo prisotnosti kemijskih zvrsti lahko določimo na podlagi njihovih *koncentracij* (glej razdelek 2.12). Opazovano fizikalno količino v gensko regulatornih sistemih zmožnih procesiranja podatkov torej predstavlja koncentracija posamezne kemijske zvrsti. Kemijske zvrsti, ki imajo vlogo nosilcev informacije, morajo biti zaradi specifične obravnavanih sistemov izbrane pri vsakem sistemu posebej.



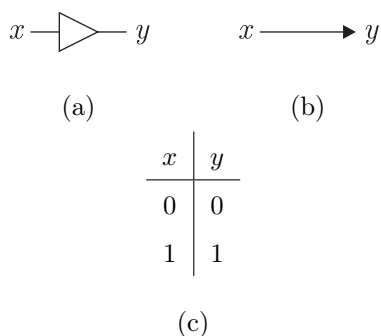
**Slika 3.1** Medsebojno povezovanje gensko regulatornih omrežij v gradnike z vnaprej določenimi funkcionalnostmi. Pri tem vlogo vhodov in izhodov iz posameznega segmenta sistema - t.j. gena (promotor s kodirajočim zaporedjem proteina), prevzemajo transkripcijski faktorji. Neprekinjene puščice predstavljajo obvezne povezave, prekinjene pa neobvezne povezave znotraj omrežja.

### 3.3 Primeri enostavnih bioloških sistemov zmožnih procesiranja podatkov

V pričujočem razdelku so opisane strukture in delovanje osnovnih bioloških sistemov zmožnih procesiranja podatkov, in sicer *biološkega gonilnika*, *biološkega negatorja*, *bioloških NOR vrat*, *preklopnega stikala* in *reprezitorja*.

#### 3.3.1 Biološki gonilnik

*Gonilnik* (angl. *driver*) je osnovni logični gradnik, ki z logičnega vidika služi ojačanju vhodnega signala. Realizira *prenosno funkcijo* (angl. *transfer function*), ki stanje vhodnega signala preslika v stanje izhodnega signala. Logični simbol in pravilnostno tabelo gonilnika prikazujeta sliki 3.2 (a) in (c).

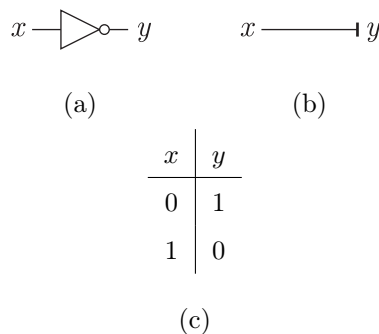


**Slika 3.2** Logična shema (a), biološka shema (b) in pravilnostna tabela (c) gonilnika.

Za realizacijo gonilnika v biološkem sistemu lahko uporabimo relativno preprosto gensko regulatorno omrežje, v katerem opazujemo dva proteina, pri čemer ima prvi protein vlogo aktivatorja gena, ki izraža drugi protein. Ob povečani koncentraciji vhodnega proteina  $x$  se sčasoma poveča tudi koncentracija izhodnega proteina  $y$ . Vhodni protein torej pospešuje (aktivira) izražanje izhodnega proteina. Zmanjšanje koncentracije vhodnega proteina privede do zmanjšanja izražanja izhodnega proteina. Zaradi razgradnje se tako sčasoma zmanjša tudi koncentracija izhodnega proteina. Shema biološkega gonilnika je prikazana na sliki 3.2 (b).

### 3.3.2 Biološki negator

*Negator* (angl. *inverter*) na izhod preslika negirano logično vrednost vhodnega signala. V dvojiški logiki aktiven vhod torej povzroči neaktivnost izhoda, neaktiven vhod pa aktivnost izhoda. Logični simbol in pravilnostno tabelo negatorja prikazujeta sliki 3.3(a) in 3.3(c).

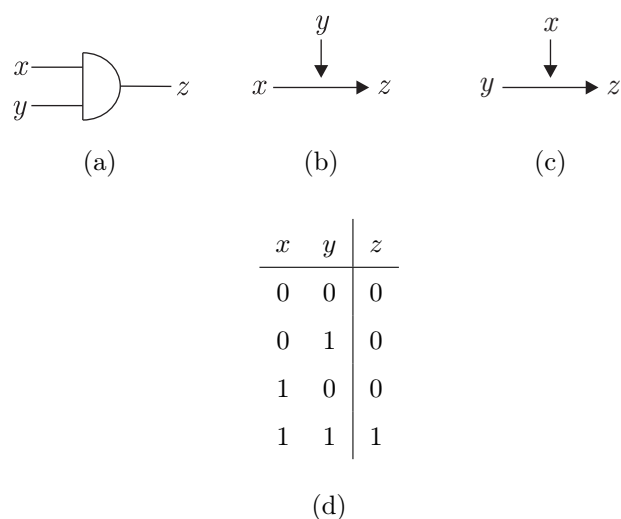


**Slika 3.3** Logična shema (a), biološka shema (b) in pravilnostna tabela (c) negatorja.

Podobno kot biološki gonilnik lahko tudi biološki negator realiziramo z genskim regulatornim omrežjem dveh proteinov, pri čemer ima prvi protein vlogo represorja gena, ki izraža drugi protein. Ob povečanju koncentracije vhodnega proteina  $x$  se koncentracija izhodnega proteina  $y$  zmanjša. Vhodni protein torej zavira (represira) izražanje izhodnega proteina. Z zmanjšanjem koncentracije vhodnega proteina se represija izhodnega proteina zmanjša. Ob odsotnosti vhodnega proteina se izhodni protein nemoteno izraža. Shema biološkega negatorja je prikazana na sliki 3.3 (b).

### 3.3.3 Biološka AND vrata

Logična *AND* vrata (angl. *AND gates*) realizirajo 2-vhodno preklopno funkcijo, pri kateri je izhod aktiven natanko takrat, ko sta aktivna oba vhodna signala (glej pravilnostno tabelo preklopne funkcije na sliki 3.4 (c)). Vrata so zanimiva zaradi dejstva, da skupaj z negacijo predstavljajo *funkcijsko poln nabor* [32, 34]. Logični simbol AND vrat prikazuje slika 3.4 (a).



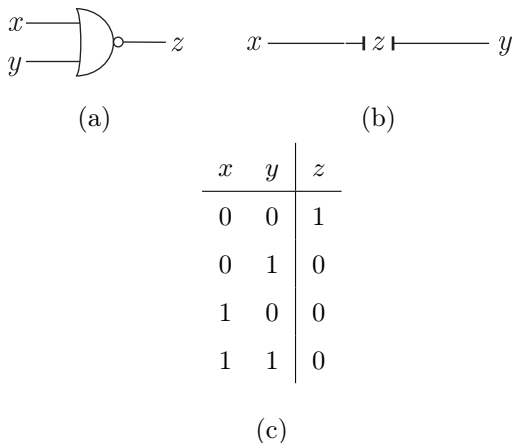
**Slika 3.4** Logična shema (a), funkcionalno ekvivalentni biološki shemi (b in c) in pravilnostna tabela (d) AND vrat, pri čemer  $x$  in  $y$  predstavljata vhoda,  $z$  pa izhod iz logičnih vrat.

Biološka AND vrata lahko realiziramo z genskim regulatornim omrežjem treh proteinov. Vhodna proteina  $x$  in  $y$  aktivirata izražanje izhodnega proteina  $z$ . Pri tem ima  $x$  vlogo aktivatorja,  $y$  pa induktorja, ki z vezavo na protein  $x$  omogoči njegovo vezavo na promotor kodirajočega zaporedja proteina  $z$ . Vlogi proteinov  $x$  in  $y$  sta pri tem lahko tudi zamenjani. Brez vezave induktorja na aktivator aktivacija promotorja ni možna. Za izražanje izhodnega proteina morata biti v sistemu prisotna oba vhodna proteina. Ob odsotnosti najmanj enega se izražanje ustavi. Shema bioloških AND vrat je prikazana na sliki 3.4 (b).

### 3.3.4 Biološka NOR vrata

Logična *NOR* vrata (angl. *NOR gates*, *negative OR*) realizirajo 2-vhodno preklopno funkcijo, pri kateri je izhod aktiven le ob neaktivnosti obeh vhodov. V vseh ostalih

primerih je izhod neaktiven (glej pravilnostno tabelo preklopne funkcije na sliki 3.5 (c)). Delovanje NOR vrat je torej enako delovanju OR vrat z negiranim izhodom. Vrata so še posebej zanimiva zaradi dejstva, da že sama po sebi predstavljajo *funkcijsko poln nabor*. Z določenim številom NOR vrat lahko torej realiziramo poljubno preklopno funkcijo. Logični simbol NOR vrat prikazuje slika 3.5 (a).



**Slika 3.5** Logična shema (a), biološka shema (b) in pravilnostna tabela (c) NOR vrat, pri čemer  $x$  in  $y$  predstavljata vhoda,  $z$  pa izhod iz logičnih vrat.

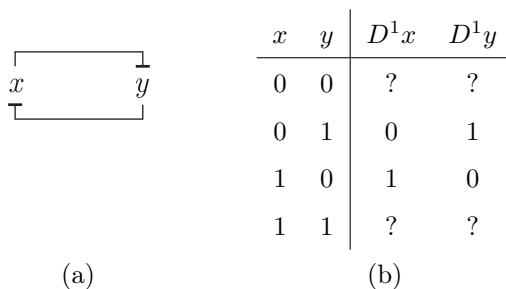
Biološka NOR vrata lahko realiziramo z genskim regulatornim omrežjem treh proteinov, pri čemer vhodna proteina nastopata v vlogi represorjev izhodnega proteina [35]. Ob odsotnosti obeh vhodnih proteinov izražanje izhodnega proteina nemoteno poteka. S povečanjem prisotnosti vsaj enega vhodnega proteina pa se izražanje izhodnega proteina zmanjša oziroma ustavi. Shema bioloških NOR vrat je prikazana na sliki 3.5 (b), pri čemer  $x$  in  $y$  predstavljata vhodna proteina,  $z$  pa izhodni protein.

### 3.3.5 Preklopno stikalo

Preklopno stikalo (angl. *toggle switch*) [10] je tipičen predstavnik bistabilnih sistemov v umetno ustvarjenih gensko regulatornih omrežjih. Bistvena značilnost bistabilnih sistemov je ta, da se večino časa nahajajo v enem od dveh stabilnih stanj. Zelo velika verjetnost je, da bo sistem po daljšem času brez spreminjanja zunanjih vplivov iz začetnega stanja prišel v eno od stabilnih stanj. Pravimo, da je *stabilno ravnovesno stanje sistema* (angl. *stable steady state*) eno od dveh stabilnih stanj, proti kateremu sistem s časom konvergira. Sistem bo torej pri idealnih pogojih po prehodu v stabilno ravnovesno

stanje v njem tudi ostal, zato lahko tak sistem uporabimo kot pomnilni element.

Bistabilno obnašanje sistema lahko dosežemo s pozitivno povratno zanko (angl. *positive feedback loop*). Osnovno različico preklopnega stikala predstavlja gensko regulatorno omrežje dveh medsebojno zavirajočih proteinov (angl. *mutual repression*). Protein, ki je v sistemu prisoten z visoko koncentracijo, torej posredno (preko represiranega proteina) aktivira samega sebe. Biološko shemo takega omrežja prikazuje slika 3.6 (a), njegovo delovanje glede na koncentracije posameznega proteina pa slika 3.6 (b).

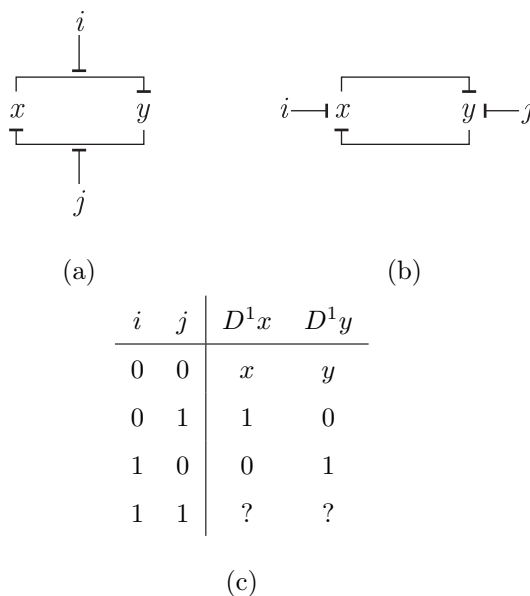


**Slika 3.6** Biološka shema (a) in pravilnostna tabela (b) preklopnega stikala, pri čemer  $x$  in  $y$  predstavljata trenutni koncentraciji opazovanih proteinov,  $D^1x$  in  $D^1y$  pa koncentraciji proteinov v naslednjem časovnem koraku. 0 določa nizko koncentracijo proteina, 1 visoko, “?” pa nedefinirano.

Analiza sistema pokaže, da ima le-ta dve stabilni ravnovesni stanji, in sicer stanje, v katerem je z visoko koncentracijo (1) prisoten protein  $x$  (aktiven je promotor, ki regulira izražanje proteina  $x$ ), z nizko (0) pa protein  $y$  in stanje, v katerem je z visoko koncentracijo (1) prisoten protein  $y$  (aktiven je promotor, ki regulira izražanje proteina  $y$ ), z nizko (0) pa protein  $x$ .

Opisani sistem je sicer sposoben pomnjenja, ne omogoča pa preklapljanja med enim in drugim stabilnim stanjem. V ta namen v sistem dodamo dva dodatna zunanja vpliva ( $i$  in  $j$ ), ki lahko vplivata na zmožnost vezave opazovanega proteina na operator drugega proteina (*induktor*) [10, 36], na hitrost izražanja proteinov ali celo na hitrost razgradnje posameznega proteina [1]. Faktorja lahko predstavljata izpostavljenost sistema določenemu svetlobnemu spektru, ki vpliva na razgradnjo posameznega proteina, spreminjanje temperature ali prisotnost določenega proteina, ki s svojo vezavo spremeni njuno strukturo ali preprosto zavira njuno izražanje. Biološko shemo omrežja z dodanima zunanjsima vplivoma v vlogi induktorjev prikazuje slika 3.7 (a), shemo omrežja z dodanima zunanjsima vplivoma, ki povečata hitrost razgradnje posameznega proteina oziroma zavirata

izražanje posameznega proteina, prikazuje slika 3.7 (b), njegovo delovanje glede na koncentracije posameznega proteina in prisotnost zunanjih vplivov pa slika 3.7 (c).



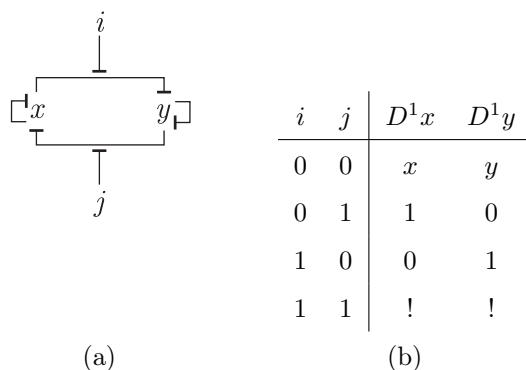
**Slika 3.7** Biološka shema preklopnega stikala z dodanima zunanjima vplivoma v vlogi induktorjev (a), z dodanima zunanjima vplivoma, ki povečata razgradnjo oziroma zavirata izražanje posameznega proteina (b) in pravilnostna tabela razširjenega preklopnega stikala z možnostjo preklapljanja med stabilnimi stanji (c). Pri tem  $x$  in  $y$  predstavljata trenutni koncentraciji opazovanih proteinov,  $i$  in  $j$  prisotnost zunanjih vplivov,  $D^1x$  in  $D^1y$  pa koncentraciji proteinov v naslednjem časovnem koraku. 0 določa nizko koncentracijo proteina, 1 visoko, “?” pa nedefinirano.

Opišimo najprej situacijo, pri kateri zunanja faktorja predstavljata t. i. induktorja. Predpostavljajmo, da se sistem nahaja v stanju z visoko koncentracijo proteina  $x$  in nizko koncentracijo proteina  $y$ . Ob vpeljavi induktorja  $i$  se le-ta veže na protein  $x$  in s tem prepreči njegovo vezavo na promotor proteina  $y$ . Represija proteina  $y$  se sčasoma ustavi, zato se lahko le-ta nemoteno izraža in njegova koncentracija naraste. Visoka koncentracija proteina  $y$  onemogoči izražanje proteinu  $x$ , ki se sčasoma razgradi. Sistem torej pristane v stanju z visoko koncentracijo proteina  $y$  in nizko koncentracijo proteina  $x$ . Stanje se ohranja tudi, ko zunanji faktor  $i$  v sistemu ni več prisoten. Do podobnega obnašanja pridemo, ko vpeljava zunanjega faktorja poveča hitrost degradacije posameznega proteina. Zopet predpostavljajmo, da se sistem nahaja v stabilnem stanju z visoko koncentracijo proteina  $x$  in nizko koncentracijo proteina  $y$ . Ob vpeljavi zunanjega faktorja  $i$  se bo koncentracija proteina  $x$  občutno zmanjšala, s čimer bo represija proteina  $y$  prekin-



jena. Posledično se bo njegovo izražanje drastično povečalo in sčasoma bo sistem pristal v stanju z visoko koncentracijo proteina  $y$  in nizko koncentracijo proteina  $x$ . Ker je to stanje stabilno, se bo ohranjalo tudi po umiku zunanjega vpliva  $i$ .

Ob odstranitvi zunanjih vplivov se protein prisoten z visoko koncentracijo izraža z maksimalno možno hitrostjo. Zgornja meja njegove koncentracije postane tako omejena zgolj s koeficientom, ki določa njegovo hitrost razgradnje. Koncentracija proteina doseže nivo, pri katerem je količina razgrajenega proteina enaka količini sintetiziranega proteina. V primeru, da želimo visoke nivoje koncentracij opazovanih proteinov dodatno omejiti in s tem zmanjšati čas za preklop iz visokega v nizko stanje, lahko preklopno stikalo razširimo z vpeljavo t. i. *samoregulacije* proteinov  $x$  in  $y$  [1, 16]. Biološko shemo razširjenega preklopnega stikala z induktorji prikazuje slika 3.8 (a), njegovo pravilnostno tabelo pa slika 3.8 (b).

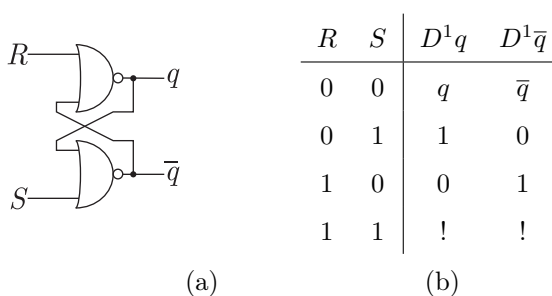


**Slika 3.8** Biološka shema (a) in pravilnostna tabela (b) razširjenega preklopnega stikala z možnostjo preklapljanja med stabilnimi stanji. Pri tem  $x$  in  $y$  predstavljata trenutni koncentraciji opazovanih proteinov,  $i$  in  $j$  prisotnost zunanjih vplivov,  $D^1x$  in  $D^1y$  pa koncentraciji proteinov v naslednjem časovnem koraku. 0 določa nizko koncentracijo proteina, 1 visoko (z upoštevanjem samoregulacije), "!" pa prepovedano kombinacijo zunanjih vplivov.

Poleg tega da posamezen protein zavira izražanje drugega proteina, zavira tudi izražanje samega sebe (*samorepresija*), pri čemer je samorepresija veliko šibkejša od represije drugega proteina, tako da pride do izraza šele pri velikih koncentracijah proteina. Opisani sistem ob nizkih koncentracijah posameznega proteina torej deluje v skladu s prejšnjima različicama, ob visokih koncentracijah pa začne aktiven protein zavirati svoje izražanje in s tem zniža zgornjo mejo svoje maksimalne koncentracije.

Delovanje preklopnega stikala je zelo podobno delovanju *RS pomnilne celice* (angl. *RS*

*flip-flop*), ki predstavlja osnovni pomnilni element v klasični elektroniki in med drugim omogoča realizacijo determinističnega končnega avtomata [34]. Podobno kot preklapno stikalo z zmožnostjo preklapljanja med stanji tudi *RS* pomnilna celica deluje pod vplivom dveh zunanjih vhodov, tj. vhoda *R*, ki notranje stanje celice deaktivira (postavi na logično 0) in *S*, ki notranje stanje celice aktivira (postavi na logično 1). Ob odsotnosti obeh vhodov se notranje stanje ohranja, hkratna prisotnost obeh pa je prepovedana. Logično shemo *RS* pomnilne celice prikazuje slika 3.9 (a), njeno pravilnostno tabelo pa 3.9(b).



**Slika 3.9** Logična shema (a) in pravilnostna tabela (b) *RS* pomnilne celice realizirane s povratno vezavo NOR (angl. *negative OR*) vrat. Pri tem *R* in *S* predstavljata zunanja vhoda v pomnilno celico, *q* in  $\bar{q}$  trenutno stanje in negirano trenutno stanje pomnilne celice,  $D^1q$  in  $D^1\bar{q}$  pa stanje in negirano stanje pomnilne celice v naslednjem časovnem koraku.

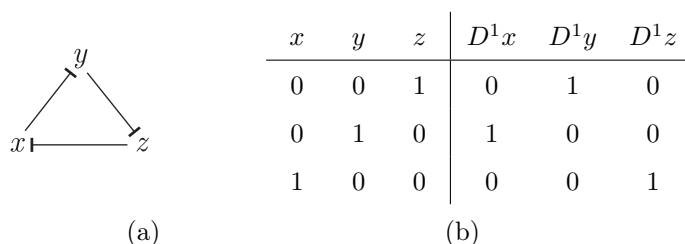
Interpretacija delovanja biološkega in logičnega gradnika je enaka v primeru, da enačimo prisotnost zunanjega vpliva *i* z aktivnostjo vhoda *R*, prisotnost zunanjega vpliva *j* z aktivnostjo vhoda *S*, prisotnost proteina *x* s stanjem pomnilne celice *q*, prisotnost proteina *y* z negiranim stanjem pomnilne celice  $\bar{q}$ , visoke koncentracije posameznega proteina z logično 1, nizke pa z logično 0.

### 3.3.6 Represilator

*Repressilator* (angl. *repressilator*) [12] je tipičen predstavnik oscilatornih sistemov v umetno ustvarjenih gensko regulatornih omrežjih. Za oscilatorne sisteme je značilno, da oscilacije vzdržujejo samostojno (brez zunanjega vzbujanja) vsaj določen čas in da se oscilacije med seboj v grobem ujema po amplitudi in frekvenci.

Oscilacije so pri represilatorju dosežene z *negativno povratno zanko* (angl. *negative feedback*). Osnovno različico sistema predstavlja gensko regulatorno omrežje treh krožno povezanih proteinov, pri čemer vsak protein zavira izražanje proteina, ki mu v sekvenci

sledi. Ker je število proteinov liho, vsak izmed njih posredno, tj. preko izražanja drugih proteinov zavira lastno izražanje. Samorepresija proteina se izvaja z zakasnitvijo, zato pride do prehajanja prisotnosti posameznega proteina iz visokih v nizke koncentracije in obratno. Sistem na ta način odraža oscilatorno vedenje, pri čemer je naenkrat z visoko koncentracijo prisoten le eden od proteinov v sekvenci. Za ustrezno delovanje je potrebno poskrbeti za inicializacijo sistema, tj. pravilno začetno stanje represilatorja (prisotnost zgolj enega od opazovanih proteinov). Biološko shemo osnovne različice represilatorja prikazuje slika 3.10 (a), njegovo pravilnostno tabelo pa 3.10(b).



**Slika 3.10** Logična shema (a) in pravilnostna tabela (b) represilatorja. Pri tem  $x$ ,  $y$  in  $z$  predstavljajo trenutne koncentracije opazovanih proteinov,  $D^1x$ ,  $D^1y$  in  $D^1z$  koncentracije proteinov v naslednjem časovnem koraku, 0 nizko koncentracijo proteina, 1 pa visoko.

### 3.4 Zaključek poglavja

Na podlagi regulacije transkripcije v gensko regulatornih omrežjih lahko vzpostavimo gensko regulatorna omrežja osnovnih gradnikov za procesiranje podatkov. To so biološki gonilnik, negator, AND vrata, NOR vrata, preklopno stikalo in represilator. V poglavju 4 predstavimo postavitev modelov teh gradnikov, analizo njihovega delovanja na podlagi vzpostavljenih metrik pa v poglavju 6. Pokažemo, da je modelirano delovanje osnovnih bioloških gradnikov ustrezno in da torej razpolagamo z elementi, ki predstavljajo funkcijsko poln nabor logičnih funkcij.



# 4 Matematični modeli za opisovanje genskih regulatornih omrežij

S povečevanjem števila kemijskih zvrsti in reakcij v bioloških sistemih njihova kompleksnost drastično narašča. Možnost napovedovanja dinamike takih sistemov brez posebnih orodij postane tako omejena zgolj na zelo preproste sisteme v smislu števila kemijskih zvrsti in števila opazovanih reakcij. Z uporabo matematičnih modelov lahko z različnimi pristopi analiziramo delovanje tudi kompleksnejših bioloških sistemov. Matematični modeli omogočajo potrditev zelenega oziroma neželenega delovanja biološkega sistema že pred samo realizacijo v celicah. Z njihovo uporabo lahko pridemo do določenih lastnosti bioloških sistemov brez eksperimentalnega dela in tako bistveno zmanjšamo stroške razvoja, poleg tega pa je iskanje potencialnih napak s pomočjo modelov bistveno lažje [16]. Matematične modele za simuliranje delovanja gensko regulatornih omrežij lahko razdelimo v tri osnovne skupine [15]:

- deterministični modeli,
- stohastični modeli,
- semikvantitativni modeli.

Deterministični modeli so relativno enostavni. Na račun tega lahko z njimi aproksimiramo le povprečen odziv celice v populaciji oziroma odziv celice po daljšem časovnem obdobju. Do modeliranja heterogenega odziva populacije enakih celic lahko pridemo s stohastičnimi modeli, ki upoštevajo stohastičnost kemijskih reakcij, njihovo reševanje pa postane veliko težje oziroma zahtevnejše za procesiranje. Kompromis med obema pristopoma je semikvantitativno modeliranje. Vsak izmed pristopov k modeliranju delovanja gensko regulatornih omrežij je podrobneje predstavljen v nadaljevanju.

Modele lahko med seboj ločimo tudi glede na stopnjo resolucije (angl. *level of resolution*), ki jo upoštevajo, oziroma glede na natančnost modeliranja kemijskih procesov. Najpreprostejši modeli naprimer opisujejo zgolj potek translacije v odvisnosti od transkripcijskih faktorjev. Kompleksnejši modeli vključujejo še vmesni korak - sintezo mRNA molekul oziroma proces transkripcije. Upoštevamo lahko tudi hitrosti vezave transkripcijskih faktorjev, različne hitrosti degradacije itd.

#### 4.1 Deterministični modeli

Deterministični modeli se uporabljajo za aproksimiranje povprečnega izražanja genov v populaciji identičnih celic oziroma v eni sami celici po daljšem časovnem obdobju. Večinoma temeljijo na *navadnih diferencialnih enačbah* (angl. *Ordinary Differential Equations - ODE*), ki izhajajo iz *kinetike zakona o vplivu koncentracij* (angl. *mass action kinetics*) ali na fenomenoloških predstavitev sistema, ki temeljijo na empiričnih opisih. Deterministični modeli predpostavljajo, da opazujemo sistem pri konstantni temperaturi. Ponavadi so podani z nizom diferencialnih enačb, ki določajo spremembo koncentracij posameznih kemijskih zvrsti po času z enačbo

$$\frac{dx_i}{dt} = f_i(\mathbf{x}(t)); 1 \leq i \leq N. \quad (4.1)$$

Pri tem je funkcija  $f_i$  odvisna od trenutnih koncentracij kemijskih zvrsti, določena na podlagi opazovanih kemijskih reakcij,  $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$  vektor nenegativnih realnih vrednosti, ki predstavljajo koncentracije posameznih kemijskih zvrsti v času  $t$ ,  $N$  pa število opazovanih kemijskih zvrsti [17].

Izgradnja determinističnih modelov lahko poteka na dva različna načina, in sicer

- izgradnja na podlagi *fenomenoloških dognanj*, ki jo lahko opredelimo kot *pristop od zgoraj navzdol* (angl. *top-down approach*),

- izgradnja na podlagi različnih procesov, ugotovljenih s pomočjo matematičnih modelov na nižjih nivojih (zakon o vplivu koncentracij, Michaelis-Menten kinetika ipd.), ki jo lahko opredelimo kot *pristop od spodaj navzgor* (angl. *bottom-up approach*).

V našem primeru je izgradnja modelov zanimiva z vidika od spodaj navzgor, saj želimo modelirati biološki sistem, ki naj bi opravljal določeno funkcijo in ga v primeru obetavnih rezultatov realizirati v živih celicah. V nadaljevanju so razloženi pristopi pri determinističnem modeliranju, ki so tudi prikazani na nekaj praktičnih primerih.

#### 4.1.1 Primerjava determinističnih in stohastičnih modelov

Deterministični modeli omogočajo natančen prikaz povprečnega odziva populacije celic, niso pa zmožni modeliranja obnašanja posameznih celic znotraj populacije, katerih odziv lahko bistveno odstopa od povprečja (z determinističnimi modeli naprimer ne moremo modelirati vpliva šuma). Ti modeli so torej še posebej neprimerni za sisteme, ki odražajo bimodalno distribucijo odziva celic (npr. bistabilen sistem, pri katerem je v določenem trenutku polovica populacije v prvem, druga polovica pa v drugem ravnovesnem stabilnem stanju). Če želimo dobiti odziv posamezne celice in s tem prikaz heterogenosti odziva populacije, moramo torej sistem opisati s stohastičnimi modeli.

Pri določenih pogojih je odziv stohastičnih in determinističnih modelov enak oziroma zelo podoben. Do takih odzivov pridemo naprimer pri obravnavanju *velikih sistemov* (velika količina izraženih mRNA molekul in proteinov, velika prostornina celic) ali sistemov s hitrim izražanjem genov, tj. s hitro kinetiko promotorjev [27]. Stohastični modeli tako odražajo relativno majhne deviacije (odstopanja odziva od povprečja) oziroma relativno majhne količine šuma. Šum se torej zmanjšuje z velikostjo sistema - pojavu pravimo *efekt končnega števila* (angl. *finite number effect*), ko manjše koncentracije opazovanih kemijskih zvrsti v sistemu povzročajo večji šum [28].

#### 4.1.2 Zakon o vplivu koncentracij

Zakon o vplivu koncentracij (angl. *law of mass action*) oziroma *ravnotežno enačbo* (angl. *equilibrium equation*) lahko razložimo z dvema vidikoma [31]:

- *ravnotežni vidik* temelji na stanju raztopin v *ravnotežju* (tj. ravnovesno stanje sistema, ki je doseženo po določenem času in se ne spreminja več oziroma se spreminja minimalno),

- *kinetični vidik* temelji na enačbah hitrosti elementarnih reakcij (glej razdelek 4.1.3).

Zakon o vplivu koncentracij izhaja iz dejstva, da se masa sistema kot celote s časom ne spreminja - *zakon o ohranitvi mase* (angl. *law of conservation of mass*). Ob konstantnih vplivih iz okolice sistem slej ko prej preide v ravnotežno stanje, v katerem se koncentracije kemijskih zvrsti ne spreminjajo več. Izračunamo lahko ravnotežno konstanto  $K$ , ki določa razmerje med ravnotežnim stanjem oziroma koncentracijo reaktantov in produktov.

Kot zgled predpostavimo, da imamo na razpolago *reaktanta*  $A$  in  $B$  in *produkta*  $C$  in  $D$ , reakcija pa lahko poteka v obe smeri. Sistem lahko tako opišemo kot



kjer so  $a$ ,  $b$ ,  $c$  in  $d$  *stehiometrični koeficienti* uravnotežene enačbe (razberemo jih lahko iz ravnotežne reakcije (angl. *equilibrium reaction*) ali ugotovimo eksperimentalno) in določajo kvantitativno povezavo med reaktanti in produkti.

Zakon o vplivu koncentracij trdi, da bo razmerje med produkti in reaktanti v ravnotežnem stanju vedno enako

$$K = \frac{[C]^c [D]^d}{[A]^a [B]^b}. \quad (4.3)$$

Pri tem  $[A]$ ,  $[B]$ ,  $[C]$  in  $[D]$  predstavljajo koncentracije *kemijskih zvrsti* (angl. *chemical species*),  $K$  pa ravnotežno konstanto (angl. *equilibrium constant*). Slednjo lahko zapišemo tudi v obliki

$$K = \frac{k_f}{k_b}, \quad (4.4)$$

kjer  $k_f$  predstavlja *hitrostno konstanto vodilne reakcije* (angl. *rate constant for forward reaction*),  $k_b$  pa *hitrostno konstantno povratne reakcije* (angl. *rate constant for backward reaction*). Definiramo lahko tudi *disociacijsko konstanto* (angl. *dissociation constant*), ki je definirana z izrazom

$$K_d = \frac{1}{K} \quad (4.5)$$

in predstavlja moč vezi produkta.

### 4.1.3 Enačba hitrosti reakcije

Enačbe hitrosti reakcije (angl. *rate equations, rate law*) povezujejo hitrost reakcije s koncentracijami reaktantov in konstantami [31]. Določimo jih z združitvijo hitrosti reakcij in masnim ravnotežjem sistema. Predpostavljamo namreč, da se masa sistema s časom



ne spreminja. Kot primer vzemimo generično reakcijo



kjer  $A$  in  $B$  predstavljata reaktanta,  $C$  produkt,  $k$  pa hitrostni koeficient (angl. *rate coefficient*) reakcije. Potem lahko definiramo hitrost reakcije (angl. *rate*) kot

$$r = k[A]^m[B]^n, \quad (4.7)$$

kjer  $[A]$  in  $[B]$  predstavljata koncentraciji reaktantov,  $m$  in  $n$  pa stehiometrična koeficienta uravnotežene enačbe. Na podoben način lahko zapišemo diferencialno enačbo, ki določa spremembo koncentracije produkta po času v odvisnosti od koncentracij reaktantov

$$\frac{dC}{dt} = k[A]^m[B]^n. \quad (4.8)$$

Pri tem so enote hitrostnega koeficienta odvisne od tipa reakcije, in sicer

- $\text{Ms}^{-1}$  za reakcije 0-tega reda ( $m + n = 0$ ;  $\emptyset \rightleftharpoons C$ ),
- $\text{s}^{-1}$  za monomolekularne reakcije - reakcije prvega reda ( $m + n = 1$ ),
- $\text{M}^{-1}\text{s}^{-1}$  za bimolekularne reakcije - reakcije drugega reda ( $m + n = 2$ ),
- ...
- $\text{M}^{-(m+n-1)}\text{s}^{-1}$  za  $m + n$ -molekularne reakcije (reakcije  $m + n$ -tega reda),

kjer  $M$  predstavlja koncentracijo podano v molih na liter ( $\frac{\text{mol}}{\text{L}}$ ).

Opozoriti je potrebno na to, da enačba hitrosti reakcije velja samo za elementarne reakcije. Enačba hitrosti reakcije za mehanizem z več koraki je lahko določena zgolj eksperimentalno, saj smo pri izpeljavi enačbe predpostavljali, da se skupna količina kemijskih zvrsti, ki nastopajo v opazovani reakciji, v sistemu ne spreminja. Kljub temu zelo pogosto enačbe hitrosti reakcije uporabljamo pri modeliranju gensko regulatornih omrežjih z neelementarnimi reakcijami [18].

Na osnovi enačbe hitrosti reakcije lahko postavimo relativno natančen sistem diferencialnih enačb, ki ustreza enačbi (4.1) in je relativno natančen, saj upošteva vse predvidene kemijske reakcije. Po drugi strani je njegova kompleksnost v primerjavi z modeli, ki temeljijo na Hillovih enačbah, velika, zato če je le upravičeno, uporabljamo le-te (glej razdelek 4.1.5).

#### 4.1.4 Michaelis-Menten kinetika

Michaelis-Menten kinetika opisuje encimsko kinetiko in se uporablja pri preprostih encimskih reakcijah (mehanizem z enim korakom) [37]. Encimsko reakcijo lahko zapišemo kot



pri čemer  $E$  predstavlja encim,  $S$  substrat,  $ES$  njuno vezavo,  $P$  pa produkt reakcije. Pri tem lahko verjetnost za prehod oziroma hitrostni koeficient iz stanja  $E + S$  v  $ES$  označimo s  $k_f$ , iz  $ES$  v  $E + S$  s  $k_b$ , iz  $ES$  v  $E + P$  pa s  $k$ .

Definiramo lahko spremembo koncentracije kompleksa  $ES$  v časovnem koraku  $dt$  kot

$$\frac{d[ES]}{dt} = k_f[E][S] - k_b[ES] - k[ES]. \quad (4.10)$$

V stabilnem stanju tako pridemo do koncentracije kompleksa  $[ES]$ , ki je enaka

$$[ES] = \frac{k_f}{k_b + k}[E][S]. \quad (4.11)$$

Potem lahko hitrost reakcije izračunamo kot

$$r = k[ES] = \frac{k[E][S]}{K_d + [S]}, \quad (4.12)$$

kjer  $K_d$  predstavlja Michaelisovo konstanto

$$K_d = \frac{k_b + k}{k_f}. \quad (4.13)$$

Ko je koncentracija substrata zelo velika (*v nasičenju*), poteka reakcija pri maksimalni hitrosti, ki je neodvisna od količine substrata (upoštevamo substrat z eksponentom 0 - *kinetika 0-tega reda*) po izrazu

$$r_{\max} = k[E]. \quad (4.14)$$

Če je substrat prisoten v majhnih količinah, je v primerjavi z velikostnim redom konstante  $K_d$  zanemarljiv in lahko enačbo hitrosti zapišemo kot

$$r = \frac{k[E][S]}{K_d}. \quad (4.15)$$

Substrat upoštevamo z eksponentom 1 - *reakcija prvega reda*.

### 4.1.5 Hillove enačbe

Hillove enačbe omogočajo deterministično modeliranje koncentracije specifičnih kemijskih zvrsti v odvisnosti od transkripcijskih faktorjev [16]. Modeliranje na osnovi Hillovih enačb je torej še posebej uporabno pri modeliranju gensko regulatornih omrežij. Časovni potek koncentracije kemijskih zvrsti je aproksimiran z monotonimi krivuljami *sigmoidne oblike* (angl. *sigmoid curves*). Glavna prednost uporabe Hillovih enačb je ta, da sistem enačb, ki predstavlja model biološkega sistema, postane enostavnejši, njegovo reševanje pa hitrejše od sistema, ki temelji na osnovnih enačbah hitrosti reakcij (glej razdelek 4.1.3).

V nadaljevanju je na primeru represibilnega promotorja razložena izpeljava Hillovih enačb, ki izhajajo iz zakona o vplivu koncentracij, enačb hitrosti reakcije in Michaelis-Menten kinetike [16].

#### Modeliranje vezave represorja na promotor

Vezavo represorja na promotor lahko opišemo z upoštevanjem kinetike zakona o vplivu koncentracij. Predpostavljamo, da represor  $R$  prepozna določen del DNA  $D$ , ki se nahaja na promotorju.  $R$  in  $D$  z vezavo formirata kompleks  $RD$ . Promotorski del  $D$  je torej lahko prost ali vezan z represorjem, pri čemer transkripcija poteka le v primeru, ko je prost. Predpostavljamo, da je skupna količina deleža promotorja, na katerega se lahko veže represor, enaka  $D_T$ . Velja torej ravnotežna enačba

$$[D_T] = [D] + [RD]. \quad (4.16)$$

Pogostost vezave med represorjem  $R$  in operatorskim delom  $D$  lahko opišemo z enačbo hitrosti reakcije

$$r = k_f[R][D], \quad (4.17)$$

kjer  $k_f$  predstavlja verjetnost vezave med  $R$  in  $D$  oziroma hitrostni koeficient (glej enačbo (4.7)). Podobno lahko definiramo verjetnost razpada kompleksa  $RD$ , ki je podana s koeficientom  $k_b$  in je odvisna od moči vezi. Spremembo koncentracije kompleksa  $RD$  v časovnem koraku  $dt$  lahko torej opišemo z enačbo

$$\frac{[RD]}{dt} = k_f[R][D] - k_b[RD]. \quad (4.18)$$

V stabilnem stanju torej pridemo do ravnotežja, ki ga lahko opišemo z enačbo

$$K_d = \frac{[R][D]}{[RD]} = \frac{k_b}{k_f}, \quad (4.19)$$

pri čemer  $K_d$  predstavlja disociacijsko konstanto (glej enačbe (4.3), (4.4) in (4.5)). Verjetnost prostega promotorja torej izrazimo kot

$$\frac{[D]}{[D_T]} = \frac{1}{1 + \frac{[R]}{K_d}} \quad (4.20)$$

in se zmanjšuje z večanjem koncentracije represorja. Verjetnost prostega promotorja lahko intepretiramo kot verjetnost transkripcije proteina. *Maksimalna hitrost transkripcije* (angl. *maximal transcription rate*) je odvisna od samega DNA zaporedja, pozicije vezavnega mesta RNA polimeraze in ostalih faktorjev ter je določena s koeficientom  $\alpha$ . Hitrost produkcije mRNA imenujemo tudi *aktivnost promotorja* (angl. *promoter activity*) in je definirana na podlagi verjetnosti prostega promotorja in maksimalni aktivnosti promotorja kot

$$\frac{\alpha}{1 + \frac{[R]}{K_d}}. \quad (4.21)$$

### Modeliranje vezave induktorja na represor

Ob vezavi represorja na promotor se transkripcija proteina ustavi. V prisotnosti induktorja ( $E$ ) se struktura represorja spremeni, tako da se verjetnost njegove vezave na operator bistveno zmanjša.

Vezavo induktorja na represor lahko opišemo s pomočjo Michaelis-Menten kinetike (glej razdelek 4.1.4), pri čemer induktor predstavlja encim, represor pa substrat. Razlika med modelom vezave induktorja na represor in predstavljeno Michaelis-Menten kinetiko je v tem, da po vezavi induktorja in represorja nimamo produkta, torej imamo možna le dva prehoda, in sicer iz  $E$ ,  $R$  v  $ER$  in obratno. Zapišemo lahko kinetiko zakona o vplivu koncentracij kot

$$\frac{d[ER]}{dt} = k_f[R][E] - k_b[ER]. \quad (4.22)$$

Koncentracija kompleksa *represor-induktor* v ravnotežnem stanju je torej

$$[ER] = \frac{[R_T][E]}{[E] + K_d}, \quad (4.23)$$

kjer  $R_T$  predstavlja skupno količino represorja (predpostavljamo, da se spreminja veliko počasneje kot proces same vezave),  $K_d$  pa disociacijsko konstanto. Funkcija ima tri

prepoznavne lastnosti, ki jih lahko opazimo tudi pri realnem sistemu represor-induktor, in sicer

- pri veliki koncentraciji induktorja  $E$  sistem preide v nasičenje,
- če velja  $[E] \ll K_d$ , koncentracija kompleksa  $[ER]$  narašča linearno s koncentracijo induktorja  $E$ ,
- v primeru, da je  $[E] = K_d$ , je delež vezanega represorja z induktorjem 50%.

Na podlagi tega lahko definiramo prosti del represorja  $R$  kot

$$[R^*] = \frac{[R_T]}{1 + \frac{[E]}{K_d}}. \quad (4.24)$$

### Modeliranje kooperativnosti induktorja

V realnosti imamo lahko opravka z represorjem, pri katerem se morajo transkripcijski faktorji vezati na več različnih vezavnih mest. V tem primeru je funkcionalnost induktorja popolna samo v primeru, da je na represor vezanih več enakih proteinskih molekul naenkrat. Predpostavljamo, da potrebujemo vezanih  $\gamma$  molekul induktorja  $E$  in da je vezava na represor  $S$  popolna. Pri tem število  $\gamma$  imenujemo *Hillov koeficient* oziroma *koeficient kooperativnosti* (angl. *cooperativity coefficient*). Torej lahko zapišemo kinetiko zakona o vplivu koncentracij z izrazom

$$\frac{d[\gamma ER]}{dt} = k_f[R][E]^\gamma - k_b[\gamma ER]. \quad (4.25)$$

Stanja, kjer je vezanih manj kot  $\gamma$  molekul na represor, v tem modelu zanemarjamo, saj je njihova verjetnost občutno manjša. Če predpostavljamo, da je skupna količina represorja konstantna, lahko zapišemo

$$R_T = [R] + \gamma[ER] \quad (4.26)$$

in tako pridemo do enačbe v stabilnem stanju

$$\frac{[\gamma ER]}{R_T} = \frac{[E]^\gamma}{K_d^\gamma + [E]^\gamma}, \quad (4.27)$$

kjer je  $K_d^\gamma$  enak

$$K_d^\gamma = \frac{k_b}{k_f}. \quad (4.28)$$

Opazimo lahko, da pri  $\gamma = 1$  pridemo do Michaelis-Menten modela. Z večanjem števila  $\gamma$  se povečuje strmina krivulje, ko sistem prehaja v nasičenje. S pomočjo tega matematičnega modela lahko pridemo do deleža prostega represorja, ki je enak

$$\frac{[R^*]}{[R_T]} = \frac{1}{1 + \left(\frac{[E]}{K_d}\right)^\gamma}. \quad (4.29)$$

Opozoriti je potrebno na to, da v primeru, da imamo eno vezavno mesto, na katerega se veže samo en transkripcijski faktor, ki je sestavljen iz več podenot (npr. dimer, tetramer ...), Hillove enačbe predpostavljajo, da reakcija vezave induktorja na represor poteče istočasno z vezavo obravnavanih podenot v popoln induktor in so zato v takem primeru manj natančne [38].

### Modeliranje vezave represorja na promotor ob odsotnosti induktorja

Z združitvijo definiranih enačb lahko pridemo do enačbe, ki določa hitrost transkripcije

$$f([E]) = \frac{\alpha}{1 + \frac{[R^*]}{K_d}} = \frac{\alpha}{1 + \frac{[R_T]}{K_d(1 + \left(\frac{[E]}{K_d}\right)^\gamma)}}, \quad (4.30)$$

kjer  $\alpha$  določa maksimalno hitrost transkripcije.

### Deterministično modeliranje translacije proteina na osnovi Hillovih enačb

#### Hilova enačba aktivatorja

Koncentracija proteina, katerega translacija poteka le ob prisotnosti določenega transkripcijskega faktorja - aktivatorja, je ob odsotnosti le-tega ničelna oziroma minimalna - vedno imamo določeno stopnjo puščanja, ki je določen z nivojem osnovnega izražanja gena (angl. *basal expression*). S prisotnostjo aktivatorja se hitrost izražanja proteina poveča, zato začne njegova koncentracija rasti in sčasoma doseže maksimalno, nasičeno vrednost. Spremembo koncentracije proteina  $P$  v času  $dt$  lahko tako opišemo z diferencialno enačbo

$$\frac{d[P]}{dt} = \frac{\alpha[A]^\gamma}{K_d^\gamma + [A]^\gamma} + \alpha_0 - \delta[P], \quad (4.31)$$

kjer  $[A]$  predstavlja koncentracijo aktivatorja,  $K_d$  koeficient aktivacije  $\left(K_d = \left(\frac{k_b}{k_f}\right)^{\frac{1}{\gamma}}\right)$ ,  $\alpha$  maksimalno izražanje,  $\gamma$  Hillov koeficient,  $\alpha_0$  osnovno izražanje gena,  $\delta$  pa koeficient razgradnje proteina. Koeficient aktivacije določa koncentracijo aktivatorja  $X$ , pri kateri je dosežena polovica maksimalnega izražanja, koeficient kooperativnosti določa število

molekul aktivatorja, ki se morajo naenkrat vezati na promotorski del za začetek transkripcije in s tem posledično strmino krivulje, koeficient osnovnega izražanja gena količino izraženega proteina v odsotnosti aktivatorja, koeficient razgradnje pa delež proteina, ki se v časovnem koraku  $dt$  razgradi.

### *Hilova enačba represorja*

Koncentracija proteina, katerega translacija poteka v odsotnosti določenega transkripcijskega faktorja - represorja, je ob odsotnosti le-tega maksimalna. S prisotnostjo začne njegova koncentracija padati in po določenem času doseže minimalno vrednost. Spremembo koncentracije proteina  $P$  v času  $dt$  lahko tako opišemo z diferencialno enačbo

$$\frac{dP}{dt} = \frac{\alpha}{1 + \left(\frac{R}{K_d}\right)^\gamma} + \alpha_0 - \delta P, \quad (4.32)$$

kjer  $R$  predstavlja koncentracijo represorja. Ostali koeficienti imajo podobno vlogo kot njihovi ekvivalenti opisani v prejšnjem zgledu.

### *Hilove enačbe kompleksnejših gradnikov*

Hilovi enačbi aktivatorja in represorja lahko tako kot tudi funkcije opisane v prejšnjih razdelkih z medsebojnim povezovanjem razširimo v modele kompleksnejših gradnikov. Na ta način pridemo do relativno preprostih modelov kompleksnejših bioloških gradnikov.

### *Slabosti modeliranja na osnovi Hilovih enačb*

Poleg osnovnih slabosti determinističnih modelov, opisanih v razdelku 4.1.1, je glavna slabost Hilovih enačb ta, da v reakcijah, pri katerih se na promotorsko mesto veže transkripcijski faktor sestavljen iz več elementarnih enot (npr. protein v dimerizirani obliki), ne vključujejo vmesnih stanj [38]. V primeru, da kot transkripcijski faktor nastopa opazovani protein v dimerizirani obliki, Hilove enačbe ne upoštevajo reakcij dimerizacije in disociacije dimera, ampak predpostavljajo, da reakcija vezave proteina na promotor poteče sočasno z dimerizacijo.

## **4.2 Stohastični modeli**

Prednost stohastičnih modelov pred determinističnimi je v tem, da omogočajo modeliranje heterogenosti populacije, ki ne izvira iz različnosti celic (vse celice v populaciji so namreč lahko identične), temveč je posledica *stohastičnosti genskega izražanja* in *suma*

pri samem izražanju (glej razdelek 2.11). Ponavadi želimo *heterogenost populacije* minimizirati. Rezultate stohastičnih modelov lahko tako uporabljamo za identifikacijo heterogenosti in posledično njeno minimiziranje. Preprosta različica stohastičnih modelov vključuje *stohastičnost kemijskih reakcij*, ki pride do izraza še posebej pri modeliranju obnašanja posamezne celice. Na heterogenost populacije pri gensko regulatornih omrežjih vplivajo tudi ostali faktorji, kot je naprimer celična delitev (angl. *cell division*), nihanja v količini RNA polimeraze, ribosomov ... Načeloma na te faktorje težko vplivamo, lahko pa vplivamo na faktorje, kot so hitrosti transkripcije in translacije posameznih vrst molekul in hitrosti degradacije. Tako lahko sistem na določenih parametrih *uglašujemo* (angl. *tune*) in s tem pridobimo obnašanje v skladu z želenim.

Pri stohastičnih modelih operiramo s stohastičnostjo specifičnih kemijskih procesov (npr. transkripcije). Iskanje rešitev teh modelov poteka numerično ali s pomočjo pristopa z glavno enačbo (angl. *master equation approach*) [17].

#### 4.2.1 Opis stohastičnosti kinetike kemijskih procesov

Velika večina stohastičnih modelov, ki se uporabljajo za modeliranje gensko regulatornih omrežij, se poleg upoštevanja določene stopnje naključnosti kemijskih reakcij od opisanih determinističnih modelov razlikujejo tudi po tem, da modeliran biološki sistem obravnavajo kot diskreten sistem, pri katerem so molekularna števila posameznih kemijskih zvrsti cela števila [17, 39]. Predpostavljajmo, da opazujemo  $N$  različnih kemijskih zvrsti  $\{S_1, \dots, S_N\}$ . Stanje sistema ob času  $t$  lahko torej opišemo z vektorjem  $\mathbf{y}(t) = (y_1(t), \dots, y_N(t))^T$ , kjer komponenta  $y_i(t)$  podaja količino kemijske zvrsti  $S_i$  v času  $t$ . Poleg tega predpostavljajmo, da imamo v sistemu  $M$  različnih reakcij  $R_j$  ( $j \in \{1, \dots, M\}$ ), ki spreminjajo stanje sistema. Pri tem reakcija  $R_j$  spremeni stanje sistema na sledeč način

$$\mathbf{y}(t + dt) \longrightarrow \mathbf{y}(t) + \nu_j, \quad (4.33)$$

kjer  $dt$  predstavlja časovni korak,  $\nu_j$  ( $j \in \{1, \dots, M\}$ ) pa  $N$ -dimenzionalni vektor, pri čemer njegove komponente vsebujejo stehiometrične koeficiente reakcije - spremembo določene zvrsti po reakciji  $R_j$ . S pomočjo vektorjev  $\nu_j$  lahko definiramo stehiometrično matriko  $S = [\nu_1, \dots, \nu_M]$ .

Ali se bo posamezna reakcija izvršila v naslednjem časovnem koraku, je odvisno od količine kemijskih zvrsti, ki nastopajo kot reaktanti te reakcije in od njene hitrosti izvedbe. Verjetnost, da se bo v časovnem koraku, ki traja  $dt$  enot, sprožila reakcija  $R_j$ ,



lahko izračunamo posredno preko funkcije  $a_j(\mathbf{y}(t))dt$  (angl. *propensity function*). Vsota teh funkcij je podana z  $a_0(\mathbf{y}(t))dt$ .

Če je  $R_j$  monomolekularna reakcija tipa  $S_i \rightarrow \text{produkti}$ , obstaja konstanta  $c_j$ , tako da velja

$$a_j(\mathbf{y}(t)) = c_j y_i(t), \quad (4.34)$$

pri čemer je  $c_j$  numerično enak konstanti, ki določa hitrost reakcije -  $k_j$  pri deterministični kinetiki

$$c_j = k_j. \quad (4.35)$$

Velja torej izraz

$$a_j(\mathbf{y}(t)) = k_j y_i(t). \quad (4.36)$$

Če je  $R_j$  bimolekularna reakcija tipa  $S_{i_1} + S_{i_2} \rightarrow \text{produkti}$ , obstaja konstanta  $c_j$ , tako da velja

$$a_j(\mathbf{y}(t)) = \begin{cases} c_j \cdot y_{i_1}(t) \cdot y_{i_2}(t); i_1 \neq i_2 \\ \frac{1}{2} c_j \cdot y_{i_1}(t) \cdot (y_{i_2}(t) - 1); i_1 = i_2 \end{cases}. \quad (4.37)$$

V prvem primeru je  $c_j$  numerično enak  $k_j/\Omega$  pri deterministični kinetiki, v drugem pa  $2k_j/\Omega$  pri deterministični kinetiki, pri čemer je  $\Omega$  velikost sistema [40]

$$c_j = \begin{cases} \frac{k_j}{\Omega}; i_1 \neq i_2 \\ \frac{2k_j}{\Omega}; i_1 = i_2 \end{cases}. \quad (4.38)$$

Velja torej

$$a_j(\mathbf{y}(t)) = \begin{cases} \frac{k_j}{\Omega} \cdot y_{i_1}(t) \cdot y_{i_2}(t); i_1 \neq i_2 \\ \frac{k_j}{\Omega} \cdot y_{i_1}(t) \cdot (y_{i_2}(t) - 1); i_1 = i_2 \end{cases}. \quad (4.39)$$

Pri bimolekularnih reakcijah z upoštevanjem prostornine upoštevamo verjetnost, da se bosta molekuli potrebni za reakcijo srečali v prostoru. Pri monomolekularnih reakcijah nam zato prostornine pri izračunu ni potrebno upoštevati.

Po [41] v splošnem velja

$$a_j(\mathbf{y}(t)) = c_j \prod_{i=1}^{n_1} \binom{y_i(t)}{l_i}, \quad (4.40)$$

kjer je  $n_1$  število različnih kemijskih zvrsti, ki nastopajo v reakciji  $R_j$  kot reaktanti,  $y_i(t)$  število molekul, ki pripadajo kemijskih zvrsti  $i$  ob času  $t$ ,  $l_i$  število porabljenih molekul ob izvedbi reakcije, ki pripadajo kemijski zvrsti  $i$ ,  $c_j$  pa konstanta, ki jo lahko določimo preko konstante, ki določa hitrost reakcije  $R_j$  pri deterministični kinetiki

$$c_j = \frac{k_j \prod_{i=1}^{n_1} l_i!}{\Omega^{n_2-1}}, \quad (4.41)$$

kjer je  $n_2$  število molekul, ki se porabijo pri izvedbi reakcije  $R_j$ . Torej je

$$a_j(\mathbf{y}(t)) = \frac{k_j}{\Omega^{n_2-1}} \prod_{i=1}^{n_1} \frac{y_i(t)!}{(y_i(t) - l_i)!}. \quad (4.42)$$

Ugotovimo lahko, da so enote za  $a_j$  ne glede na tip reakcije vedno podane v  $\frac{mol}{s}$ .

### 4.2.2 Pristop z glavno enačbo

Pristop z glavno enačbo (angl. *master equation approach*) je analitičen pristop, ki služi reševanju stohastičnih modelov - natančneje iščemo verjetnost, da se ob času  $t$  sistem nahaja v določenem stanju  $\mathbf{y}(t)$ , pri čemer lahko postavimo glavno enačbo (angl. *Chemical Master Equation - CME*), ki določa spremembo vektorja  $\mathbf{y}(t)$  v odvisnosti od časa

$$\dot{p}(\mathbf{y}(t)) = -p(\mathbf{y}(t)) \sum_{j=1}^M a_j(\mathbf{y}(t)) + \sum_{j=1}^M p(\mathbf{y}(t) - \nu_j; t) a_j(\mathbf{y}(t) - \nu_j). \quad (4.43)$$

Glavna enačba predstavlja sistem navadnih diferencialnih enačb, s pomočjo katerih lahko na spreminjanje stanja sistema (evolucijo sistema) gledamo kot na Markovski proces (angl. *Markov process*), tj. stanje sistema v naslednjem časovnem koraku je odvisno le od trenutnega stanja sistema [40, 41].

Analitično reševanje enačbe nas tako sicer pripelje do eksaktne rešitve verjetnosti določenega stanja pri predpostavki, da je sistem ob času  $t_0$  v začetnem stanju  $\mathbf{y}(t_0)$ . Analitično reševanje je zaradi računske kompleksnosti možno le pri zelo preprostih sistemih. Problem reševanja glavne enačbe lahko hitro postane neobvladljiv zaradi potencialno velikega prostora stanj. Slednjega lahko do določene mere omejimo z upoštevanjem *modularnosti* (izoliramo manjše module) in *specifičnosti* (določen protein se veže le na majhno število drugih molekul).

### 4.2.3 Gillespiejev algoritem za stohastično simuliranje

Gillespiejev algoritem za stohastično simuliranje (angl. *Stochastic simulation algorithm - SSA*) temelji na Monte Carlo simulaciji in služi numeričnemu reševanju stohastičnih modelov [42, 43]. S pomočjo algoritma pridemo do ene od možnih trajektorij vektorja  $\mathbf{y}(t)$  skozi čas. Algoritem je določen s sledečimi koraki:

- *inicijalizacija*: določimo parametre in začetno stanje sistema,
- *Monte Carlo korak*: naključno določimo naslednjo reakcijo in časovni korak (glej enačbi (4.45) in (4.46)),

- *posodobitev*: povečamo časovni korak in posodobimo število molekul glede na rezultate Monte Carlo koraka, zatem se vrnemo na Monte Carlo korak.

Simulacijo je smiselno izvajati toliko časa, dokler sistem ne preide v ravnovesno stanje. Ponavadi se upošteva aproksimacija, da je ravnovesno stanje doseženo v času, ki je enak 10-kratniku razpolovne dobe proteina, katerega pretvorba poteka najpočasneje [25].

Reakcijo in dolžino časovnega koraka izberemo s pomočjo reakcijske gostote verjetnosti (angl. *reaction probability density function*)

$$P_{\mathbf{y}(t)}(\tau, j) = a_j(\mathbf{y}(t))e^{-\sum_{j=1}^M a_j(\mathbf{y}(t))\tau}, \quad (4.44)$$

kjer  $\tau$  predstavlja trajanje časovnega koraka ( $dt = \tau$ ). Naslednjo reakcijo in časovni korak lahko izračunamo tako, da najprej izberemo dve naključni števili  $r_1$  in  $r_2$  iz intervala  $[0, 1]$  glede na uniformno porazdelitev in uporabimo enačbi

$$\tau = \frac{1}{a_0(\mathbf{y}(t))} \ln \left( \frac{1}{r_1} \right), \quad (4.45)$$

$$\sum_{j'=1}^{j-1} a_{j'}(\mathbf{y}(t)) \leq r_2 a_0(\mathbf{y}(t)) < \sum_{j'=j}^N a_{j'}(\mathbf{y}(t)). \quad (4.46)$$

Prednost te metode je ta, da prostor rešitev ne raste eksponentno s številom kemijskih zvrsti  $N$ . Poleg tega je dobljena trajektorija sistema eksakten posnetek ene izmed možnih dinamik sistema in ne zgolj njena aproksimacija. Slabost metode je počasnost, počasna konvergenca in slab vpogled v sistem. Počasnost izvira iz člena  $\frac{1}{a_0(\mathbf{y}(t))}$  v enačbi (4.45), ki je pri velikih populacijah izredno majhen. Zaradi tega pri kompleksnejših sistemih uporabljamo metode, ki aproksimirajo trajektorije na podoben način kot SSA.

#### 4.2.4 Metoda $\tau$ -skokov

Metoda  $\tau$ -skokov (angl.  *$\tau$ -leaping*) [17, 44] je alternativa stohastičnemu simuliranju (SSA) in temelji na fiksnem časovnem koraku  $\tau$ , v katerem se lahko sproži več kot ena reakcija (angl. *leap condition*). Izbrani čas mora biti dovolj majhen, da se stanje sistema v njem bistveno ne spremeni. V tem primeru velja, da se bo reakcija  $R_j$  sprožila približno  $\mathcal{P}_j(a_j(\mathbf{y}(t)), \tau)$ -krat v časovnem intervalu  $[t, t + \tau]$ . Na podlagi te metode lahko spremembo stanja izračunamo kot

$$\mathbf{y}(t + \tau) \approx \mathbf{y}(t) + \sum_{j=1}^M \nu_j \mathcal{P}_j(a_j(\mathbf{y}(t)), \tau), \quad (4.47)$$

kjer je  $\mathcal{P}_j(a_j(\mathbf{y}(t)), \tau)$  naključna Poissonova spremenljivka, ki določa število dogodkov v času  $\tau$  pri predpostavki, da je verjetnost dogodka v infinitezimalnem času  $dt$  enaka  $a \cdot dt$ ,  $a$  pa določa poljubno pozitivno konstanto. Načeloma je ta metoda hitrejša od metode SSA, potrebna pa je previdnost pri izbiri koraka  $\tau$ . V primeru, da velikost koraka limitiramo proti ničli, postane metoda natančnejša od metode SSA, po drugi strani pa postane neučinkovita, saj se reakcije ne prožijo več, čas pa se ustavi. Metoda je prav tako neučinkovita, če je dolžina koraka istega reda kot  $\frac{1}{a_0(\mathbf{y}(t))}$ , saj je v tem primeru bolj smiselna uporaba natančnejše metode SSA. Po drugi strani pridemo pri preveliki vrednosti koraka  $\tau$  do napačnih rezultatov (pridemo lahko celo do negativnih molekularnih števil).

#### Primer: modeliranje preprostega sistema na podlagi $\tau$ -skokov

Primer je povzet po [36]. Predpostavljajmo, da opazujemo dve reakciji, in sicer



Z uporabo Poissonove metode  $\tau$ -skokov izračunamo spremembo števila molekul  $S_3$  v intervalu  $[t, t + \tau]$

$$y_3(t + \tau) = y_3(t) + \mathcal{P}[c_1 y_1(t) y_2(t) \tau] - \mathcal{P}[c_2 y_3(t) \tau], \quad (4.50)$$

kjer  $y_i$  določa število molekul kemijske zvrsti  $i$ . Povprečno število molekul posamezne kemijske zvrsti ( $\bar{y}_i$ ) lahko tako izračunamo kot matematično upanje  $E(y_i)$  izraza

$$\frac{dy_3}{dt} = c_1 \bar{y}_1 \bar{y}_2 - c_2 \bar{y}_3. \quad (4.51)$$

Časovni potek koncentracij opazovanih kemijskih zvrsti dobimo z uporabe Eulerjeve metode za reševanje navadnih diferencialnih enačb

$$\bar{y}_3(t + \tau) = \bar{y}_3(t) + \tau [c_1 \bar{y}_1(t) \bar{y}_2(t) - c_2 \bar{y}_3(t)]. \quad (4.52)$$

#### 4.2.5 Langevinova enačba in enačba hitrosti reakcije

Predpostavljajmo, da velja naslednji pogoj pri izbiri dolžine koraka

$$a_j(\mathbf{y}(t))\tau \gg 1, \forall j = 1, \dots, M. \quad (4.53)$$

S tem v določenih primerih pride do kršitve pogoja skokov. V primerih, ko držita oba pogoja, lahko enačbo (4.47) poenostavimo [17, 45]. Poenostavitev upošteva dejstvo, da je Poissonova naključna spremenljivka  $\mathcal{P}(a, \tau)$  lahko aproksimirana z naključno normalno spremenljivko, ko je  $a\tau \gg 1$

$$\mathcal{P}_j(a_j(\mathbf{y}(t)), \tau) \approx \mathcal{N}_j(a_j(\mathbf{y}(t))\tau, a_j(\mathbf{y}(t))\tau) = a_j(\mathbf{y}(t))\tau + (a_j(\mathbf{y}(t))\tau)^{1/2} \mathcal{N}_j(0, 1), \quad (4.54)$$

kjer  $\mathcal{N}(m, \sigma^2)$ , predstavlja normalno porazdelitev s sredino  $m$  in varianco  $\sigma^2$ .

Zapišemo lahko *Langevinovo enačbo* (angl. *Langevin leaping formula*), ki določa stanje sistema v časovnem koraku  $t + \tau$

$$\mathbf{y}(t + \tau) \approx \mathbf{y}(t) + \sum_{j=1}^M \nu_j a_j(\mathbf{y}(t))\tau + \sum_{j=1}^M \nu_j \sqrt{a_j(\mathbf{y}(t))} \mathcal{N}_j(0, 1) \sqrt{\tau}. \quad (4.55)$$

Približek velja samo v primeru, da je korak  $\tau$  *dovolj majhen* (velja pogoj skokov) in hkrati tudi dovolj velik (velja enačba (4.53)). Če z obeh strani odštejemo  $\mathbf{y}(t)$  in enačbo delimo s  $\tau$ , pridemo do *kemijske Langevinove enačbe* (angl. *chemical Langevin equation* - *CLE*)

$$\frac{d\mathbf{y}(t)}{dt} \approx \sum_{j=1}^M \nu_j a_j(\mathbf{y}(t)) + \sum_{j=1}^M \nu_j \sqrt{a_j(\mathbf{y}(t))} \Gamma_j(t), \quad (4.56)$$

pri čemer  $\Gamma_j(t)$  predstavlja statistično neodvisen beli šum.

Aproksimacija še posebej dobro opisuje makroskopske sisteme. Molekularni sistemi postanejo makroskopski, ko postane razmerje med prostornino sistema in populacijo kemijskih zvrsti konstantno.

#### 4.2.6 Stohastično modeliranje s Poissonovo porazdelitvijo

Pristop je relativno enostaven, saj nam omogoča preprost prehod iz determinističnih modelov na stohastične [15, 36]. Od ostalih stohastičnih modelov se bistveno razlikuje po tem, da ne temelji na diskretnih stanjih sistema, ampak izhaja iz zveznih determinističnih modelov. Deterministične opise bioloških procesov namreč pretvorimo v stohastične tako, da predpostavljamo, da so porazdeljeni po Poissonovi naključni porazdelitvi.

Kot primer predpostavimo, da imamo deterministični model

$$\frac{d\bar{x}_i}{dt} = f_i(\bar{x}_1(t), \dots, \bar{x}_N(t)) - g_i(\bar{x}_1(t), \dots, \bar{x}_N(t)), i = 1, \dots, N, \quad (4.57)$$

kjer  $f_i(\bar{x}_1(t), \dots, \bar{x}_N(t))$  in  $g_i(\bar{x}_1(t), \dots, \bar{x}_N(t))$  določata spremembo vrednosti  $\bar{x}_i(t)$  za kemijsko zvrst  $S_i$ . Predpostavimo lahko, da se število molekul  $x_i(t)$  v majhnem intervalu  $[t, t +$

$\tau]$  spreminja s Poissonovo porazdelitvijo s srednjima vrednostima  $f_i(\bar{x}_1(t), \dots, \bar{x}_N(t)) \cdot \tau$  in  $g_i(\bar{x}_1(t), \dots, \bar{x}_N(t)) \cdot \tau$ . Tako lahko opišemo spremembo sistema z enačbo

$$x_i(t + \tau) = x_i(t) \mathcal{P}[f_i(\bar{x}_1, \dots, \bar{x}_N) \cdot \tau] - \mathcal{P}[g_i(\bar{x}_1, \dots, \bar{x}_N) \cdot \tau], i = 1, \dots, N. \quad (4.58)$$

### 4.3 Semikvantitativni modeli

Medtem ko so stohastični modeli zelo splošni in so primerni za modeliranje v kakršnih koli celicah in okoljih, so kvantitativni modeli bolj specifični in s tem tudi enostavnejši. Izdelamo lahko kvantitativne modele za specifične module (npr. TetR in LacI represorja), pri čemer obstaja veliko nedoločljivih parametrov, zaradi katerih so potrebne poenostavitve. Semikvantitativni modeli poenostavijo kvantitativne modele tako, da v model vključujejo le najpomembnejša stanja sistema in predvidevajo, da prehodi med njimi sledijo zakonu o vplivu koncentracij. S tem pridemo do modelov, ki imajo dodatne stopnje prostosti (angl. *degree of freedom*), s čimer lahko njihovo delovanje prilagajamo eksperimentalnim podatkom.

### 4.4 Primeri modeliranja enostavnih bioloških sistemov zmožnih procesiranja podatkov

Deterministične modele lahko vzpostavimo na podlagi navadnih diferencialnih enačb, ki so podane v obliki enačb hitrosti reakcije (glej razdelek 4.1.3) ali v obliki Hillovih enačb (glej razdelek 4.1.5). Pri stohastičnih modelih izhajamo iz vzpostavitve glavne enačbe (glej razdelek 4.2.2), katere reševanje lahko poteka z uporabo različnih postopkov. Za reševanje so bili implementirani algoritmi opisani v razdelku 4.2, nadaljnje analize pa temeljijo zgolj na reševanju glavne enačbe z algoritmom SSA, ki ima za naše potrebe najboljše razmerje med natančnostjo rezultatov in hitrostjo delovanja. Pravilnost delovanja implementiranih algoritmov smo potrdili z medsebojnim ujemanjem rezultatov in ujemanjem rezultatov dobljenih z uporabo referenčnega orodja *SGN Sim* [46].

#### 4.4.1 Izhodišča pri vzpostavitvi modelov

Pri vseh vzpostavljenih modelih v okviru celotnega dela bomo izhajali iz parametrov in reakcij prisotnih v bakteriofagu  $\lambda$  opisanih v virih [1, 47]. Izhajali bomo iz lastnosti operatorjev  $O_{R2}$  in  $O_{R3}$ , ki določata izražanje gena proteina Cro. Poljubno operatorsko

mesto lahko iz bakteriofaga  $\lambda$  z mutacijo iz sistema izločimo, tako kot izločimo operatorsko mesto  $O_{R1}$  [48]. Predpostavljamo, da gen, ki se izraža ob aktivnosti promotorja, tj. gen proteina Cro, zamenjamo za gen nekega drugega, *izhodnega proteina*. Prav tako predpostavljamo, da lahko določimo, kateri protein (*vhodni protein*) se bo kot transkripcijski faktor vezal na posamezni del promotorja in kakšna bo njegova vloga (represor ali aktivator). Vhodni protein se na enak način kot transkripcijski faktor CI na promotor veže v dimerizirani obliki. Ker je razmerje med hitrostjo vezave in razcepom dimera CI relativno veliko ( $1,8 \cdot 10^8$  M) [49], prav tako pa je relativno velika hitrost dimerizacije v primerjavi z vezavo dimera CI na operator, bomo pri modeliranju združili reakcijo dimerizacije vhodnega proteina z reakcijo vezave dimera vhodnega proteina na operator in reakcijo disociacije dimera na dva monomera proteina z reakcijo disociacije dimera proteina z operatorja.

### Izhodiščne reakcije

Pri vzpostavitvi modela osnovnih bioloških sistemov zmožnih procesiranja podatkov bomo izhajali iz reakcij, ki so predstavljene v tabeli 4.1.

Pri tem predpostavljamo, da je lahko zaseden le en operator istega promotorja naenkrat (bodisi operator  $O_{R2}$  bodisi operator  $O_{R3}$ ). Zaradi tega v nadaljevanju namesto notacije operatorjev, tj.  $O_{R2y}$  in  $O_{R3y}$ , uporabljamo notacijo  $DNA_y$ , pri čemer se  $y$  nanaša na izhodni protein, ki je sintetiziran, ko je promotor v aktivnem stanju ( $DNA_y^*$ ). Pri tem imajo parametri reakcij sledeče vrednosti:

- vezava proteina  $x$  na operator  $O_{R2}$ :  $k_{v_2} = 0,012 \text{ s}^{-1} \text{ nM}^{-2}$ ,
- disociacija proteina  $x$  z operatorja  $O_{R2}$ :  $k_{c_2} = 0,4791 \text{ s}^{-1}$ ,
- vezava proteina  $x$  na operator  $O_{R3}$ :  $k_{v_3} = 0,00012 \text{ s}^{-1} \text{ nM}^{-2}$ ,
- disociacija proteina  $x$  z operatorja  $O_{R3}$ :  $k_{c_3} = 0,4791 \text{ s}^{-1}$ ,
- vezava induktorja  $i$  na protein  $x^*$ :  $k_{v_i} = 0,00012 \text{ s}^{-1} \text{ nM}^{-2}$ ,
- disociacija kompleksa induktor-protein na protein  $x$  in induktor  $i$ :  $k_{c_i} = 0,4791 \text{ s}^{-1}$ ,
- transkripcija mRNA proteina  $y$ :  $k_{trs} = 0,0715 \text{ s}^{-1}$ ,
- translacija proteina  $y$ :  $k_{trl} = 0,043 \text{ s}^{-1}$ ,

	reakcija	pomen
1:	$2x + O_R2_y \xrightarrow{k_{v2}} x_2O_R2_y$	vezava dimera proteina $x$ na operator $O_R2$ promotorja proteina $y$
2:	$2x + O_R3_y \xrightarrow{k_{v3}} x_2O_R3_y$	vezava dimera proteina $x$ na operator $O_R3$ promotorja proteina $y$
3:	$x_2O_R2_y \xrightarrow{k_{c2}} 2x + O_R2_y$	disociacija dimera proteina $x$ z operatorja $O_R2$ promotorja proteina $y$
4:	$x_2O_R3_y \xrightarrow{k_{c3}} 2x + O_R3_y$	disociacija dimera proteina $x$ z operatorja $O_R3$ promotorja proteina $y$
5:	$x_2O_R2_y \xrightarrow{k_{deg}} O_R2_y$	razgradnja dimera proteina $x$ vezanega na operator $O_R2$ promotorja proteina $y$
6:	$x_2O_R3_y \xrightarrow{k_{deg}} O_R3_y$	razgradnja dimera proteina $x$ vezanega na operator $O_R3$ promotorja proteina $y$
7:	$x \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $x$
8:	$i + x \xrightarrow{k_{vi}} x^*$	vezava induktorja $i$ na transkripcijski faktor $x$
9:	$x^* \xrightarrow{k_{ci}} i + x$	disociacija kompleksa induktor-protein na induktor $i$ in protein $x$
10:	$DNA_y^* \xrightarrow{k_{trs}} DNA_y^* + mRNA_y$	transkripcija mRNA proteina $y$
11:	$mRNA_y \xrightarrow{k_{trt}} mRNA_y + y$	translacija proteina $y$
12:	$y \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $y$
13:	$mRNA_y \xrightarrow{k_{degm}} \emptyset$	razgradnja mRNA proteina $y$ ,

**Tabela 4.1** Reakcije, iz katerih bomo izhajali pri vzpostavitvi modelov osnovnih bioloških sistemov sposobnih procesiranja podatkov. Pri tem se  $x$  nanaša na vhodni protein, ki nastopa kot transkripcijski faktor (protein CI v bakteriofagu  $\lambda$ ),  $y$  pa na izhodni protein (protein Cro v bakteriofagu  $\lambda$ ).

- razgradnja proteinov  $x$  in  $y$ :  $k_{deg} = 0,0007 \text{ s}^{-1}$  (upoštevamo, da je hitrost razgradnje vezanega dimera enaka hitrosti razgradnje monomera [50]),
- razgradnja mRNA proteina  $y$ :  $k_{degm} = 0,0039 \text{ s}^{-1}$ .

Vrednosti so povzete po literaturi in so bile v določeni meri eksperimentalno izmerjene na realnih sistemih, v določeni meri pa ocenjene na podlagi ujemanj eksperimentalnih rezultatov in rezultatov modeliranja obravnavanega sistema [1, 48].



### Izhodišča pri vzpostavitvi determinističnih modelov

Pri determinističnih modelih bomo izhajali iz Hilllove enačbe represorja oziroma aktivatorja, pri kateri je kooperativnost vezave transkripcijskega faktorja na promotor enaka 2 (glej razdelek 4.1.5). V primeru, da vezava transkripcijskega faktorja na operator zavira aktivnost promotora, lahko transkripcijo mRNA modeliramo s Hillovo enačbo represorja

$$\frac{dm}{dt} = \frac{\alpha}{1 + \left(\frac{x}{K_d}\right)^\gamma} + \alpha_0 - \delta_m \cdot m. \quad (4.59)$$

V primeru, da vezava transkripcijskega faktorja na operator povečuje aktivnost promotora, lahko transkripcijo mRNA modeliramo s Hillovo enačbo aktivatorja

$$\frac{dm}{dt} = \alpha \frac{x^\gamma}{K_d^\gamma + x^\gamma} + \alpha_0 - \delta_m \cdot m. \quad (4.60)$$

Represorski oziroma aktivatorski protein lahko obravnavamo v kontekstu z induktorjem. Pri tem lahko vezavo induktorja na transkripcijski faktor modeliramo z enačbo

$$x^* = x \cdot \frac{y^{\gamma_i}}{y^{\gamma_i} + K_{d_i}^{\gamma_i}}. \quad (4.61)$$

Translacijo izhodnega proteina lahko opišemo z izrazom

$$\frac{dy}{dt} = \beta \cdot m - \delta_y \cdot y. \quad (4.62)$$

Pri tem  $x$  predstavlja koncentracijo vhodnega proteina,  $i$  koncentracijo induktorja,  $x^*$  vhodnega proteina vezanega z induktorjem,  $y$  koncentracijo izhodnega proteina,  $m$  pa njegove mRNA. Glede na uporabljeno notacijo parametrov priredimo sledeče vrednosti:

- maksimalna hitrost transkripcije:  $\alpha = k_{trs} = 0,0715 \text{ s}^{-1}$ ,
- hitrost translacije:  $\beta = k_{trl} = 0,043 \text{ s}^{-1}$ ,
- osnovno izražanje gena:  $\alpha_0 = 0 \text{ s}^{-1}$ ,
- Hillov koeficient:  $\gamma = 2$ ,
- Hillov koeficient induktorja:  $\gamma_i = 1$ ,
- koeficient aktivacije za operator  $O_{R2}$ :  
 $K_d = K_{d_2} = \left(\frac{k_{c_2}}{k_{v_2}}\right)^{\frac{1}{\gamma}} = 6,3168$ ,
- koeficient aktivacije za operator  $O_{R3}$ :  
 $K_d = K_{d_3} = \left(\frac{k_{c_3}}{k_{v_3}}\right)^{\frac{1}{\gamma}} = 63,1862$ ,

- koeficient aktivacije pri vezavi induktorja na aktivator:

$$K_d = K_{d_i} = \left( \frac{k_{c_i}}{k_{v_i}} \right)^{\frac{1}{\gamma_i}} = 3992,5,$$

- hitrost razgradnje proteina  $y$ :  $\delta_y = k_{deg} = 0,0007 \text{ s}^{-1}$ ,
- hitrost razgradnje mRNA:  $\delta_m = k_{deg_m} = 0,0039 \text{ s}^{-1}$ .

### Izhodišča pri vzpostavitvi stohastičnih modelov

Stohastični modeli bodo vzpostavljeni z glavno enačbo, pri čemer bodo upoštevane reakcije, ki so v sistemu predvidene. Glavno enačbo bomo reševali z metodo SSA (glej razdelek 4.2.3). Poleg tega je za postavitev modela potrebno določiti stohastične koeficiente hitrosti reakcij ( $c_j$ ), preko katerih računamo verjetnosti izvedbe posamezne reakcije v posameznem časovnem koraku ( $a_j(\mathbf{y}(t))dt$ ). Pri tem stohastične koeficiente hitrosti reakcij določamo po izrazih (4.35), (4.38) in (4.41). Predpostavljajmo, da je prostornina celice konstantna ( $V = 2 \cdot 10^{-15} \text{ L}$ ) in da se v njej nahaja 1 molekula DNA s pripadajočim promotorjem, operatorjem(a) in kodirajočim zaporedjem proteina  $y$ . Z večanjem števila molekul DNA bi sicer lahko do določene mere zmanjšali vpliv stohastičnosti oziroma šuma na obnašanje sistema [28]. Glede na reakcije predstavljene v tabeli 4.1 imajo stohastični koeficienti hitrosti reakcij sledeče vrednosti:

- vezava dimera proteina  $x$  na operator  $O_{R2}$ :  $c_{v_2} = 2 \cdot \frac{k_{v_2}}{\Omega^2} = 0,0165 \text{ s}^{-1}$ ,
- disociacija dimera proteina  $x$  z operatorja  $O_{R2}$ :  $c_{c_2} = k_{c_2} = 0,4791 \text{ s}^{-1}$ ,
- vezava proteina  $x$  na operator  $O_{R3}$ :  $c_{v_3} = 2 \cdot \frac{k_{v_3}}{\Omega^2} = 1,654410 \cdot 10^{-4} \text{ s}^{-1}$ ,
- disociacija proteina  $x$  z operatorja  $O_{R3}$ :  $c_{c_3} = k_2 = 0,4791 \text{ s}^{-1}$ ,
- transkripcija mRNA proteina  $y$ :  $c_{tr_s} = k_{tr_s} = 0,0715 \text{ s}^{-1}$ ,
- translacija proteina  $y$ :  $c_{tr_l} = k_{tr_l} = 0,043 \text{ s}^{-1}$ ,
- razgradnja proteina  $y$ :  $c_{deg} = k_{deg} = 0,0007 \text{ s}^{-1}$ ,
- razgradnja mRNA proteina  $y$ :  $c_{deg_m} = k_{deg_m} = 0,0039 \text{ s}^{-1}$ .

pri čemer je  $\Omega$  velikost sistema, ki je določena preko prostornine celice (glej izraz (2.7)).

### 4.4.2 Modeliranje biološkega gonilnika

Biološki gonilnik lahko realiziramo z gensko regulatornim omrežjem, v katerem sta prisotna dva proteina, pri čemer ima prvi vlogo aktivatorja drugega proteina (glej razdelek 3.3.1). Model biološkega gonilnika je postavljen na osnovi parametrov in reakcij prisotnih v povezavi z operatorjem  $O_{R3}$  bakteriofaga  $\lambda$  [1, 47], pri čemer vezava proteina CI v dimerizirani obliki na operator aktivira izražanje gena, ki določa zapis proteina Cro [48]. V našem primeru predpostavljamo, da gen, ki se izraža ob aktivnosti promotorja, tj. gen proteina Cro, zamenjamo za gen proteina  $y$ . V nadaljevanju za protein CI uporabljamo oznako  $x$ . Model se torej ujema s sliko 3.2.

Predpostavljamo, da so v sistemu prisotne reakcije predstavljene v tabeli 4.2.

	reakcija	pomen
1:	$2x + \text{DNA} \xrightarrow{k_{v3}} x_2\text{DNA}$	vezava dimera proteina $x$ na operator $O_{R3}$
2:	$x_2\text{DNA} \xrightarrow{k_{e3}} 2x + \text{DNA}$	disociacija dimera proteina $x$ z operatorja $O_{R3}$
3:	$x_2\text{DNA} \xrightarrow{k_{tr3}} \text{mRNA} + x_2\text{DNA}$	transkripcija mRNA proteina $y$
4:	$\text{mRNA} \xrightarrow{k_{tr1}} \text{mRNA} + y$	translacija proteina $y$
5:	$\text{mRNA} \xrightarrow{k_{deg1}} \emptyset$	razgradnja mRNA proteina $y$
6:	$y \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $y$

Tabela 4.2 Opazovane kemijske reakcije v modelu biološkega gonilnika.

Koncentracije proteina  $y$  in njegove mRNA se s časom spreminjajo v odvisnosti od koncentracije proteina  $x$ , za katero pa predpostavljamo, da je konstantna. Pri modeliranju poteka koncentracij v odvisnosti od časa izhajamo iz začetnega stanja, ko je koncentracija proteina  $y$  in njegove mRNA nizka ( $y(0) = 0$  nM,  $\text{mRNA}(0) = 0$  nM).

#### Deterministično modeliranje biološkega gonilnika

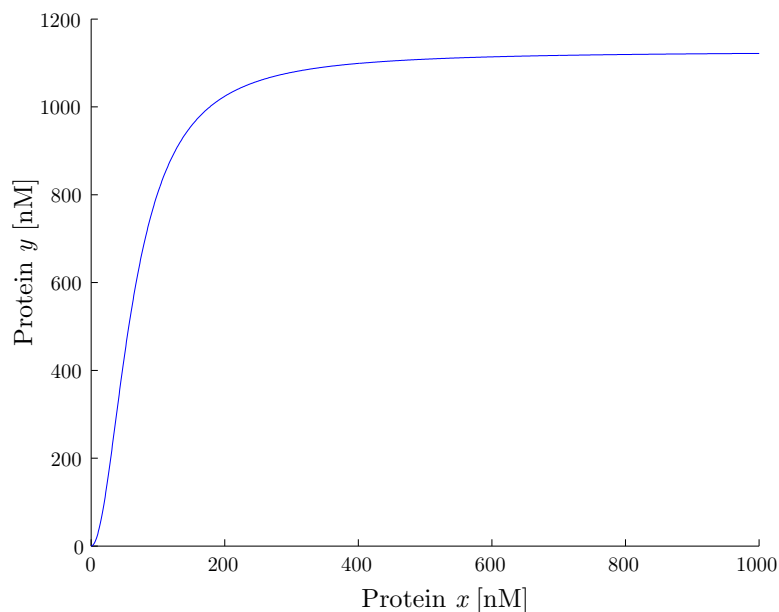
Izhajamo iz Hillove enačbe aktivatorja, pri kateri je kooperativnost vezave transkripcijskega faktorja na promotor enaka 2 (glej razdelek 4.1.5). Dinamiko sistema opišemo z dvema diferencialnima enačbama, in sicer

$$\frac{dm}{dt} = \alpha \frac{x^\gamma}{K_{d3}^\gamma + x^\gamma} + \alpha_0 - \delta_m \cdot m \quad (4.63)$$

in

$$\frac{dy}{dt} = \beta \cdot m - \delta_y \cdot y. \quad (4.64)$$

Opazujemo obnašanje sistema v odvisnosti od koncentracije vhodnega proteina  $x$ . Delovanje sistema lahko analiziramo na podlagi računanja ravnovesnih stanj (koncentracije proteina  $y$ ) v odvisnosti od koncentracije vhodnega proteina  $x$  (glej sliko 4.1). Problem takega prikaza je v tem, da nimamo vpogleda v časovni potek koncentracij izhodnega proteina. Po drugi strani lahko opazujemo časovni potek koncentracije izhodnega pro-

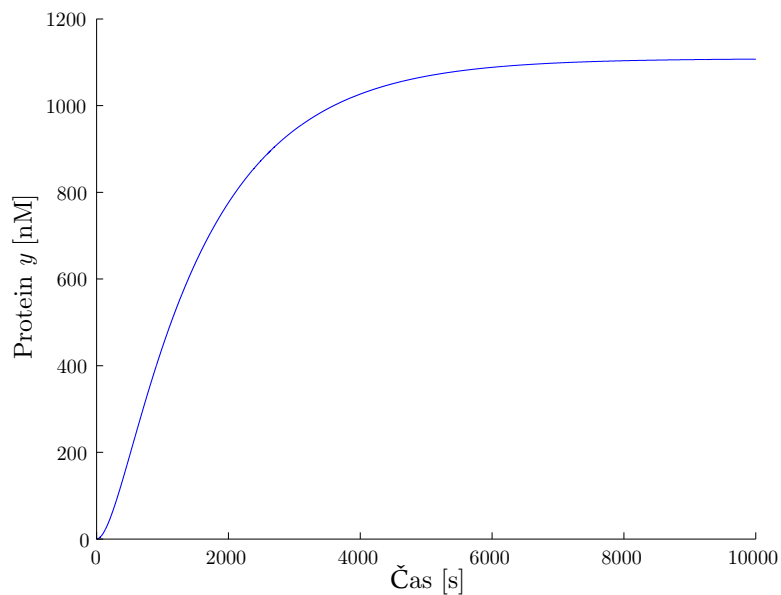


**Slika 4.1** Deterministično modeliranje biološkega gonilnika na podlagi Hillovih enačb, pri čemer graf prikazuje koncentracijo izhodnega proteina  $y$  v odvisnosti od koncentracije vhodnega proteina  $x$ , ko sistem doseže ravnovesno stanje.

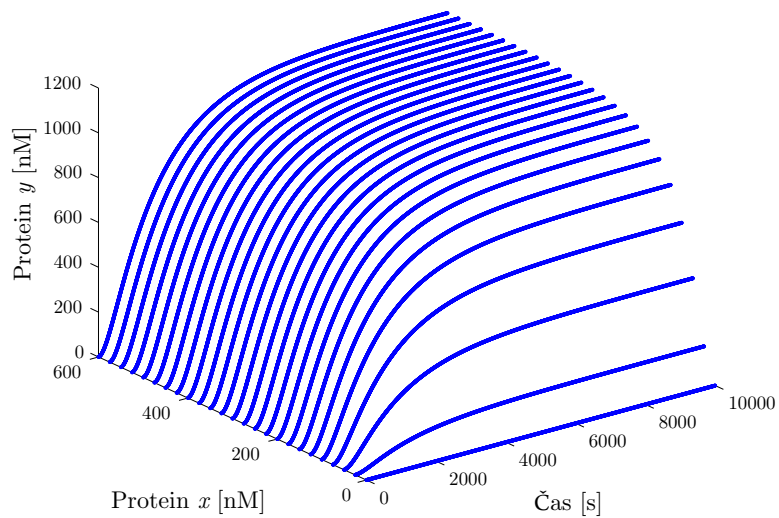
teina  $y$  pri določeni koncentraciji vhodnega proteina  $x$ . Problem takega prikaza je v tem, da lahko opazujemo časovni potek izhodnega proteina le pri določeni količini vhodnega proteina (slika 4.2 prikazuje časovni potek pri konstantni koncentraciji vhodnega proteina, in sicer  $500nM$ ). Oba prikaza lahko združimo v skupen graf, ki prikazuje časovni potek koncentracij izhodnega proteina pri različnih koncentracijah vhodnega proteina (glej sliko 4.3).

### Stohastično modeliranje biološkega gonilnika

Stohastičen model biološkega gonilnika lahko vzpostavimo s pomočjo glavne enačbe, pri čemer upoštevamo reakcije opisane v tabeli 4.2. Glavno enačbo rešujemo z metodo SSA (glej razdelek 4.2.3).



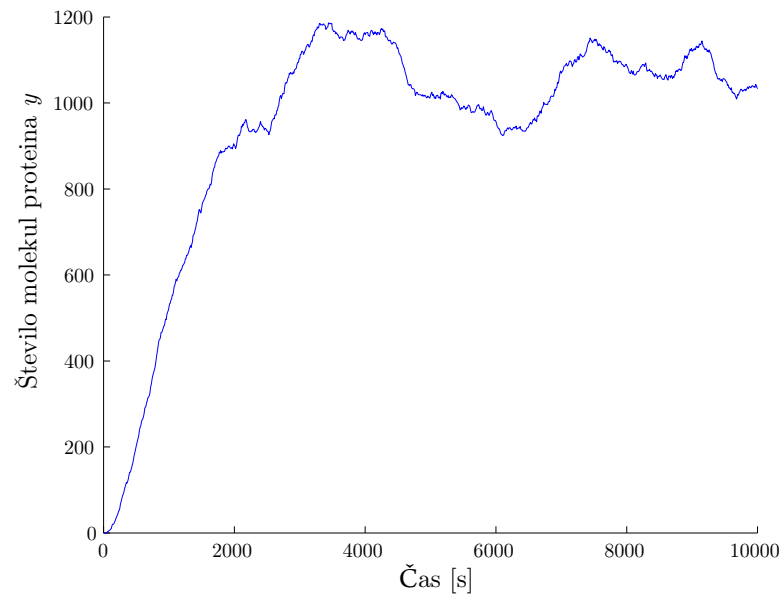
**Slika 4.2** Časovni potek koncentracij izhodnega proteina  $y$  pri konstantni koncentraciji vhodnega proteina  $x$  ( $500\text{nM}$ ) na podlagi determinističnega modela biološkega gonilnika (Hillove enačbe).



**Slika 4.3** Časovni potek koncentracij izhodnega proteina  $y$  pri različnih koncentracijah vhodnega proteina  $x$  na podlagi determinističnega modela biološkega gonilnika (Hillove enačbe).

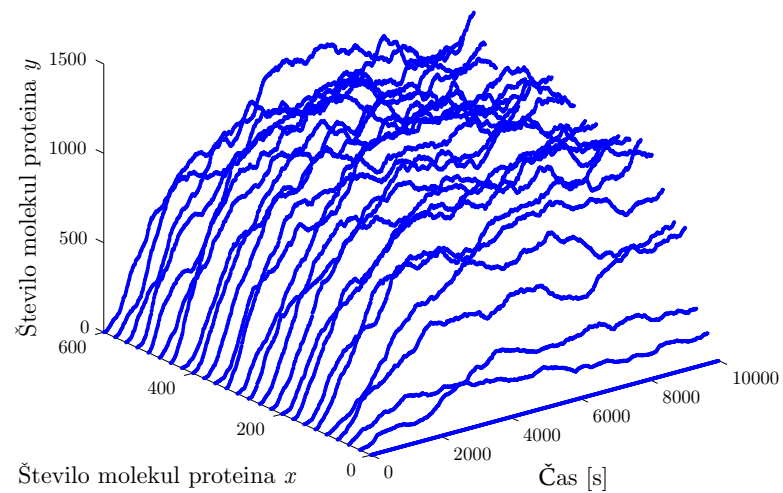
Časovni potek koncentracije izhodnega proteina  $y$  opazujemo pri konstantni količini vhodnega proteina  $x$  (tj.  $500\text{nM} \approx 600$  molekul proteina  $x$  v celici - za izračun glej izraz (2.6)). Potek koncentracij v stohastičnem modelu biološkega gonilnika prikazuje

slika 4.4. Prikazujemo lahko časovni potek koncentracij izhodnega proteina pri različnih



**Slika 4.4** Časovni potek koncentracij izhodnega proteina  $y$  pri konstantnem številu vhodnega proteina  $x$  (600 molekul) na podlagi stohastičnega modela biološkega gonilnika.

koncentracijah vhodnega proteina (glej sliko 4.5).



**Slika 4.5** Časovni potek koncentracij izhodnega proteina  $y$  pri različnih koncentracijah vhodnega proteina  $x$  na podlagi stohastičnega modela biološkega gonilnika.

### 4.4.3 Modeliranje biološkega negatorja

Biološki negator lahko realiziramo s podobnim gensko regulatornim omrežjem kot biološki gonilnik, le da vhodni protein  $x$  v tem sistemu z vezavo na operator zavira izražanje gena, ki določa zapis izhodnega proteina  $y$ . Model se torej ujema s sliko 3.3.

Predpostavljamo, da so v sistemu prisotne reakcije predstavljene v tabeli 4.3. Model

	reakcija	pomen
1:	$2x + \text{DNA} \xrightarrow{k_{v3}} x_2 \text{DNA}$	vezava proteina $x$ na operator $O_R3$
2:	$x_2 \text{DNA} \xrightarrow{k_{c3}} 2x + \text{DNA}$	disociacija proteina $x$ z operatorja $O_R3$
3:	$\text{DNA} \xrightarrow{k_{trs}} \text{mRNA} + \text{DNA}$	transkripcija mRNA proteina $y$
4:	$\text{mRNA} \xrightarrow{k_{trt}} \text{mRNA} + y$	translacija proteina $y$
5:	$\text{mRNA} \xrightarrow{k_{deg^m}} \emptyset$	razgradnja mRNA proteina $y$
6:	$y \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $y$

**Tabela 4.3** Opazovane kemijske reakcije v modelu biološkega negatorja.

negatorja se od modela gonilnika razlikuje zgolj po reakciji 3, saj transkripcija poteka le v primeru, da je operator prost.

Prav tako kot pri modelu gonilnika se koncentracije proteina  $y$  in njegove mRNA s časom spreminjajo v odvisnosti od koncentracije proteina  $x$ , za katero predpostavljamo, da je konstantna. Pri modeliranju poteka koncentracij v odvisnosti od časa izhajamo iz začetnega stanja, ko je koncentracija proteina  $y$  in njegove mRNA visoka ( $y(0) = 1\,000$  nM,  $\text{mRNA}(0) = 20$  nM).

#### Deterministično modeliranje biološkega negatorja

Izhajamo iz Hillove enačbe represorja s kooperativnostjo vezave transkripcijskega faktorja na promotor 2 (glej razdelek 4.1.5). Dinamiko sistema lahko opišemo z dvema diferencialnima enačbama, in sicer

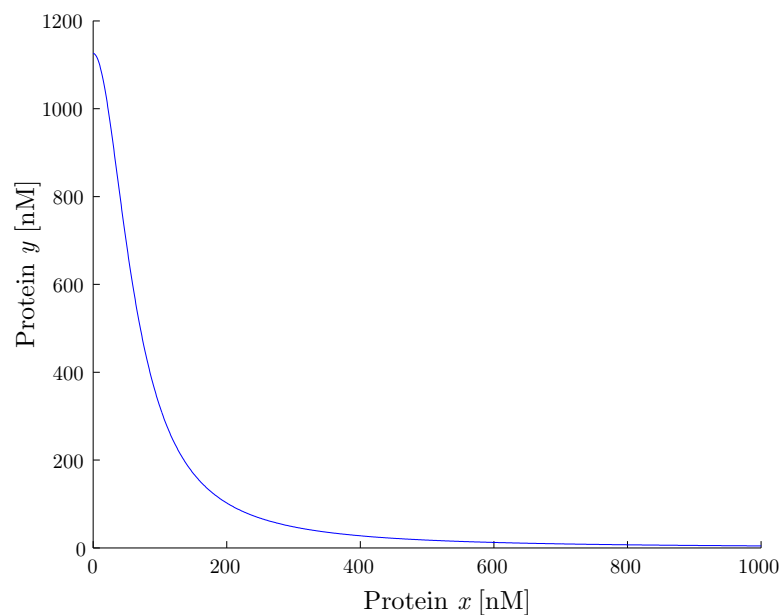
$$\frac{dm}{dt} = \frac{\alpha}{1 + \left(\frac{x}{K_{d3}}\right)^\gamma} + \alpha_0 - \delta_m \cdot m \quad (4.65)$$

in

$$\frac{dy}{dt} = \beta \cdot m - \delta_y \cdot y. \quad (4.66)$$

Opazujemo obnašanje sistema v odvisnosti od koncentracije vhodnega proteina  $x$ . Delovanje sistema lahko analiziramo na podlagi računanja ravnovesnih stanj koncentracije

proteina  $y$  v odvisnosti od koncentracije vhodnega proteina  $x$  (glej sliko 4.6). Po drugi



**Slika 4.6** Deterministično modeliranje biološkega negatorja na podlagi Hillovih enačb, pri čemer graf prikazuje koncentracijo izhodnega proteina  $y$  v odvisnosti od koncentracije vhodnega proteina  $x$ , ko sistem doseže ravnovesno stanje.

strani lahko opazujemo časovni potek koncentracije izhodnega proteina  $y$  pri določeni koncentraciji vhodnega proteina  $x$  (slika 4.7 prikazuje časovni potek pri konstantni koncentraciji vhodnega proteina, tj. 500 nM). Oba prikaza lahko združimo v skupen graf, ki prikazuje časovni potek koncentracij izhodnega proteina pri različnih koncentracijah vhodnega proteina (glej sliko 4.8).

### Stohastično modeliranje biološkega negatorja

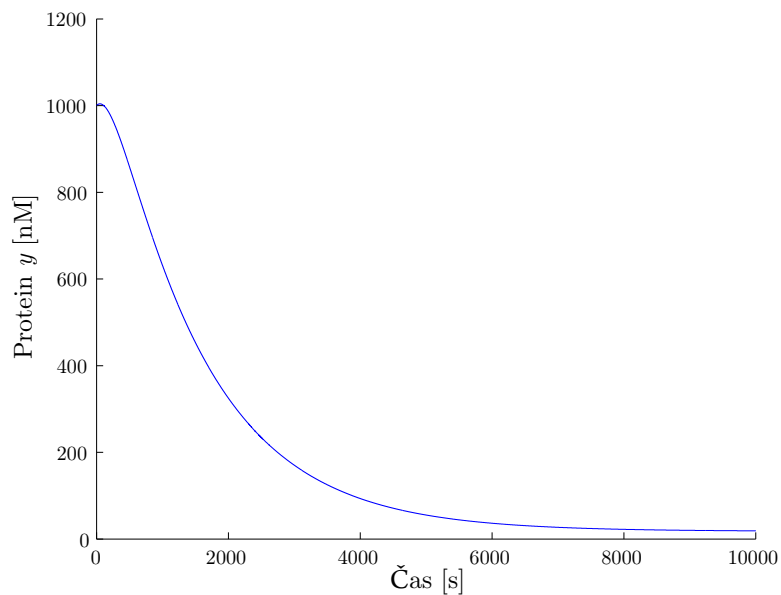
Stohastičen model biološkega negatorja lahko postavimo na podlagi reakcij opisanih v tabeli 4.3. Glavno enačbo rešujemo z metodo SSA (glej razdelek 4.2.3).

Časovni potek koncentracije izhodnega proteina  $y$  opazujemo pri konstantnem številu vhodnega proteina  $x$  (tj. 500 nM  $\approx$  600 molekul proteina  $x$  v celici - za izračun glej izraz (2.6)). Pri tem predpostavljamo, da je začetno število molekul proteina  $y$  v celici 1200 (približno 1000 nM), začetno število molekul mRNA pa 25 (približno 20 nM). Potek koncentracij v stohastičnem modelu biološkega gonilnika prikazuje slika 4.9.

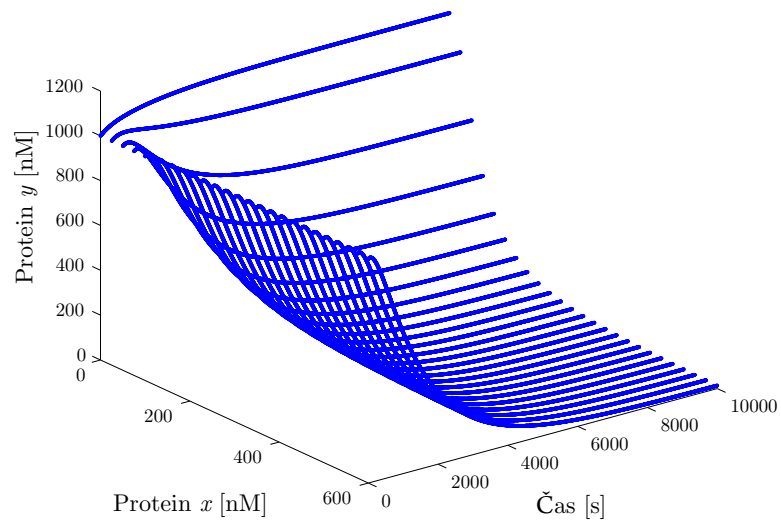
Prikazujemo lahko tudi časovni potek koncentracij izhodnega proteina pri različnih



#### 4.4 PRIMERI MODELIRANJA ENOSTAVNIH BIOLOŠKIH SISTEMOV ZMOŽNIH PROCESIRANJA PODATKOV61

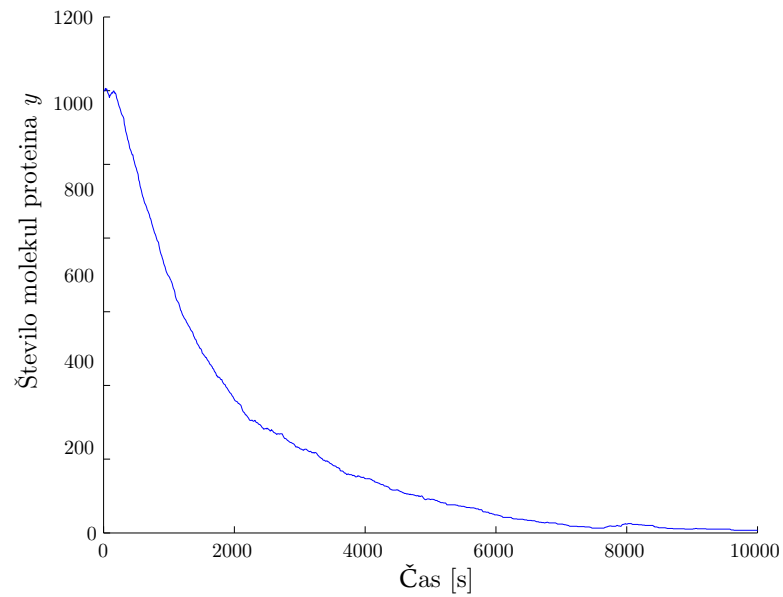


**Slika 4.7** Časovni potek koncentracij izhodnega proteina  $y$  pri konstantni koncentraciji vhodnega proteina  $x$  (500 nM) na podlagi determinističnega modela biološkega negatorja (Hillove enačbe).

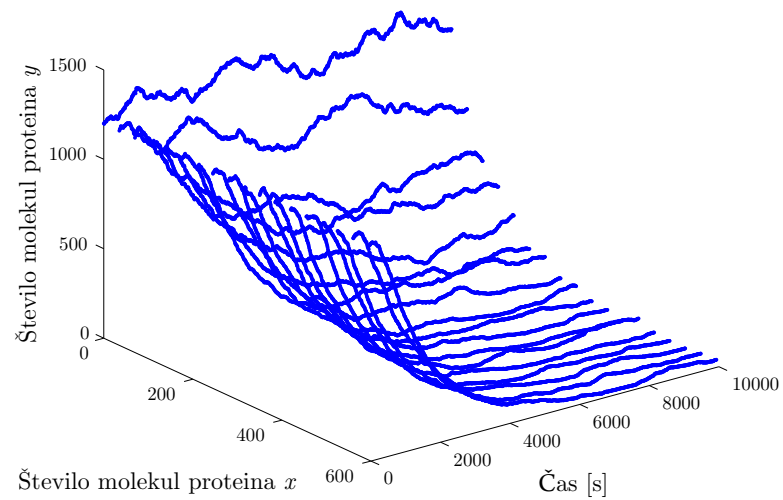


**Slika 4.8** Časovni potek koncentracij izhodnega proteina  $y$  pri različnih koncentracijah vhodnega proteina  $x$  na podlagi determinističnega modela biološkega negatorja (Hillove enačbe).

koncentracijah vhodnega proteina (glej sliko 4.10).



**Slika 4.9** Časovni potek koncentracij izhodnega proteina  $y$  pri konstantnem številu vhodnega proteina  $x$  (600 molekul) na podlagi stohastičnega modela biološkega negatorja.



**Slika 4.10** Časovni potek koncentracij izhodnega proteina  $y$  pri različnih koncentracijah vhodnega proteina  $x$  na podlagi stohastičnega modela biološkega negatorja.

#### 4.4.4 Modeliranje bioloških AND vrat

Biološka AND vrata lahko realiziramo z genskim regulatornim omrežjem z dvema transkripcijskima faktorjema, in sicer aktivatorjem  $x$  in induktorjem  $y$ . Ob vezavi induktorja

na aktivator se le-ta lahko veže na promotor, s čimer aktivira transkripcijo izhodnega proteina  $y$ . Model sistema lahko postavimo na podlagi reakcij, ki so predstavljene v tabeli 4.4.

	reakcija	pomen
1:	$x + y \xrightarrow{k_{v_i}} x^*$	vezava proteina induktorja $y$ na aktivator $x$
2:	$x^* \xrightarrow{k_{e_i}} x + y$	disociacija kompleksa induktor-protein
3:	$2x^* + \text{DNA} \xrightarrow{k_{v_3}} x_2^* \text{DNA}$	vezava aktivatorja na operator $O_{R3}$
4:	$x_2^* \text{DNA} \xrightarrow{k_{e_3}} 2x^* + \text{DNA}$	disociacija aktivatorja z operatorja $O_{R3}$
5:	$x_2^* \text{DNA} \xrightarrow{k_{tr_3}} \text{mRNA} + x_2^* \text{DNA}$	transkripcija mRNA proteina $z$
6:	$\text{mRNA} \xrightarrow{k_{tr_l}} \text{mRNA} + z$	translacija proteina $z$
7:	$\text{mRNA} \xrightarrow{k_{deg_m}} \emptyset$	razgradnja mRNA proteina $z$
8:	$z \xrightarrow{k_{deg_z}} \emptyset$	razgradnja proteina $z$

**Tabela 4.4** Opazovane kemijske reakcije v modelu bioloških AND vrat.

Predpostavljamo, da se koncentracije proteina  $z$  in njegove mRNA s časom spreminjajo v odvisnosti od koncentracije vhodnih proteinov  $x$  in  $y$ . Pri modeliranju poteka koncentracij v odvisnosti od časa izhajamo iz začetnega stanja, ko je koncentracija proteina  $z$  in njegove mRNA nizka ( $y(0) = 0$  nM,  $\text{mRNA}(0) = 0$  nM).

#### Deterministično modeliranje bioloških AND vrat

Izhajamo iz Hillove enačbe vezave induktorja na aktivator. Dinamiko sistema lahko opišemo s sledečimi diferencialnimi enačbami:

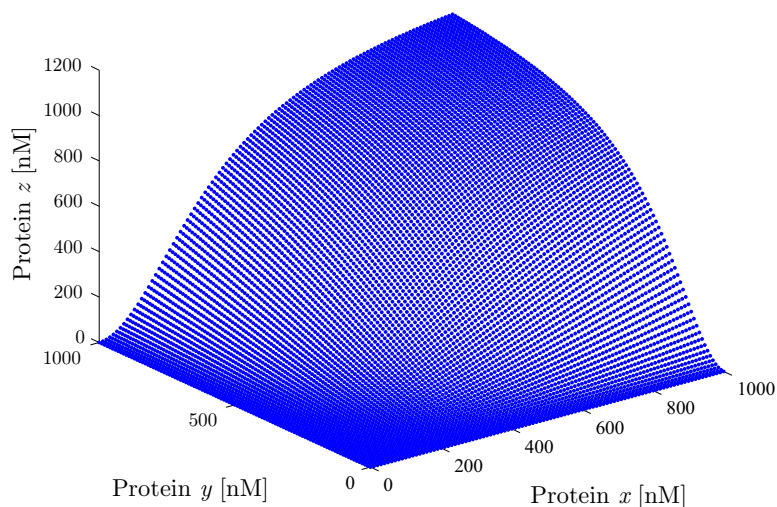
$$x^* = x \cdot \frac{y^{\gamma_i}}{y^{\gamma_i} + K_{d_i}^{\gamma_i}}, \quad (4.67)$$

$$\frac{dm}{dt} = \frac{\alpha \cdot x^{*\gamma}}{K_{d_3}^{\gamma} + x^{*\gamma}} + \alpha_0 - \delta_m \cdot m \quad (4.68)$$

in

$$\frac{dz}{dt} = \beta \cdot m - \delta_z \cdot z. \quad (4.69)$$

Opazujemo obnašanje sistema pri konstantnih koncentracijah vhodnih proteinov  $x$  in  $y$ . Delovanje sistema lahko analiziramo na podlagi računanja ravnovesnih stanj koncentracije proteina  $z$  v odvisnosti od koncentracije vhodnih proteinov  $x$  in  $y$  (glej sliko 4.11).



**Slika 4.11** Deterministično modeliranje bioloških AND vrat na podlagi Hillovih enačb, pri čemer graf prikazuje koncentracijo izhodnega proteina  $z$  v odvisnosti od aktivatorja  $x$  in induktorja  $y$ , ko sistem doseže ravnovesno stanje.

Po drugi strani lahko opazujemo časovni potek koncentracije izhodnega proteina  $z$  pri določenih koncentracijah vhodnih proteinov  $x$  in  $y$  (slika 4.12 prikazuje časovni potek pri štirih različnih koncentracijah vhodnih proteinov).

#### Stohastično modeliranje bioloških AND vrat

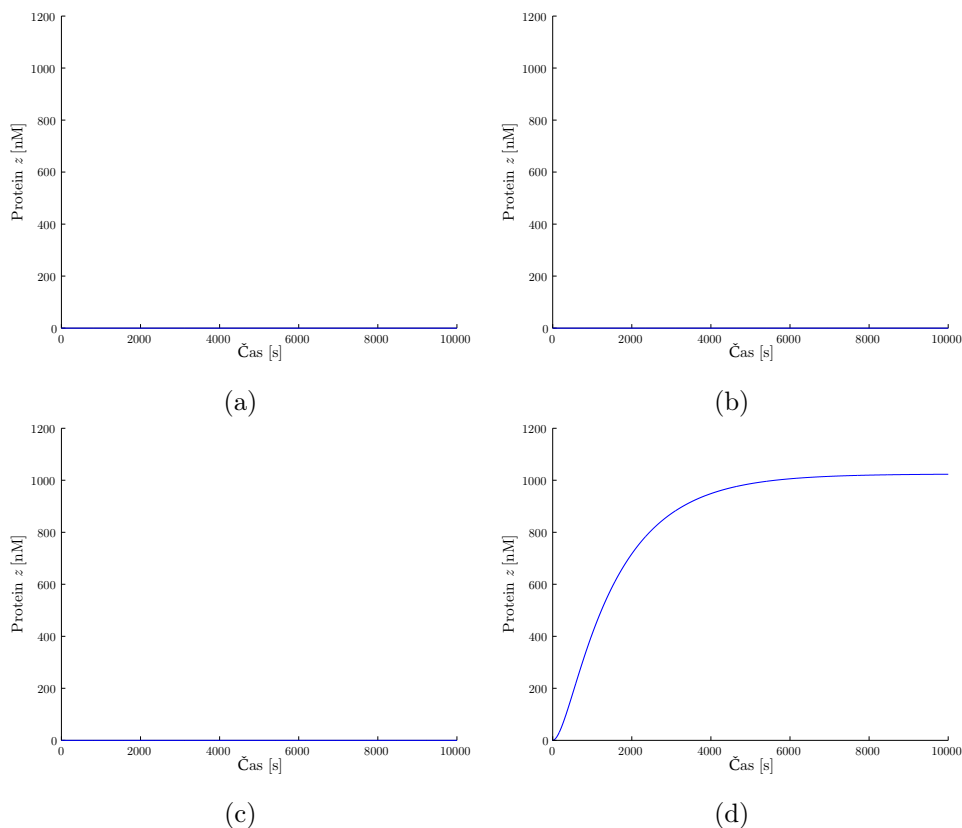
Stohastičen model bioloških AND vrat lahko postavimo na podlagi reakcij opisanih v tabeli 4.4. Glavno enačbo rešujemo z metodo SSA (glej razdelek 4.2.3).

Časovni potek koncentracije izhodnega proteina  $z$  opazujemo pri konstantnem številu molekul vhodnih proteinov  $x$  in  $y$ . Pri tem predpostavljamo, da je začetno število molekul proteina  $z$  in njegovih mRNA v celici enako 0. Časovni potek koncentracij izhodnega proteina  $z$  pri štirih različnih koncentracijah vhodnih proteinov v stohastičnem modelu bioloških AND vrat prikazuje slika 4.13.

#### 4.4.5 Modeliranje bioloških NOR vrat

Biološka NOR vrata lahko realiziramo z razširitvijo biološkega negatorja, tako da prisotnost kateregakoli od vhodnih proteinov  $x$  in  $y$  zavira izražanje proteina  $z$ . Model se torej ujema s sliko 3.5. Predpostavljamo, da so v sistemu prisotne reakcije predstavljene v tabeli 4.5. Koncentracije proteina  $z$  in njegove mRNA se s časom spreminjajo v odvis-

#### 4.4 PRIMERI MODELIRANJA ENOSTAVNIH BIOLOŠKIH SISTEMOV ZMOŽNIH PROCESIRANJA PODATKOV 65



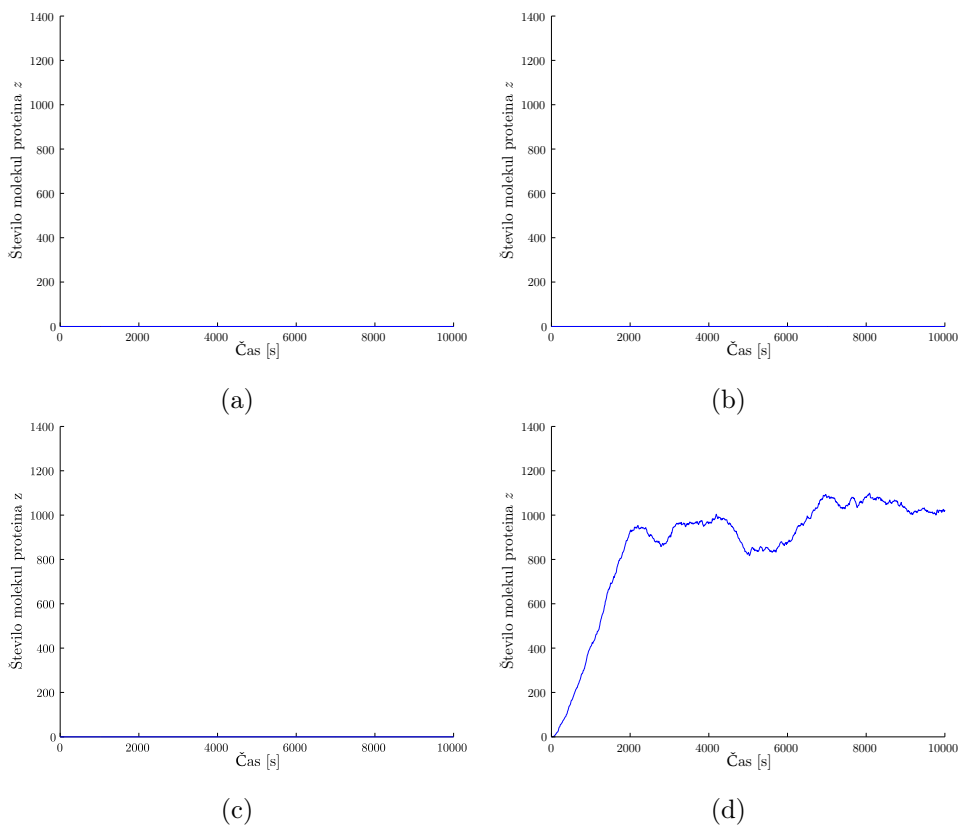
**Slika 4.12** Časovni potek koncentracij izhodnega proteina  $z$  pri štirih različnih koncentracijah aktivatorja  $x$  in induktorja  $y$  na podlagi determinističnega modela bioloških AND vrat (Hillove enačbe), in sicer  $x, y = (0 \text{ nM}, 0 \text{ nM})$  na sliki (a),  $x, y = (0 \text{ nM}, 1000 \text{ nM})$  na sliki (b),  $x, y = (1000 \text{ nM}, 0 \text{ nM})$  na sliki (c) in  $x, y = (1000 \text{ nM}, 1000 \text{ nM})$  na sliki (d).

nosti od koncentracije vhodnih proteinov  $x$  in  $y$ . Pri modeliranju poteka koncentracij v odvisnosti od časa izhajamo iz začetnega stanja, ko je koncentracija proteina  $z$  in njegove mRNA visoka ( $z(0) = 1000 \text{ nM}$ ,  $\text{mRNA}(0) = 20 \text{ nM}$ ).

#### Deterministično modeliranje bioloških NOR vrat

Izhajamo iz Hillove enačbe represorja s kooperativnostjo vezave transkripcijskega faktorja na promotor 2 (glej razdelek 4.1.5), pri čemer predpostavljamo, da v sistemu nastopata dva represorja. Dinamiko sistema lahko opišemo z dvema diferencialnima enačbama, in sicer

$$\frac{dm}{dt} = \frac{\alpha}{1 + \left(\frac{x}{K_{d_3}}\right)^\gamma + \left(\frac{y}{K_{d_3}}\right)^\gamma} + \alpha_0 - \delta_m \cdot m \quad (4.70)$$



**Slika 4.13** Časovni potek koncentracij izhodnega proteina  $z$  pri štirih različnih koncentracijah vhodnih proteinov na podlagi stohastičnega modela bioloških AND vrat, in sicer  $x, y = (0, 0)$  na sliki (a),  $x, y = (0, 1200)$  na sliki (b),  $x, y = (1200, 0)$  na sliki (c) in  $x, y = (1200, 1200)$  na sliki (d).

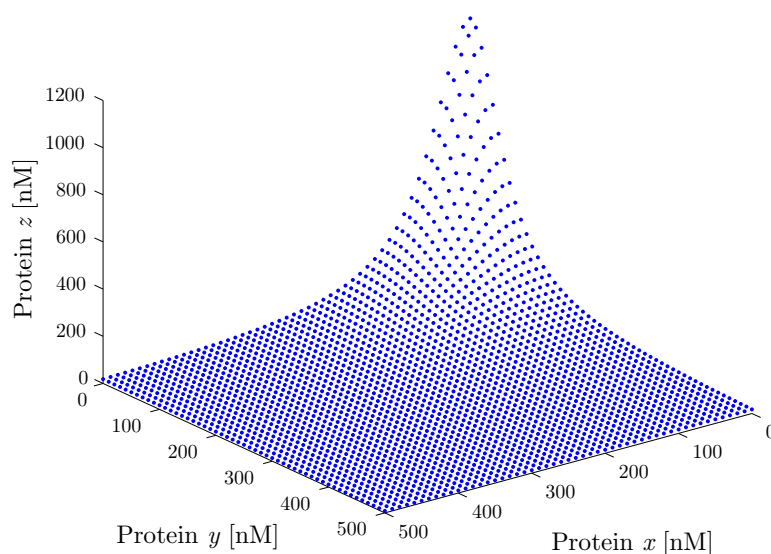
	reakcija	pomen
1:	$2x + \text{DNA} \xrightarrow{k_{v3}} x_2 \text{DNA}$	vezava proteina $x$ na operator $O_{R3}$
2:	$x_2 \text{DNA} \xrightarrow{k_{e3}} 2x + \text{DNA}$	disociacija proteina $x$ z operatorja $O_{R3}$
3:	$2y + \text{DNA} \xrightarrow{k_{v3}} y_2 \text{DNA}$	vezava proteina $y$ na operator $O_{R3}$
4:	$y_2 \text{DNA} \xrightarrow{k_{e3}} 2y + \text{DNA}$	disociacija proteina $y$ z operatorja $O_{R3}$
5:	$\text{DNA} \xrightarrow{k_{tr3}} \text{mRNA} + \text{DNA}$	transkripcija mRNA proteina $z$
6:	$\text{mRNA} \xrightarrow{k_{tr1}} \text{mRNA} + z$	translacija proteina $z$
7:	$\text{mRNA} \xrightarrow{k_{degm}} \emptyset$	razgradnja mRNA proteina $z$
8:	$z \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $z$

**Tabela 4.5** Opazovane kemijske reakcije v modelu bioloških NOR vrat.

in

$$\frac{dz}{dt} = \beta \cdot m - \delta_z \cdot z. \quad (4.71)$$

Opazujemo obnašanje sistema v odvisnosti od koncentracij vhodnih proteinov  $x$  in  $y$ . Delovanje sistema lahko analiziramo na podlagi računanja ravnovesnih stanj koncentracije proteina  $z$  v odvisnosti od koncentracije vhodnih proteinov  $x$  in  $y$  (glej sliko 4.14).



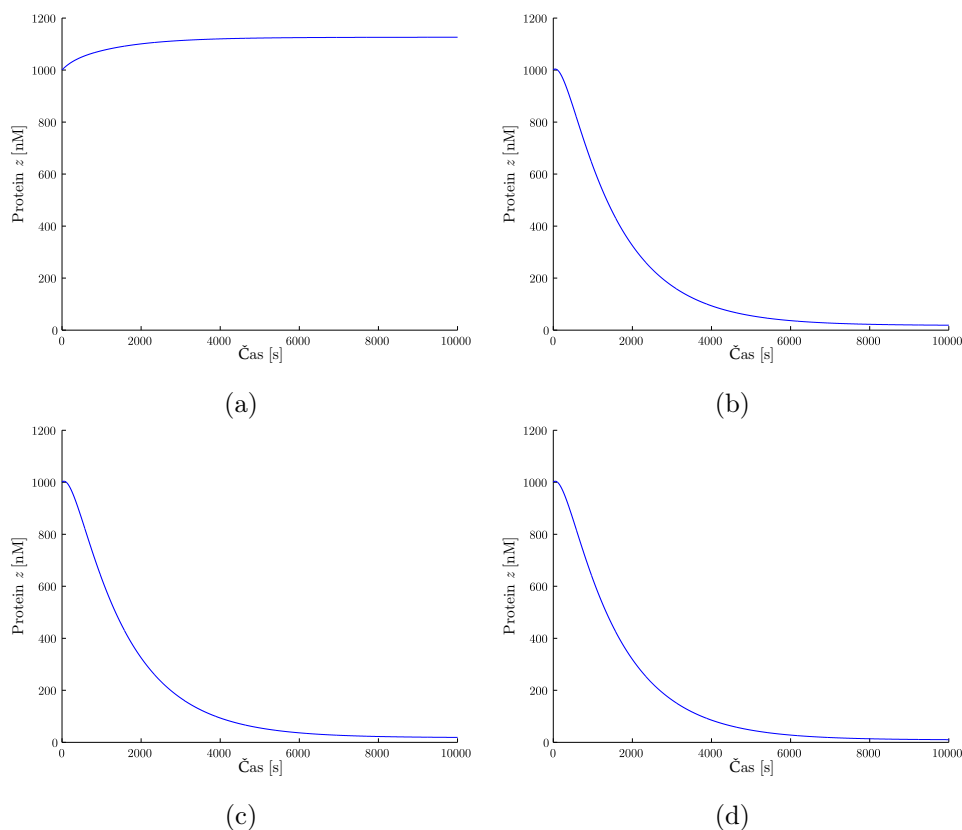
**Slika 4.14** Deterministično modeliranje bioloških NOR vrat na podlagi Hillovih enačb, pri čemer graf prikazuje koncentracijo izhodnega proteina  $z$  v odvisnosti od koncentracij vhodnih proteinov  $x$  in  $y$ , ko sistem doseže ravnovesno stanje.

Po drugi strani lahko opazujemo časovni potek koncentracije izhodnega proteina  $z$  pri določenih koncentracijah vhodnih proteinov  $x$  in  $y$  (slika 4.15 prikazuje časovni potek pri štirih različnih koncentracijah vhodnih proteinov).

### Stohastično modeliranje bioloških NOR vrat

Stohastičen model bioloških NOR vrat lahko postavimo na podlagi reakcij opisanih v tabeli 4.5. Glavno enačbo rešujemo z metodo SSA (glej razdelek 4.2.3).

Časovni potek koncentracije izhodnega proteina  $z$  opazujemo v odvisnosti od koncentracij vhodnih proteinov  $x$  in  $y$ . Pri tem predpostavljajmo, da je začetno število molekul proteina  $z$  v celici 1200 (približno 1000 nM), začetno število molekul mRNA pa 25 (približno 20 nM). Časovni potek koncentracij izhodnega proteina  $z$  pri štirih različnih



**Slika 4.15** Časovni potek koncentracij izhodnega proteina  $z$  pri štirih različnih koncentracijah vhodnih proteinov na podlagi determinističnega modela bioloških NOR vrat (Hillove enačbe), in sicer  $x, y = (0 \text{ nM}, 0 \text{ nM})$  na sliki (a),  $x, y = (0 \text{ nM}, 500 \text{ nM})$  na sliki (b),  $x, y = (500 \text{ nM}, 0 \text{ nM})$  na sliki (c) in  $x, y = (500 \text{ nM}, 500 \text{ nM})$  na sliki (d).

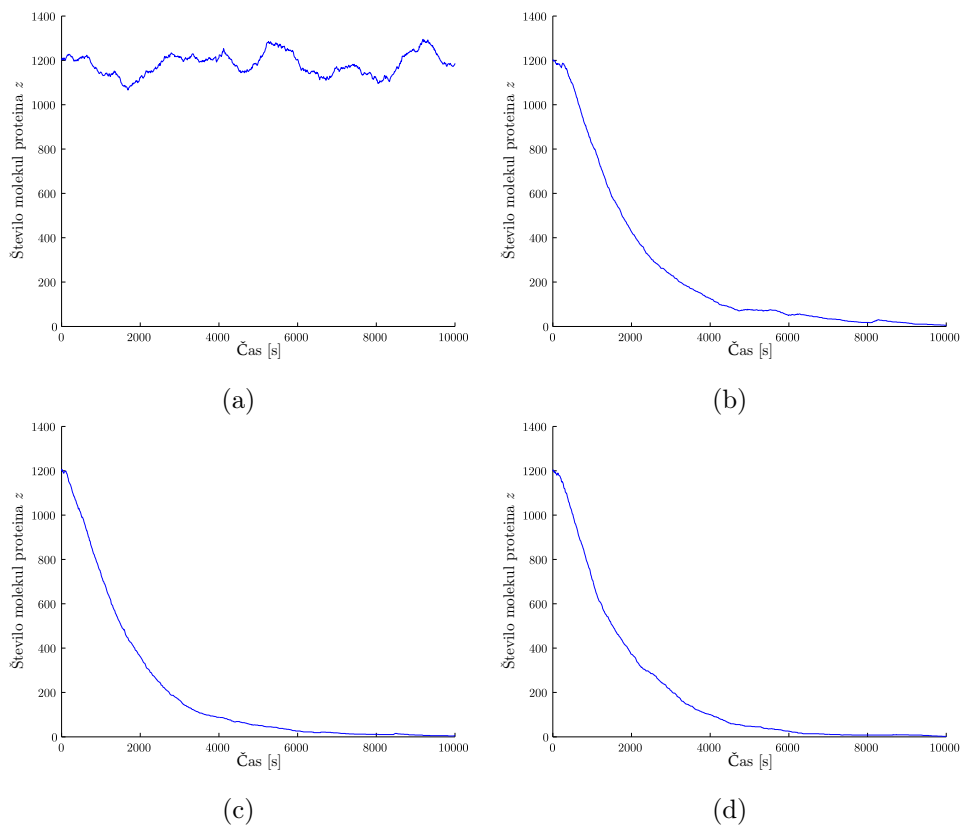
koncentracijah vhodnih proteinov v stohastičnem modelu bioloških NOR vrat prikazuje slika 4.16.

#### 4.4.6 Modeliranje preklopnega stikala

Izhajamo iz modificiranega modela preklopnega mehanizma *bakteriofaga*  $\lambda$ , ki določa ali je virus v stanju *lizogenije* ali *lize* [1, 47]. Posamezen protein zavira ekspresijo drugega proteina preko vezave na njegovo operatorsko mesto  $OR_2$ . Poleg tega v modelu upoštevamo prisotnost zunanjih vplivov v obliki transkripcijskih faktorjev, ki zavirata izražanje posameznega proteina preko vezave na operatorsko mesto  $OR_3$ . Pri tem predpostavljamo, da transkripcijski faktor  $i$  zavira izražanje proteina  $x$ , transkripcijski faktor  $j$  pa izražanje proteina  $y$ . Modeliran sistem se torej ujema s preklopnim stikalom z zunan-



#### 4.4 PRIMERI MODELIRANJA ENOSTAVNIH BIOLOŠKIH SISTEMOV ZMOŽNIH PROCESIRANJA PODATKOV 69



**Slika 4.16** Časovni potek koncentracij izhodnega proteina  $z$  pri štirih različnih koncentracijah vhodnih proteinov na podlagi stohastičnega modela bioloških NOR vrat, in sicer  $x, y = (0, 0)$  na sliki (a),  $x, y = (0, 600)$  na sliki (b),  $x, y = (600, 0)$  na sliki (c) in  $x, y = (600, 600)$  na sliki (d).

jima vplivoma, ki zavirata izražanje posameznega proteina (glej sliko 3.7 (b)). Reakcije upoštevane v modelu preklopnega stikala so navedene v tabeli 4.6.

#### Deterministično modeliranje preklopnega stikala

Izhajamo iz Hillove enačbe represorja s kooperativnostjo vezave transkripcijskega faktorja na promotor 2 (glej razdelek 4.1.5), pri čemer predpostavljamo, da v sistemu nastopata dva represorja. Dinamiko sistema lahko opišemo s sledečimi diferencialnimi enačbami:

$$\frac{dm_x}{dt} = \frac{\alpha_x}{1 + \left(\frac{y}{K_{d2}}\right)^\gamma + \left(\frac{i}{K_{d3}}\right)^\gamma} + \alpha_{0_x} - m_x \cdot \delta_{m_x}, \quad (4.72)$$

$$\frac{dm_y}{dt} = \frac{\alpha_y}{1 + \left(\frac{x}{K_{d2}}\right)^\gamma + \left(\frac{j}{K_{d3}}\right)^\gamma} + \alpha_{0_y} - m_y \cdot \delta_{m_y}, \quad (4.73)$$

	reakcija	pomen
1:	$2x + \text{DNA}_y \xrightarrow{k_{v2}} x_2 \text{DNA}_y$	vezava dimera proteina $x$ na operator $O_{R2}$ proteina $y$
2:	$2y + \text{DNA}_x \xrightarrow{k_{v2}} y_2 \text{DNA}_x$	vezava dimera proteina $y$ na operator $O_{R2}$ proteina $x$
3:	$x_2 \text{DNA}_y \xrightarrow{k_{c2}} 2x + \text{DNA}_y$	disociacija dimera proteina $x$ z operatorja $O_{R2}$ proteina $y$
4:	$y_2 \text{DNA}_x \xrightarrow{k_{c2}} 2y + \text{DNA}_x$	disociacija dimera proteina $y$ z operatorja $O_{R2}$ proteina $x$
5:	$2i + \text{DNA}_x \xrightarrow{k_{v3}} i_2 \text{DNA}_x$	vezava dimera proteina $i$ na operator $O_{R3}$ proteina $x$
6:	$2j + \text{DNA}_y \xrightarrow{k_{v3}} j_2 \text{DNA}_y$	vezava dimera proteina $j$ na operator $O_{R3}$ proteina $y$
7:	$i_2 \text{DNA}_x \xrightarrow{k_{c3}} 2i + \text{DNA}_x$	disociacija dimera proteina $i$ z operatorja $O_{R3}$ proteina $x$
8:	$j_2 \text{DNA}_y \xrightarrow{k_{c3}} 2j + \text{DNA}_y$	disociacija dimera proteina $j$ z operatorja $O_{R3}$ proteina $y$
9:	$\text{DNA}_x \xrightarrow{k_{tr_s}} \text{DNA}_x + \text{mRNA}_x$	transkripcija mRNA proteina $x$
10:	$\text{DNA}_y \xrightarrow{k_{tr_s}} \text{DNA}_y + \text{mRNA}_y$	transkripcija mRNA proteina $y$
11:	$\text{mRNA}_x \xrightarrow{k_{tr_l}} \text{mRNA}_x + x$	translacija proteina $x$
12:	$\text{mRNA}_y \xrightarrow{k_{tr_l}} \text{mRNA}_y + y$	translacija proteina $y$
13:	$x \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $x$
14:	$y \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $y$
15:	$x_2 \text{DNA}_y \xrightarrow{k_{deg}} \text{DNA}_y$	razgradnja dimera proteina $x$ vezanega na operator $O_{R3}$ proteina $y$
16:	$y_2 \text{DNA}_x \xrightarrow{k_{deg}} \text{DNA}_x$	razgradnja dimera proteina $y$ vezanega na operator $O_{R3}$ proteina $x$
17:	$\text{mRNA}_x \xrightarrow{k_{deg}^m} \emptyset$	razgradnja mRNA proteina $x$
18:	$\text{mRNA}_y \xrightarrow{k_{deg}^m} \emptyset$	razgradnja mRNA proteina $y$

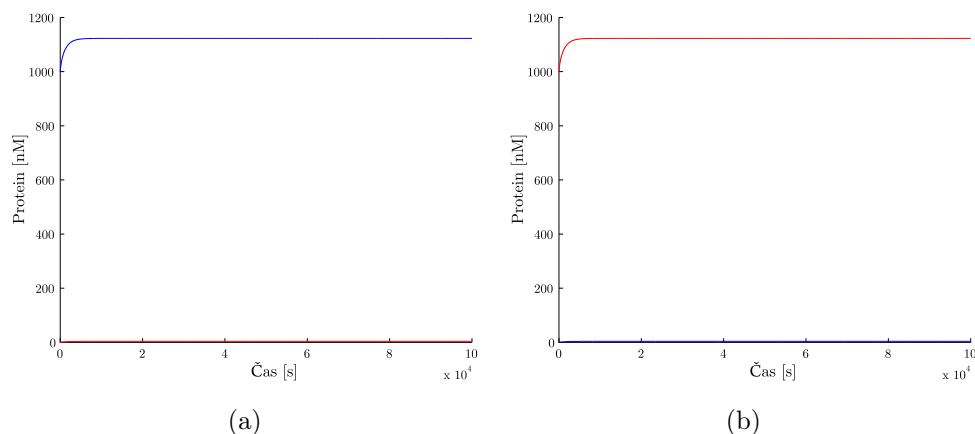
Tabela 4.6 Opazovane kemijske reakcije v modelu preklopnega stikala.

$$\frac{dx}{dt} = \beta_x \cdot m_x - \delta_x \cdot x, \quad (4.74)$$

$$\frac{dy}{dt} = \beta_y \cdot m_y - \delta_y \cdot y, \quad (4.75)$$

kjer  $i$  predstavlja koncentracijo vhodnega represorja, ki zavira izražanje proteina  $x$ ,  $j$  pa vhodnega represorja, ki zavira izražanje proteina  $y$ .

Opazujemo časovni potek koncentracij proteinov  $x$  in  $y$ , pri čemer izhajamo iz dveh različnih veljavnih začetnih stanj. V prvem primeru predpostavljajmo, da je koncentracija proteina  $x$  visoka (1000 nM), koncentracija proteina  $y$  pa nizka (0 nM) - glej sliko 4.17 (a), v drugem pa, da je koncentracija proteina  $x$  nizka (0 nM) in koncentracija proteina  $y$  visoka (1000 nM) - glej sliko 4.17 (b). V obeh primerih smo predpostavili, da vhodna proteina  $i$  in  $j$  v sistemu nista prisotna.

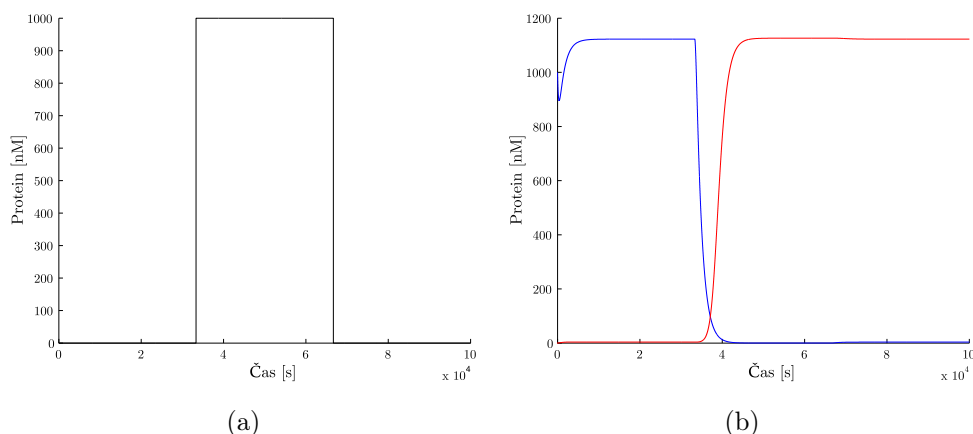


**Slika 4.17** Deterministično modeliranje biološkega preklopnega stikala na podlagi Hillovih enačb, pri čemer grafa prikazujeta koncentraciji proteinov  $x$  (modra) in  $y$  (rdeča) pri dveh različnih začetnih stanjih.

Opazujemo lahko tudi časovni potek koncentracij proteinov v odvisnosti od prisotnosti zunanjih proteinov. Slika 4.18 (a) prikazuje prisotnost zunanjega proteina  $i$  (predpostavljajmo odsotnost zunanjega proteina  $j$ ), slika 4.18 (b) pa časovni potek koncentracij proteinov v odvisnosti od prisotnosti zunanjih proteinov.

### Stohastično modeliranje preklopnega stikala

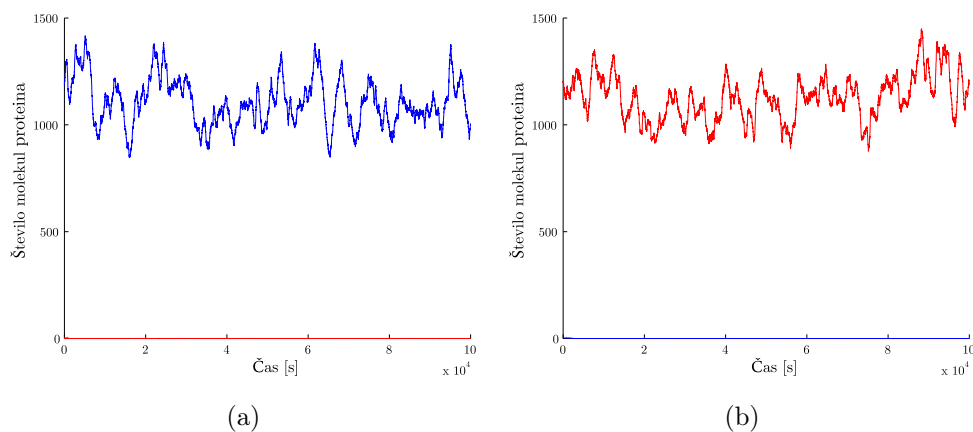
Stohastičen model razširjenega biološkega preklopnega stikala lahko postavimo na osnovi reakcij predstavljenih v tabeli 4.6. Vzpostavljeno glavno enačbo rešujemo z metodo SSA



**Slika 4.18** Deterministično modeliranje biološkega preklopnega stikala na podlagi Hillovih enačb z upoštevanjem prisotnosti zunanjih proteinov  $i$  in  $j$ . Slika (a) prikazuje prisotnost zunanjega proteina  $i$  ( $j$  je cel čas simuliranja odsoten), slika (b) pa časovni potek koncentracij proteinov  $x$  (modra) in  $y$  (rdeča).

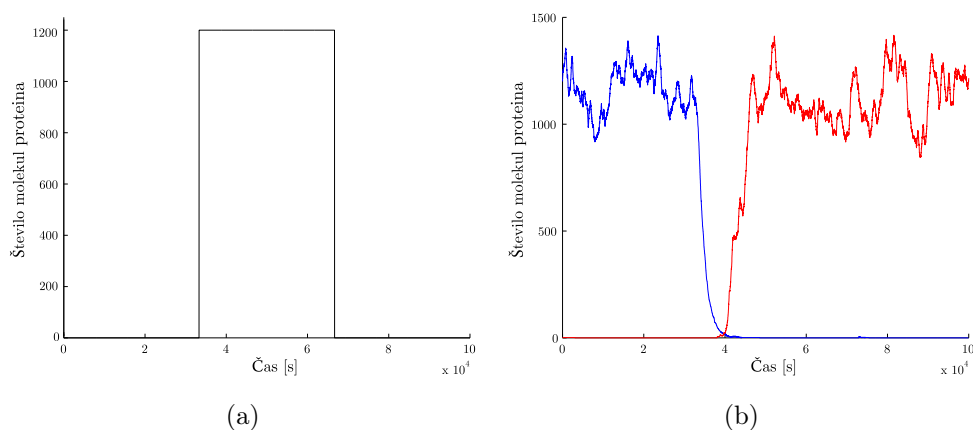
(glej razdelek 4.2.3).

Opazujemo časovni potek koncentracij proteinov  $x$  in  $y$ , pri čemer izhajamo iz dveh različnih veljavnih začetnih stanj. V prvem primeru predpostavljamo, da je koncentracija proteina  $x$  visoka (1200 molekul proteina  $x$ ), koncentracija proteina  $y$  pa nizka (0 molekul proteina  $y$ ) - glej sliko 4.19 (a), v drugem pa, da je koncentracija proteina  $x$  nizka (0 molekul proteina  $x$ ) in koncentracija proteina  $y$  visoka (1200 molekul proteina  $y$ ) - glej sliko 4.19 (b). V obeh primerih predpostavljamo, da vhodna proteina  $i$  in  $j$  v sistemu nista prisotna.



**Slika 4.19** Časovni potek koncentracij proteinov  $x$  (modra) in  $y$  (rdeča) pri dveh različnih začetnih stanjih brez prisotnosti zunanjih proteinov  $i$  in  $j$  na podlagi stohastičnega modeliranja biološkega preklopnega stikala.

Časovni potek koncentracij proteinov v odvisnosti od prisotnosti zunanjih proteinov v stohastičnem modelu prikazuje slika 4.20. Pri tem slika (a) prikazuje prisotnost zunanjega proteina  $i$  (predpostavljajmo odsotnost zunanjega proteina  $j$ ), slika (b) pa časovni potek koncentracij proteinov v odvisnosti od zunanjih proteinov.



**Slika 4.20** Stohastično modeliranje biološkega preklopnega stikala z upoštevanjem prisotnosti zunanjih proteinov  $i$  in  $j$ . Slika (a) prikazuje prisotnost zunanjega proteina  $i$  ( $j$  je cel čas simuliranja odsoten), slika (b) pa časovni potek koncentracij proteinov  $x$  (modra) in  $y$  (rdeča).

#### 4.4.7 Modeliranje razširjenega preklopnega stikala

Prav tako kot v razdelku 4.4.6 tudi tu izhajamo iz modificiranega modela preklopnega mehanizma *bakteriofaga*  $\lambda$ , ki določa, ali je virus v stanju lizogenije ali lize [1, 47]. Posamezni protein zavira ekspresijo drugega proteina preko vezave na njegovo operatorsko mesto  $O_{R2}$  in ekspresijo samega sebe preko vezave na svoje operatorsko mesto  $O_{R3}$ . Izhajamo torej iz sistema predstavljenega na sliki 3.8. Modeliran sistem lahko predstavimo z reakcijami v tabeli 4.7. V modelu upoštevamo prisotnost zunanjih vplivov in s tem zmožnost simuliranja preklopa sistema iz enega stabilnega stanja v drugo oziroma osveževanja trenutnega stanja, če je to potrebno z vpeljavo zunanjih vplivom, ki povečata hitrost razgradnje opazovanih proteinov. Predpostavljamo, da signala  $i$  in  $j$  predstavljata prisotnost zunanjega vpliva (npr. svetlobe), ki poveča hitrost razgradnje posameznega proteina na  $1s^{-1}$ , in sicer signal  $i$  poveča hitrost razgradnje proteinu  $x$ , signal  $j$  pa hitrost razgradnje proteinu  $y$ .

### Deterministično modeliranje razširjenega preklopnega stikala

Izhajamo iz Hilllove enačbe represorja, pri čemer upoštevamo, da se izražanje posameznega gena ustavi ob vezavi transkripcijskega faktorja bodisi na operator  $O_{R2}$  bodisi na operator  $O_{R3}$ . Deterministično dinamiko sistema lahko opišemo z enačbami:

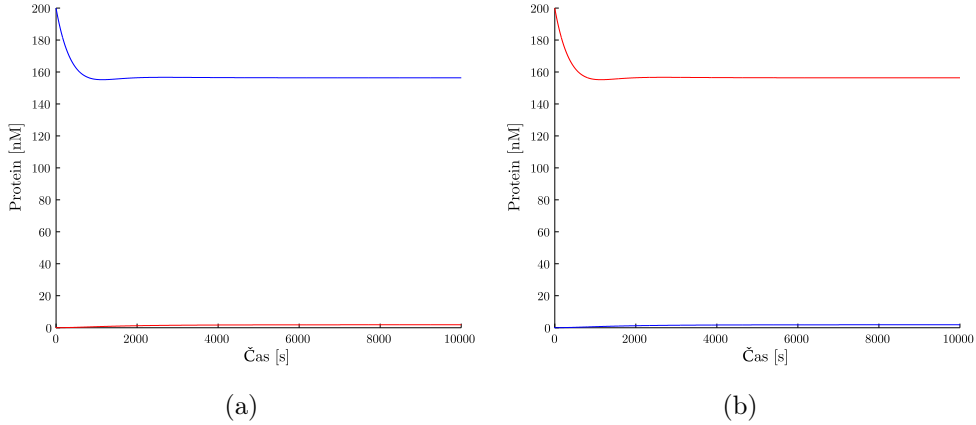
$$\frac{dm_x}{dt} = \frac{\alpha_x}{1 + \left(\frac{y}{K_{d2}}\right)^\gamma + \left(\frac{x}{K_{d3}}\right)^\gamma} + \alpha_{0x} - m_x \cdot \delta_{m_x}, \quad (4.76)$$

$$\frac{dm_y}{dt} = \frac{\alpha_y}{1 + \left(\frac{x}{K_{d2}}\right)^\gamma + \left(\frac{y}{K_{d3}}\right)^\gamma} + \alpha_{0y} - m_y \cdot \delta_{m_y}, \quad (4.77)$$

$$\frac{dx}{dt} = \beta_x \cdot m_x - \delta_x \cdot x, \quad (4.78)$$

$$\frac{dy}{dt} = \beta_y \cdot m_y - \delta_y \cdot y. \quad (4.79)$$

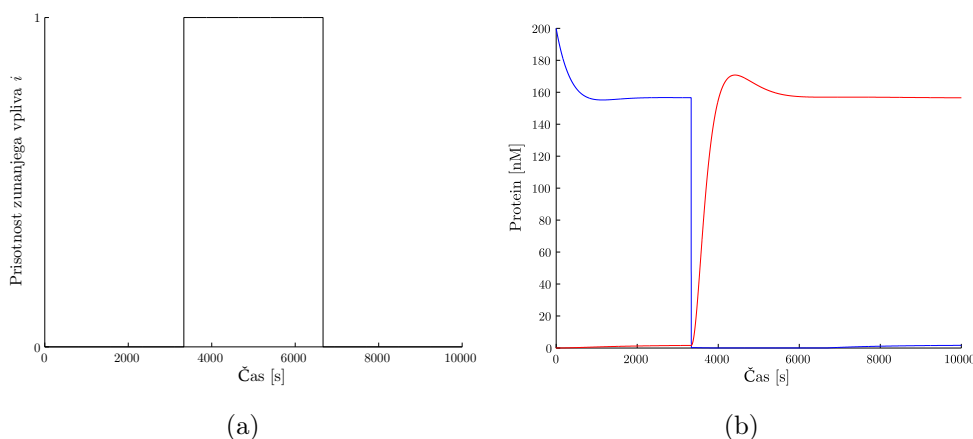
Opazujemo časovni potek koncentracij proteinov  $x$  in  $y$ , pri čemer izhajamo iz dveh različnih veljavnih začetnih stanj. V prvem primeru predpostavljajmo, da je koncentracija proteina  $x$  visoka (200 nM), koncentracija proteina  $y$  pa nizka (0 nM) - glej sliko 4.21 (a), v drugem pa, da je koncentracija proteina  $x$  nizka (0 nM) in koncentracija proteina  $y$  visoka (200 nM) - glej sliko 4.21 (b). Opazujemo lahko tudi časovni potek koncentracij



**Slika 4.21** Deterministično modeliranje razširjenega preklopnega stikala na podlagi Hillovih enačb, pri čemer grafa prikazujeta koncentraciji proteinov  $x$  (modra) in  $y$  (rdeča) pri dveh različnih začetnih stanjih.

proteinov v odvisnosti od prisotnosti zunanjih vplivov. Slika 4.22 (a) prikazuje prisotnost zunanjega vpliva  $i$  (predpostavljajmo odsotnost zunanjega vpliva  $j$ ), pri čemer vrednost

0 predstavlja odsotnost zunanje vpliva, vrednost 1 pa njegovo prisotnost, slika 4.22 (b) pa časovni potek koncentracij proteinov v odvisnosti od zunanje vpliva.

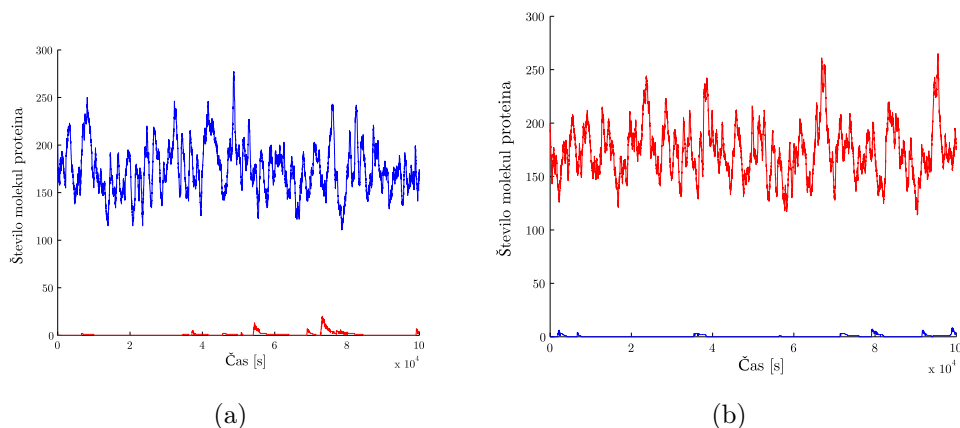


**Slika 4.22** Deterministično modeliranje razširjenega preklopnega stikala na podlagi Hillovih enačb z upoštevanjem prisotnosti zunanjih vplivov  $i$  in  $j$ . Slika (a) prikazuje prisotnost zunanje vpliva  $i$  ( $j$  je cel čas simuliranja odsoten), slika (b) pa časovni potek koncentracij proteinov  $x$  (modra) in  $y$  (rdeča).

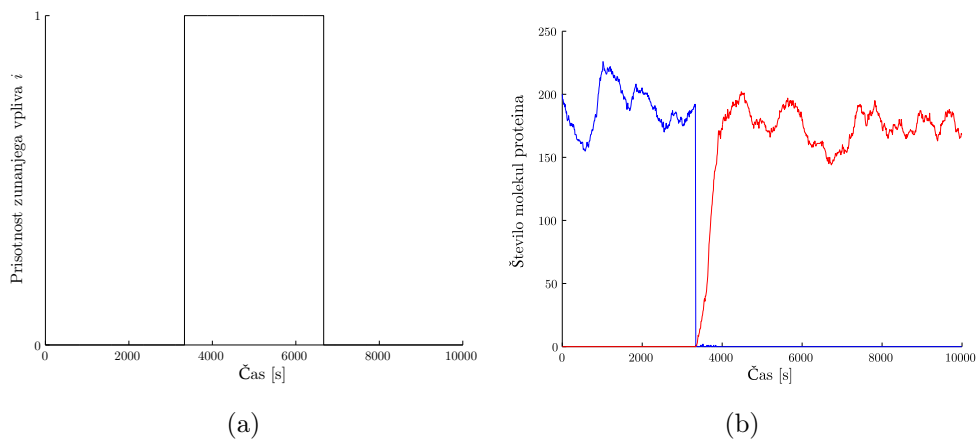
### Stohastično modeliranje razširjenega preklopnega stikala

Stohastičen model razširjenega preklopnega stikala lahko postavimo na osnovi reakcij predstavljenih v tabeli 4.7. Vzpostavljeno glavno enačbo rešujemo z metodo SSA (glej razdelek 4.2.3).

Opazujemo časovni potek koncentracij proteinov  $x$  in  $y$ , pri čemer izhajamo iz dveh različnih veljavnih začetnih stanj. V prvem primeru predpostavljamo, da je koncentracija proteina  $x$  visoka (200 molekul proteina  $x$ ), koncentracija proteina  $y$  pa nizka (0 molekul proteina  $y$ ) - glej sliko 4.23 (a), v drugem pa, da je koncentracija proteina  $x$  nizka ((0 molekul proteina  $x$ )) in koncentracija proteina  $y$  visoka (200 molekul proteina  $y$ ) - glej sliko 4.23 (b). Časovni potek koncentracij proteinov v odvisnosti od prisotnosti zunanjih vplivov v stohastičnem modelu prikazuje slika 4.24, pri čemer slika (a) prikazuje prisotnost zunanje vpliva  $i$  (predpostavljamo odsotnost zunanje vpliva  $j$ ), pri čemer vrednost 0 predstavlja odsotnost zunanje vpliva, vrednost 1 pa njegovo prisotnost, slika (b) pa časovni potek koncentracij proteinov v odvisnosti od zunanje vpliva.



**Slika 4.23** Časovni potek koncentracij proteinov  $x$  (modra) in  $y$  (rdeča) pri dveh različnih začetnih stanjih na podlagi stohastičnega modeliranja razširjenega preklopnega stikala.



**Slika 4.24** Stohastično modeliranje razširjenega preklopnega stikala z upoštevanjem prisotnosti zunanjih vplivov  $i$  in  $j$ . Slika (a) prikazuje prisotnost zunanjega vpliva  $i$  ( $j$  je cel čas simuliranja odsoten), slika (b) pa časovni potek koncentracij proteinov  $x$  (modra) in  $y$  (rdeča).

#### 4.4.8 Modeliranje represilatorja

Negativno povratno zanko represilatorja lahko dosežemo z gensko regulatornim omrežjem treh krožno povezanih promotorjev, pri čemer je ob aktivnosti posameznega promotorja sintetiziran protein, ki zavira aktivnost promotorja, ki mu v genskem omrežju sledi (glej sliko 3.10). Do modela represilatorja lahko torej pridemo z razširitvijo modela negatorja (glej razdelek 4.4.3), in sicer tako, da v negativno povratno zanko vežemo tri negatorje.

Izhajamo iz osnovnega modela negatorja, le da tokrat predpostavljamo vezavo repre-



sorjev na operatorje  $O_R2$ . Podobno kot v razdelku 4.4.2 tudi pri modelu razširjenega preklopnega stikala reakcije dimerizacije združimo z reakcijami vezave dimera na operator, reakcije disociacije dimera na monomera pa z reakcijami disociacije dimera z operatorja.

Modeliran sistem lahko predstavimo z reakcijami v tabeli 4.8.

### Deterministično modeliranje represilatorja

Spet izhajamo iz Hillove enačbe represorja, pri čemer se na operator veže dimerizirana oblika proteina, ki je produkt prejšnjega gena represilatorja. Deterministično dinamiko sistema lahko opišemo z enačbami:

$$\frac{dm_x}{dt} = \frac{\alpha_x}{1 + \left(\frac{z}{K_{d_2}}\right)^\gamma} + \alpha_{0_x} - m_x \cdot \delta_{m_x}, \quad (4.80)$$

$$\frac{dm_y}{dt} = \frac{\alpha_y}{1 + \left(\frac{x}{K_{d_2}}\right)^\gamma} + \alpha_{0_y} - m_y \cdot \delta_{m_y}, \quad (4.81)$$

$$\frac{dm_z}{dt} = \frac{\alpha_z}{1 + \left(\frac{y}{K_{d_2}}\right)^\gamma} + \alpha_{0_z} - m_z \cdot \delta_{m_z}, \quad (4.82)$$

$$\frac{dx}{dt} = \beta_x \cdot m_x - \delta_x \cdot x, \quad (4.83)$$

$$\frac{dy}{dt} = \beta_y \cdot m_y - \delta_y \cdot y, \quad (4.84)$$

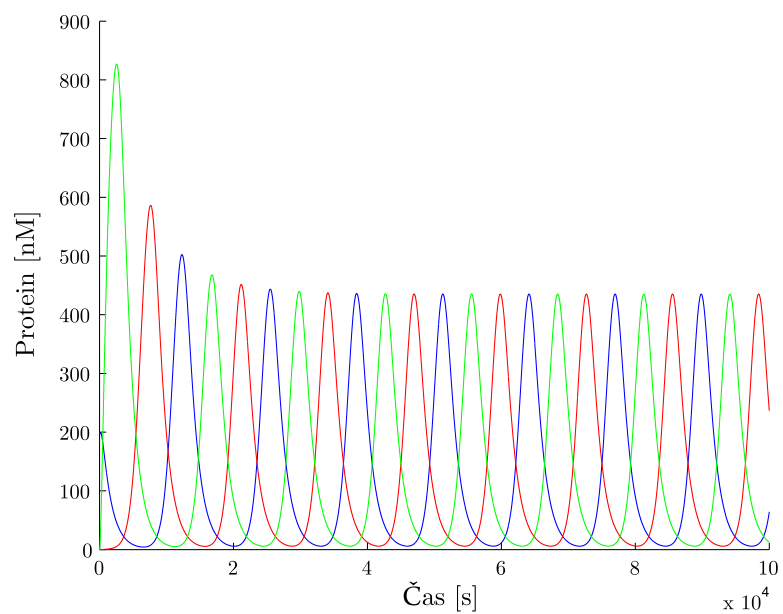
$$\frac{dz}{dt} = \beta_z \cdot m_z - \delta_z \cdot z. \quad (4.85)$$

Opazujemo časovni potek koncentracij proteinov  $x$ ,  $y$  in  $z$ . Predpostavljajmo, da je začetna koncentracija proteina  $x$  visoka (200 nM), koncentracija proteinov  $y$  in  $z$  pa nizka (0 nM). Časovni potek koncentracij proteinov prikazuje slika 4.25.

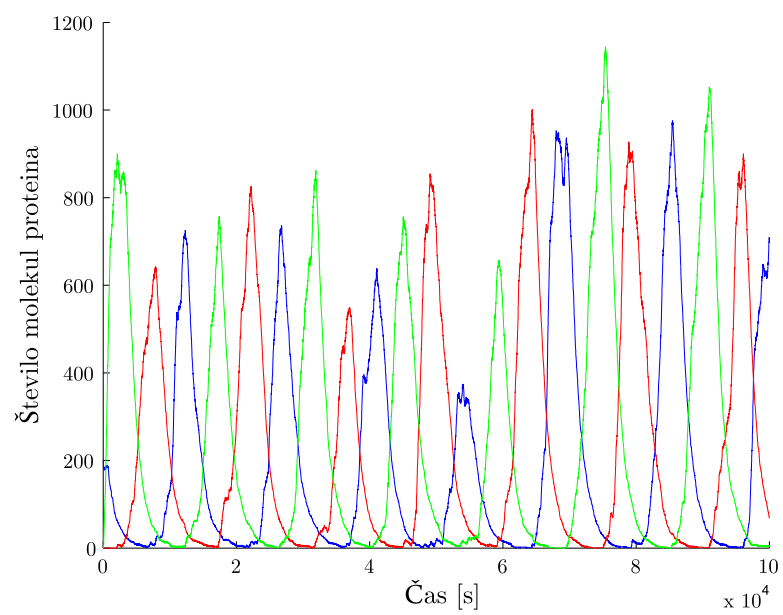
### Stohastično modeliranje represilatorja

Stohastičen model represilatorja lahko postavimo na osnovi reakcij predstavljenimi v tabeli 4.8. Vzpostavljeno glavno enačbo rešujemo z metodo SSA (glej razdelek 4.2.3).

Opazujemo časovni potek koncentracij proteinov  $x$ ,  $y$  in  $z$ , pri čemer izhajamo iz enakega začetnega stanja kot pri determinističnem modelu represilatorja (200 molekul proteina  $x$ ). Rezultati stohastičnega modela so prikazani na sliki 4.26.



**Slika 4.25** Deterministično modeliranje represilatorja na podlagi Hillovih enačb, pri čemer graf prikazuje potek koncentracij proteinov  $x$  (modra)  $y$  (rdeča) in  $z$  (zelena).



**Slika 4.26** Časovni potek koncentracij proteinov  $x$  (modra),  $y$  (rdeča) in  $z$  (zelena) v stohastičnem modelu represilatorja.

## 4.5 Zaključek poglavja

Tako z determinističnim kot tudi s stohastičnim modeliranjem odziva bioloških sistemov lahko pridemo do rezultatov, ki sovpadajo z realnim odzivom modeliranih bioloških sistemov. Kljub temu za naše potrebe potrebujemo rezultate obeh. Deterministični modeli po eni strani omogočajo analizo dinamike modeliranega sistema na podlagi vzpostavljenih sistemov diferencialnih enačb, pri čemer na celoten biološki sistem gledamo kot na nelinearen dinamičen sistem. V nadaljevanju predstavimo t. i. stabilnostno analizo, ki temelji na rezultatih determinističnih modelov in je opisana v poglavju 5. Zaradi določenih slabosti determinističnih modelov je po drugi strani prav tako nepogrešljiva vzpostavitev stohastičnih modelov, ki omogočajo modeliranje stohastičnosti kemijskih reakcij. Pri ocenjevanju primernosti bioloških sistemov za procesiranje podatkov v poglavju 6 predstavimo vzpostavitev metrik, ki temeljijo tako na rezultatih determinističnega kot tudi na rezultatih stohastičnega modeliranja.

	reakcija	pomen
1:	$2x + \text{DNA}_y \xrightarrow{k_{v2}} x_2 \text{DNA}_y$	vezava dimera proteina $x$ na operator $O_{R2}$ proteina $y$
2:	$2y + \text{DNA}_x \xrightarrow{k_{v2}} y_2 \text{DNA}_x$	vezava dimera proteina $y$ na operator $O_{R2}$ proteina $x$
3:	$x_2 \text{DNA}_y \xrightarrow{k_{c2}} 2x + \text{DNA}_y$	disociacija dimera proteina $x$ z operatorja $O_{R2}$ proteina $y$
4:	$y_2 \text{DNA}_x \xrightarrow{k_{c2}} 2y + \text{DNA}_x$	disociacija dimera proteina $y$ z operatorja $O_{R2}$ proteina $x$
5:	$2x + \text{DNA}_x \xrightarrow{k_{v3}} x_2 \text{DNA}_x$	vezava dimera proteina $x$ na operator $O_{R3}$ proteina $x$
6:	$2y + \text{DNA}_y \xrightarrow{k_{v3}} y_2 \text{DNA}_y$	vezava dimera proteina $y$ na operator $O_{R3}$ proteina $y$
7:	$x_2 \text{DNA}_x \xrightarrow{k_{c3}} 2x + \text{DNA}_x$	disociacija dimera proteina $x$ z operatorja $O_{R3}$ proteina $x$
8:	$y_2 \text{DNA}_y \xrightarrow{k_{c3}} 2y + \text{DNA}_y$	disociacija dimera proteina $y$ z operatorja $O_{R3}$ proteina $y$
9:	$\text{DNA}_x \xrightarrow{k_{tr_s}} \text{DNA}_x + \text{mRNA}_x$	transkripcija mRNA proteina $x$
10:	$\text{DNA}_y \xrightarrow{k_{tr_s}} \text{DNA}_y + \text{mRNA}_y$	transkripcija mRNA proteina $y$
11:	$\text{mRNA}_x \xrightarrow{k_{tr_l}} \text{mRNA}_x + x$	translacija proteina $x$
12:	$\text{mRNA}_y \xrightarrow{k_{tr_l}} \text{mRNA}_y + y$	translacija proteina $y$
13:	$x \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $x$
14:	$y \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $y$
15:	$x_2 \text{DNA}_y \xrightarrow{k_{deg}} \text{DNA}_y$	razgradnja dimera proteina $x$ vezanega na operator $O_{R2}$ proteina $y$
16:	$y_2 \text{DNA}_x \xrightarrow{k_{deg}} \text{DNA}_x$	razgradnja dimera proteina $y$ vezanega na operator $O_{R2}$ proteina $x$
17:	$x_2 \text{DNA}_x \xrightarrow{k_{deg}} \text{DNA}_x$	razgradnja dimera proteina $x$ vezanega na operator $O_{R3}$ proteina $x$
18:	$y_2 \text{DNA}_y \xrightarrow{k_{deg}} \text{DNA}_y$	razgradnja dimera proteina $y$ vezanega na operator $O_{R3}$ proteina $y$
19:	$\text{mRNA}_x \xrightarrow{k_{deg_m}} \emptyset$	razgradnja mRNA proteina $x$
20:	$\text{mRNA}_y \xrightarrow{k_{deg_m}} \emptyset$	razgradnja mRNA proteina $y$

Tabela 4.7 Opazovane kemijske reakcije v modelu razširjenega preklopnega stikala.

	reakcija	pomen
1:	$2x + \text{DNA}_y \xrightarrow{k_{v2}} x_2 \text{ DNA}_y$	vezava dimera proteina $x$ na operator $O_R2$ proteina $y$
2:	$2y + \text{DNA}_z \xrightarrow{k_{v2}} y_2 \text{ DNA}_z$	vezava dimera proteina $y$ na operator $O_R2$ proteina $z$
3:	$2z + \text{DNA}_x \xrightarrow{k_{v2}} z_2 \text{ DNA}_x$	vezava dimera proteina $z$ na operator $O_R2$ proteina $x$
4:	$x_2 \text{ DNA}_y \xrightarrow{k_{c2}} 2x + \text{DNA}_y$	disociacija dimera proteina $x$ z operatorja $O_R2$ proteina $y$
5:	$y_2 \text{ DNA}_z \xrightarrow{k_{c2}} 2y + \text{DNA}_z$	disociacija dimera proteina $y$ z operatorja $O_R2$ proteina $z$
6:	$z_2 \text{ DNA}_x \xrightarrow{k_{c2}} 2z + \text{DNA}_x$	disociacija dimera proteina $z$ z operatorja $O_R2$ proteina $x$
7:	$\text{DNA}_x \xrightarrow{k_{tr\bar{s}}} \text{DNA}_x + \text{mRNA}_x$	transkripcija mRNA proteina $x$
8:	$\text{DNA}_y \xrightarrow{k_{tr\bar{s}}} \text{DNA}_y + \text{mRNA}_y$	transkripcija mRNA proteina $y$
9:	$\text{DNA}_z \xrightarrow{k_{tr\bar{s}}} \text{DNA}_z + \text{mRNA}_z$	transkripcija mRNA proteina $z$
10:	$\text{mRNA}_x \xrightarrow{k_{tr\bar{l}}} \text{mRNA}_x + x$	translacija proteina $x$
11:	$\text{mRNA}_y \xrightarrow{k_{tr\bar{l}}} \text{mRNA}_y + y$	translacija proteina $y$
12:	$\text{mRNA}_z \xrightarrow{k_{tr\bar{l}}} \text{mRNA}_z + z$	translacija proteina $z$
13:	$x \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $x$
14:	$y \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $y$
15:	$z \xrightarrow{k_{deg}} \emptyset$	razgradnja proteina $z$
16:	$x_2 \text{ DNA}_y \xrightarrow{k_{deg}} \text{DNA}_y$	razgradnja dimera proteina $x$ vezanega na operator $O_R2$ proteina $y$
17:	$y_2 \text{ DNA}_z \xrightarrow{k_{deg}} \text{DNA}_z$	razgradnja dimera proteina $y$ vezanega na operator $O_R2$ proteina $z$
18:	$z_2 \text{ DNA}_x \xrightarrow{k_{deg}} \text{DNA}_x$	razgradnja dimera proteina $z$ vezanega na operator $O_R2$ proteina $x$
19:	$\text{mRNA}_x \xrightarrow{k_{deg^m}} \emptyset$	razgradnja mRNA proteina $x$
20:	$\text{mRNA}_y \xrightarrow{k_{deg^m}} \emptyset$	razgradnja mRNA proteina $y$
21:	$\text{mRNA}_z \xrightarrow{k_{deg^m}} \emptyset$	razgradnja mRNA proteina $z$

Tabela 4.8 Opazovane kemijske reakcije v modelu represilatorja.



# 5 Analiza stabilnosti bioloških sistemov

Stanje opazovanega biološkega sistema lahko predstavimo s koncentracijami opazovanih kemijskih zvrsti ob določenem času z vektorjem pozitivnih realnih števil

$$\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t)), \quad (5.1)$$

pri čemer je  $x_i(t)$  ( $i = 1, 2, \dots, n$ ) koncentracija kemijske zvrsti  $i$  ob času  $t$ ,  $n$  pa število opazovanih kemijskih zvrsti. Na podlagi rezultatov postavljenih modelov lahko za posamezen biološki sistem opazujemo časovno evolucijo stanja sistema (tj. časovni potek opazovanih koncentracij) pri postavljenih začetnih pogojih. V določenih primerih nas zanima *asimptotično obnašanje sistema*, tj. stanje sistema po daljšem času ( $t \rightarrow \infty$ ) brez spreminjanja zunanjih vplivov. V takem primeru lahko sistem odraža obnašanje, ki je v skladu z eno od sledečih možnosti [51]:

1. sistem sčasoma preide v *ravnovesno stanje* (angl. *steady state*),
2. sistem sčasoma preide v oscilatorno delovanje - sistem odraža *stabilne periodične oscilacije*,
3. sistem ne konvergira proti nobeni stabilni rešitvi,

4. sistem odraža kaotično obnašanje (minimalna sprememba v začetnih pogojih povzroči velike spremembe v obnašanju sistema [52]).

Poleg tega, da do obnašanja bioloških sistemov, ki sovpadajo z zadnjima dvema možnostima, pride zelo redko, sta le-ti za nas nezanimivi, saj ne omogočata predvidljivega obnašanja, na podlagi katerega bi lahko gradili kompleksnejše logične gradnike zmožne procesiranja podatkov. Po drugi strani so lahko sistemi iz skupine 1 osnova za izgradnjo logičnih gradnikov, sistemi iz skupine 2 pa osnova oscilatornih elementov znotraj bioloških vezij.

V nadaljevanju sledi podrobnejša predstavitev analize za nas zanimivih sistemov, tj. sistemov, ki sčasoma preidejo v neko ravnovesno stanje (skupina 1), in sistemov, ki izkazujejo stabilne periodične oscilacije (skupina 2). Poleg tega predstavimo analizo parametrskih vrednosti, pri katerih sistemi pripadajo posamezni skupini, t. i. *bifurkacijsko analizo*.

## 5.1 Analiza bioloških sistemov z ravnovesnimi stanji

Spremembe koncentracij posameznih kemijskih zvrsti v biološkem sistemu lahko opišemo z determinističnim modelom, ki temelji na sistemu navadnih diferencialnih enačb (glej razdelek 4.1)

$$\frac{dx_i(t)}{dt} = f(x_1(t), x_2(t), \dots, x_n(t)), \quad (5.2)$$

pri čemer  $x_i(t)$  predstavlja koncentracijo kemijske zvrsti  $i$  v času  $t$  ( $i = 1, \dots, n$ ),  $n$  pa število opazovanih zvrsti v sistemu.

*Ravnovesno stanje* sistema je določeno s stanjem sistema  $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$ , v katerem so vse nadaljnje spremembe koncentracij enake nič, tj. vsi odvodi, ki določajo spremembo koncentracij po času, so enaki nič [53]:

$$\frac{dx_i(t)}{dt} = 0, i = 1, \dots, n. \quad (5.3)$$

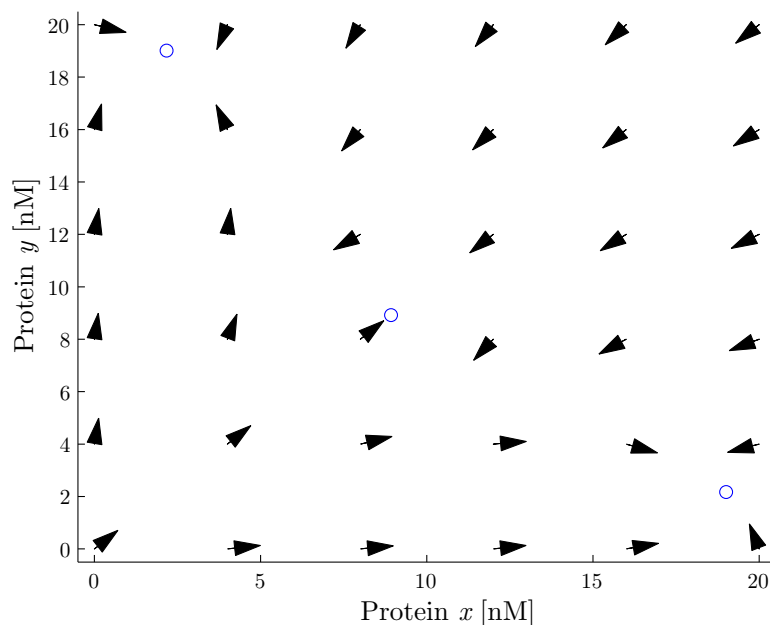
V splošnem poznamo dva tipa ravnovesnih stanj sistema, in sicer

- *stabilno stanje*, proti kateremu sistem s časom konvergira (se mu približuje),
- *nestabilno stanje*, od katerega sistem s časom divergira (se od njega oddaljuje).

Gradienti sprememb koncentracij kažejo v smeri stabilnih stanj in stran od nestabilnih stanj. Verjetnost, da bo sistem po daljšem času ( $t \rightarrow \infty$ ) pristal v stabilnem stanju, je zelo velika. Po drugi strani je verjetnost, da bo pristal v nestabilnem stanju, zelo majhna.



Gradiente grafično ponazorimo s t. i. *smernim poljem* (angl. *direction field*). Primer smernega polja za sistem s tremi ravnovesnimi stanji, pri čemer je eno nestabilno, dve pa stabilni, prikazuje slika 5.1. Stabilnost posameznega ravnovesnega stanja lahko določimo



**Slika 5.1** Primer smernega polja gradientov za sistem s tremi ravnovesnimi stanji, ki so označena z modrimi krogi. Srednje ravnovesno stanje je nestabilno, saj nekateri okoliški gradienti kažejo stran od njega, robni pa stabilni, saj vsi okoliški gradienti kažejo proti njima.

s postopkom navedenim v nadaljevanju.

### 5.1.1 Izračun Jacobijeve matrice

Jacobijeva matrika je sestavljena iz parcialnih odvodov diferencialnih enačb, ki predstavljajo spremembe koncentracij kemijskih zvrsti v sistemu. Določena je z izrazom

$$J(\mathbf{x}(t)) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}(t))}{\partial x_1}, & \frac{\partial f_1(\mathbf{x}(t))}{\partial x_2}, & \dots & \frac{\partial f_1(\mathbf{x}(t))}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x}(t))}{\partial x_1}, & \frac{\partial f_2(\mathbf{x}(t))}{\partial x_2}, & \dots & \frac{\partial f_2(\mathbf{x}(t))}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n(\mathbf{x}(t))}{\partial x_1}, & \frac{\partial f_n(\mathbf{x}(t))}{\partial x_2}, & \dots & \frac{\partial f_n(\mathbf{x}(t))}{\partial x_n} \end{bmatrix}. \quad (5.4)$$

### 5.1.2 Določitev stabilnosti ravnovesne točke

Naj stanje koncentracij opazovanih kemijskih zvrsti  $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$  predstavlja eno od ravnovesnih točk sistema. Stabilnost posamezne ravnovesne točke  $\mathbf{x}^*$  določimo tako,

da izračunamo lastne vrednosti Jacobijeve matrike pri le-tej po izrazu

$$|J(\mathbf{x}^*) - \lambda I| = 0, \quad (5.5)$$

kjer je  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$  vektor lastnih vrednosti,  $I$  enotska matrika, operator  $|\cdot|$  pa določa determinanto matrike  $\cdot$ .

Ravnovesna točka  $\mathbf{x}^*$  predstavlja stabilno točko, če imajo vse lastne vrednosti  $\lambda_1, \lambda_2, \dots, \lambda_n$  v tej točki realni del manjši od nič. Za stabilne točke mora torej veljati pogoj

$$\Re(\lambda_i) < 0, \forall i \in 1 \dots n. \quad (5.6)$$

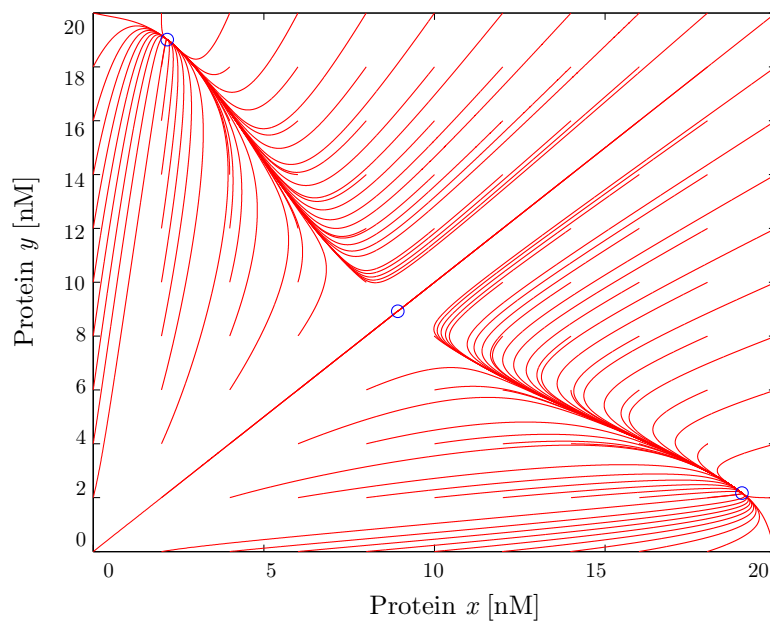
### 5.1.3 Prikazovanje stabilnosti

Evolucijo koncentracij opazovanih kemijskih zvrsti ponavadi prikazujemo na grafu v odvisnosti od časa, pri analizi stabilnosti oziroma prikazovanju stabilnih točk pa je bolj smiselno, da na osi grafa nanašamo koncentracije kemijskih zvrsti pri najpogostejših trajektorijah, ki se pojavljajo v evoluciji sistema. Takemu prikazu pravimo prikaz v t. i. *faznem prostoru* (angl. *phase space diagram*), oziroma na *fazni ravnini* (angl. *phase plane*), če opazujemo zgolj 2 dimenziji [54]. Točke, proti katerim konvergira stanje sistema, so iz takega prikaza zelo dobro razvidne.

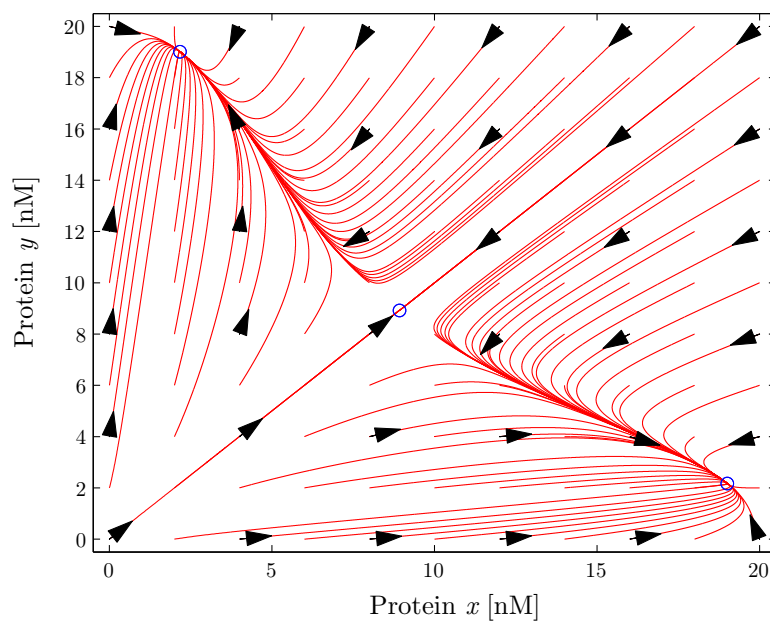
Primer prikaza obnašanja sistema na fazni ravnini predstavlja slika 5.2. Konvergenco koncentracij lahko prikazujemo tudi na že prej omenjenem smernem grafu (glej sliko 5.1). Smiselno je, da prikaza združimo v skupen graf, ki ga prikazuje slika 5.3.

## 5.2 Analiza oscilatornih sistemov

Oscilacije so v sistemu pogojene z obstojem *limitnega cikla* (angl. *limit cycle*). Limitni cikel lahko v fazni ravnini predstavimo z izolirano zaprto trajektorno krivuljo brez lastnih presečišč (angl. *isolated simple oriented closed curve trajectory*), ki ne vsebuje stabilnih ravnovesnih točk, proti katerim bi stanje sistema lahko konvergiralo, tj. *singularnih točk*. Limitni cikel je poleg tega pogojen s tem, da sosednje trajektorije konvergirajo proti njemu bodisi z naraščanjem časa ( $t \rightarrow \infty$ ) bodisi z zmanjševanjem časa ( $t \rightarrow -\infty$ ). V prvem primeru je tak limitni cikel *stabilen*, v drugem pa *nestabilen*. Analizo obstoja oscilacij v nekem biološkem sistemu lahko torej prevedemo na analizo obstoja limitnega cikla [53]. Za nas najbolj zanimiv je obstoj stabilnega limitnega cikla, saj povzroči konvergenco stanja sistema proti njemu in posledično samostojno vzdrževanje oscilacij.

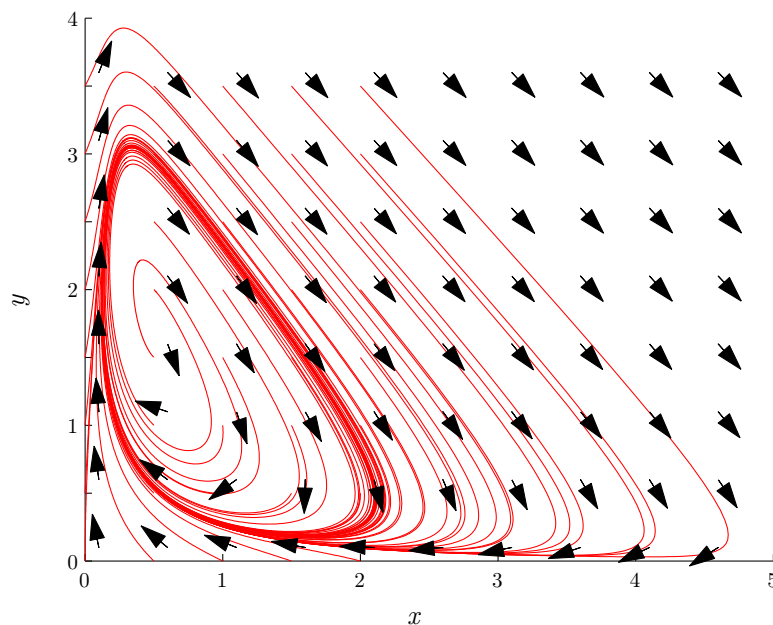


**Slika 5.2** Primer prikaza evolucije sistema s tremi ravnovesnimi stanji na fazni ravnini. Ravnovesna stanja sistema so označena z modrimi krogi. Iz diagrama je očitno, da večina začetnih koncentracij skonvergira v robni stanji, za kateri analiza pokaže, da sta stabilni ravnovesni stanji, medtem ko je sredinsko ravnovesno stanje nestabilno.



**Slika 5.3** Združen prikaz smernega grafa in evolucije sistema na fazni ravnini.

Primer limitnega cikla na fazni ravnini z označenim smernim poljem prikazuje slika 5.4. Iz slike je razvidno, da proti limitnemu ciklu konvergirajo tako zunanje kot tudi notranje trajektorije. Ker je problem obstoja limitnega cikla analitično v splošnem nerešljiv, obstoj



Slika 5.4 Limitni cikel glikolitičnega sistema [54] na fazni ravnini z označenim smernim poljem.

limitnih ciklov preverjamo z računalniškim preiskovanjem prostora. Glavni problem tega pristopa je določitev območja, na katerem preverjamo obstoj limitnih ciklov. V našem primeru je prostor omejen z minimalnimi in maksimalnimi koncentracijami opazovanih kemijskih zvrsti, zato določitev območja ne predstavlja problema.

### 5.3 Bifurkacijska analiza

Kvalitativno obnašanje sistema v smislu obstoja ravnovesnih točk, limitnih ciklov in njihove stabilnosti je večinoma odvisno od vrednosti parametrov, ki opisujejo opazovane kemijske zvrsti in s tem posredno celotni sistem. Z njihovim spreminjanjem lahko tako opazujemo prehode med različnimi načini obnašanja sistema. Prehode med različnimi načini imenujemo *bifurkacije* (angl. *bifurcations*), točke pri katerih pride do sprememb pa *bifurkacijske točke* (angl. *bifurcation points*) [55]. Kvalitativno obnašanje sistema v odvisnosti od vrednosti parametrov, ki opisujejo sistem, prikazujemo na bifurkacijskem diagramu (angl. *bifurcation diagram*), iz katerega je razviden obstoj in tip (stabilen ali

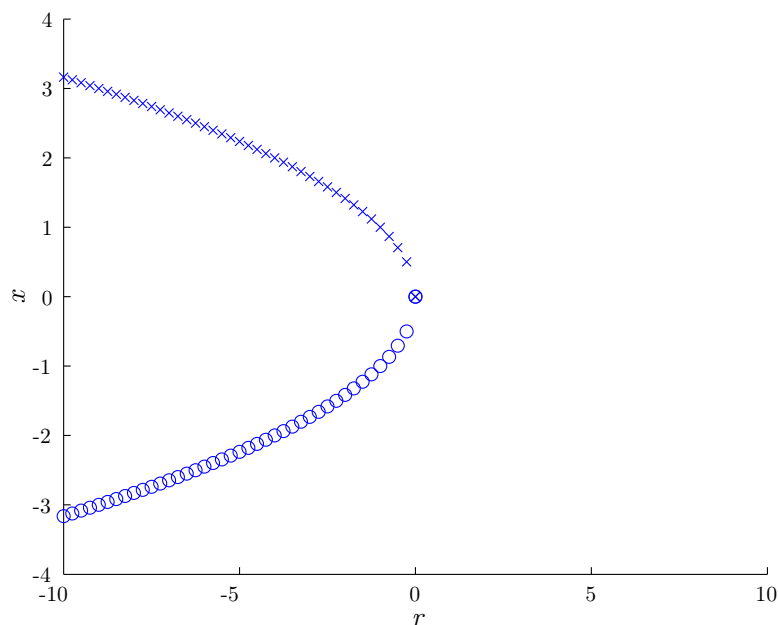
nestabilen) ravnovesnih stanj in limitnih ciklov pri posamezni vrednosti bifurkacijskega parametra.

### 5.3.1 Tipi bifurkacij

Bifurkacije lahko razdelimo na nekaj osnovnih tipov, ki so opisani v nadaljevanju razdelka [54]. Obstajajo sicer še drugi tipi, ki pa jih v tem delu ne bomo obravnavali, saj jih pri analizi obravnavanih bioloških sistemov le redko srečamo.

#### Sedelne bifurkacije

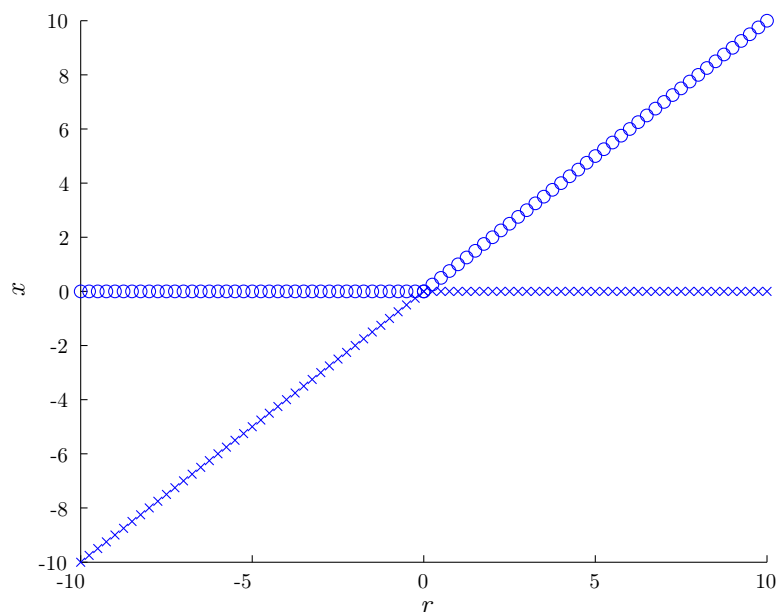
*Sedelne bifurkacije* (angl. *saddle-node bifurcations*) so bifurkacije, pri katerih se s spreminjanjem vrednosti opazovanega parametra spreminja število in stabilnost ravnovesnih točk. Bifurkacijski diagram v tem primeru prikazuje število, lokacije in tipe (stabilna ali nestabilna) ravnovesnih točk v odvisnosti od vrednosti parametra. Primer sedelne bifurkacije prikazuje slika 5.5.



**Slika 5.5** Primer sedelne bifurkacije pri sistemu, ki ga določa enačba  $\frac{dx}{dt} = r + x^2$ . Stabilna ravnovesna stanja so označena s krožcem, nestabilna pa s križcem. Pri vrednostih parametra  $r > 0$  sistem nima več ravnovesnih stanj.

### Transkritične bifurkacije

Pri *transkritičnih bifurkacijah* (angl. *transcritical bifurcations*) ravnovesna točka obstaja pri vseh vrednostih opazovanega parametra, njena stabilnost pa se spreminja. Bifurkacijski diagram je v tem primeru podoben kot pri sedelnih bifurkacijah, le da je število ravnovesnih točk vedno enako. Primer transkritične bifurkacije prikazuje slika 5.6.



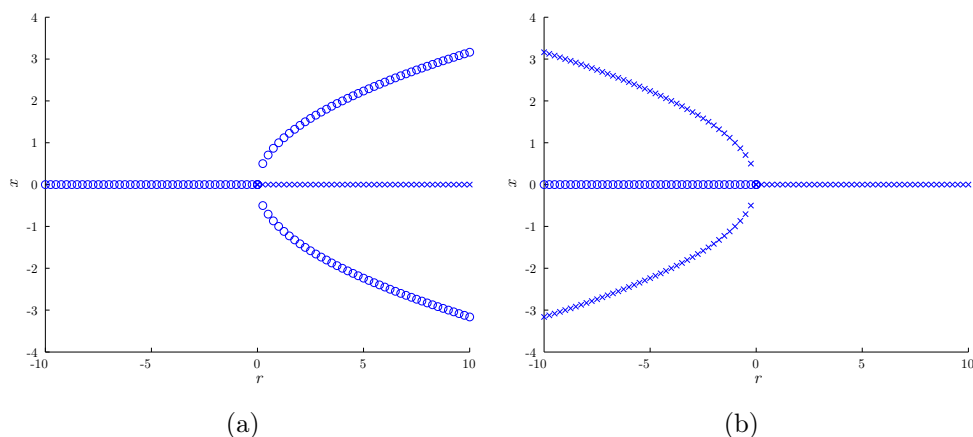
**Slika 5.6** Primer transkritične bifurkacije pri sistemu, ki ga določa enačba  $\frac{dx}{dt} = r \cdot x - x^2$ . Stabilna ravnovesna stanja so označena s krogcem, nestabilna pa s križcem.

### Viličaste bifurkacije

*Viličaste bifurkacije* (angl. *pitchfork bifurcations*) so značilne za sisteme, ki odražajo simetrično obnašanje. Pri takih sistemih se ravnovesne točke pojavljajo in izginjajo simetrično.

Glede na obliko ločimo *superkritične* (angl. *supercritical*) in *subkritične* (angl. *subcritical*) bifurkacije tega tipa. Pri superkritičnih viličastih bifurkacijah se pri spreminjanju vrednosti opazovanega parametra eno stabilno stanje razveji v dve stabilni simetrični stanji in eno vmesno nestabilno stanje (primer prikazuje slika 5.7(a)), pri subkritičnih pa iz dveh nestabilnih simetričnih stanj in vmesnega stabilnega dobimo eno nestabilno stanje (primer prikazuje slika 5.7(b)). Superkritične bifurkacije torej stabilizirajo sistem,

subkritične pa ga destabilizirajo.



**Slika 5.7** Primer superkritične viličaste bifurkacije pri sistemu  $\frac{dx}{dt} = r \cdot x - x^3$  prikazuje slika (a), primer subkritične viličaste bifurkacije pri sistemu  $\frac{dx}{dt} = r \cdot x + x^3$  pa slika (b). Stabilna ravnovesna stanja so označena s krogcem, nestabilna pa s križcem.

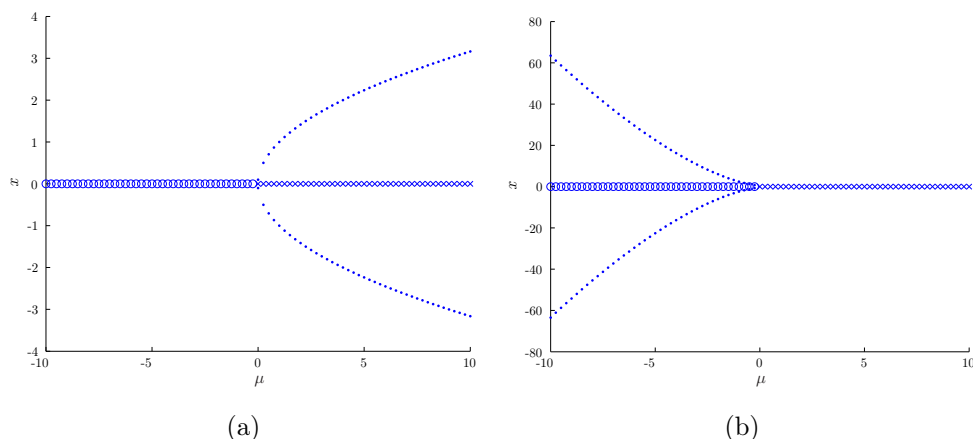
### Hopfove bifurkacije

*Hopfove bifurkacije* (angl. *Hopf bifurcations*) obstajajo samo v dvo- ali večdimenzionalnih sistemih. Prav tako kot viličaste bifurkacije lahko tudi Hopfove bifurkacije razdelimo na *superkritične* in *subkritične*. Do superkritične Hopfove bifurkacije pride, ko spreminjanje parametra povzroči spremembo stabilnega ravnovesnega stanja sistema v nestabilno ravnovesno stanje, poleg tega pa pri spremembi nastane stabilen limitni cikel (primer prikazuje slika 5.8(a)<sup>1</sup>). Iz konvergence proti stabilnemu stanju sistem torej preide v konvergenco proti oscilatornemu obnašanju. Pri subkritični Hopfovi bifurkaciji spreminjanje parametra povzroči spremembo nestabilnega ravnovesnega stanja v stabilno ravnovesno stanje, poleg tega pa pri spremembi nastane nestabilen limitni cikel (primer prikazuje slika 5.8(b)).

Pri analizi Hopfovih bifurkacij si lahko pomagamo s *Hopfovimi bifurkacijskim teoremom* (angl. *Hopf Bifurcation Theorem*) [53, 55]:

**Teorem 1:** (*Hopfov bifurkacijski teorem*) Opazujemo bifurkacije glede na nek parameter  $\chi$ . Naj ima sistem pri vrednosti opazovanega parametra  $\chi = \chi^*$  izolirano ravnovesno

<sup>1</sup>Primer prikazuje obnašanje sistema, ki je zaradi krajšega zapisa predstavljen v polarnih koordinatah, pri čemer  $r$  določa polmer,  $\theta$  pa polarni kot. Pretvorbo v kartezične koordinate lahko naredimo s pomočjo enakosti  $x = r \cdot \cos(\theta)$  in  $y = r \cdot \sin(\theta)$ .



**Slika 5.8** Primer superkritične Hopfove bifurkacije pri sistemu, ki ga določata enačbi  $\frac{dr}{dt} = \mu \cdot r - r^3$  in  $\frac{d\theta}{dt} = 2 + r^2$  prikazuje slika (a), primer subkritične Hopfove bifurkacije pri sistemu, ki ga določata enačbi  $\frac{dx}{dt} = \mu \cdot x - y + x \cdot y^2$  in  $\frac{dy}{dt} = x + \mu \cdot y + y^3$  pa slika (b). Stabilna ravnovesna stanja so označena s krogcem, nestabilna s križcem, limitni cikel pa s pikami. Pri tem je limitni cikel na sliki (a) stabilen (sistem proti njem konvergira pri  $t \rightarrow \infty$ ), na sliki (b) pa nestabilen (sistem proti njem konvergira pri  $t \rightarrow -\infty$ ).

točko  $x^*$ . Naj obstajata lastni vrednosti Jacobijeve matrike v tej točki, ki imata obliko  $\lambda^* = a^* \pm ib^*$ , pri čemer je  $a^* = 0$ ,  $b^* \neq 0$  in  $\frac{da^*}{d\chi} \neq 0$  (realni del lastne vrednosti v tej točki pri spreminjanju vrednosti parametra  $\chi$  spremeni predznak). Ostale lastne vrednosti v tej točki imajo realni del različen od 0. V takem primeru velja, da je  $\chi^*$  točka bifurkacije, pri kateri lahko pride bodisi do superkritične bodisi do subkritične Hopfove bifurkacije:

- *Superkritična Hopfova bifurkacija:* Obstaja neka vrednost opazovanega parametra  $c$ , pri kateri velja, da je  $\chi^* < \chi < c$ . Sistem ima pri teh vrednostih parametra  $\chi$  nestabilno ravnovesno stanje in stabilen limitni cikel s premerom, ki je proporcionalen  $\sqrt{|\chi - \chi^*|}$ .
- *Subkritična Hopfova bifurkacija:* Obstaja neka vrednost opazovanega parametra  $d$ , pri kateri velja, da je  $d < \chi < \chi^*$ . Sistem ima pri teh vrednostih parametra  $\chi$  stabilno ravnovesno stanje in nestabilen limitni cikel s premerom, ki je proporcionalen  $\sqrt{|\chi - \chi^*|}$ .

Do bifurkacije torej pride, ko se spremeni stabilnost izolirane ravnovesne točke.

### 5.3.2 Zgledi analize bifurkacij



### Bifurkacijska analiza bistabilnega sistema

Bistabilen sistem je določen z dvema stabilnima stanjema. Pri določeni vrednosti parametrov se lahko delovanje takega sistema spremeni iz bistabilnega v monostabilno - stabilni stanji sistema se združita v eno.

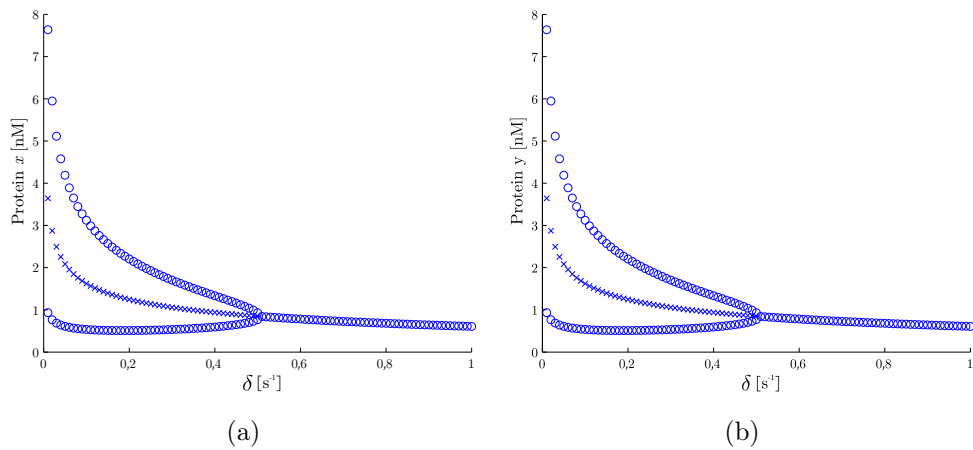
Bifurkacijsko analizo bomo izvedli na tipičnem predstavniku bistabilnih sistemov v gensko regulatornih omrežjih, tj. na preklopnem stikalu (angl. *toggle switch*) [10]. Poenostavljen model preklopnega stikala je podan s sistemom dveh diferencialnih enačb, ki določata spremembo koncentracij dveh opazovanih proteinov

$$\frac{dx}{dt} = \alpha_0 - \delta \cdot x + \frac{\alpha}{1 + \left(\frac{y}{K_1}\right)^\gamma + \left(\frac{x}{K_2}\right)^\gamma}, \quad (5.7)$$

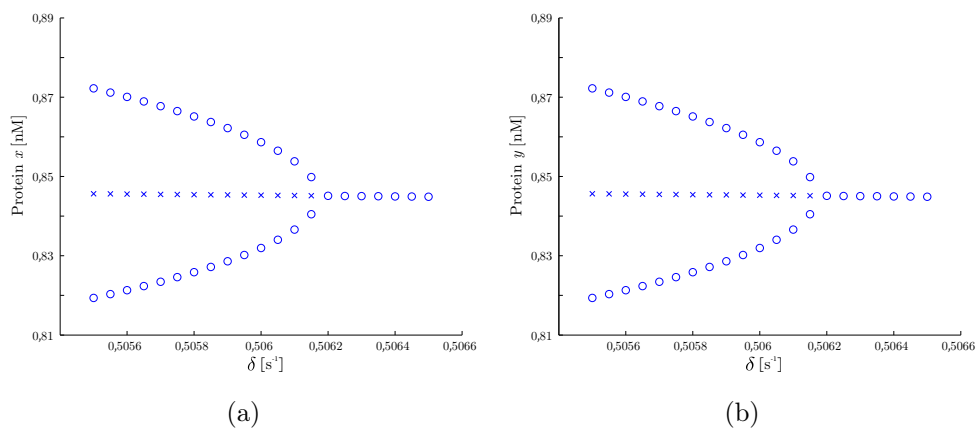
$$\frac{dy}{dt} = \alpha_0 - \delta \cdot y + \frac{\alpha}{1 + \left(\frac{x}{K_1}\right)^\gamma + \left(\frac{y}{K_2}\right)^\gamma}, \quad (5.8)$$

pri čemer  $x$  in  $y$  predstavljata trenutni koncentraciji opazovanih proteinov v času, ostali koeficienti pa parametre sistema, in sicer  $\alpha_0$  koeficient puščanja ( $0\text{ s}^{-1}$ ),  $\alpha$  maksimalno hitrost ekspresije genov ( $1,1\text{ s}^{-1}$ ),  $K_1$  disociacijsko konstanto vezave proteina na operator kompetentnega proteina ( $0,7071$ ),  $K_2$  disociacijsko konstanto vezave proteina na lasten operator ( $2,2361$ ),  $\gamma$  Hillov koeficient ( $2$ ),  $\delta$  pa koeficient razgradnje preko katerega delamo analizo bifurkacij. Zanima nas torej kvalitativno obnašanje sistema v odvisnosti od koeficienta razgradnje.

Analizo ravnovesnih stanj sistema izvajamo pri posamezni vrednosti koeficienta razgradnje po postopku predstavljenem v razdelku 5.1. Na podlagi tega lahko narišemo bifurkacijska diagrama (glej sliko 5.9), ki predstavljata ravnovesna stanja proteina  $X$  in ravnovesna stanja proteina  $Y$  pri posamezni vrednosti koeficienta razgradnje. Bifurkacijska diagrama poleg tega prikazujeta tudi tipe ravnovesnih stanj (tj. stabilno ali nestabilno ravnovesno stanje). Iz slike 5.9 je razvidno, da s spreminjanjem koeficienta razgradnje pri preklopnem stikalu pride do t. i. *superkritične viličaste bifurkacije*, saj iz enega stabilnega ravnovesnega stanja dobimo dve stabilni in eno nestabilno. Natančnejši pogled na sistem prikazuje slika 5.10, iz katere je razvidno, da je bifurkacijska točka sistema v okolici vrednosti  $\delta = 0,50615\text{ s}^{-1}$ . Časovni potek koncentracije opazovanih proteinov pred in po točki bifurkacije prikazuje slika 5.11, iz katere je razvidno, da koncentraciji obeh proteinov pri vrednostih bifurkacijskega parametra, ki so manjše od bifurkacijske, s časom konvergirata v dve različni stabilni stanji (bistabilno delovanje), pri

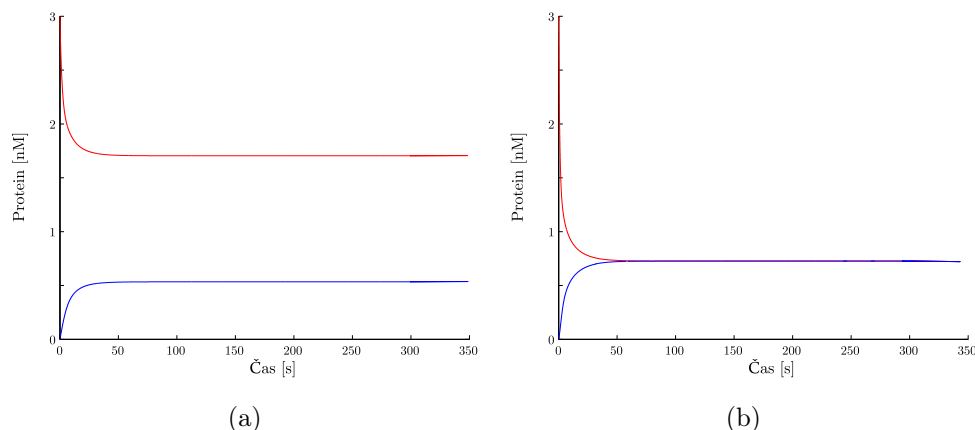


**Slika 5.9** Bifurkacijski diagram preklonnega stikala glede na hitrost (koeficient) razgradnje ( $\delta$ ), pri čemer slika (a) prikazuje ravnovesna stanja proteina  $x$ , slika (b) pa ravnovesna stanja proteina  $Y$  glede na vrednost koeficienta razgradnje ( $\delta$ ). Stabilna ravnovesna stanja so označena s krogcem, nestabilna pa s križcem.



**Slika 5.10** Natančnejši pogled na točko bifurkacije preklonnega stikala glede na hitrost (koeficient) razgradnje. Slika (a) prikazuje ravnovesna stanja proteina  $x$ , slika (b) pa ravnovesna stanja proteina  $y$  glede na vrednost koeficienta razgradnje ( $\delta$ ). Stabilna ravnovesna stanja so označena s krogcem, nestabilna pa s križcem.

vrednostih, ki so večje od bifurkacijske, pa v isto stabilno stanje (monostabilno delovanje). Če želimo doseči robustno bistabilnost sistema, se moramo čimbolj oddaljiti od bifurkacijske točke. Po drugi strani bomo v tem primeru za preklon porabili več energije, kot če si izberemo vrednost koeficienta razgradnje, ki je bližji bifurkacijski točki.



**Slika 5.11** Časovni potek koncentracij opazovanih proteinov, ki določata stanje preklopnega stikala. Slika (a) prikazuje potek koncentracij pri vrednosti koeficienta razgradnje, za katero je bifurkacijska analiza pokazala obstoj dveh različnih stabilnih stanj, kar določa bistabilno delovanje ( $\delta = 0,3 \text{ s}^{-1}$ ), slika (b) pa pri vrednosti, za katero je bifurkacijska analiza pokazala obstoj enega stabilnega stanja, kar določa monostabilno delovanje ( $\delta = 0,7 \text{ s}^{-1}$ ).

### Bifurkacijska analiza oscilatornega sistema

Oscilatorni sistem je določen z nedušenim nihanjem. Oscilacije se pri takem sistemu ohranjajo brez zunanjega vzbujanja, kar je pogojeno z obstojem stabilnega limitnega cikla. Pri določenih vrednostih parametrov lahko stabilen limitni cikel preide v stabilno ravnovesno točko. Ta pojav lahko opazujemo z bifurkacijsko analizo na podlagi Hopfovega bifurkacijskega teorema.

Bifurkacijsko analizo bomo izvedli na tipičnem predstavniku oscilatornih sistemov v gensko regulatornih omrežjih, tj. na represinatorju (angl. *repressilator*) [10]. Poenostavljen model represilatorja lahko podamo s sledečim sistemom diferencialnih enačb

$$\frac{dm_x}{dt} = -m_x + \frac{\alpha}{1 + z^\gamma} + \alpha_0, \quad (5.9)$$

$$\frac{dm_y}{dt} = -m_y + \frac{\alpha}{1 + x^\gamma} + \alpha_0, \quad (5.10)$$

$$\frac{dm_z}{dt} = -m_z + \frac{\alpha}{1 + y^\gamma} + \alpha_0, \quad (5.11)$$

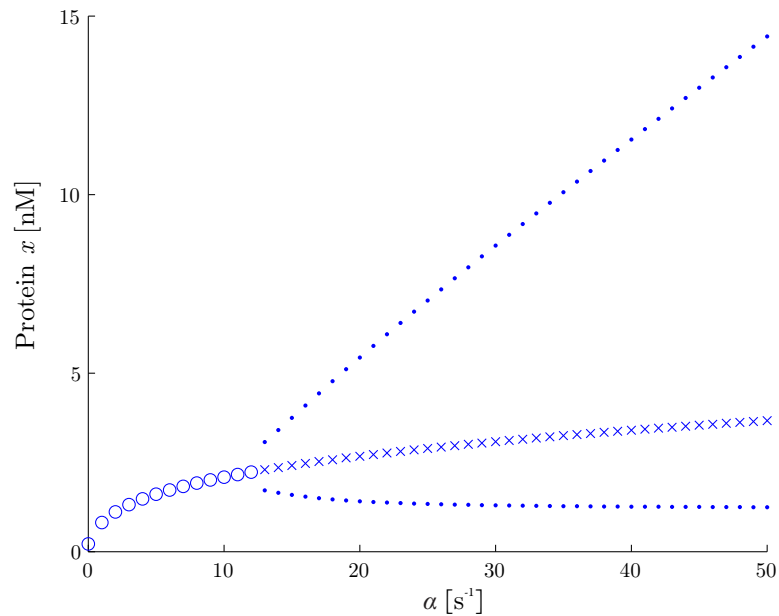
$$\frac{dx}{dt} = -\beta \cdot (x - m_x), \quad (5.12)$$

$$\frac{dy}{dt} = -\beta \cdot (y - m_y), \quad (5.13)$$

$$\frac{dz}{dt} = -\beta \cdot (z - m_z), \quad (5.14)$$

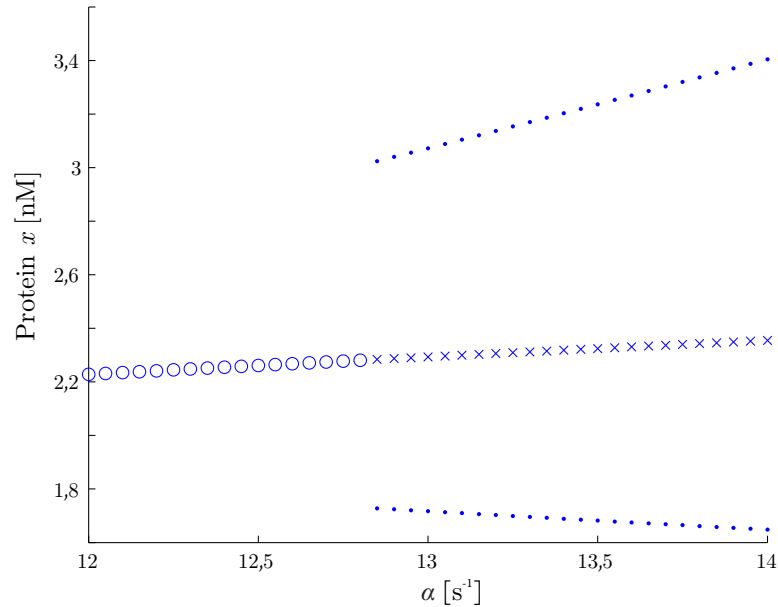
pri čemer  $x$ ,  $y$  in  $z$  predstavljajo trenutne koncentracije opazovanih proteinov,  $m_x$ ,  $m_y$  in  $m_z$  trenutne koncentracije mRNA molekul, ki pripadajo posameznemu proteinu,  $\alpha$  maksimalno hitrost transkripcije, preko katere bomo delali bifurkacijsko analizo,  $\alpha_0$  koeficient puščanja ( $0,216 \text{ s}^{-1}$ ),  $\beta$  hitrost translacije mRNA molekul v protein, ki jo enačimo s hitrostjo razgradnje proteina ( $5 \text{ s}^{-1}$ ),  $\gamma$  pa Hillov koeficient (2). Zanima nas kvalitativno obnašanje sistema v odvisnosti od maksimalne hitrosti transkripcije  $\alpha$ .

Analizo ravnovesnih stanj sistema izvajamo pri posameznih vrednostih koeficienta, ki določa maksimalno hitrost transkripcije s pomočjo Hopfovega teorema (glej razdelek 5.3.1), saj v bifurkacijski točki pričakujemo prehod izoliranega stabilnega stanja v stabilen limitni cikel - pričakujemo torej *superkritično Hopfovo bifurkacijo*. Bifurkacijski diagram za poenostavljen model represilatorja glede na maksimalno hitrost transkripcije prikazuje slika 5.12. Natančnejši pogled na točko bifurkacije sistema prikazuje slika 5.13, iz katere je



**Slika 5.12** Bifurkacijski diagram represilatorja glede na maksimalno hitrost transkripcije ( $\alpha$ ). Stabilna ravnovesna stanja so označena s praznim krogcem, nestabilna s križcem, stabilen limitni cikel pa s pikami.

razvidno, da je bifurkacijska točka sistema, pri kateri le-ta preide v oscilatorno delovanje, v okolici vrednosti  $\alpha = 12,9\text{s}^{-1}$ . Časovni potek koncentracije enega od opazovanih

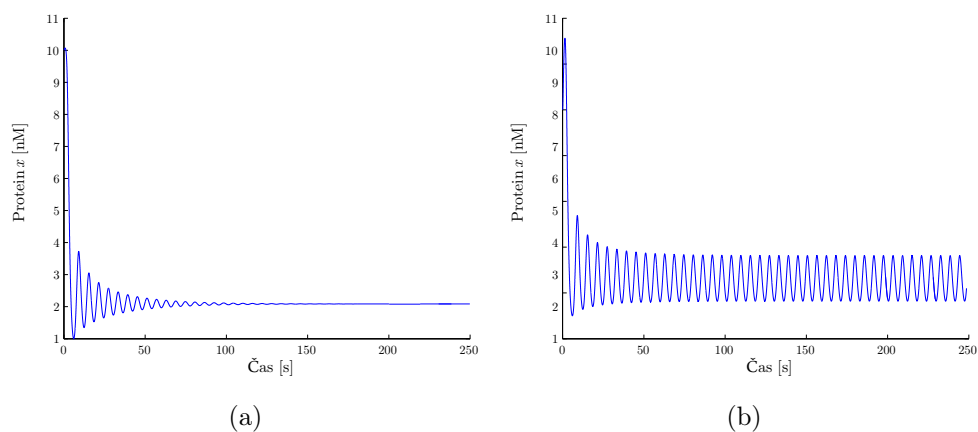


**Slika 5.13** Natančnejši pogled na točko bifurkacije represilatorja glede na maksimalno hitrost transkripcije ( $\alpha$ ). Stabilna ravnovesna stanja so označena s praznim krogcem, nestabilna s križcem, stabilen limitni cikel pa s pikami.

proteinov pred in po točki bifurkacije prikazujeta sliki 5.14 (a) in 5.14 (b), iz katerih je razvidno, da represilator pri vrednostih, ki so manjše od bifurkacijske, s časom konvergira v stabilno stanje, pri vrednostih, ki so večje od bifurkacijske pa v stabilen limitni cikel - samostojno vzdrževanje oscilacij.

## 5.4 Zaključek poglavja

Z determinističnimi modeli, ki jih določajo sistemi navadnih diferencialnih enačb, podajamo opis gensko regulatornih omrežij na enak način kot sicer opisujemo nelinearne dinamične sisteme. Pri analizi asimptotičnega obnašanja gensko regulatornih omrežij si torej lahko pomagamo s teorijo nelinearnih dinamičnih sistemov. Pri tem je za nas še posebej zanimiva analiza obstoja ravnovesnih stanj, analiza obstoja limitnih ciklov in bifurkacijska analiza. Na podlagi rezultatov le-teh lahko pridemo do pomembnih ugotovitev, kot so kakšno je kvalitativno obnašanje biološkega sistema (sistem z enim stabilnim stanjem, sistem z več stabilnimi stanji ali oscilatorni sistem), kakšna je raz-



**Slika 5.14** Časovni potek koncentracije enega od opazovanih proteinov ( $x$ ). Slika (a) prikazuje potek koncentracije pri hitrosti transkripcije, za katero je bifurkacijska analiza pokazala obstoj stabilnega stanja ( $\alpha = 10 \text{ s}^{-1}$ ), slika (b) pa pri hitrosti transkripcije, za katero je bifurkacijska analiza pokazala obstoj stabilnega limitnega cikla ( $\alpha = 15 \text{ s}^{-1}$ ).

dalja med morebitnimi stabilnimi ravnovesnimi stanji, kakšna je velikost morebitnega limitnega cikla in pri kakšnih vrednostih parametrov sistem odraža določene obnašanje. Pri ocenjevanju primernosti bioloških sistemov za procesiranje podatkov v poglavju 6 predstavimo vzpostavitev metrik, ki te analize upoštevajo.

# 6 Metrike

Merilo oziroma mero, s katero merimo posamezne zmogljivostne veličine računalniškega ali katerega koli drugega sistema, imenujemo *metrika* [56]. Ovrednotene metrike so številčni pokazatelji zmogljivostnih lastnosti sistema, na podlagi katerih lahko objektivno ocenjujemo in med seboj primerjamo primernosti in zmogljivosti sistemov glede na njihove predvidene funkcije.

Po [57] je za objektivno določitev ustreznosti zmogljivosti sistema potrebna vzpostavitev metrik in izbira ustrezne *metode* za njihovo vrednotenje. Zaradi narave našega pristopa k problemu je izbira metode vrednotenja metrik relativno enostavna. Ker izhajamo iz modelov bioloških gradnikov, bodo vzpostavljene metrike merjene na simulacijskih rezultatih postavljenih modelov. Uporabljene metode vrednotenja metrik bodo torej v našem primeru *računalniške simulacije*. Izbira ustreznih metrik je v primeru bioloških sistemov po drugi strani kompleksnejša od izbire same metode.

## 6.1 Izhodišča pri izbiri metrik

Pri ocenjevanju primernosti bioloških sistemov za procesiranje podatkov lahko izhajamo iz metrik, ki se uporabljajo v klasični elektroniki [33, 58–60], in sicer pri sistemih, ki so po obnašanju sorodni modeliranim biološkim sistemom. Po drugi strani lahko za ocenjevanje stabilnosti sistema uporabimo metrike, ki izhajajo iz matematičnega področja nelinearnih dinamičnih sistemov in kaosa [53–55] (glej poglavje 5).

### 6.1.1 Metrike za ocenjevanje zmogljivosti digitalnih vezij

Ocenjevanje zmogljivosti digitalnih vezij poteka na podlagi metrik, ki izhajajo iz samih karakteristik vezij. Bistvena lastnost le-teh je ta, da so relativno enostavno merljive in omogočajo objektivno primerjavo med različnimi elementi s podobnimi ali enakimi funkcionalnostmi. Primeri karakteristik, s katerimi lahko ocenjujemo zmogljivosti digitalnih vezij, oziroma elektronskega elementa, so po [33, 58–60]:

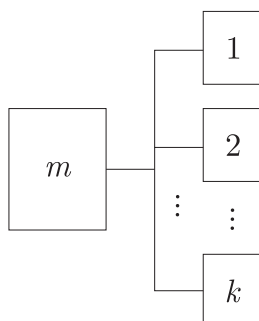
- razvejitev elementa (angl. *fan-out*),
- napetostni nivoji elementa (angl. *voltage levels*),
- šumne meje (angl. *noise margins*),
- širina prepovedanega področja,
- preklopni časi (angl. *switching times*),
- maksimalna frekvenca delovanja (angl. *maximal frequency*),
- čas osveževanja (angl. *refresh rate*).

Obstajajo še številne druge karakteristike, vendar se bomo v našem delu omejili le na zgoraj naštet.

#### Razvejitev izhoda elementa

*Razvejitev izhoda elementa* (angl. *fan-out*) predstavlja maksimalno število elementov iste družine, na katere lahko vežemo izhod iz opazovanega elementa, pri čemer napetostni nivoji vhodnih signalov še vedno ostajajo znotraj dovoljenega področja (glej sliko 6.1). Večja razvejitev torej pomeni, da je izhodni signal iz elementa močnejši in ga tako lahko uporabimo kot vhod v več drugih elementov znotraj določenega sistema. Z elementi z večjo razvejitvijo lahko tako zmanjšamo število enakih elementov znotraj vezja.





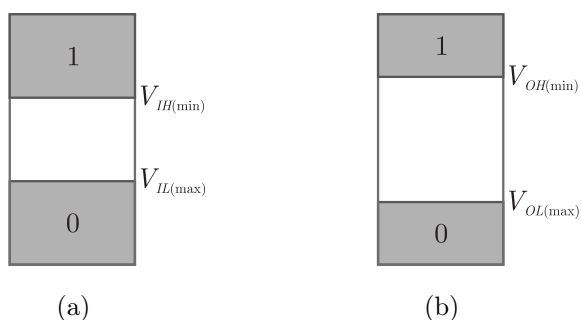
Slika 6.1 Razvejitev izhoda elementa  $m$  na  $k$  vhodov.

### Napetostni nivoji

Informacija je v digitalnih vezjih večinoma predstavljena z napetostnimi nivoji signalov. Ustrezna vrednostna interpretacija signala je lahko narejena zgolj z vnaprej postavljenimi napetostnimi nivoji, ki opredeljujejo območje nizkega logičnega stanja signala (logična 0) in visokega logičnega stanja signala (logična 1). Pri tem ločimo štiri različne meje, s katerimi so napetostni nivoji definirani, in sicer:

- Visoki nivo vhodne napetosti (angl. *high-level input voltage*):  $V_{IH(\min)}$  določa minimalno napetost vhodnega signala, ki ga vezje še prepozna kot signal v visokem logičnem stanju (logična 1).
- Visoki nivo izhodne napetosti (angl. *high-level output voltage*):  $V_{OH(\min)}$  določa minimalno napetost izhodnega signala, ki ga vezje lahko da na svoj izhod kot signal v visokem logičnem stanju (logična 1).
- Nizki nivo vhodne napetosti (angl. *low-level input voltage*):  $V_{IL(\max)}$  določa maksimalno napetost vhodnega signala, ki ga vezje še prepozna kot signal v nizkem logičnem stanju (logična 0).
- Nizki nivo izhodne napetosti (angl. *low-level output voltage*):  $V_{OL(\max)}$  določa maksimalno napetost izhodnega signala, ki ga vezje lahko da na svoj izhod kot signal v nizkem logičnem stanju (logična 0).

Meje za vhodne napetosti prikazuje slika 6.2 (a), meje za izhodne napetosti pa slika 6.2 (b).



**Slika 6.2** Meje, ki določajo napetostne nivoje, pri čemer (a) prikazuje meje za vhodne napetosti, (b) pa meje za izhodne napetosti.

### Šumne meje

Vhodna napetost ne sme pasti pod  $V_{IH(\min)}$  ali zrasti nad  $V_{IL(\max)}$ , ko je signal v mirovanju. Signal se lahko nahaja na intervalu  $]V_{IL(\max)}, V_{IH(\min)}[$ , tj. *prepovedanem področju* (angl. *noise margin*) samo pri preklapljanju iz nizkega logičnega stanja v visoko logično stanje signala ali obratno, pa še tu želimo, da je ta čas kolikor se le da kratek.

V splošnem velja, da so vrednosti vhodnih napetosti bližje prepovedanem območju, saj se lahko nivo signala med prenosom od izhoda iz posameznega elementa do vhoda v drug element še dodatno poslabša (zaradi vpliva šuma, neidealnih karakteristik prenosnih linij, itd.). Veljata torej relaciji

$$V_{IH(\min)} < V_{OH(\min)} \quad (6.1)$$

in

$$V_{IL(\max)} > V_{OL(\max)}. \quad (6.2)$$

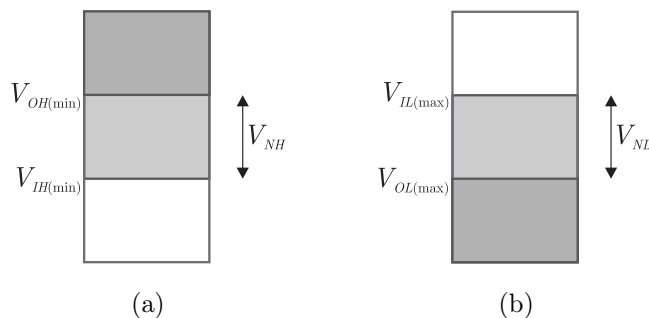
*Šumno mejo za visoko logično stanje* (angl. *high-state noise margin*) lahko definiramo kot razliko med najmanjšo dovoljeno napetostjo, ki še določa visok izhod in najmanjšo dovoljeno napetostjo, ki še določa visok vhod

$$V_{NH} = V_{OH(\min)} - V_{IH(\min)}. \quad (6.3)$$

Po drugi strani lahko definiramo *šumno mejo za nizko logično stanje* (angl. *low-state noise margin*) kot razliko med največjo dovoljeno napetostjo, ki še določa nizek vhod in največjo dovoljeno napetostjo, ki še določa nizek izhod

$$V_{NL} = V_{IL(\max)} - V_{OL(\max)}. \quad (6.4)$$

Šumne meje za visoko logično stanje prikazuje slika 6.3 (a), meje za nizko logično stanje pa slika 6.3 (b).



Slika 6.3 Šumne meje, pri čemer (a) prikazuje meje za visoko logično stanje, (b) pa za nizko logično stanje.

### Prepovedano področje

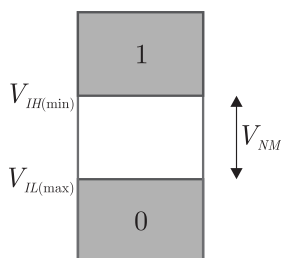
Ker se želimo izogniti napačnim interpretacijam signala zaradi vpliva šuma (tj. neželenim preklpom iz enega logičnega stanja v drugo), so vrednosti  $V_{IL(\max)}$  in  $V_{IH(\min)}$  določene tako, da maksimalni pričakovani šum ne more povzročiti preklopa iz enega logičnega stanja v drugo (glej sliko 6.4). Velja torej

$$V_{IH(\min)} > V_{IL(\max)}. \quad (6.5)$$

Prepovedano področje lahko torej definiramo kot področje na intervalu

$$NM = ]V_{IL(\max)}, V_{IH(\min)}[. \quad (6.6)$$

Z večanjem razlike med mejama večamo širino prepovedanega področja in s tem robustnost sistema oziroma njegovo odpornost na šum. Po drugi strani širše prepovedano področje povzroči daljše preklopne čase in poveča količino energije potrebne za sprožitve preklopa sistema iz enega logičnega stanja v drugo. Iščeemo torej kompromis med preklpnim časom oziroma potrebno energijo za preklop in odpornostjo sistema na šum.

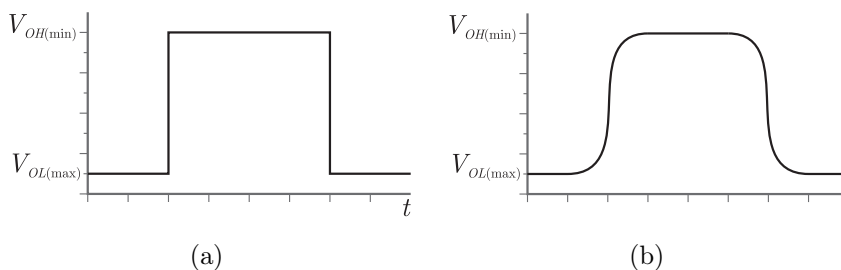


Slika 6.4 Prepovedano področje med nizkim (0) in visokim logičnim stanjem (1).

### Preklopni časi

V digitalnih sistemih se srečujemo z dvema tipoma prekopov signala, in sicer s prekopom iz nizkega v visoko logično stanje (angl. *low to high*) in s prekopom iz visokega v nizko logično stanje (angl. *high to low*). V idealnem primeru bi bila oba tipa prekopov takojšnja, tj. preklopni časi (angl. *switching times*) bi bili enaki 0 (glej sliko 6.5 (a)), česar pa ne moremo trditi za realna vezja (glej sliko 6.5 (b)). V povezavi s prekopmi med logičnimi stanji signala torej ločimo dva časovna parametra, in sicer

- čas za prekop signala iz visokega v nizko logično stanje (angl. *fall time*) in
- čas za prekop signala iz nizkega v visoko logično stanje (angl. *rise time*).

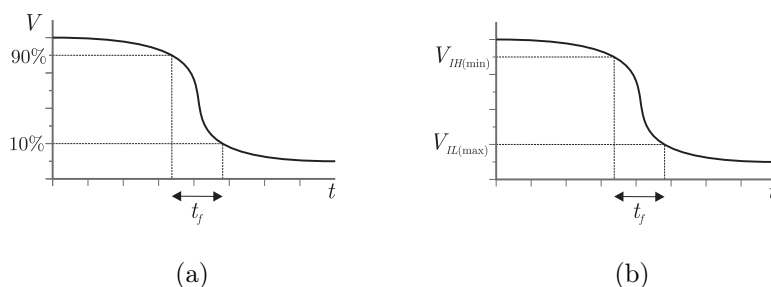


Slika 6.5 Prekop signala iz nizkega v visoko in iz visokega v nizko logično stanje. Slika (a) prikazuje idealen prekop, slika (b) pa zakasnen prekop realnega signala.

Čas za prekop signala iz visokega v nizko logično stanje ( $t_f$ ) lahko merimo na dva načina, in sicer

- merimo čas od točke, ko napetost signala doseže 90% maksimalne vrednosti, do točke, ko napetost signala doseže 10% njene maksimalne vrednosti (glej sliko 6.6 (a)) [33],

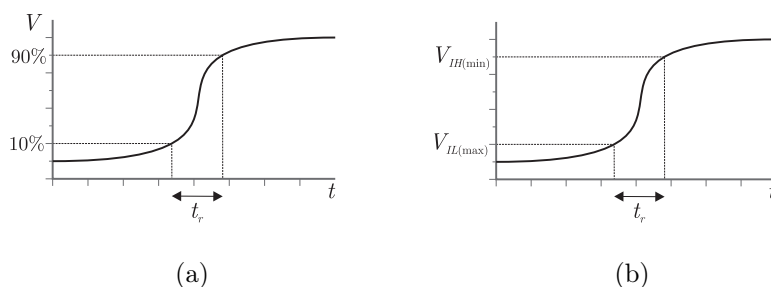
- merimo čas zadrževanja signala v prepovedanem območju, torej čas od takrat, ko napetost signala doseže vrednost  $V_{IH(\min)}$ , do časa, ko napetost signala doseže vrednost  $V_{IL(\max)}$  (glej sliko 6.6 (b)).



**Slika 6.6** Merjenje časa za preklon signala iz visokega v nizko logično stanje, pri čemer slika (a) prikazuje merjenje na podlagi nivoja maksimalne napetosti, slika (b) pa na podlagi časa nahajanja signala v prepovedanem področju.

Podobno lahko na dva načina merimo čas za preklon signala iz nizkega v visoko logično stanje ( $t_r$ ), in sicer

- merimo čas od točke, ko napetost signala doseže 10% maksimalne vrednosti, do točke, ko napetost signala doseže 90% njene maksimalne vrednosti (glej sliko 6.7 (a)) [34],
- merimo čas, ko je signal v prepovedanem območju, torej čas od takrat, ko napetost signala doseže vrednost  $V_{IL(\max)}$ , do časa, ko napetost signala doseže vrednost  $V_{IH(\min)}$  (glej sliko 6.7 (b)) [34].



**Slika 6.7** Merjenje časa za preklon signala iz nizkega v visoko logično stanje, pri čemer slika (a) prikazuje merjenje na podlagi nivoja maksimalne napetosti, slika (b) pa na podlagi časa nahajanja signala v prepovedanem področju.

Praviloma časovna parametra, ki določata preklopne čase, nista enaka. Velja, da je preklopni čas

$$t_s = \max(t_f, t_r). \quad (6.7)$$

Definiramo lahko tudi mero *absolutni minimalni preklopni čas*, ki določa čas za prekop signala iz nizkega v visoko stanje in nazaj v nizko stanje

$$t_{\min} = t_r + t_f. \quad (6.8)$$

Preklopni časi v elektronskih digitalnih vezjih se danes gibljejo v rangu nanosekund (ns).

### Maksimalna frekvenca delovanja

Maksimalna frekvenca delovanja (angl. *maximum output switching frequency*) določa število prekopov, ki jih lahko element izvede v eni sekundi. Današnja tehnologija izdelave digitalnih vezij podpira delovanje le-teh v rangu gigahertzov (GHz). Izračunamo jo lahko kot

$$f_{\max} = \frac{1}{t_r + t_f} = \frac{1}{t_{\min}}. \quad (6.9)$$

### Čas osveževanja

Čas osveževanja (angl. *refresh rate*) v digitalnih vezjih omenjamo v povezavi z dinamičnimi pomnilniki (npr. DRAM pomnilniki), ki so zmožni začasnega pomnjenja logičnega nivoja shranjenega signala. Pomnjenje je namreč omejeno z določenim časom, po preteku katerega ni več zagotovljeno, da bo shranjen signal še imel veljavno logično interpretacijo. Brez osveževanja se njihova vsebina po tem času izgubi. Ker je ta čas ponavadi relativno kratek in želimo razpolagati s pomnjenjem na daljše časovno obdobje, je potrebno vsebino dinamičnega pomnilnika periodično osveževati. Maksimalni čas, po katerem je posamezni element potrebno osvežiti, imenujemo čas osveževanja. Vezje torej periodično osvežujemo s periodo osveževanja  $t_{REF}$ . Na podlagi tega lahko izračunamo frekvenco osveževanja (angl. *refresh frequency*) kot

$$f_{REF} = \frac{1}{t_{REF}}. \quad (6.10)$$

Za namen pomnjenja na daljše časovno obdobje dinamični pomnilniki vsebujejo posebno vezje, ki njihovo vsebino osvežuje s frekvenco, ki mora biti večja ali enaka frekvenci  $f_{REF}$ .

### 6.1.2 Ocenjevanje stabilnosti nelinearnih dinamičnih sistemov

Z matematično teorijo s področja nelinearnih dinamičnih sistemov in kaosa [53–55] (glej poglavje 5) si lahko pomagamo pri analizi asimptotičnega obnašanja sistemov. Izvedba analize poteka na podlagi navadnih diferencialnih enačb, torej determinističnih modelov bioloških sistemov. Vzpostavitev metrik in njihovo analizo lahko ločimo glede na število stabilnih ravnovesnih stanj obravnavanega sistema. V nadaljevanju namesto besedne zveze *stabilno ravnovesno stanje* uporabljamo zgolj besedno zvezo *ravnovesno stanje*.

#### Sistemi z enim stabilnim stanjem

Zanima nas lokacija stabilnega stanja pri določenih pogojih (npr. pri določenem stanju vhodnih signalov). Na podlagi tega lahko določimo koncentracije proteinov, ki določajo nizko logično stanje (logična 0) in visoko logično stanje (logična 1).

#### Sistemi z več stabilnimi stanji

Pri sistemih z več stabilnimi stanji, lahko ugotovljamo njihovo robustnost na dva načina, in sicer kot *razdaljo med stabilnimi stanji* in kot *razdaljo od točke bifurkacije*.

##### *Razdalja med stabilnimi stanji*

Večja je razdalja med stabilnimi stanji, bolj je sistem robusten, posledično pa je z večanjem razdalje v preklope med stanji potrebno vložiti več energije.

##### *Razdalja od točke bifurkacije*

Opazujemo lahko razdaljo od točke bifurkacije glede na posamezni parameter. Večja je razdalja od točke bifurkacije, bolj je sistem robusten. Opazujemo torej oddaljenost sistema od točke bifurkacije glede na posamezni parameter. Pri tem je potrebno upoštevati najšibkejši člen, ni pa potrebno upoštevati vseh parametrov - upoštevamo lahko le tiste, ki od celice do celice variirajo bodisi zaradi vpliva šuma bodisi zaradi uporabljenih približkov v modelu.

#### Sistemi brez ravnovesnih stanj

V tej skupini sistemov so za nas zanimivi predvsem oscilatorni sistemi, pri katerih lahko podobno kot pri sistemih z več stabilnimi stanji opazujemo robustnost oscilacij, tj. oddaljenost od točke bifurkacije, v kateri preidemo iz periodičnega spreminjanja koncentracij kemijskih zvrsti, ki predstavljajo nosilce informacije (oscilatorno delovanje sistema), v konvergenco teh koncentracij proti enemu stabilnemu stanju (sistem z enim stabilnim

stanjem). Podobno kot pri sistemih z več stabilnimi stanji lahko tudi pri teh sistemih definiramo metriko *razdalja od točke bifurkacije*  $R_b$ , namesto metrike *razdalja med stabilnimi stanji* pa tu definiramo *velikost stabilnega limitnega cikla*  $R_c$ , ki določa amplitudo oscilacij.

## 6.2 Metrike za vrednotenje bioloških sistemov

Sledeči razdelek opisuje vzpostavitev metrik za vrednotenje bioloških sistemov kot gradnikov za procesiranje informacij. Izhajali bomo iz metrik za ocenjevanje zmogljivosti digitalnih sistemov (glej razdelek 6.1.1) in metrik za ocenjevanje stabilnosti nelinearnih dinamičnih sistemov (glej razdelek 6.1.2). Za vrednotenje določenih metrik bomo morali najprej izbrati primerne nosilca informacije in določiti šum, vsebovan v izbranem signalu. V nadaljevanju razdelka sledi postopek določanja metrik, na podlagi katerih lahko izvedemo analizo sistema.

### 6.2.1 Izbira nosilca informacije

Pod pojmom nosilca informacije smatramo merljivo fizikalno količino, ki odraža stanje modeliranega sistema in jo tako lahko uporabimo kot njegov izhod. Stanje sistema se torej odraža preko stanja nosilca informacije. Izbrana fizikalna količina naj bi torej imela podobno vlogo, kot jo ima napetost v klasičnih digitalnih vezjih in s tem odražala stanje modeliranega sistema. V tej vlogi v našem sistemu nastopa koncentracija posamezne kemijske zvrsti, saj odraža trenutno stanje sistema in je relativno enostavno merljiva. Koncentracijo kot indikator stanja sistema je poleg tega potrebno obravnavati v kontekstu izbranih kemijskih zvrsti. Pri tem mora biti dinamika koncentracij izbranih kemijskih zvrsti skladna z dinamiko, ki naj bi jo opazovani sistem odražal (npr. stanje vhoda se mora prenesti na stanje izhoda v primeru biološkega gonilnika). Ostale lastnosti izbrane kemijske zvrsti lahko vrednotimo s strani ostalih metrik za vrednotenje bioloških sistemov.

### 6.2.2 Določitev šuma

V modeliranem sistemu opazujemo šum, ki je posledica notranjih nihanj (angl. *fluctuations*) in sledi iz stohastične narave kemijskih procesov [27]. Določanje šuma bo potekalo zgolj na rezultatih stohastičnega modela. Ker želimo na podlagi meritev šuma določati



širino prepovedanega področja, je bila za mero uporabljena maksimalna absolutna vrednost šuma, ki jo merimo kot maksimalni odklon od povprečne koncentracije opazovane kemijske zvrsti pri posameznem logičnem stanju signala.

### 6.2.3 Določitev logičnih nivojev

Predpostavimo, da visoka koncentracija opazovane kemijske zvrsti predstavlja visoko logično stanje (logična 1), nizka koncentracija pa nizko logično stanje (logična 0). Orientacijsko lahko logične nivoje določimo na podlagi stabilnih stanj determinističnega modela biološkega sistema. Natančneje lahko logične nivoje našega sistema opredelimo na podoben način, kot so okarakterizirana klasična digitalna vezja (glej razdelek 6.1.1). Zaradi specifičnosti obravnavanih sistemov se določanje logičnih nivojev nekoliko razlikuje od določanja le-teh pri digitalnih elektronskih vezjih. Medtem ko so klasična digitalna elektronska vezja zgrajena iz standardnih gradnikov, npr. tranzistorjev in diod, pri čemer imajo vsi ti gradniki enake karakteristike, biološka vezja gradimo iz različnih komponent, ki jih lahko bodisi najdemo v naravi ali ustvarimo s postopki sintezne biologije. Lastnosti uporabljenih gradnikov (npr. hitrost razgradnje proteina, afiniteta promotorja itd.) posledično določajo tudi logične nivoje zgrajenega vezja, zato je treba le-te določati pri vsakem vezju posebej. Prav tako je potrebno posebej določati vhodne in izhodne nivoje koncentracij, ki niso nujno enake. Tako je potrebno ovrednotiti sledeče karakteristike:

- visoki nivo izhodne koncentracije:  $C_{OH}$  določa koncentracijo izhodnega signala v visokem logičnem stanju (logična 1),
- spodnjo mejo visokega nivoja izhodne koncentracije:  $C_{OH(\min)}$  določa koncentracijo izhodnega signala, ki še predstavlja signal v visokem logičnem stanju (logična 1),
- spodnjo mejo visokega nivoja vhodne koncentracije:  $C_{IH(\min)}$  določa koncentracijo vhodnega signala v visokem logičnem stanju (logična 1), ki izhodni signal še pripelje v veljavno logično stanje,
- nizki nivo izhodne koncentracije:  $C_{OL}$  določa koncentracijo izhodnega signala v nizkem logičnem stanju (logična 0),
- zgornjo mejo nizkega nivoja izhodne koncentracije:  $C_{OL(\max)}$  določa koncentracijo izhodnega signala, ki še predstavlja signal v nizkem logičnem stanju (logična 0),

- zgornjo mejo nizkega nivoja vhodne koncentracije:  $C_{IL(\max)}$  določa koncentracijo vhodnega signala v nizkem logičnem stanju (logična 0), ki izhodni signal še pripelje v veljavno logično stanje.

Nivoje koncentracij izhodnega signala ( $C_{OH}$  in  $C_{OL}$ ) določimo z merjenjem koncentracij takoj po preklopu bodisi iz nizkega v visoko logično stanje pri določanju  $C_{OH}$  bodisi iz visokega v nizko stanje pri določanju  $C_{OL}$ . Na podlagi več meritev tako določimo minimalno koncentracijo, ki še predstavlja visoko logično stanje izhoda in maksimalno koncentracijo, ki še predstavlja nizko logično stanje izhoda.

Skrajne meje nivojev izhodnih koncentracij (tj. spodnjo mejo visokega nivoja izhodne koncentracije in zgornjo mejo nizkega nivoja izhodne koncentracije) lahko določimo posredno preko že določenih nivojev koncentracij izhodnega signala ( $C_{OH}$  in  $C_{OL}$ ) z upoštevanjem šuma oziroma *šumnih mej*, ki morajo biti definirane na tak način, da maksimalni pričakovani šum signala ne prestavi iz dovoljenega v prepovedano področje. Osnovo za določanje teh nivojev predstavljata enačbi

$$C_{OH(\min)} < C_{OH} - N_H, \quad (6.11)$$

$$C_{OL(\max)} > C_{OL} + N_L, \quad (6.12)$$

kjer  $N_H$  predstavlja maksimalen pričakovani šum pri signalu v visokem logičnem stanju,  $N_L$  pa maksimalen pričakovani šum pri signalu v nizkem logičnem stanju.

Pri določanju vhodnih koncentracij določamo koncentracije, ki izhodni signal še pripeljejo v veljavno logično stanje. Pri tem določamo spodnjo mejo visokega nivoja vhodne koncentracije  $C_{IH(\min)}$ , ki določa najnižjo dovoljeno koncentracijo vhodnega signala v visokem logičnem stanju (logična 1), in zgornjo mejo nizkega nivoja vhodne koncentracije  $C_{IL(\max)}$ , ki določa najvišjo dovoljeno koncentracijo vhodnega signala v nizkem logičnem stanju (logična 0).

#### 6.2.4 Določitev prepovedanega področja

Zaradi ločenega določanja izhodnih in vhodnih logičnih nivojev tudi prepovedano področje določamo posebej za izhodne signale in posebej za vhodne signale.

Območje oziroma širino izhodnega prepovedanega področja lahko določimo neposredno preko postavljenih izhodnih logičnih nivojev

$$NM_O = ]C_{OL(\max)}, C_{OH(\min)}[. \quad (6.13)$$

Vhodno prepovedano področje določimo na podoben način preko postavljenih vhodnih logičnih nivojev

$$NM_I = ]C_{IL(\max)}, C_{IH(\min)}[. \quad (6.14)$$

### 6.2.5 Določitev preklonnih časov

Preklone čase lahko določamo na enak način kot v klasičnih digitalnih elektronskih vezjih, in sicer jih bomo merili kot čase, ko se signal nahaja v prepovedanem področju. Pri tem lahko definiramo dva različna časovna parametra. Pri prehodu signala iz nizkega (logična 0) v visoko logično stanje (logična 1) merimo *čas vzpona* (angl. *rise time*) -  $t_r$ , pri prehodu signala iz visokega (logična 1) v nizko logično stanje (logična 0) pa *čas padca* (angl. *fall time*) -  $t_f$ . Na podlagi preklonnih časov lahko določimo čas preklopa  $t_s$  (glej enačbo 6.7).

### 6.2.6 Določitev porabe vhoda

Koncentracija kemijske zvrsti, ki v posamezno biološko vezje vstopa kot vhod, se lahko za potrebe procesiranja v samem vezju občutno zmanjša. Vpeljemo metriko *poraba vhoda*, ki določa maksimalno zmanjšanje koncentracije kemijske zvrsti, ki nastopa kot vhod v obravnavano vezje. Vpeljana metrika je posredno povezana z razvejitvijo izhoda elementa (glej razdelek 6.1.1), saj večja poraba vhoda povzroči zmanjšanje razvejitve izhodov. Za metriko poraba vhoda vpeljemo oznako  $C_I$ .

### 6.2.7 Določitev periode oscilacij

Periode oscilacij določamo pri oscilatornih vezjih. V tem primeru merimo čas, ki je potreben za eno oscilacijo. Amplitude oscilacij pri tovrstnih vezjih opazujemo posredno preko logičnih nivojev koncentracij nosilcev informacije.

### 6.2.8 Določitev časa veljavnosti izhodnega signala

Čas veljavnosti izhodnega signala določa čas po preklopu oziroma osvežitvi nivoja izhodnega signala, ko ima le-ta še veljavno interpretacijo. Čas veljavnosti izhodnega signala določamo le v primeru, če se koncentracija izhodnega signala po določenem času brez spreminjanja zunanjih vplivov spremeni na tak način, da se pomakne v prepovedano področje. Pri tem določamo *čas veljavnosti izhodnega signala v visokem stanju* ( $t_{vH}$ )

in čas veljavnosti izhodnega signala v nizkem stanju ( $t_{v_L}$ ) posebej. Določimo lahko čas veljavnosti izhodnega signala kot  $t_v = \min(t_{v_H}, t_{v_L})$ .

### 6.2.9 Določitev časa osveževanja

Biološka vezja z več stabilnimi stanji lahko v določenih primerih uporabimo kot pomnilna vezja, pri čemer eno stabilno stanje predstavlja nizko logično stanje (logična 0), eno pa visoko logično stanje (logična 1). Podobno kot pri ostalih vezjih določen del vmesnega področja med stanjema predstavlja prepovedano področje, ki nima veljavne logične interpretacije. V določenih primerih lahko začne stanje z veljavno logično interpretacijo sčasoma zaradi šuma prehajati v prepovedano področje, kljub temu da rezultati determinističnega modela na to ne kažejo. Informacija, ki je bila v vezje zapisana, se v tem primeru s časom izgubi. Z namenom ohranitve zapisane informacije je tovrstna vezja potrebno periodično osveževati na podoben način kot dinamična pomnilniška vezja, še preden koncentracije nosilcev informacije preidejo v prepovedano področje. V ta namen je potrebno določiti periodo osveževanja  $t_{REF}$ , ki predstavlja minimalni čas, po katerem se informacija lahko izgubi. Čas osveževanja je pri takih vezjih enak času veljavnosti izhodnega signala  $t_v$  (glej razdelek 6.2.8).

### 6.2.10 Določitev robustnosti

Robustnost sistemov lahko določamo na več različnih načinov. V našem primeru bomo v ta namen uporabili dve različni metriki, in sicer *razdaljo med stabilnimi stanji* ( $R_d$ ) in *razdaljo od točke bifurkacije* ( $R_b$ ).

#### Razdalja med stabilnimi stanji

Pri sistemih, ki imajo stabilna ravnovesna stanja, lahko ocenjujemo njihovo robustnost glede na razmerje med razdaljo med stabilnimi stanji in maksimalnim izmerjenim šumom  $N$ , ki posledično določa verjetnost neželenih preklpov iz enega stabilnega stanja v drugega. Z upoštevanjem šuma določenega na podlagi meritev stohastičnega modela in ravnovesnih stabilnih stanj določenih na podlagi determinističnega modela lahko definiramo metriko *robustnost glede na razdaljo med stabilnimi stanji* po izrazu

$$R_d = \min_{i \neq j} \frac{d_{i,j}}{N}, \quad (6.15)$$

pri čemer je  $d_{i,j}$  evklidska razdalja med stabilnima stanjema  $i$  in  $j$ . V primeru, da je vrednost  $R_d$  manjša od 1, je verjetnost, da bo pri sistemu prihajalo do neželenih

preklopov zelo velika. V primeru, da je metrika večja od 1, do neželenih preklopov ne bi smelo prihajati.

#### Velikost stabilnega limitnega cikla

Pri oscilatornih sistemih namesto razdalje med stabilnimi stanji uporabimo metriko velikost limitnega cikla. Njihovo robustnost lahko ocenjujemo glede na razmerje med velikostjo limitnega cikla  $d_c$  in maksimalnim izmerjenim šumom  $N$  z metriko *robustnost glede na velikost limitnega cikla* po izrazu

$$R_c = \frac{d_c}{N}. \quad (6.16)$$

V primeru, da je metrika manjša od 1, je verjetnost, da bodo oscilacije sčasoma iznihale, zelo velika. V primeru, da je metrika večja od 1, je verjetnost, da bo sistem odražal oscilatorno vedenje brez dodatnega zunanje vzbujanja zelo velika.

#### Razdalja od točke bifurkacije

Definiramo lahko metriko *robustnost glede na bifurkacijske točke* po izrazu

$$R_b = \min_i \frac{d_i}{b_i}, \quad (6.17)$$

pri čemer  $b_i$  določa lokacijo bifurkacijske točke za parameter  $i$ ,  $d_i$  pa oddaljenost parametra od bifurkacijske točke za parameter  $i$ . Pri računanju metrike ni potrebno upoštevati vseh parametrov, ampak lahko upoštevamo le tiste, ki od celice do celice variirajo bodisi zaradi vpliva šuma bodisi zaradi uporabljenih približkov v modelu.

### 6.3 Analiza osnovnih bioloških sistemov, zmožnih procesiranja podatkov

Analiza bioloških sistemov na podlagi vzpostavljenih metrik bo predstavljena na simulacijskih rezultatih modelov osnovnih bioloških sistemov zmožnih procesiranja podatkov predstavljenih v razdelku 3.3. Rezultati vrednotenja se nanašajo na vrednosti parametrov in rezultate modelov, ki so predstavljeni v razdelku 4.4. Zaradi medsebojne odvisnosti je vrstni red določanja metrik pomemben, zato jih bomo določali v sledečem zaporedju:

- izbira nosilca informacije,
- določitev nivojev izhodnih koncentracij nosilca informacije,

- določitev šuma v izbranem nosilcu informacije,
- določitev časa osveževanja pri pomnilnih vezjih,
- določitev skrajnih mej nivojev izhodnih koncentracij nosilca informacije,
- določitev izhodnega prepovedanega področja,
- določitev skrajnih mej nivojev vhodnih koncentracij nosilca informacije,
- določitev vhodnega prepovedanega področja,
- določitev porabe vhoda,
- določitev preklopnih časov,
- določitev časa veljavnosti izhodnega signala,
- določitev periode oscilacij pri oscilatornih vezjih.

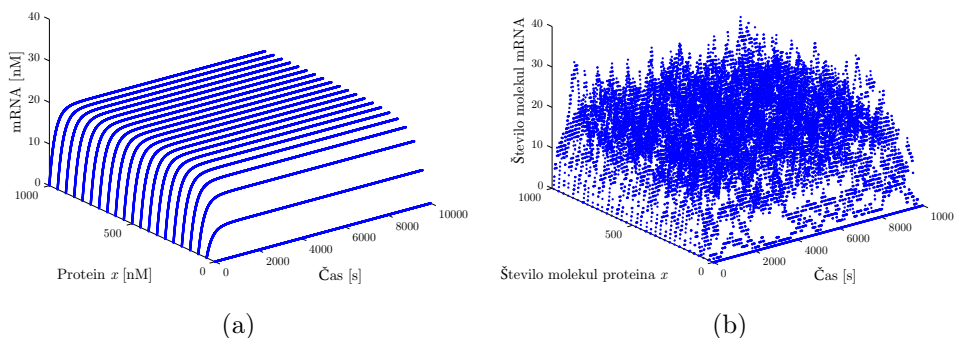
Metrike za analizo robustnosti lahko določamo na podlagi izbranega nosilca informacije in šuma prisotnega v signalu, torej ju lahko določamo kadarkoli po koraku 2.

### 6.3.1 Analiza biološkega gonilnika

Določanje vrednosti metrik na primeru biološkega gonilnika bo razloženo podrobneje, saj je to prvi primer določanja vrednosti, ki ga navajamo v tem delu.

#### Biološki gonilnik: izbira nosilca informacije

Nosilca informacije izbiramo med kemijskimi zvrstmi, ki lahko nastopajo kot izhod iz sistema. V primeru biološkega gonilnika torej izbiramo zgolj med izhodnim proteinom  $y$  in mRNA izhodnega proteina. Poleg tega, da so koncentracije mRNA izhodnega proteina zelo nizke ne glede na koncentracijo vhodnega proteina (glej sliko 6.8), je izhodni protein veliko bolj primeren nosilec informacije oziroma veliko bolj primerna izhodna kemijska zvrst sistema, saj lahko igra vlogo transkripcijskega faktorja v nekem drugem transkripcijsko regulatornem omrežju. Potek koncentracij izhodnega proteina je prikazan na slikah 4.3 in 4.5.



**Slika 6.8** Potek koncentracij mRNA izhodnega proteina v determinističnem (a) in stohastičnem (b) modelu biološkega gonilnika v odvisnosti od koncentracije vhodnega proteina.

### Biološki gonilnik: določitev nivoja izhodnih koncentracij nosilca informacije

Orientacijsko lahko logične nivoje izhoda določimo na podlagi stabilnih stanj determinističnega modela biološkega gonilnika. Logične nivoje določamo posebej pri nizkem stanju vhodnega signala  $x$  ( $C_x \approx 0$ ) in posebej pri visokem stanju vhodnega signala  $x$  ( $C_x \gg 0$ ) (glej sliko 4.3):

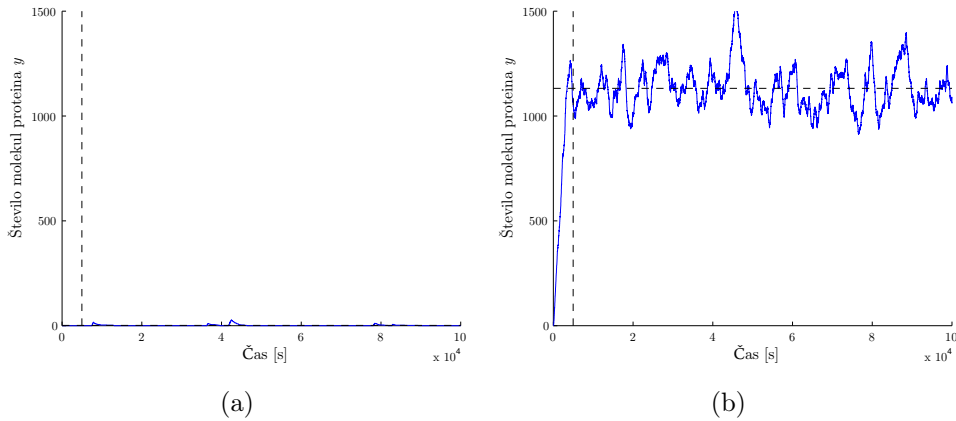
$$C_L \approx 0, \quad (6.18)$$

$$C_H \approx 1350 \approx 1120 \text{ nM}. \quad (6.19)$$

Pri natančnejši analizi je potrebno določiti vrednosti koncentracij, ki predstavljajo maksimalne koncentracije izhodnega signala v nizkem logičnem stanju ( $C_{OH}$ ) in minimalne koncentracije izhodnega signala v visokem logičnem stanju ( $C_{OL}$ ). Koncentracije torej določamo pri dveh različnih začetnih pogojih, in sicer pri konstantni vrednosti koncentracij vhodnega proteina  $x$ , ki sistem pripeljejo v nizko logično stanje ( $C_x \approx 0$ ) in pri konstantni vrednosti koncentracij vhodnega proteina  $x$ , ki sistem pripeljejo v visoko logično stanje ( $C_x \gg 0$ ).

Nivoje koncentracij določamo kot povprečne vrednosti signala, ko se le-ta stabilizira. V našem primeru z računanjem povprečja začnemo po 5000 sekundah od začetka simulacije, pri čemer je ob inicializaciji ( $t = 0$  s) koncentracija nosilca informacije enaka 0 (glej sliko 6.9). Vrednosti so bile merjene na podlagi desetih ponovitev simulacij. Na podlagi simulacijskih rezultatov velja, da je

$$C_{OL} = 5 \approx 5 \text{ nM}, \quad (6.20)$$



**Slika 6.9** Merjenje nivojev izhodnih koncentracij signala v stohastičnem modelu biološkega gonilnika. Slika (a) prikazuje potek nizkega logičnega nivoja izhodnih koncentracij, slika (b) pa visokega logičnega nivoja izhodnih koncentracij. Merjenje poteka od časovne točke  $t = 5000$ s naprej (vertikalna črta) kot merjenje povprečne vrednosti signala v visokem oziroma nizkem logičnem stanju (horizontalna črta).

$$C_{OH} = 1100 \approx 915\text{nM}. \quad (6.21)$$

### Biološki gonilnik: določitev šuma

Šum določamo pri dveh različnih začetnih pogojih, in sicer pri koncentracijah vhodnega proteina  $x$ , ki sistem pripeljejo v nizko logično stanje ( $C_x \approx 0$ ), in pri koncentracijah vhodnega proteina  $x$ , ki sistem pripeljejo v visoko logično stanje ( $C_x \gg 0$ ). Merimo torej šum pri signalu v nizkem logičnem stanju ( $N_L$ ) in šum pri signalu v visokem logičnem stanju ( $N_H$ ). Zanima nas višja vrednost, ki jo tudi upoštevamo pri določanju ostalih metrik

$$N = \max(N_L, N_H). \quad (6.22)$$

Šum merimo kot odstopanje signala od povprečnega signala po določenem času, tj. po času, ki je potreben za stabilizacijo signala pri določenih vhodnih pogojih. V našem primeru šum merimo po 5000 sekundah od začetka simulacije, pri čemer je začetna koncentracija nosilca informacije enaka 0. Meritve potekajo torej na podoben način kot pri določanju nivojev izhodnih koncentracij izbranega nosilca informacije (glej sliko 6.9).

Na podlagi simulacijskih rezultatov lahko ugotovimo, da maksimalni šum v visokem logičnem stanju ne preseže vrednosti 400 ( $N_H < 400$ ), maksimalni šum v nizkem logičnem



stanju pa vrednosti 20 ( $N_L < 20$ ). Ker nas zanima višja vrednost, v nadaljevanju upoštevamo maksimalni šum signala v visokem logičnem stanju. Velja torej

$$N = 400 \approx 335 \text{ nM}. \quad (6.23)$$

**Biološki gonilnik: določitev skrajnih mej nivojev izhodnih koncentracij nosilca informacije**

Skrajne meje nivojev izhodnih koncentracij nosilca informacije določamo na podlagi izhodnih koncentracij ( $C_{OH}$ ,  $C_{OL}$ ) in šuma v nizkem in visokem stanju ( $N_L$  in  $N_H$ ) na podlagi izrazov

$$C_{OL(\max)} = C_{OL} + N_L = 5 + 20 = 25 \approx 20 \text{ nM}, \quad (6.24)$$

$$C_{OH(\min)} = C_{OH} - N_H = 1100 - 400 = 700 \approx 580 \text{ nM}. \quad (6.25)$$

**Biološki gonilnik: določitev izhodnega prepovedanega področja**

Prepovedano področje je interval, znotraj katerega koncentracije nosilca informacije nimajo veljavne interpretacije. Prepovedano področje se nahaja na intervalu

$$NM_O = ]C_{OL(\max)}, C_{OH(\min)}[ = ]25, 700[ \approx ]20 \text{ nM}, 580 \text{ nM}[. \quad (6.26)$$

Njegova širina je torej enaka  $675 = 560 \text{ nM}$ .

**Biološki gonilnik: določitev skrajnih mej nivojev vhodnih koncentracij nosilca informacije**

Določiti je potrebno koncentracije, ki izhodni signal še pripeljejo v veljavno logično stanje, in sicer spodnjo mejo visokega nivoja vhodne koncentracije  $C_{IH(\min)}$ , ki izhod biološkega gonilnika še pripelje v veljavno visoko logično stanje ( $C > C_{OH(\min)}$ ), in zgornjo mejo nizkega nivoja vhodne koncentracije  $C_{IL(\max)}$ , ki izhod biološkega gonilnika še pripelje v veljavno nizko logično stanje ( $C < C_{OL(\max)}$ ). Pri tem so bile na podlagi simulacijskih rezultatov določene vrednosti

$$C_{IL(\max)} < 15 \approx 10 \text{ nM}, \quad (6.27)$$

$$C_{IH(\min)} > 150 \approx 125 \text{ nm}. \quad (6.28)$$

### Biološki gonilnik: določitev vhodnega prepovedanega področja

Vhodno prepovedano področje določimo na podoben način kot izhodno prepovedano področje po izrazu

$$NM_I = ]C_{IL(\max)}, C_{IH(\min)}[ = ]15, 150[ \approx ]10 \text{ nM}, 125 \text{ nM}[. \quad (6.29)$$

### Biološki gonilnik: določitev porabe vhoda

Koncentracija vhodne kemijske zvrsti, ki jo biološki gonilnik v najslabšem primeru porabi, je določena z metriko *poraba vhoda*. Glede na reakcije v tabeli 4.2 se vhodni protein v dimerizirani obliki veže na operator  $O_{R3}$ . Ker predpostavljamo, da je v celici prisotna le ena molekula DNA za en gonilnik, je maksimalna poraba vhoda enaka  $C_I = 2$ .

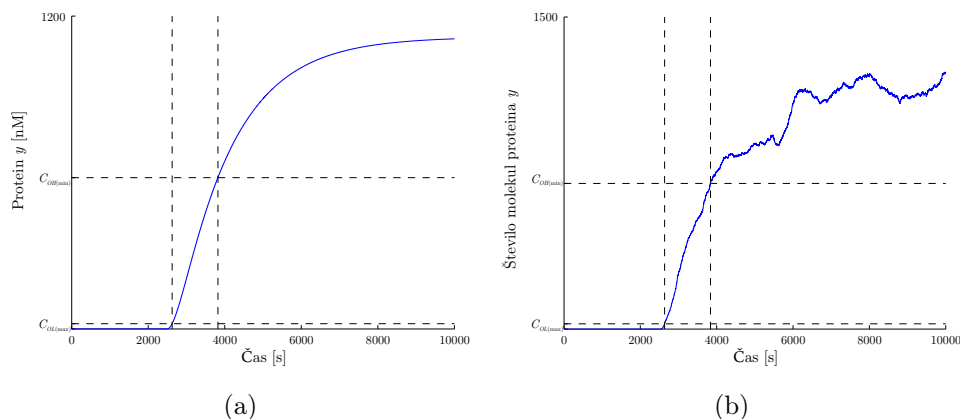
### Biološki gonilnik: določitev preklonih časov

Preklonni čas je čas, ki je potreben za preklon iz koncentracij, ki predstavljajo eno veljavno logično stanje, v koncentracije, ki predstavljajo drugo veljavno logično stanje. Merilo ga bomo kot čas, ko se signal oziroma koncentracija nosilca informacije nahajajo znotraj prepovedanega področja po spremembi vhodnega signala, ki sistem pripelje iz enega stanja v drugo. Merimo lahko torej dva tipa preklonih časov, in sicer čas vzpona ( $t_r$ ) in čas padca ( $t_f$ ).

*Čas vzpona* (angl. *rise time*) merimo kot čas potreben za preklon signala iz nizkega (logična 0) v visoko logično stanje (logična 1). Merimo ga kot čas, ko se signal po sprožitvi preklopa nahaja znotraj prepovedanega področja. Čas vzpona merimo tako na determinističnem kot tudi na stohastičnem modelu, pri čemer upoštevamo najdaljši izmerjeni čas. Primer merjenja časa vzpona na podlagi simulacijskih rezultatov prikazuje slika 6.10 (slika ne prikazuje najslabšega primera, ampak le eno od izvedenih meritev). Ker se rezultati stohastičnega modela razlikujejo, je bilo merjenje časa vzpona opravljeno na podlagi desetih izvedb simulacij. Na podlagi tega je bil določen maksimalni čas vzpona

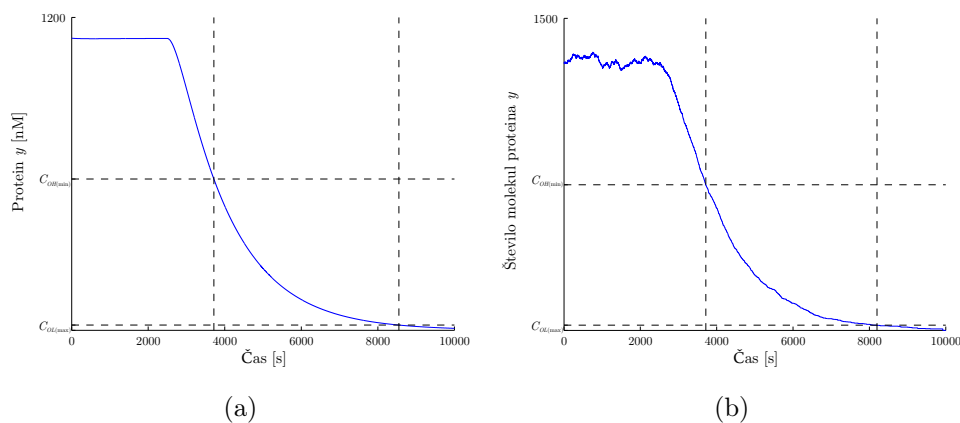
$$t_r = t_{r(\max)} \approx 2500 \text{ s}. \quad (6.30)$$

*Čas padca* (angl. *fall time*) merimo kot čas potreben za preklon signala iz visokega (logična 1) v nizko logično stanje (logična 0). Podobno kot čas vzpona merimo čas, ko se signal po sprožitvi preklopa nahaja znotraj prepovedanega področja tako na determinističnem kot tudi na stohastičnem modelu, pri čemer upoštevamo najdaljši izmerjeni



**Slika 6.10** Primer merjenja časa vzpona signala ( $t_r$ ) na podlagi simulacijskih rezultatov determinističnega (a) in stohastičnega (b) modela. Čas vzpona merimo kot čas, ko se signal nahaja znotraj prepovedanega področja po sprožitvi preklopa iz nizkega v visoko logično stanje signala.

čas. Primer merjenja časa padca na podlagi simulacijskih rezultatov prikazuje slika 6.11 (slika ne prikazuje najslabšega primera, ampak le eno od izvedenih meritev). Ker se rezul-



**Slika 6.11** Primer merjenje časa padca signala ( $t_f$ ) na podlagi simulacijskih rezultatov determinističnega (a) in stohastičnega (b) modela. Čas padca merimo kot čas, ko se signal nahaja znotraj prepovedanega področja po sprožitvi preklopa iz visokega v nizko logično stanje signala.

tati stohastičnega modela razlikujejo, je bilo merjenje časa padca opravljeno na podlagi desetih simulacij. Na podlagi tega je bil določen maksimalni čas padca

$$t_f = t_f(\max) \approx 5000 \text{ s.} \quad (6.31)$$

Glede na enačbo (6.7) lahko določimo preklopni čas kot

$$t_s = 5000 \text{ s.} \quad (6.32)$$

### Biološki gonilnik: določitev časa veljavnosti izhodnega signala

Ker čas veljavnosti izhodnega signala določamo le v primeru, ko brez zunanjih vplivov po določenem času signal izgubi oziroma spremeni svojo logično interpretacijo, ga pri biološkem gonilniku ne določamo oziroma je enak  $t_v = \infty$ .

### Biološki gonilnik: določitev robustnosti

Glede na že določene vrednosti  $C_H$ ,  $C_L$  in  $N$  (glej razdelek 6.3.1) lahko po enačbi (6.15) določimo *robustnost glede na razdaljo med stabilnimi stanji* sistema

$$R_d = \left| \frac{1\,120\text{ nM}}{335\text{ nM}} \right| \approx 3,34. \quad (6.33)$$

Ker ima sistem samo eno stabilno stanje, *robustnosti glede na razdaljo od bifurkacijskih točk* ne računamo.

### 6.3.2 Analiza biološkega negatorja

Ker je določanje vrednosti metrik pri biološkemu negatorju zelo podobno določanju metrik biološkega gonilnika, v tem razdelku navaajamo zgolj rezultate meritev pridobljene s simulacijskimi rezultati postavljenih modelov biološkega negatorja. Izbiro nosilca informacije lahko argumentiramo na podoben način kot njegovo izbiro pri biološkem gonilniku. Meritve torej delamo na poteku koncentracij izhodnega proteina  $y$  v odvisnosti od koncentracij vhodnega proteina  $x$ . Vrednosti metrik so v tem primeru sledeče:

- $C_L = 5\text{ nM} \approx 5$ ,
- $C_H = 1\,120\text{ nM} \approx 1350$ ,
- $C_{OL} = 10 \approx 8\text{ nM}$ ,
- $C_{OH} = 1100 \approx 913\text{ nM}$ ,
- $N_L < 50$ ,
- $N_H < 400$ ,
- $N = 350 \approx 290\text{ nM}$ ,
- $C_{OL(\max)} = 10 + 50 = 60 \approx 50\text{ nM}$ ,
- $C_{OH(\min)} = 1100 - 350 = 750 \approx 620\text{ nM}$ ,

- $NM_O = ]60, 750[ \approx ]50 \text{ nM}, 620 \text{ nM}[$ ,
- $C_{IL(\max)} < 60 \approx 50 \text{ nM}$ ,
- $C_{IH(\min)} > 400 \approx 330 \text{ nM}$ ,
- $NM_I = ]60, 400[ \approx ]50 \text{ nM}, 330 \text{ nM}[$ ,
- $C_I = 2$ ,
- $t_r = 1\,500 \text{ s}$ ,
- $t_f = 4\,500 \text{ s}$ ,
- $t_s = 4\,500 \text{ s}$ ,
- $t_v = \infty$ ,
- $R_d = \left| \frac{1\,120 \text{ nM}}{335 \text{ nM}} \right| \approx 3, 34$ .

### 6.3.3 Analiza bioloških AND vrat

Ker je struktura bioloških AND vrat nekoliko drugačna od biološkega gonilnika, je njihova analiza opisana podrobneje.

#### Biolška AND vrata: izbira nosilca informacije

Biolška AND vrata predstavljajo razširitev biološkega gonilnika v smislu delitve vhoda na dva dela. Na promotor se namreč lahko veže še kompleks, ki je sestavljen iz obeh vhodnih proteinov. Zaradi podobnosti z biološkim gonilnikom lahko izbiro nosilca informacije argumentiramo na enak način kot pri le-tem. Opazujemo torej koncentracijo izhodnega proteina  $y$  (glej sliki 4.12 in 4.13).

#### Biolška AND vrata: določitev nivojev izhodnih koncentracij nosilca informacije

Orientacijsko lahko logične nivoje določimo kot ravnovesna stanja, v katera sistem pride pri nizkem vhodu ( $x = 0, y = 0$ ) in visokem vhodu ( $x \gg 0, y \gg 0$ ) po izrazih

$$C_L \approx 0 \text{ nM}, \quad (6.34)$$

$$C_H \approx 1353 \approx 1\,123 \text{ nM}. \quad (6.35)$$

Vrednosti, ki predstavljajo koncentracije izhodnega signala v nizkem logičnem stanju ( $C_{OH}$ ) in koncentracije izhodnega signala v visokem logičnem stanju ( $C_{OL}$ ), določamo kot povprečje koncentracij signala, ko se le-ta po preklopu stabilizira, po izrazu

$$C_{OH} \approx 1015 \approx 840 \text{ nM}, \quad (6.36)$$

$$C_{OL} \approx 0. \quad (6.37)$$

#### Biolška AND vrata: določitev šuma

Šum je bil merjen na signalu v visokem in v nizkem logičnem stanju kot odstopanje od njegove povprečne vrednosti. Na podlagi simulacijskih rezultatov so bile določene sledeče vrednosti

$$N_H < 400, \quad (6.38)$$

$$N_L < 5, \quad (6.39)$$

pri čemer  $N_H$  predstavlja šum pri signalu v visokem,  $N_L$  pa šum pri signalu v nizkem logičnem stanju. V nadaljevanju upoštevamo maksimalni šum. Velja torej

$$N = \max(N_H, N_L) = 400 \approx 335 \text{ nM}. \quad (6.40)$$

#### Biolška AND vrata: določitev skrajnih mej nivojev izhodnih koncentracij nosilca informacije

Upoštevamo enačbi (6.24) in (6.25). Velja torej

$$C_{OL(\max)} = 0 + 5 = 5 \approx 4 \text{ nM} \quad (6.41)$$

in

$$C_{OH(\min)} = 1015 - 400 = 615 \approx 510 \text{ nM}. \quad (6.42)$$

#### Biolška AND vrata: določitev izhodnega prepovedanega področja

Izhodno prepovedano področje se nahaja na intervalu

$$NM_O = ]C_{OL(\max)}, C_{OH(\min)}[ = ]5, 615[ \approx ]4 \text{ nM}, 510 \text{ nM}[. \quad (6.43)$$

**Biolška AND vrata: določitev skrajnih mej nivojev vhodnih koncentracij nosilca informacije**

Najprej določimo zgornjo mejo nizkega nivoja vhodne koncentracije ( $C_{IL(\max)}$ ) proteinov  $x$  in  $y$ , tako da ju postopoma povečujemo. Ugotovimo lahko, da je

$$C_{IL(\max)} < 120 \approx 100 \text{ nM}. \quad (6.44)$$

Podobno določimo spodnjo mejo visokega nivoja vhodne koncentracije ( $C_{IH(\min)}$ )

$$C_{IH(\min)} > 600 \approx 500 \text{ nM}. \quad (6.45)$$

**Biolška AND vrata: določitev vhodnega prepovedanega področja**

Vhodno prepovedano področje se nahaja na intervalu

$$NM_I = ]120, 600[ \approx ]100 \text{ nM}, 500 \text{ nM}[. \quad (6.46)$$

**Biolška AND vrata: določitev porabe vhodov**

Vežje pri enaki koncentraciji vhodnih proteinov  $x$  in  $y$  v kompleks  $x^*$  pretvori celotno količino vhodnih proteinov (glej tabelo 4.4). V najslabšem primeru je poraba vhoda bioloških AND vrat torej enaka  $C_{I_x} = C_{I_y} = \infty$ . Na posamezni izhod lahko torej vežemo zgolj ena biološka AND vrata.

**Biolška AND vrata: določitev preklonih časov**

Preklonni čase merimo kot čas po sprožitvi preklopa, ko se kateri koli od obeh signalov nahaja znotraj prepovedanega področja. Na podlagi opravljenih simulacij je bil izmerjen maksimalen čas vzpona  $t_r = 1\,500$  s in maksimalen čas padca  $t_f = 5\,000$  s. Maksimalen preklonni čas je torej  $t_s = 5\,000$  s.

**Biolška AND vrata: določitev časa veljavnosti izhodnega signala**

Čas veljavnosti izhodnega signala je pri bioloških AND vratih enak  $t_v = \infty$ .

**Biolška AND vrata: določitev robustnosti**

Določimo lahko *robustnost glede na razdaljo med stabilnimi stanji* sistema

$$R_d = \left| \frac{1\,123 \text{ nM}}{335 \text{ nM}} \right| \approx 3,35. \quad (6.47)$$

Ker ima sistem samo eno stabilno stanje, *robustnosti glede na razdaljo od bifurkacijskih točk* ne računamo.

### 6.3.4 Analiza bioloških NOR vrat

Biološka NOR vrata izhajajo iz biološkega negatorja in so njegova razširitev. Izkaže se, da sta si sistema medsebojno tako sorodna, da so vrednosti metrik bioloških NOR vrat enake kot vrednosti metrik biološkega negatorja. Njihovih vrednosti zato tu ponovno ne navajamo.

Vrednosti se razlikujejo le za metriko zgornja meja nizkega nivoja vhodne koncentracije  $C_{IL(\max)}$ , saj imajo biološka NOR vrata dva vhodna signala in je zato meja nizkega nivoja vhodne koncentracije prej presežena. Do neke mere sta namreč tudi v nizkem stanju prisotna oba vhodna proteina. Na podlagi meritev lahko ovrednotimo, da je  $C_{IL(\max)} = 40 \approx 35$  nM. AND vrata se razlikujejo torej tudi v vhodnem prepovedanem področju, ki se nahaja na intervalu  $NM_I = ]40, 400[ \approx ]35$  nM, 330 nM[.

### 6.3.5 Analiza preklopnega stikala

#### Preklopno stikalo: izbira nosilca informacije

Izbiri nosilca informacije lahko argumentiramo na podoben način kot njegovo izbiro pri biološkem gonilniku in negatorju. Razlika je v tem, da tokrat analiziramo dve kemijski zvrsti, in sicer proteina  $x$  in  $y$ , pri čemer eden predstavlja trenutno stanje sistema, drugi pa njegovo negirano vrednost. Potek koncentracij proteinov in njunih mRNA v stohastičnem modelu prikazuje slika 6.12.

#### Preklopno stikalo: določitev nivojev izhodnih koncentracij nosilca informacije

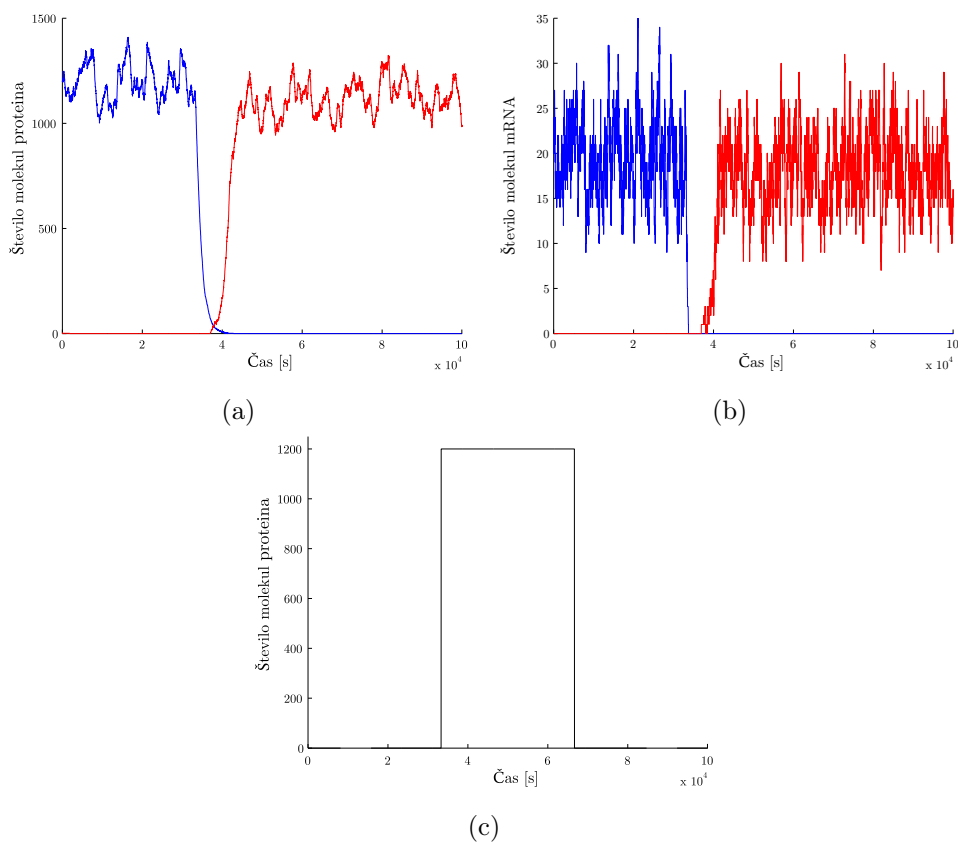
Orientacijsko lahko logične nivoje določimo na podlagi stabilnih stanj determinističnega modela preklopnega stikala, pri čemer predpostavljamo odsotnost zunanjih represorjev, saj nas zanimajo logični nivoji, ko se stanje sistema ohranja. Izmerimo lahko, da je

$$C_L \approx 4 \text{ nM} \approx 5, \quad (6.48)$$

$$C_H \approx 1123 \text{ nM} \approx 1353. \quad (6.49)$$

Podobno kot pri biološkemu gonilniku in negatorju določamo vrednosti, ki predstavljajo koncentracije izhodnega signala v nizkem logičnem stanju ( $C_{OH}$ ) in koncentracije izhodnega signala v visokem logičnem stanju ( $C_{OL}$ ). Vrednosti določamo pri signalu v





**Slika 6.12** Časovni potek števila molekul proteinov  $x$  in  $y$  (a) in njihovih mRNA (b) v odvisnosti od prisotnosti represorja  $i$  (c) v stohastičnem modelu preklopnega stikala, pri čemer modra barva prikazuje kemijske zvrsti, ki se nanašajo na protein  $x$ , rdeča pa na protein  $y$ . Predpostavljajmo, da represor  $j$  v sistemu ni prisoten.

visokem in signalu v nizkem logičnem stanju kot njegovo povprečje, ko se po preklopu stabilizira po izrazih

$$C_{OH} \approx 1115 \approx 925 \text{ nM}, \quad (6.50)$$

$$C_{OL} \approx 0. \quad (6.51)$$

### Preklopno stikalo: določitev šuma

Šum je bil merjen na signalu v visokem in nizkem logičnem stanju kot odstopanje od povprečne vrednosti signala. Izmerjene so bile vrednosti po izrazih

$$N_H < 400, \quad (6.52)$$

$$N_L < 5, \quad (6.53)$$

pri čemer  $N_H$  predstavlja šum pri signalu v visokem,  $N_L$  pa šum pri signalu v nizkem logičnem stanju. V nadaljevanju upoštevamo maksimalni šum. Velja torej

$$N = \max(N_H, N_L) = 400 \approx 335 \text{ nM}. \quad (6.54)$$

#### Preklopno stikalo: določitev časa osveževanja

Predpostavljajmo, da je potrebno vezje osvežiti, ko postane razdalja med signalom v visokem in signalom v nizkem logičnem stanju premajhna. Iz simulacijskih rezultatov je razvidno, da se stanje obravnavanega vezja ohranja brez osveževanja (glej poglavje 4.4.6). Čas osveževanja je torej enak  $t_{ref} = \infty$ .

#### Preklopno stikalo: določitev skrajnih mej nivojev izhodnih koncentracij nosilca informacije

Skrajne meje nivojev izhodnih koncentracij določamo na podlagi koncentracij  $C_{OH}$  in  $C_{OL}$  in šuma v nizkem in visokem stanju ( $N_L$  in  $N_H$ ). Pri tem upoštevamo enačbi (6.24) in (6.25). V našem primeru velja, da je

$$C_{OL(\max)} = 0 + 5 = 5 \approx 5 \text{ nM} \quad (6.55)$$

in

$$C_{OH(\min)} = 1115 - 400 = 715 \approx 590 \text{ nM}. \quad (6.56)$$

#### Preklopno stikalo: določitev izhodnega prepovedanega področja

Izhodno prepovedano področje se nahaja na intervalu

$$NM_O = ]C_{OL(\max)}, C_{OH(\min)}[ = ]5, 715[ \approx ]5 \text{ nM}, 590 \text{ nM}[. \quad (6.57)$$

#### Preklopno stikalo: določitev skrajnih mej nivojev vhodnih koncentracij nosilca informacije

Določiti je potrebno koncentracije, ki oba izhodna signala, tj.  $x$  in  $y$ , še pripeljejo v veljavno logično stanje, in sicer moramo določiti spodnjo mejo visokega nivoja vhodne koncentracije  $C_{IH(\min)}$ , ki izhod biološkega gonilnika še pripelje v veljavno visoko logično stanje ( $C > C_{OH(\min)}$ ), in zgornjo mejo nizkega nivoja vhodne koncentracije  $C_{IL(\max)}$ , ki izhod biološkega gonilnika še pripelje v veljavno nizko logično stanje ( $C < C_{OL(\max)}$ ).

Najprej določimo zgornjo mejo nizkega nivoja vhodne koncentracije ( $C_{IL(\max)}$ ) proteina  $i$ . Pri tem predpostavljamo, da je koncentracija proteina  $j$  enaka 0. Zanima nas, do kakšne mere lahko povečamo koncentracijo vhodnega represorja, pri čemer se mora stanje ohranjati (ne sme priti do preklopa), signala pa morata ostati zunaj prepovedanega področja. Na podlagi simulacijskih rezultatov velja sledeči izraz

$$C_{IL(\max)} < 50 \approx 40 \text{ nM}. \quad (6.58)$$

V nadaljevanju določimo spodnjo mejo visokega nivoja vhodne koncentracije  $C_{IH(\min)}$  proteina  $j$ , pri čemer predpostavljamo, da je koncentracija proteina  $i$  enaka  $C_{IL(\max)}$  (najslabši možni primer). Zanima nas, do kakšne mere lahko zmanjšamo koncentracijo vhodnega represorja, pri čemer mora priti do preklopa, signala pa morata ostati zunaj prepovedanega področja. Na podlagi simulacijskih rezultatov smo določili, da je

$$C_{IH(\min)} > 300 \approx 250 \text{ np}. \quad (6.59)$$

**Preklopno stikalo: določitev vhodnega prepovedanega področja**

Vhodno prepovedano področje se nahaja na intervalu

$$NM_I = ]50, 300[ \approx ]40 \text{ nM}, 250 \text{ nM}[. \quad (6.60)$$

**Preklopno stikalo: določitev porabe vhodov**

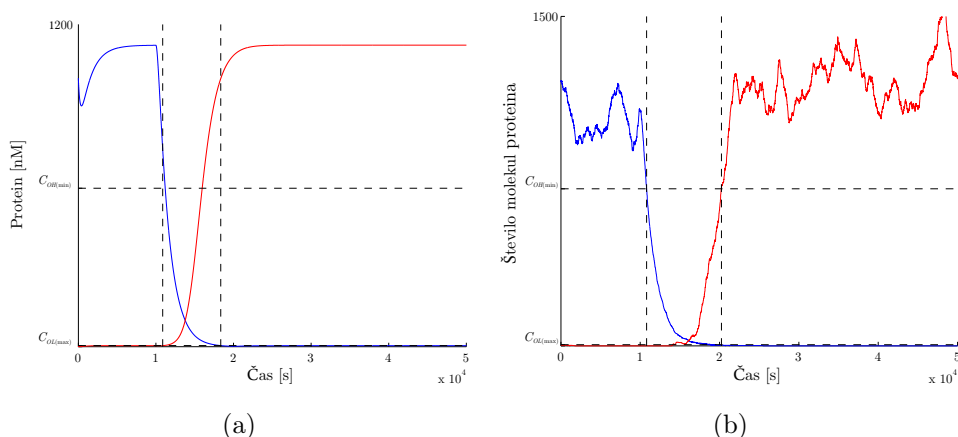
Glede na upoštevane reakcije (glej tabelo 4.6) je maksimalna poraba obeh vhodnih represorjev pri eni molekuli DNA  $C_{I_i} = C_{I_j} = 2$ .

**Preklopno stikalo: določitev preklopnih časov**

Preklopne čase merimo kot čas po sprožitvi preklopa, ko se kateri koli od obeh signalov nahaja znotraj prepovedanega področja (glej sliko 6.13). Na podlagi opravljenih meritev je bil izmerjen maksimalen preklopni čas  $t_s = 10\,000 \text{ s} \approx 3 \text{ h}$ .

**Preklopno stikalo: določitev časa veljavnosti izhodnega signala**

Čas veljavnosti izhodnega signala je pri preklopnemu stikalu enak času osveževanja, in sicer  $t_v = \infty$ .



**Slika 6.13** Merjenje časa preklopa pri preklopnem stikalu na podlagi simulacijskih rezultatov determinističnega (a) in stohastičnega (b) modela. Čas preklopa merimo kot čas, ko se kateri koli od koncentracij proteinov  $x$  ali  $y$  nahaja znotraj prepovedanega področja po sprožitvi preklopa.

### Preklopno stikalo: določitev robustnosti

#### *Robustnost glede na razdaljo med stabilnimi stanji*

Stabilnostno analizo delamo na sistemu, ko le-ta pomni svoje stanje - koncentracija zunanjih represorjev je enaka 0. Stabilnostna analiza preklopnega stikala pokaže, da ima dve stabilni in eno nestabilno ravnovesno stanje. Ravnovesne točke sistema so:

- $\mathbf{x}_1 = (3'556\text{nM}, 1122'634\text{nM}),$
- $\mathbf{x}_2 = (1122'634\text{nM}, 3'556\text{nM}),$
- $\mathbf{x}_3 = (157\text{nM}, 157\text{nM}),$

pri čemer nadaljnja analiza stabilnosti pokaže, da sta točki  $\mathbf{x}_1$  in  $\mathbf{x}_2$  stabilni, točka  $\mathbf{x}_3$  pa nestabilna. Na podlagi razdalje med točkama lahko izračunamo *robustnost glede na razdaljo med stabilnimi stanji*

$$R_d = \frac{d_{1,2}}{N} = \frac{1\,582,6\text{ nM}}{335\text{ nM}} \approx 4,72. \quad (6.61)$$

#### *Robustnost glede na točke bifurkacije*

Preko bifurkacijske analize lahko pridemo do območij vrednosti posameznega parametra, znotraj katerih sistem odraža bistabilno delovanje. Pri tem spreminjamo samo en parameter naenkrat (npr. v primeru spreminjanja hitrosti razgradnje proteina, spreminjamo samo hitrost razgradnje za enega od opazovanih proteinov). Bifurkacijsko analizo

delamo le nad parametri, za katere predpostavljamo variiranje (za Hillov koeficient je ne delamo). Vrednosti parametrov so sledeče:

- hitrost transkripcije:  $\alpha \in [0,02922656\text{s}^{-1}, 0,4161850\text{s}^{-1}]$  ( $d_1 = 0,04227344\text{s}^{-1}$ ),
- hitrost puščanja:  $\alpha_0 \in [0\text{s}^{-1}, 0,00371759\text{s}^{-1}]$  ( $d_2 = 0,00371759\text{s}^{-1}$ ),
- hitrost translacije:  $\beta \in [0,01757656\text{s}^{-1}, 0,2502937\text{s}^{-1}]$  ( $d_3 = 0,02542344\text{s}^{-1}$ ),
- razmerje med disociacijo in vezavo zunanega represorja na operator proteina:  $K_{d1}$  ne določamo, ker vpliva zgolj na preklon in ne na bistabilnost,
- razmerje med disociacijo in vezavo na operator nasprotnega proteina:  $K_{d2} \in [10,85533, 154,5721]$  ( $d_4 = 52,33087$ ),
- razgradnja mRNA:  $\delta_m \in [0,00066875\text{s}^{-1}, 0,009541174\text{s}^{-1}]$  ( $d_5 = 0,005641174\text{s}^{-1}$ ),
- razgradnja proteina:  $\delta \in [0,00011875\text{s}^{-1}, 0,0017125\text{s}^{-1}]$  ( $d_6 = 0,00058125\text{s}^{-1}$ ).

Na podlagi določenih vrednosti lahko ovrednotimo robustnost glede na točke bifurkacije

$$R_b = \min \left( \frac{d_1}{\alpha}, \frac{d_2}{\alpha_0}, \frac{d_3}{\beta}, \frac{d_4}{K_{d2}}, \frac{d_5}{\delta_m}, \frac{d_6}{\delta} \right) \approx 0,59124. \quad (6.62)$$

Najšibkejši (minimalni) člen v našem primeru prispevata bifurkacijski razdalji glede na hitrost transkripcije in hitrost translacije.

### 6.3.6 Analiza razširjenega preklopnega stikala

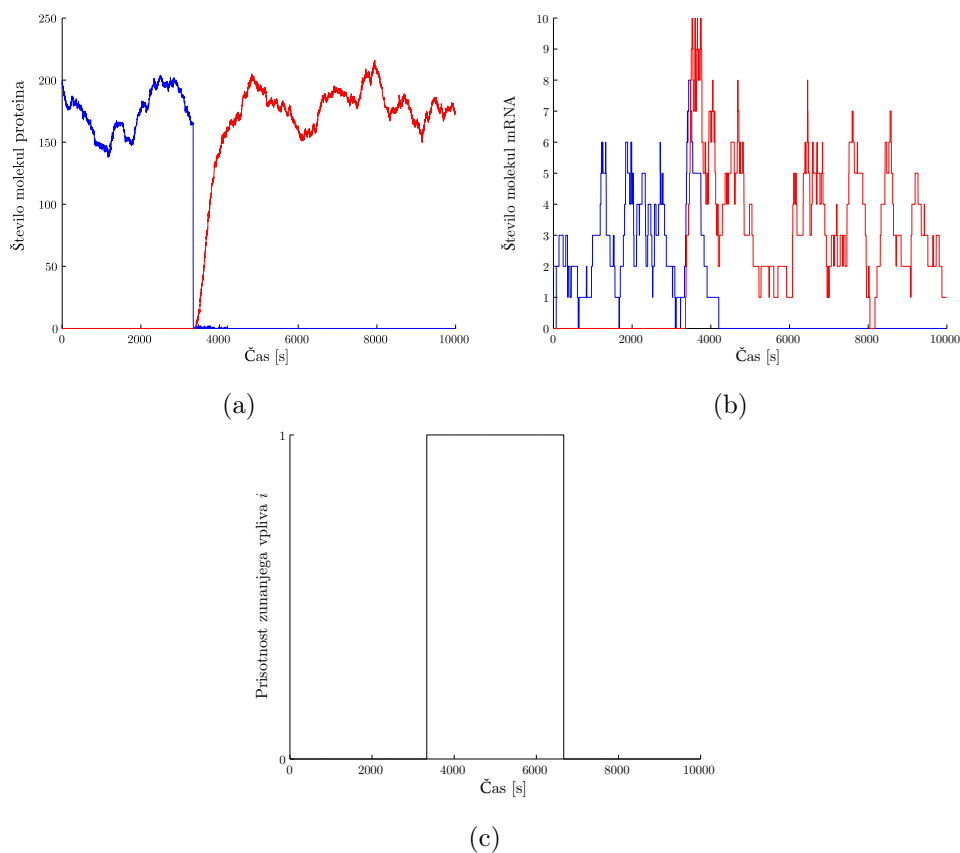
**Razširjeno preklopno stikalo: izbira nosilca informacije**

Izbiri nosilca informacije lahko argumentiramo na enak način kot pri že obravnavanih sistemih. Potek koncentracij proteinov  $x$  in  $y$  in njunih mRNA v stohastičnem modelu prikazuje slika 6.14.

**Razširjeno preklopno stikalo: določitev nivojev izhodnih koncentracij nosilca informacije**

Ker je določanje nivojev izhodnih koncentracij enako kot pri preklopnem stikalu analiziranem v razdelku 6.3.5, navajamo zgolj rezultate meritev, pridobljene na podlagi simulacij razširjenega preklopnega stikala:

$$C_L \approx 2\text{ nM} \approx 2, \quad (6.63)$$



**Slika 6.14** Časovni potek števila molekul proteinov  $x$  in  $y$  (a) in njihovih mRNA (b) v odvisnosti od prisotnosti induktorja  $i$  (c) v stohastičnem modelu razširjenega preklopnega stikala, pri čemer modra barva prikazuje kemijske zvrsti, ki se nanašajo na protein  $x$ , rdeča pa na protein  $y$ .

$$C_H \approx 155 \text{ nM} \approx 190, \quad (6.64)$$

$$C_{OH} \approx 175 \approx 145 \text{ nM}, \quad (6.65)$$

$$C_{OL} \approx 5 \approx 5 \text{ nM}. \quad (6.66)$$

### Razširjeno preklopno stikalo: določitev šuma

Šum je bil merjen na signalu v visokem in v nizkem logičnem stanju kot odstopanje od povprečne vrednosti signala. Izmerjene so bile vrednosti po izrazih

$$N_H < 85, \quad (6.67)$$

$$N_L < 5. \quad (6.68)$$

V nadaljevanju upoštevamo maksimalni šum. Velja izraz

$$N = \max(N_H, N_L) = 85 \approx 70 \text{ nM}. \quad (6.69)$$

#### Razširjeno preklopno stikalo: določitev časa osveževanja

Iz simulacijskih rezultatov je razvidno, da se stanje obravnavanega vezja ohranja brez osveževanja.

#### Razširjeno preklopno stikalo: določitev skrajnih mej nivojev izhodnih koncentracij nosilca informacije

Ker je določanje skrajnih mej nivojev izhodnih koncentracij enako kot pri preklopnem stikalu analiziranem v razdelku 6.3.5, navajamo zgolj rezultate

$$C_{OL(\max)} = 5 + 5 = 10 \approx 8 \text{ nM}, \quad (6.70)$$

$$C_{OH(\min)} = 175 - 85 = 90 \approx 75 \text{ nM}. \quad (6.71)$$

#### Razširjeno preklopno stikalo: določitev izhodnega prepovedanega področja

Izhodno prepovedano področje se nahaja na intervalu

$$NM_O = ]10, 85[ \approx ]8 \text{ nM}, 70 \text{ nM}[. \quad (6.72)$$

#### Razširjeno preklopno stikalo: določitev skrajnih mej nivojev vhodnih koncentracij nosilca informacije

Ker predpostavljamo, da kot vhodni nosilec informacije v sistem vstopa zunanji vpliv v obliki določenega svetlobnega spektra, njegove prisotnosti ne merimo s koncentracijo ampak le z dvema nivojema, in sicer *ni prisoten* (0) in *je prisoten* (1). Določanje mej vhodnih koncentracij torej v primeru tovrstnega sistema ni potrebno.

#### Razširjeno preklopno stikalo: določitev vhodnega prepovedanega področja

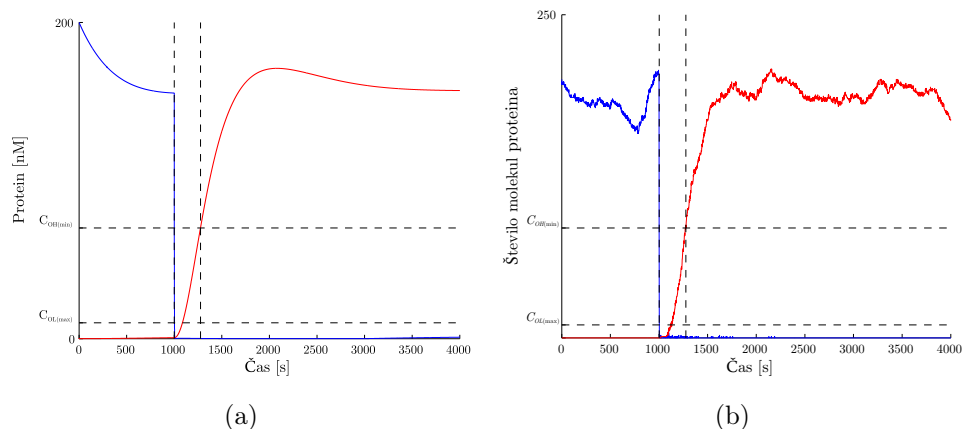
Ker ne določamo nivojev vhodnih koncentracij, ne določamo niti vhodnega prepovedanega področja.

#### Razširjeno preklopno stikalo: določitev porabe vhodov

Ker predpostavljamo, da kot vhodni nosilec informacije v sistem vstopa zunanji vpliv v obliki določenega svetlobnega spektra, o njegovi porabi težko govorimo oziroma je ničelna. Velja torej  $C_{I_i} = C_{I_j} = 0$ .

### Razširjeno preklopno stikalo: določitev preklonih časov

Na podlagi opravljenih meritev (glej sliko 6.15) je bil izmerjen maksimalen preklonni čas  $t_s = 350$  s.



**Slika 6.15** Merjenje časa preklopa pri razširjenem preklopnem stikalu na podlagi simulacijskih rezultatov determinističnega (a) in stohastičnega (b) modela. Čas preklopa merimo kot čas, ko se kateri koli od koncentracij proteinov  $x$  ali  $y$  nahaja znotraj prepovedanega področja po sprožitvi preklopa.

### Razširjeno preklopno stikalo: določitev časa veljavnosti izhodnega signala

Čas veljavnosti izhodnega signala je pri razširjenem preklopnem stikalu enak času osveževanja, in sicer  $t_v = \infty$ .

### Razširjeno preklopno stikalo: določitev robustnosti

#### *Robustnost glede na razdaljo med stabilnimi stanji*

Stabilnostna analiza preklopnega stikala pokaže, da ima dve stabilni in eno nestabilno ravnovesno stanje. Ravnovesne točke sistema so:

- $\mathbf{x}_1 = (1'8372\text{nM}, 156'3151\text{nM}),$
- $\mathbf{x}_2 = (156'3151\text{nM}, 1'8372\text{nM}),$
- $\mathbf{x}_3 = (35'0697\text{nM}, 35'0697\text{nM}),$

pri čemer nadaljnja analiza stabilnosti pokaže, da sta točki  $\mathbf{x}_1$  in  $\mathbf{x}_2$  stabilni, točka  $\mathbf{x}_3$  pa nestabilna. Na podlagi razdalje med točkama lahko izračunamo *robustnost glede na*



razdaljo med stabilnimi stanji

$$R_d = \frac{d_{1,2}}{N} = \frac{218,5 \text{ nM}}{70 \text{ nM}} \approx 3,12. \quad (6.73)$$

*Robustnost glede na točke bifurkacije*

Vrednosti bifurkacijskih točk parametrov so sledeče:

- hitrost transkripcije:  $\alpha \in [0,01895684\text{s}^{-1}, 0,3186481\text{s}^{-1}]$  ( $d_1 = 0,05254316\text{s}^{-1}$ ),
- hitrost puščanja:  $\alpha_0 \in [0\text{s}^{-1}, 0,0009216309\text{s}^{-1}]$  ( $d_2 = 0,0009216309\text{s}^{-1}$ ),
- hitrost translacije:  $\beta \in [0,01140015\text{s}^{-1}, 0,1916351\text{s}^{-1}]$  ( $d_3 = 0,03159985\text{s}^{-1}$ ),
- razmerje med disociacijo in vezavo na operator nasprotnega proteina:  
 $K_{d1} \in [1,417955, 13,40188]$  ( $d_4 = 4,900645$ ),
- razmerje med disociacijo in vezavo na svoj operator:  
 $K_{d2} \in [12,86555, \infty]$  ( $d_5 = 50,32065$ ),
- razgradnja mRNA:  $\delta_m \in [0,000875\text{s}^{-1}, 0,01471094\text{s}^{-1}]$  ( $d_6 = 0,01081094\text{s}^{-1}$ ),
- razgradnja proteina:  $\delta \in [0,00015625\text{s}^{-1}, 0,002640918\text{s}^{-1}]$  ( $d_7 = 0,00054375\text{s}^{-1}$ ).

Na podlagi določenih vrednosti lahko ovrednotimo robustnost glede na točke bifurkacije

$$R_b = \min \left( \frac{d_1}{\alpha}, \frac{d_2}{\alpha_0}, \frac{d_3}{\beta}, \frac{d_4}{K_{d1}}, \frac{d_5}{K_{d2}}, \frac{d_6}{\delta_m}, \frac{d_7}{\delta} \right) \approx 0,73488. \quad (6.74)$$

Najšibkejši (minimalni) člen v našem primeru prispeva bifurkacijska razdalja glede na hitrost transkripcije.

### 6.3.7 Analiza represilatorja

Evalvacija vrednosti določenih metrik je pri represilatorju zelo podobna evalvaciji metrik že analiziranih bioloških sistemov, zato zanje v tem razdelku navajamo zgolj rezultate meritev pridobljene s simulacijami postavljenih modelov represilatorja. Po drugi strani je evalvacija določenih metrik specifična zgolj za oscilatorne sisteme. Postopek določanja le-teh navajamo posebej.

#### Represilator: izbira nosilca informacije

Izbira nosilca informacije lahko argumentiramo na podoben način kot njegovo izbiro pri ostalih že analiziranih sistemih. Opazujemo torej koncentracijo proteinov  $x$ ,  $y$  in  $z$  (glej razdelek 4.4.8).

### Represilator: določitev nivojev izhodnih koncentracij nosilca informacije

V grobem nivoje koncentracij določimo kot ekstreme koncentracij, tj. minimalne in maksimalne koncentracije opazovanih proteinov v determinističnem modelu

$$C_L = 6 \text{ nM} \approx 7, \quad (6.75)$$

$$C_H = 435 \text{ nM} \approx 524. \quad (6.76)$$

Natančneje nivoje izhodnih koncentracij, ki določajo visoko stanje signala, določamo kot povprečje maksimalne vrednosti koncentracije opazovanega proteina v eni periodi represilatorja, koncentracij, ki določajo nizko stanje signala pa kot povprečje minimalne vrednosti koncentracije opazovanega proteina v eni periodi represilatorja (glej sliko 6.16).

Na ta način pridemo do vrednosti iz izraza

$$C_{OL} = 5 \approx 5 \text{ nM}. \quad (6.77)$$

$$C_{OH} = 800 \approx 665 \text{ nM}. \quad (6.78)$$

### Represilator: določitev šuma

Šum določamo kot odstopanje ekstremnih koncentracij v eni periodi represilatorja od izhodnih nivojev koncentracij. Pri tem posebej računamo šum pri nizkem signalu ( $N_L$ ) in posebej pri visokem signalu ( $N_H$ ):

$$N_L < 10, \quad (6.79)$$

$$N_H < 400. \quad (6.80)$$

Določimo lahko maksimalni šum po izrazu

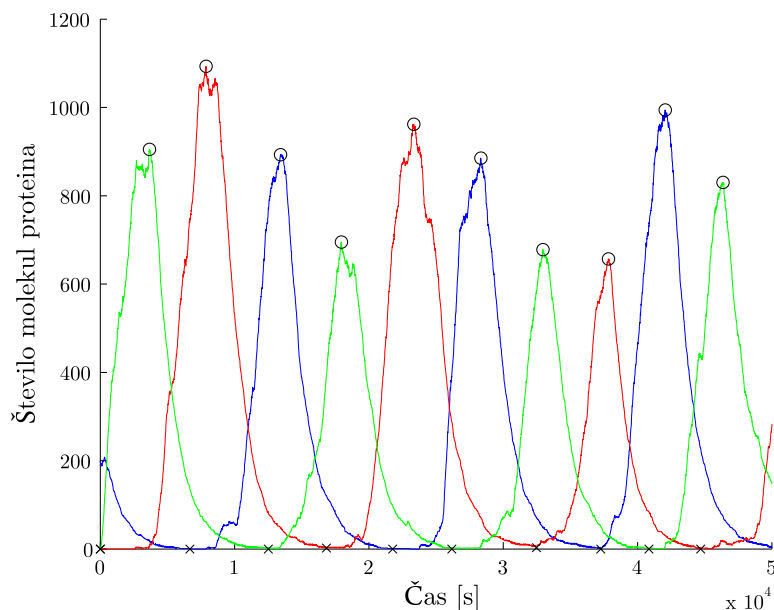
$$N = \max(N_H, N_L) = 400 \approx 330 \text{ nM}. \quad (6.81)$$

### Represilator: določitev skrajnih mej nivojev izhodnih koncentracij nosilca informacije

Skrajne meje nivojev izhodnih koncentracij določamo na enak način kot pri preklopnem stikalu, in sicer glede na enačbi (6.24) in (6.25). Veljata torej izraza

$$C_{OL(\max)} = 5 + 10 = 15 \approx 12,5 \text{ nM}, \quad (6.82)$$

$$C_{OH(\min)} = 800 - 400 = 400 \approx 330 \text{ nM}. \quad (6.83)$$



**Slika 6.16** Nivoje izhodnih koncentracij merimo kot ekstreme koncentracij dosežene v eni periodi represilatorja. Točke, na podlagi katerih računamo visoki nivo izhodnih koncentracij, so označene s krogcem. Točke, na podlagi katerih računamo nizki nivo izhodnih koncentracij, so označene s križcem. Pri tem modra barva prikazuje potek koncentracije proteina  $x$ , rdeča proteina  $y$ , zelena pa proteina  $z$ .

#### Represilator: določitev izhodnega prepovedanega področja

Na podlagi skrajnih mej nivojev izhodnih koncentracij lahko določimo izhodno prepovedano področje kot

$$NM_O = ]15,400[ \approx ]12,5 \text{ nM}, 330 \text{ nM}[. \quad (6.84)$$

#### Represilator: določitev skrajnih mej nivojev vhodnih koncentracij nosilca informacije

Ker represilator predstavlja vezje brez zunanjsih vhodov, določanje nivojev vhodnih koncentracij ni potrebno.

#### Represilator: določitev vhodnega prepovedanega področja

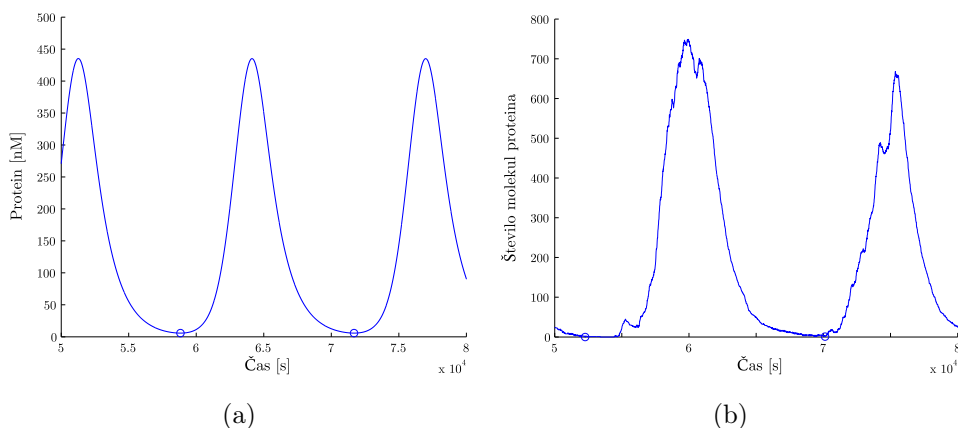
Ker ne določamo nivojev vhodnih koncentracij, ne določamo niti vhodnega prepovedanega področja.

#### Represilator: določitev porabe vhodov

Represilator deluje brez zunanjsih vhodov, zato njihove porabe ne določamo.

### Represilator: določitev periode oscilacij

Periode oscilacij represilatorja merimo kot povprečni čas, ki je potreben, da koncentracija opazovanega proteina spet doseže minimalno vrednost, pri čemer izvajanje meritev začnemo v minimalni vrednosti koncentracije proteina. Meritve izvajamo na obeh modelih (glej sliko 6.17). Pri merjenju periode oscilacij na determinističnem modelu vselej dobimo



**Slika 6.17** Merjenje periode oscilacij na koncentraciji proteina  $x$  v determinističnem (a) in stohastičnem (b) modelu represilatorja, pri čemer krogca prikazujeta začetek in konec ene periode.

enak rezultat po izrazu

$$t_p(det) \approx 12\,850\text{ s.} \quad (6.85)$$

Po drugi strani periode oscilacij pri stohastičnem modelu variirajo, zato vzamemo njihovo povprečje

$$t_p(stoh) \approx 13\,500\text{ s.} \quad (6.86)$$

### Represilator: določitev časa veljavnosti izhodnega signala

Čas veljavnosti izhodnega signala določamo na podlagi ene periode represilatorja. Čas veljavnosti izhodnega signala v visokem stanju  $t_{v_H}$  merimo kot čas, ko je koncentracija signala večja od  $C_{OH(\min)}$ . Čas veljavnosti izhodnega signala v nizkem stanju  $t_{v_L}$  merimo kot čas, ko je koncentracija signala manjša od  $C_{OL(\max)}$ . Čase merimo tako na determinističnem kot tudi na stohastičnem modelu. Pridemo do sledečih rezultatov po izrazih

$$t_{v_H}(det) = 2\,500\text{ s,} \quad (6.87)$$

$$t_{v_H}(stoh) = 3\,500\text{ s,} \quad (6.88)$$

$$t_{v_L}(det) = 3\,000\text{ s}, \quad (6.89)$$

$$t_{v_L}(stoh) = 4\,000\text{ s}. \quad (6.90)$$

Čas veljavnosti lahko določimo kot minimalni čas izmerjenih časov po izrazu

$$t_v = \min(t_{v_H}(det), t_{v_H}(stoh), t_{v_L}(det), t_{v_L}(stoh)) = 2\,500\text{ s}. \quad (6.91)$$

### Represilator: določitev robustnosti

#### *Robustnost glede na velikost stabilnega limitnega cikla*

Stabilnostna analiza represilatorja s podanimi parametri pokaže, da ima sistem eno nestabilno ravnovesno stanje in en stabilen limitni cikel. Pri tem se nestabilno stanje nahaja v točki  $\mathbf{x} = (35\,2\text{nM}, 35\,2\text{nM})$ , limitni cikel z velikostjo  $d_c = 429\,4\text{nM}$  pa na intervalu  $[5\,8\text{nM}, 435\,2\text{nM}]$ . Na podlagi velikosti limitnega cikla določenega na podlagi determinističnega modela in šuma izmerjenega na podlagi stohastičnega modela, lahko določimo *robustnost glede na velikost stabilnega limitnega cikla* po izrazu

$$R_c = \frac{d_c}{N} = \frac{429,4\text{ nM}}{330} \approx 1,3. \quad (6.92)$$

#### *Robustnost glede na točke bifurkacije*

Preko bifurkacijske analize lahko na podoben način kot pri preklopnem stikalu pridemo do območij vrednosti posameznega parametra, znotraj katerih sistem odraža oscilatorno delovanje. Pri tem spreminjamo samo en parameter naenkrat. Bifurkacijsko analizo delamo le nad parametri, za katere predpostavljamo variiranje. Vrednosti parametrov so sledeče:

- hitrost transkripcije:  $\alpha \in [0\,002802734\text{s}^{-1}, 51\,62438\text{s}^{-1}]$  ( $d_1 = 0,068697266\text{s}^{-1}$ ),
- hitrost puščanja:  $\alpha_0 \in [0\text{s}^{-1}, 0\,001216943\text{s}^{-1}]$  ( $d_2 = 0,001216943\text{s}^{-1}$ ),
- hitrost translacije:  $\beta \in [0\,001685791\text{s}^{-1}, 31\,104683\text{s}^{-1}]$  ( $d_3 = 0,041314209\text{s}^{-1}$ ),
- razmerje med disociacijo in vezavo na operator sledečega proteina:  
 $K_d \in [0\,00875, 161\,2472]$  ( $d_4 = 6,30985$ ),
- razgradnja mRNA:  $\delta_m \in [0\,00010125\text{s}^{-1}, 0\,07876709\text{s}^{-1}]$  ( $d_5 = 0,00379875\text{s}^{-1}$ ),
- razgradnja proteina:  $\delta \in [0\,00014\text{s}^{-1}, 0\,01158135\text{s}^{-1}]$  ( $d_6 = 0,00056\text{s}^{-1}$ ).

Na podlagi določenih vrednosti lahko ovrednotimo robustnost glede na točke bifurkacije po izrazu

$$R_b = \min \left( \frac{d_1}{\alpha}, \frac{d_2}{\alpha_0}, \frac{d_3}{\beta}, \frac{d_4}{K_d}, \frac{d_5}{\delta_m}, \frac{d_6}{\delta} \right) \approx 0,8. \quad (6.93)$$

Najšibkejši (minimalni) člen v našem primeru prispeva bifurkacijska razdalja glede na hitrost razgradnje proteina.

## 6.4 Zaključek poglavja

Vrednotenje metrik na posameznem sistemu oziroma na simulacijskih rezultatih njegovih modelov nam omogoča vzpostavitev objektivne ocene dinamike preklopa sistema za potrebe procesiranja podatkov. V množici gradnikov z na videz podobnimi funkcionalnostmi lahko tako izberemo najprimernejšega v smislu prehajanja med logičnimi nivoji, koncentracij, ki določajo logične nivoje, časovnih karakteristik sistema in robustnosti delovanja sistema. Pri načrtovanju kompleksnejših sistemov, sestavljenih iz več osnovnih gradnikov, lahko na podlagi vzpostavljenih metrik analiziramo njihovo medsebojno kompatibilnost in tako k njihovi gradnji pristopimo na bolj sistematičen način v nasprotju z uveljavljenim *ad-hoc* pristopom. Zgled takega načrtovanja je predstavljen v poglavju 7.

Predstavljene metrike lahko poleg tega uporabljamo za vzpostavitev kriterijev, ki jih je potrebno definirati pri avtomatskem načrtovanju bioloških sistemov. Metrike imajo v tem primeru vlogo vhodov v t. i. *fitnes funkcije*, na podlagi katerih ocenjujemo uspešnost oziroma ustreznost delovanja posameznega sistema. S tem lahko pri načrtovanju bioloških sistemov uporabimo različne metahevristične pristope za preiskavo prostora rešitev, kot so genetski algoritmi, simulirano ohlajanje, evolucijsko programiranje in drugi [61]. Glede na podano topologijo gensko regulatornega omrežja lahko z naštetimi pristopi uglasimo delovanje sistema s spreminjanjem vrednosti parametrov reakcij, količine plazmidov, koncentracij vhodnih kemijskih zvrsti ipd.

Poudariti je potrebno, da so vrednosti metrik strogo odvisne od samega modela, ki mora do zadostne mere odražati ujemanje z dejanskim delovanjem opazovanega biološkega sistema. Večino vzpostavljenih metrik lahko sicer ovrednotimo neposredno na rezultatih laboratorijskih eksperimentov, v primeru analize na podlagi simulacijskih rezultatov pa je pogoj za uporabnost dobljenih vrednosti dovolj dober model, ki v danih pogojih odraža dinamiko primerljivo z dejanskim delovanjem biološkega sistema.

# 7 Izgradnja kompleksnejšega biološkega sistema

Predpostavimo, da kompleksnejši sistem gradimo s povezovanjem osnovnih bioloških sistemov. Izgradnja torej poteka modularno, pri čemer si lahko pomagamo z metrikami, ki so bile vzpostavljene v poglavju 6. Analiza možnosti povezovanja gradnikov oziroma analiza njihove kompatibilnosti je predstavljena v pričujočem poglavju, ki mu sledi zgled izgradnje modela kompleksnejšega biološkega sistema zmožnega procesiranja podatkov sestavljenega iz preprostejših sistemov, ki so bili opisani v razdelku 3.3 in okarakterizirani v razdelku 6.3.

## 7.1 Povezovanje osnovnih gradnikov

Pri povezovanju osnovnih gradnikov moramo še posebej upoštevati naslednje zahteve:

- ortogonalnost sestavnih delov,
- kompatibilnost v smislu logičnih nivojev,
- kompatibilnost v smislu časovnih karakteristik,
- ohranjanje logičnih nivojev.

### 7.1.1 Ortogonalnost sestavnih delov

Dva sestavna dela, npr. proteina, sta med seboj *ortogonalna*, če neposredno ne vplivata na izražanje drug drugega [62]. Zagotoviti je potrebno, da so proteini, ki v usmerjenem grafu genskega regulatornega omrežja niso povezani (glej sliki 2.3 in 2.4), med seboj ortogonalni. V našem primeru ortogonalnosti proteinov ne bomo preverjali, saj predpostavljamo, da imamo na razpolago zadostno količino medsebojno ortogonalnih proteinov za izgradnjo načrtovanega sistema.

### 7.1.2 Kompatibilnost gradnikov v smislu logičnih nivojev

Gradnika sta *kompatibilna v smislu logičnih nivojev*, če se vhodni logični nivoji vhodnega gradnika ujemaajo z izhodnimi logičnimi nivoji izhodnega gradnika. Veljati morata torej relaciji

$$C_{OOL(\max)} \leq C_{IIL(\max)} \quad (7.1)$$

in

$$C_{OOH(\min)} \geq C_{I IH(\min)}. \quad (7.2)$$

Pri tem vrednost  $C_{OOL(\max)}$  določa zgornjo mejo nizkega nivoja izhodne koncentracije izhodnega gradnika,  $C_{IIL(\max)}$  zgornjo mejo nizkega nivoja vhodne koncentracije vhodnega gradnika,  $C_{OOH(\min)}$  spodnjo mejo visokega nivoja izhodne koncentracije izhodnega gradnika,  $C_{I IH(\min)}$  pa spodnjo mejo visokega nivoja vhodne koncentracije vhodnega gradnika. Pogoji lahko zapišemo tudi v sledeči obliki

$$NM_{Oo} \subseteq NM_{Ii}, \quad (7.3)$$

pri čemer  $NM_{Ii}$  določa vhodno prepovedano področje vhodnega gradnika,  $NM_{Oo}$  pa izhodno prepovedano področje izhodnega gradnika.

Kompatibilnost gradnikov v smislu logičnih nivojev mora veljati zgolj za gradnike, ki so med seboj *sosejni*, tj. za gradnike, ki so neposredno povezani preko skupnega proteina, ki v izhodnem gradniku nastopa kot izhod, v vhodnem gradniku pa kot vhod.

### 7.1.3 Kompatibilnost gradnikov v smislu časovnih karakteristik

Gradnika sta *kompatibilna v smislu časovnih karakteristik*, če je čas veljavnosti izhodnega signala izhodnega gradnika večji ali enak od časa preklopa signala vhodnega gradnika.



Veljati mora torej pogoj

$$t_{v_O} \geq t_{s_I}, \quad (7.4)$$

pri čemer  $t_{v_O}$  določa minimalni čas veljavnosti signala izhodnega gradnika,  $t_{s_I}$  pa čas preklopa vhodnega gradnika.

Medtem, ko mora kompatibilnost v smislu logičnih nivojev veljati le za sosednje gradnike, mora kompatibilnost v smislu časovnih karakteristik veljati za vse gradnike, ki so direktno ali preko vmesnih gradnikov vezani na izhod izhodnega gradnika s kritično vrednostjo časa  $t_{v_O}$ . Čas veljavnosti signala se namreč prenese na vse gradnike, ki mu v sekvenci sledijo.

#### 7.1.4 Ohranjanje logičnih nivojev

Upoštevati moramo dejstvo, da vsak vhod porabi določeno količino izhodnega proteina določeno z metriko  $C_I$  (glej razdelek 6.2.6). V določenih primerih je poraba zanemarljiva, v določenih primerih pa ne, zato jo je potrebno dosledno upoštevati.

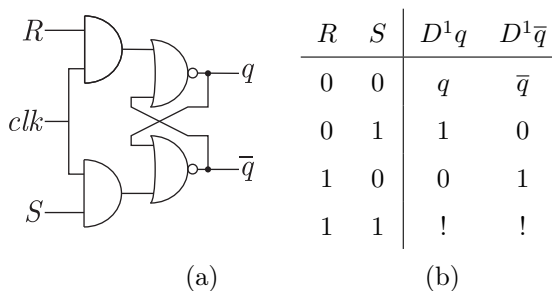
## 7.2 Zgled kompleksnejšega biološkega sistema zmožnega procesiranja podatkov

Izgradnjo kompleksnejšega biološkega sistema zmožnega procesiranja podatkov oziroma izgradnjo njegovega modela bomo prikazali na zgledu izgradnje modela *sinhronnega preklonnega stikala*, ki je po funkcionalnosti enak *RS* pomnilni celici s sinhronizacijo. V nadaljevanju najprej opišemo želeno delovanje biološkega sistema, nato pa sledi opis izgradnje modela biološkega sistema na podlagi razpoložljivih osnovnih bioloških gradnikov zmožnih procesiranja podatkov. Na koncu podamo rezultate modeliranja zgrajenega sistema.

### 7.2.1 Sinhronska *RS* pomnilna celica

*RS* pomnilna celica, ki je bila opisana v razdelku 3.3.5 je t. i. asinhronska pomnilna celica, saj se preklopi med logičnimi nivoji prožijo zgolj v odvisnosti od prisotnosti vhodov *R* in *S*. Vpeljemo lahko t. i. sinhronizacijski signal (*urin signal*), ki periodično spreminja svoje stanje iz nizkega logičnega stanja v visoko in obratno. Z vpeljavo urinega signala lahko sinhroniziramo vhoda *R* in *S*, kar pomeni, da se prehodi prožijo v odvisnosti tako od vhodov *R* in *S* kot tudi od prisotnosti urinega signala. Preklop notranjega

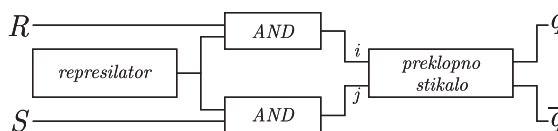
stanja  $q$  pomnilne celice se torej lahko izvrši samo ob točno določenih trenutkih. Na ta način lahko dosežemo, da se vsi preklopi v obravnavanem vezju začnejo istočasno - vezje sinhroniziramo z urinim signalom [32]. Logično shemo in pravilnosto tabelo sinhronne  $RS$  pomnilne celice prikazuje slika 7.1.



**Slika 7.1** Logična shema (a) in pravilnostna tabela (b) sinhronne  $RS$  pomnilne celice. Pri tem  $R$  in  $S$  predstavljata zunanja vhoda v pomnilno celico,  $clk$  urin signal,  $q$  in  $\bar{q}$  trenutno stanje in negirano trenutno stanje pomnilne celice,  $D^1q$  in  $D^1\bar{q}$  pa stanje in negirano stanje pomnilne celice v naslednjem časovnem koraku.

## 7.2.2 Model sinhronskega preklopnega stikala

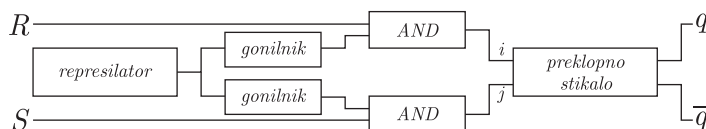
Ker je delovanje preklopnega stikala kot biološkega sistema sorodno delovanju  $RS$  pomnilne celice kot elektronskega elementa, ga lahko podobno kot  $RS$  celico z vpeljavo sinhronizacijskega signala razširimo v *sinhronsko preklopno stikalo*. Kot generator urinega signala uporabimo represilator. Vhoda  $i$  in  $j$  v preklopno stikalo preko bioloških AND vrat tako povežemo z zunanjim proteinom in izhodom represilatorja. Pri tem upoštevamo osnovni model preklopnega stikala, ki ga povežemo z izhodi drugih transkripcijskih regulatornih omrežji. Pri modelu opisanem v razdelku 4.4.7 to ni mogoče, saj predpostavljamo, da prelope proži zunanji vpliv v obliki svetlobe. Shemo opisanega vezja prikazuje slika 7.2.



**Slika 7.2** Shema biološkega sinhronskega preklopnega stikala, pri čemer  $R$  predstavlja reset vhod,  $S$  set vhod,  $q$  in  $\bar{q}$  pa trenutno stanje in negirano trenutno stanje preklopnega stikala.

Z analizo zasnove biološkega gradnika lahko takoj ugotovimo, da zahteva o ohranjanju logičnih nivojev ni izpolnjena, saj je poraba vhoda bioloških AND vrat enaka

$C_I(AND) = \infty$ . Ker represilator izhod, ki vodi v biološka AND vrata, uporablja kot povraten vhod, poleg tega pa je uporabljen pri dveh bioloških AND vratih, delovanje vezja pri danih parametrih zagotovo ne bo pravilno. Problem lahko rešimo tako, da med izhod iz represilatorja in AND vrata vežemo biološki gonilnik, ki zmanjša porabo izhoda na  $C_I(driver) = 2$ . Vrednost je v primerjavi z izhodnimi logičnimi nivoji represilatorja določenimi v razdelku 6.3.7 zanemarljiva. Poleg tega je poraba vhoda s strani bioloških AND vrat dopustna, saj predstavlja edini gradnik, ki uporablja gonilnikov izhod. Shema izboljšane vezja prikazuje slika 7.3



**Slika 7.3** Shema razširjenega biološkega sinhronskega preklopnega stikala, pri čemer  $R$  predstavlja *reset* vhod,  $S$  set vhod,  $q$  in  $\bar{q}$  pa trenutno stanje in negirano trenutno stanje preklopnega stikala.

V tem primeru problem predstavlja kompatibilnost gradnikov v smislu časovnih karakteristik. Ker izhod iz represilatorja posredno ali neposredno vežemo na vhode vseh ostalih gradnikov v vezju, bi moral veljati sledeč pogoj

$$t_v(\text{represilator}) \geq \min(t_s(\text{gonilnik}), t_s(AND), t_s(\text{stikalo})), \quad (7.5)$$

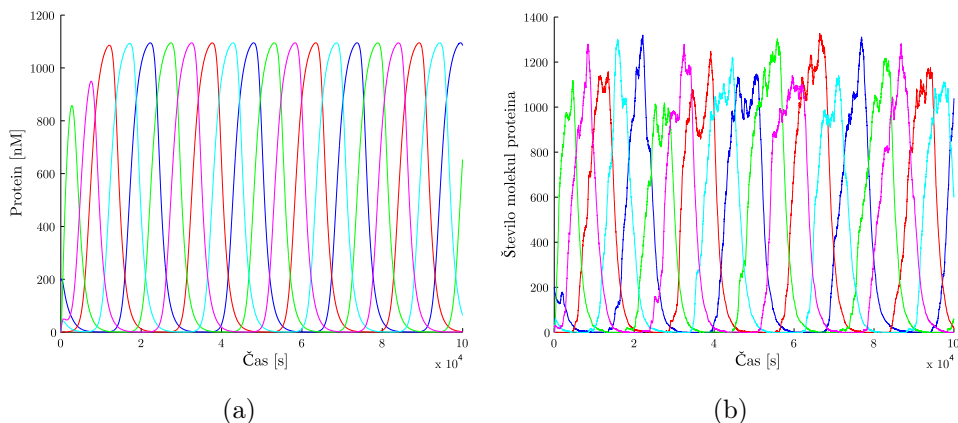
pri čemer je  $t_v(\text{represilator})$  veljavnost izhoda represilatorja (2 500 s),  $t_s(\text{gonilnik})$  čas preklopa biološkega gonilnika (5 000 s),  $t_s(AND)$  čas preklopa bioloških AND vrat (5 000 s) in  $t_s(\text{stikalo})$  čas preklopa preklopnega stikala (10 000 s). Glede na vrednosti določene v poglavju 6, represilator z ostalimi gradniki ni kompatibilen v smislu časovnih karakteristik. Za zagotovitev medsebojne povezljivosti je torej potrebno izboljšati preklone karakteristike represilatorja natančneje veljavnost izhoda represilatorja.

Periodo oscilacij represilatorja in s tem čas veljavnosti njegovega izhoda lahko izboljšamo s povečanjem števila proteinov, ki so krožno povezani v represorskem gensko regulatornem omrežju. Da dosežemo zakasnjeno represijo lastnega izražanja, mora biti število liho [63]. Izkaže se, da pri represilatorju s petimi represorji dosežemo ustrezno veljavnost izhodnega signala, ki je enaka izrazu

$$t_v(\text{represilator}_e) = 10\,000\text{ s}. \quad (7.6)$$

Poleg tega se poveča tudi robustnost sistema (izboljšajo se logični nivoji, zmanjšata se

šum in širina prepovedanega področja). Potek koncentracij proteinov v determinističnem in stohastičnem modelu represilatorja s petimi proteini prikazuje slika 7.4.



**Slika 7.4** Časovni potek števila molekul proteinov v determinističnem (a) in stohastičnem modelu (b) razširjenega represilatorja, pri čemer vsaka barva prikazuje časovni potek koncentracij drugega proteina.

### 7.2.3 Analiza rezultatov modela sinhronskega preklopnega stikala

Pred izgradnjo modela sinhronskega preklopnega stikala preverimo še ostale zahteve za medsebojno povezljivost osnovnih gradnikov uporabljenih v vezju sinhronskega preklopnega stikala (glej sliko 7.3). Pri tem upoštevamo vrednosti parametrov predstavljene v tabeli 7.1.

gradnik	$C_{OL(max)}$	$C_{OH(min)}$	$C_{IL(max)}$	$C_{IH(min)}$	$t_s$	$C_I$	$t_v$
represilator	15	400					$10^4s$
gonilnik	25	700	15	150	5 000 s	2	$\infty$
AND vrata	5	615	120	600	5 000 s	$\infty$	$\infty$
prekl. stikalo	5	715	50	300	$10^4s$		$\infty$

**Tabela 7.1** Vrednosti parametrov uporabljenih gradnikov povzete po poglavju 6.

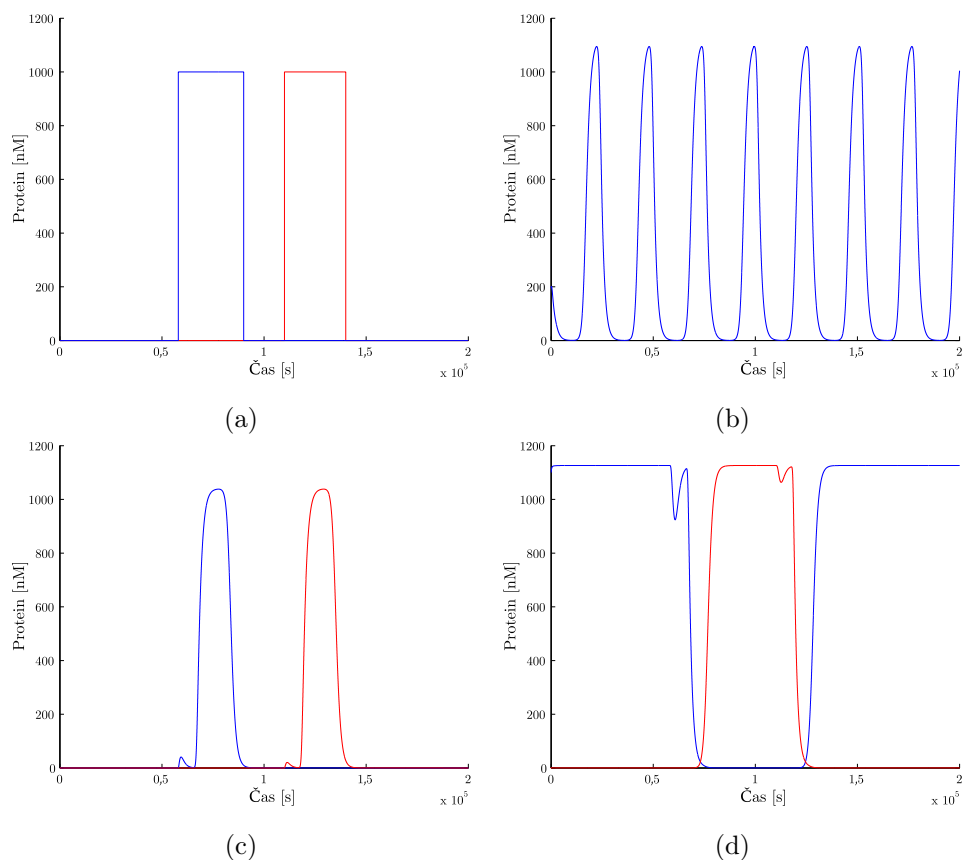
Ugotovimo lahko, da so gradniki glede na zahteve podane v razdelku 7.1 in vrednosti parametrov podane v tabeli 7.1 povezljivi v načrtovano vezje. Torej je postavitve modela sinhronskega preklopnega stikala smotrna.

Model sinhronskega preklopnega stikala je bil postavljen na podlagi povezovanja modelov osnovnih gradnikov v kompleksnejšo celoto, ki je predstavljena na sliki 7.3. Pričako-

vano delovanje sinhronskega preklopnega stikala je zelo podobno običajnemu preklopnemu stikalu. Razlika je v tem, da sta pogoja za preklon iz enega stanja v drugo dva, in sicer *prisotnost zunanjega vhoda*, tj. represorja  $i$  ali  $j$ , in prisotnost *urinega oziroma sinhronizacijskega signala*. Pri preverjanju pravilnosti delovanja smo aktivirali določen zunanji vhod pred aktivnostjo sinhronizacijskega signala. Potek zunanjih vhodov prikazuje slika 7.5(a) in slika 7.6(a), potek sinhronizacijskega signala pa slika 7.5(b) in slika 7.6(b). Zunanji vhod sam po sebi ne sme sprožiti preklopa. Prav tako ne sme preklopa sam po sebi sprožiti sinhronizacijski signal. Preklope iz enega stanja preklopnega stikala v drugo torej prožimo preko operacije AND med prisotnostjo zunanjega vhoda in prisotnostjo sinhronizacijskega signala (glej sliki 7.5(c) in 7.6(c)). Preklon se izvrši šele, ko sta prisotna oba signala. Rezultati determinističnega modeliranja so prikazani na sliki 7.5, stohastičnega pa na sliki 7.6.

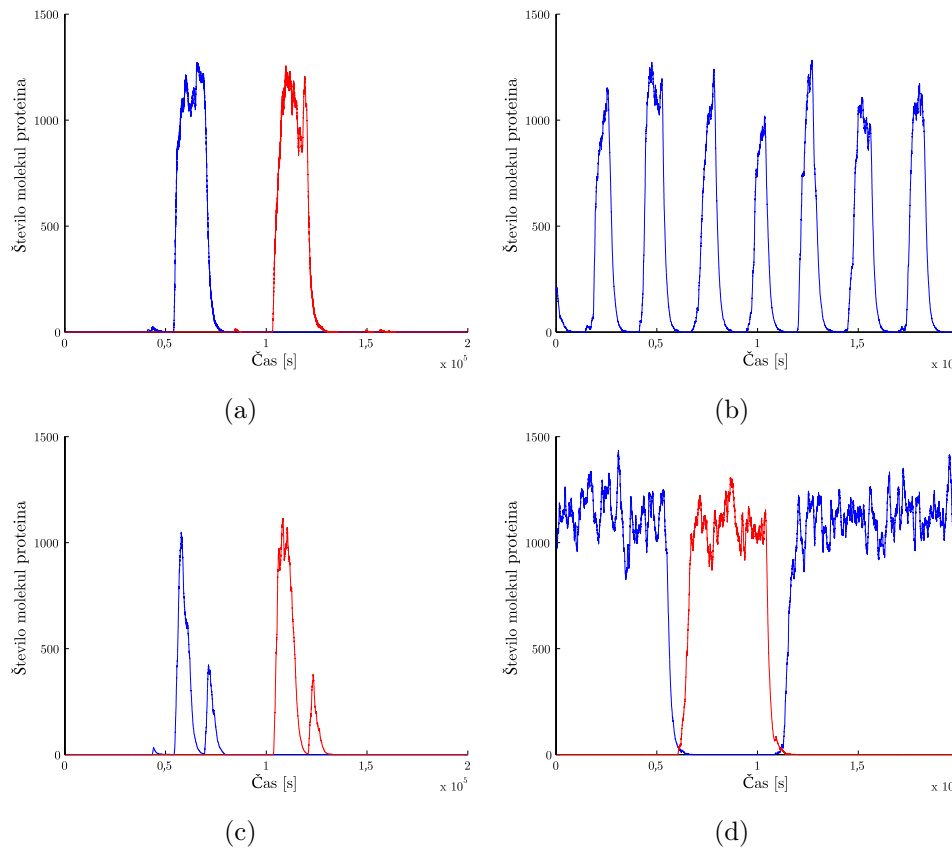
### 7.3 Zaključek poglavja

Biološko računalništvo je trenutno v fazi razvoja, v kateri je možna realizacija zgolj preprostejših aplikacij. Sistem opisan v pričujočem poglavju naprimer predstavlja nekakšno trenutno zgornjo mejo kompleksnosti gradnika zmožnega procesiranja podatkov, ki ga še lahko umetno realiziramo znotraj nekega organizma. Digitalno vezje z enakimi funkcionalnostmi, tj. RS sinhronska pomnilna celica, po drugi strani predstavlja enega od osnovnih gradnikov na področju elektronskih digitalnih vezij. Biološki sistemi zmožni procesiranja večje količine podatkov tako še niso bili razviti, kljub temu da razpolagamo z osnovnimi gradniki, ki predstavljajo funkcijsko poln nabor in s tem teoretično omogočajo izgradnjo poljubnega sistema. Glavni problem, ki ga vidimo, je na tej točki problem razširljivosti. V že prej omenjenem sinhronskem preklopnem stikalu je brez sinhronizacijskega proteina prisotnih 8 proteinov. Če bi želeli količino pomnjene informacije povečati, bi ugotovili, da nabor proteinov dostopnih v naravi, ki so med seboj ortogonalni, zelo hitro izčrpamo. Poleg tega morajo biti uporabljeni proteini ortogonalni z gostiteljem, znotraj katerega poteka realizacija našega sistema, saj sicer pride do nenačrtovane interference med gostiteljem in vnesenim sistemom. S sintezo umetnih transkripcijskih faktorjev, s katerimi lahko rešimo problem ortogonalnosti, se v tem trenutku v svoji doktorski nalogi ukvarja Rok Gaber pod mentorstvom prof. dr. Romana Jerale z ljubljanskega Kemijskega inštituta [64].



**Slika 7.5** Časovni potek števila molekul proteinov v determinističnem modelu sinhronskega preklopnega stikala, pri čemer slika (a) prikazuje potek koncentracij vhodnih signalov  $R$  (modra) in  $S$  (rdeča), slika (b) potek izhodnega signala iz represilatorja, slika (c) vhodnih signalov v preklopno stikalo  $i$  (modra) in  $j$  (rdeča), slika (d) pa potek izhodnih signalov iz preklopnega stikala  $q$  (modra) in  $\bar{q}$  (rdeča).

Če obravnavane biološke gradnike sposobne procesiranja podatkov primerjamo z elektronskimi gradniki uporabljenimi v sodobnih računalnikih, lahko ugotovimo, da slednji bistveno vodijo v številnih karakteristikah, kot so naprimer že omenjena *skalabilnost*, *frekvenca* in *zanesljivost delovanja*. V mislih pa moramo imeti tudi to, da se je razvoj elektronskega računalništva začel že v prvi polovici dvajsetega stoletja, biološko računalništvo pa je relativno mlado področje, ki je še v povojih. Digitalna vezja bodo kmalu dosegla mejo tako v smislu velikosti kot tudi hitrosti, razvoj umetno ustvarjenih bioloških sistemov pa se šele dobro začinja. Kljub temu biološki sistemi zmožni procesiranja podatkov že zdaj prednjačijo na določenih področjih, kot je recimo njihova implicitna masovna paralelnost, majhna poraba energije, zmožnost samopodvojevanja in zmožnost



**Slika 7.6** Časovni potek števila molekul proteinov v stohastičnem modelu sinhronskega preklopnega stikala, pri čemer slika (a) prikazuje potek koncentracij vhodnih signalov  $R$  (modra) in  $S$  (rdeča), slika (b) potek izhodnega signala iz represilatorja, slika (c) vhodnih signalov v preklopno stikalo  $i$  (modra) in  $j$  (rdeča), slika (d) pa potek izhodnih signalov iz preklopnega stikala  $q$  (modra) in  $\bar{q}$  (rdeča).

združevanja umetno sintetiziranih z naravnimi biološkimi sistemi [65]. Na področju biološkega računalništva se moramo torej zavedati tako prednosti kot tudi slabosti bioloških sistemov in to skrbno izkoristiti pri načrtovanju potencialnih aplikacij za uporabo le-teh.





## 8 Zaključek

Med predstavljenimi pristopi matematičnega modeliranja bioloških sistemov smo izbrali najustreznejše in z njimi vzpostavili modele osnovnih bioloških gradnikov zmožnih procesiranja podatkov, ki v logičnem smislu predstavljajo funkcijsko poln nabor logičnih funkcij. Dinamiko teh gradnikov smo ovrednotili na podlagi vzpostavljenih metrik po zgledu elektronskih vezij in po zgledu nelinearnih dinamičnih sistemov. Odprli smo različne možnosti za uporabo vzpostavljenih metrik, in sicer za identifikacijo gradnikov z najustreznejšimi dinamičnimi lastnostmi, za namene avtomatskega načrtovanja bioloških sistemov z zmožnostjo procesiranja podatkov in za analizo kompatibilnosti med osnovnimi gradniki pri gradnji kompleksnejših bioloških sistemov zmožnih procesiranja podatkov. Slednje smo demonstrirali na zgledu modela gensko regulatornega omrežja sinhronske *RS* pomnilne celice. S tem smo potrdili vse zastavljene hipoteze (glej poglavje 1.1). Rezultati, do katerih smo prišli tekom izdelave tega dela, so bili objavljeni v mednarodnih revijah in na mednarodnih konferencah [66–71].

V delu smo se osredotočili na izgradnjo bioloških sistemov zmožnih procesiranja podatkov na podlagi gensko regulatornih omrežij, saj so bila na področju biološkega raču-

nalništva do zdaj daleč najbolj raziskana in je tako v povezavi z njimi na voljo največ podatkov in literature. Z vpeljanimi metrikami bi bilo procesne zmožnosti smiselno analizirati tudi na alternativnih bioloških sistemih, s katerimi bi bilo ravno tako možno realizirati gradnike za procesiranje podatkov. Primer take *platforme* so biosignalne poti (angl. *signal transduction networks*), ki se v primerjavi z gensko regulatornimi omrežji na spremembe v okolju odzivajo veliko hitreje [16], poleg tega pa so določeni gradniki zmožni procesiranja podatkov z njimi že bili realizirani [72]. V nadaljevanju bi bilo zanimivo preučiti možnosti povezovanja analiziranih bioloških sistemov zmožnih procesiranja podatkov v skupno platformo [73].

Obstoječe analize so bile narejene zgolj na podlagi t. i. matematičnih modelov (angl. *mathematical models*) bioloških sistemov. V prihodnosti se nameravamo lotiti alternativnega načina modeliranja, t. i. *računskega modeliranja* (angl. *computational modeling*) [74, 75]. Modeli v tem primeru niso postavljeni na podlagi matematičnih enačb, ki aproksimirajo biološke procese, ampak na podlagi računalniških algoritmov, katerih delovanje posnemajo delovanje bioloških procesov [76]. V to področje spada naprimer modeliranje z Booleovimi mrežami (angl. *Boolean networks*) [77], Petrijevim mrežami (angl. *Petri nets*) [78] in končnimi avtomati (angl. *interacting state machine models*). Ti modeli so za razliko od matematičnih veliko bolj specifični, po drugi strani pa v določenih primerih veliko bolj učinkoviti tako v smislu natančnosti kot tudi hitrosti. Poleg tega za tovrstno modeliranje bioloških procesov določena orodja že obstajajo [79, 80].

Glede na to, da je uporabnost metrik strogo povezana z natančnostjo uporabljenih modelov, bo delo v prihodnosti potekalo na izboljševanju natančnosti le-teh. Natančnost modelov je v veliki meri strogo odvisna od vrednosti parametrov, ki opisujejo njihovo dinamiko. V določeni meri lahko te parametre izmerimo na realnih sistemih, sicer pa moramo njihove vrednosti oceniti na podlagi eksperimentalnih rezultatov. Ocenjevanje vrednosti parametrov je trenutno zelo aktualno podpodročje sintezne biologije [61, 81].

Že uveljavljene pristope k modeliranju nameravamo obogatiti tudi z uporabo mehke logike. Le-ta bi z mehkim obravnavanjem pogojev za proženje in samih rezultatov proženja kemijskih reakcij po naših predvidevanjih omogočala postavitev relativno natančnih modelov kljub določeni stopnji nepoznavanja natančnih vrednosti parametrov. S tem lahko do neke mere obidemo dejstvo nepoznavanja parametrov, ki je sicer trenutno še neizogibno.

Nenazadnje je naše delo do sedaj temeljilo zgolj na modeliranju, pri čemer se je potrebno zavedati, da lahko preko rezultatov modelov velikokrat pridemo do izkrivljene slike realnosti. Pravilnost modelov in z njimi verifikacijo vzpostavljenih sistemov in ovrednotenih metrik želimo potrditi s samo realizacijo analiziranih sistemov v živih celicah in primerjavo njihovih rezultatov z rezultati dobljenimi z uporabo modelov.



## LITERATURA

- [1] H. S. Booth, C. J. Burden, M. Hegland, L. Santoso, *Mathematical Modeling of Biological Systems, Volume I*, A Stochastic Model of Gene Regulation Using the Chemical Master Equation, str. 71–81. Birkhauser Boston, 2007.
- [2] L. M. Adleman, “Molecular computation of solutions to combinatorial problems,” *Science*, št. 266, str. 1021–1024, 1994.
- [3] L. M. Adleman, “Computing with DNA,” *Scientific American*, št. 279, str. 54–62, 1998.
- [4] S. A. Benner, “Act natural,” *Nature*, št. 421, str. 118, 2003.
- [5] S. A. Benner, A. M. Sismour, “Synthetic biology,” *Nature Reviews*, št. 6, str. 533–543, 2005.
- [6] W. A. Yeh, Brian J.; Lim, “Synthetic biology: lessons from the history of synthetic organic chemistry,” *Nature Chemical Biology*, št. 3, str. 521–525, 2007.
- [7] R. Weiss, G. Homsy, T. Knight, “Toward in vivo digital circuits,” v zborniku *Proceedings of the Dimacs Workshop on Evolution as Computation*, 1999.
- [8] R. Weiss, “Cellular computation and communications using engineered genetic regulatory networks,” doktorska disertacija, Massachusetts Institute of Technology, 2001.
- [9] R. Weiss, S. Basu, “The device physics of cellular logic gates,” v zborniku *The First Workshop of Non-Silicon Computing*, 2002.
- [10] T. S. Gardner, C. R. Cantor, J. J. Collins, “Construction of a genetic toggle switch in *Escherichia coli*,” *Nature*, št. 403, str. 339–342, 2000.

- [11] G. Fritz, N. E. Buchler, T. Hwa, U. Gerland, “Designing sequential transcription logic: a simple genetic circuit for conditional memory,” *Syst Synth Biol*, št. 1, str. 89–98, 2007.
- [12] M. B. Elowitz, S. Leibler, “A synthetic oscillatory network of transcriptional regulators,” *Nature*, št. 403, str. 335–338, 2000.
- [13] A. E. Friedland, T. K. Lu, X. Wang, D. Shi, G. Church, J. J. Collins, “Synthetic gene networks that count,” *Science*, št. 324, str. 1199–1202, 2009.
- [14] S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, W. Ron, “A synthetic multicellular system for programmed pattern formation,” *Nature*, št. 434, str. 1130–1134, 2005.
- [15] M. Kaern, W. J. Blake, J. Collins, “The engineering of gene regulatory networks,” *Annu. Rev. Biomed. Eng.*, št. 5, str. 179–206, 2003.
- [16] U. Alon, *An Introduction to Systems Biology*. Chapman & Hall, 2007.
- [17] H. El Samad, M. Khammash, L. Petzold, D. Gillespie, “Stochastic modeling of gene regulatory networks,” *International Journal of Robust and Nonlinear Control*, št. 15, str. 691–711, 2005.
- [18] T. F. Knight, G. J. Sussman, “Cellular gate technology,” v zborniku *Unconventional Models of Computation*, 1998.
- [19] R. Weiss, S. Basu, S. Hooshangi, A. Kalmbach, D. Karig, R. Mehreja, I. Netravali, “Genetic circuit building blocks for cellular computation, communications, and signal processing,” *Natural Computing*, št. 2, str. 47–84, 2003.
- [20] R. Unger, J. Moulton, “Towards computing with proteins,” *Proteins: Structure, Function, and Bioinformatics*, št. 63, str. 53–64, 2006.
- [21] N. Ramakrishnan, U. S. Bhalla, J. J. Tyson, “Computing with proteins,” *Computer*, št. 42, str. 47–56, 2009.
- [22] M. Khammash, H. El Samad, “Systems biology: From physiology to gene regulation,” *IEEE Control Systems Magazine*, št. 24, str. 62–76, 2004.

- [23] R. McDaniel, R. Weiss, "Advances in synthetic biology: on the path from prototypes to applications," *Current Opinion in Biotechnology*, št. 16, str. 476–483, 2005.
- [24] D. Greber, M. Fussenegger, "Mammalian synthetic biology: Engineering of sophisticated gene networks," *Journal of Biotechnology*, št. 130, str. 329–345, 2007.
- [25] M. Thattai, A. van Oudenaarden, "Intrinsic noise in gene regulatory networks," *Proceedings of the National Academy of Sciences*, št. 98, str. 8614–8619, 2001.
- [26] P. R. Patnaik, "External, extrinsic and intrinsic noise in cellular systems: analogies and implications for protein synthesis," *Biotechnology and Molecular Biology Review*, št. 1, str. 121–127, 2006.
- [27] M. Kaern, T. C. Elston, W. J. Blake, J. J. Collins, "Stochasticity in gene expression: From theories to phenotypes," *Nature Reviews*, št. 6, str. 451–464, 2005.
- [28] W. K. Smits, O. P. Kuipers, J.-W. Veening, "Phenotypic variation in bacteria: the role of feedback regulation," *Nature Reviews Microbiology*, št. 4, str. 259–271, 2006.
- [29] J. M. G. Vilar, H. Y. Kueh, N. Barkai, S. Leibler, "Mechanisms of noise-resistance in genetic oscillators," *Proceedings of the National Academy of Sciences*, št. 99, str. 5988–5992, 2002.
- [30] J. Zhang, "Numerical methods for the chemical master equation," doktorska disertacija, Faculty of the Virginia Polytechnic Institute and State University, 2009.
- [31] J. McMurry, R. C. Fay, *Chemistry, 4th edition*. Prentice Hall, 2007.
- [32] J. Virant, *Logične osnove odločanja in pomnjenja v računalniških sistemih*. Založba FER, 1996.
- [33] J. F. Wakerly, *Digital Design: Principles and Practices Package, 4th Edition*. Prentice Hall International, Inc., 2005.
- [34] Z. Kohavi, N. K. Jha, *Switching and Finite Automata Theory*. Cambridge University Press, 2009.
- [35] A. Tamsir, J. J. Tabor, C. A. Voigt, "Robust multicellular computing using genetically encoded nor gates and chemical 'wires'," *Nature*, št. 469, str. 212–215, 2010.

- [36] T. Tian, K. Burrage, "Stochastic models for regulatory networks of the genetic toggle switch," *Proceedings of the National Academy of Sciences*, št. 103, str. 8372–8377, 2006.
- [37] D. L. Nelson, M. M. Cox, *Lehninger Principles of Biochemistry, Fourth Edition*. W. H. Freeman, 2004.
- [38] J. N. Weiss, "The hill equation revisited: uses and misuses," *FASEB Journal*, št. 11, str. 835–841, 1997.
- [39] R. Khanin, D. J. Higham, "Chemical master equation and langevin regimes for a gene transcription model," *Theoretical Computer Science*, št. 408, str. 31–40, 2008.
- [40] D. T. Gillespie, *Markov Processes: An Introduction for Physical Scientists*. Academic Press: San Diego, 1992.
- [41] V. Wolf, G. Rushil, M. Mateescu, T. A. Henzinger, "Solving the chemical master equation using sliding windows," *BMC Systems Biology*, št. 4, str. 1–19, 2010.
- [42] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of Computational Physics*, št. 22, str. 403–434, 1976.
- [43] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The Journal of Physical Chemistry*, št. 81, str. 2340–2361, 1977.
- [44] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," *Journal of Chemical Physics*, št. 115, str. 1716–1733, 2001.
- [45] D. T. Gillespie, "The chemical langevin equation," *Journal of Chemical Physics*, št. 113, str. 297–306, 2000.
- [46] A. S. Ribeiro, J. Lloyd-Price, "SGN Sim, a stochastic genetic networks simulator," *Bioinformatics*, št. 23, str. 777–779, 2007.
- [47] M. Ptashne, *A Genetic Switch: Phage Lambda Revisited, 3rd edition*. Cold Spring Harbor Laboratory Press, 2004.
- [48] J. Goutsias, "Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems," *The Journal of Chemical Physics*, št. 122, str. 1–15, 2005.



- [49] D. Burz, D. Beckett, N. Benson, G. Ackers, “Self assembly of bacteriophage  $\lambda$  CI repressor: Effects of single site mutations on monomer-dimer equilibrium,” *Biochemistry*, št. 33, str. 8399–8405, 1994.
- [50] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, J. Hasty, “A fast, robust and tunable synthetic gene oscillator,” *Nature*, št. 456, str. 516–520, 2008.
- [51] P. A. Iglesias, B. P. Ingalls, eds., *Control theory and systems biology*. MIT Press, 2010.
- [52] G. W. Flake, *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation*. The MIT Press, 2000.
- [53] L. Edelstein-Keshet, *Mathematical Models in Biology*. Siam, Philadelphia, 2005.
- [54] S. Strogatz, *Non-linear Dynamics and Chaos: With applications to Physics, Biology, Chemistry and Engineering*. Perseus Books, 2000.
- [55] P. Fall, E. S. Marland, J. M. Wagner, J. J. Tyson, *Computational Cell Biology (Interdisciplinary Applied Mathematics)*. Springer, 2002.
- [56] N. Zimic, M. Mraz, *Temelji zmogljivosti računalniških sistemov*. Založba FE in FRI, 2006.
- [57] R. Jain, *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*. John Wiley & Sons, New York, 1991.
- [58] R. F. Tinder, *Engineering Digital Design, Revised Second Edition*. Academic Press, 2000.
- [59] A. K. Maini, *Digital Electronics: Principles, Devices and Applications*. John Wiley & Sons, 2007.
- [60] A. P. Godse, D. A. Godse, *Digital Logic Circuits*. Technical Publications Pune, India, 2010.
- [61] J. Sun, J. M. Garibaldi, C. Hodgman, “Parameter estimation using metaheuristics in systems biology: A comprehensive review,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, št. 9, str. 185–202, 2012.

- [62] V. de Lorenzo, A. Danchin, "Synthetic biology: discovering new worlds and new words," *EMBO reports*, št. 9, str. 822–827, 2008.
- [63] K.-I. Goh, B. Kahng, K.-H. Cho, "Sustained oscillations in extended genetic oscillatory systems," *Biophysical Journal*, št. 94, str. 4270–4276, 2008.
- [64] B. Šter, R. Gaber, M. Avbelj, R. Jerala, A. Dobnikar, "Design of information processing in cells using artificial gene repressors," *Przegląd Elektrotechniczny (Electrical Review)*, št. 88, str. 105–109, 2012.
- [65] H. Kobayashi, M. Kaern, M. Araki, K. Chung, T. S. Gardner, C. R. Cantor, J. J. Collins, "Programmable cells: Interfacing natural and engineered gene networks," *Proceedings of the National Academy of Sciences*, št. 101, str. 8414–8419, 2004.
- [66] M. Moškon, M. Ciglič, R. Jerala, N. Zimic, M. Mraz, "Model realizacije funkcionalnosti RS-pomnilne celice v biološkem sistemu," *Elektrotehniški vestnik*, št. 76, str. 223–228, 2009.
- [67] M. Moškon, M. Ciglič, N. Zimic, M. Mraz, "Toward in vivo digital synchronous sequential circuits," *WSEAS Trans. Circuits*, št. 8, str. 301–310, 2009.
- [68] M. Moškon, M. Ciglič, N. Zimic, M. Mraz, "Towards the realization of synchronous sequential biological circuits," v zborniku *Recent advances in mathematics and computers in biology and chemistry : proceedings of the 10th WSEAS International Conference on Mathematics and Computers in Biology and Chemistry (MCBC'09)*, 2009.
- [69] M. Moškon, M. Mraz, "Analyzing the information processing capabilities of gene expression based logic," v zborniku *IEEE ICCSIT 2011: 2011 the 4th IEEE International Conference on Computer Science and Information Technology (IEEE ICCSIT 2011)*, 2011.
- [70] M. Moškon, M. Mraz, "Modeling as the essential step in the construction of biological computer structures," *International journal of information and education technology*, št. 1, str. 185–189, 2011.
- [71] M. Moškon, M. Mraz, "Modelling and analysing the information processing capabilities of simple biological systems," *Mathematical Modelling and Analysis*, št. 17, str. 467–484, 2012.

- [72] Z. Han, T. M. Vondriska, L. Yang, W. R. MacLellan, J. N. Weiss, Z. Qu, "Signal transduction network motifs and biological memory," *Journal of Theoretical Biology*, št. 246, str. 755–761, 2007.
- [73] I. Ndiaye, M. Chaves, J. L. Gouze, "Oscillations induced by different timescales in signal transduction modules regulated by slowly evolving protein–protein interactions," *IET Systems Biology*, št. 4, str. 263–268, 2010.
- [74] L. Laursen, "Biological logic," *Nature*, št. 462, str. 408–410, 2009.
- [75] A. Garg, "Implicit methods for modeling gene regulatory networks," doktorska disertacija, Ecole Polytechnique Federale de Lausanne, 2009.
- [76] J. Fisher, T. A. Henzinger, "Executable cell biology," *Nature Biotechnology*, št. 25, str. 1239–1249, 2007.
- [77] R. Albert, "Boolean modeling of genetic regulatory networks," *Lect. Notes Phys.*, št. 650, str. 459–481, 2004.
- [78] L. J. Steggles, R. Banks, O. Shaw, A. Wipat, "Qualitatively modelling and analysing genetic regulatory networks: a petri net approach," *Bioinformatics*, št. 23, str. 336–343, 2007.
- [79] F. Fages, S. Soliman, N. Chabrier-Rivier, "Modelling and querying interaction networks in the biochemical abstract machine biocham," *Journal of Biological Physics and Chemistry*, št. 4, str. 64–73, 2004.
- [80] R. A. Banks, "Qualitatively modelling genetic regulatory networks: Petri net techniques and tools," doktorska disertacija, Newcastle University, 2009.
- [81] G. Lillacci, M. Khammash, "A distribution-matching method for parameter estimation and model selection in computational biology," *Int. J. Robust. Nonlinear Control*, št. 22, str. 1065–1081, 2012.



## PRILOGA

Delu so v digitalni obliki priložene sledeče datoteke z izhodiščnimi modeli:

- ssa.m: implementacija algoritma SSA,
- driver\_det.m: implementacija determinističnega modela biološkega gonilnika,
- driver\_ssa.m: implementacija stohastičnega modela biološkega gonilnika,
- negator\_det.m: implementacija determinističnega modela biološkega negatorja,
- negator\_ssa.m: implementacija stohastičnega modela biološkega negatorja,
- and\_det.m: implementacija determinističnega modela bioloških AND vrat,
- and\_ssa.m: implementacija stohastičnega modela bioloških AND vrat,
- nor\_det.m: implementacija determinističnega modela bioloških NOR vrat,
- nor\_ssa.m: implementacija stohastičnega modela bioloških NOR vrat,
- switch\_det.m: implementacija determinističnega modela preklopnega stikala,
- switch\_ssa.m: implementacija stohastičnega modela preklopnega stikala,
- switch\_ext\_det.m: implementacija determinističnega modela razširjenega preklopnega stikala,
- switch\_ext\_ssa.m: implementacija stohastičnega modela razširjenega preklopnega stikala,
- repressilator\_det.m: implementacija determinističnega modela represilatorja,
- repressilator\_ssa.m: implementacija stohastičnega modela represilatorja.



## IZJAVA

*Izjavljam, da sem doktorsko disertacijo izdelal samostojno pod vodstvom mentorja izr. prof. dr. Mihe Mraza. Izkazano pomoč drugim sodelavcem sem v celoti navedel v zahvali.*

— Miha Moškon, Ljubljana, 2012.