



УНИВЕРЗИТЕТ У КРАГУЈЕВЦУ  
ФАКУЛТЕТ ЗА МАШИНСТВО И  
ГРАЂЕВИНАРСТВО У КРАЉЕВУ

Владимир С. Ђорђевић

**ИНТЕЛИГЕНТНИ РЕГУЛАТОРИ  
ЗАСНОВАНИ НА АДАПТИВИНОМ  
ДИНАМИЧКОМ ПРОГРАМИРАЊУ**

Докторска дисертација

Краљево, 2023



UNIVERSITY OF KRAGUJEVAC  
FACULTY OF MECHANICAL AND CIVIL  
ENGINEERING IN KRALJEVO

VLADIMIR S. ĐORĐEVIĆ

**INTELLIGENT CONTROLLERS BASED  
ON ADAPTIVE DYNAMIC  
PROGRAMMING**

Doctoral Dissertation

Kraljevo, 2023

<b>Аутор</b>
Име и презиме: Владимир Ђорђевић
Датум и место рођења: 18. фебруар 1985. године, Крушевац
Садашње запослење: Асистент, Факултет за машинство и грађевинарство у Краљеву Универзитета у Крагујевцу
<b>Докторска дисертација</b>
Наслов: Интелигентни регулатори засновани на адаптивном динамичком програмирању
Број страница: 111
Број слика: 77
Број библиографских података: 145
Установа и место где је рад израђен: Факултет за машинство и грађевинарство у Краљеву
Научна област (УДК): 681.513.7 (Интеллигентни управљачки системи)
<b>Ментор:</b> др Владимир Стојановић, ванредни професор, Факултет за машинство и грађевинарство у Краљеву Универзитета у Крагујевцу
<b>Докторска дисертација</b>
Датум пријаве теме: 03.12.2021.
Број одлуке и датум прихватања теме докторске дисертације:
Одлука Већа за техничко-технолошке науке Универзитета у Крагујевцу број IV-04-275/8 од 13.04.2022.
Комисија за оцену научне заснованости теме и испуњености услова кандидата:
1. др Драган Пршић, ванредни професор Факултета за машинство и грађевинарство у Краљеву Универзитета у Крагујевцу, ужа научна област: Аутоматско управљање и флуидна техника, председник Комисије
2. др Радиша Јовановић, редовни професор Машинског факултета Универзитета у Београду, ужа научна област: Аутоматско управљање, члан Комисије
3. др Љубиша Дубоњић, доцент Факултета за машинство и грађевинарство у Краљеву Универзитета у Крагујевцу, ужа научна област: Аутоматско управљање и флуидна техника, члан Комисије
Комисија за оцену и одбрану докторске дисертације:
1. др Драган Пршић, редовни професор Факултета за машинство и грађевинарство у Краљеву Универзитета у Крагујевцу, ужа научна област: Аутоматско управљање и флуидна техника, председник Комисије
2. др Радиша Јовановић, редовни професор Машинског факултета Универзитета у Београду, ужа научна област: Аутоматско управљање, члан Комисије
3. др Љубиша Дубоњић, ванредни професор Факултета за машинство и грађевинарство у Краљеву Универзитета у Крагујевцу, ужа научна област: Аутоматско управљање и флуидна техника, члан Комисије
Датум одбране дисертације:

# Захвалности

Ова докторска дисертација представља резултат вишегодишњег образовања и научно-истраживачког рада на Катедри за аутоматско управљање и флуидну механику, Факултета за машинство и грађевинарство у Краљеву.

Велику захвалност дуђујем ментору др Владимиру Стојановићу, ванредном професору Факултета за машинство и грађевинарство у Краљеву Универзитета у Краљевцу, на великој подршци, сујеснијама и саветима током израде ове дисертације.

Осим тога, велику захвалност дуђујем и члановима Комисије за оцену и одбрану докторске дисертације на корисним сујеснијама које су допринеле квалитету ове дисертације.

Захваљујем се Факултету за машинство и грађевинарство у Краљеву на подршци у досадашњем раду.

На крају, захвалио бих се својој породици на подршци и разумевању током целог школовања.

# Резиме

Одређивање тачног модела је веома изазован и дуготрајан посао у већини практичних случајева, што алгоритме управљања који су засновани на моделу чини неизводљивим. Чак и ако се може добити апроксимирани модел система, динамичка неодређеност, проузрокована неслагањем апроксимације и реалног система, ће деградирати перформансе управљања класично пројектованих регулатора. Стога је пројектовање интелигентних регулатора вођених подацима, који су робусни, адаптивни и поуздани у присуству различитих неодређености и непознате динамике, од примарног значаја.

У овој тези се најпре предлажу интелигентни регулатори засновани на АДП-у линеарних система. Код класичних алгоритама управљања заснованих на АДП-у, уобичајено се претпоставља да су променљиве величине стања мерљиве. Међутим, ова претпоставка је генерално рестриктивна, јер у многим инжењерским применама често недостаје потпуна информација о стању, и само је излаз доступан за мерење. У условима непознате динамике система и немерљивих стања, од великог интереса је коришћење техника управљања заснованих на улазно/излазним мерењима, које се још називају и технике са повратном спрегом по излазу. Такође, стабилност динамичких система се постиже путем повратне спреге по излазу, при чему се за реконструкцију непознатих величина стања користе улазна и излазна мерења.

Да би се испунили захтеви перзистентне побуде, потребно је додати неки истраживачки шум управљачком улазу. Применом теорије планирања експеримента, у тези се примењује оптимални истраживачки шум у виду суме синусоида, који ће омогућити да излаз система носи максимум информације о систему. Тиме ће бити скраћено време учења, односно убрзан процес пројектовања интелигентних регулатора.

Имплементацијом стратегије управљања који су покренути догађајима смањује се број укључивања предложених интелигентних регулатора, чиме се штеде енергетски, рачунски и комуникациони ресурси. Предложена стратегија управљања ће унапредити алгоритме управљања толерантне на отказе, чиме су побољшане перформансе система управљања у случају појаве отказа.

Практична применљивост предложених интелигентних регулатора је повећана чињеницом да се нелинеарни системи могу верно представити линеарним моделама са естимацијом динамике у реалном времену. Пројектовани регулатори су способни да управљају линеарним и нелинеарним системима са потпуно непознатим матрицама система, а експоненцијална конвергенција се постиже док је услов перзистентне побуде задовољен. Поред теоријских оквира, практични аспект добијених резултата је верификован на комплексним инжењерским проблемима, као што је управљање хидрауличког серво актуатора и самобалансирајућег робота. Валидност и ефективност предложених интелигентних регулатора заснованих на АДП-у

се евалуира кроз интензивне симулације.

**Кључне речи:** интелигентни регулатори, адаптивно динамичко програмирање, непозната динамика, управљање покренуто догађајима, управљање толерантно на отказе

# Abstract

Determining an accurate model is very challenging and time-consuming in most practical cases, making model-based control algorithms infeasible. Even if an approximate model of the system can be obtained, a dynamic uncertainty caused by the mismatch between the approximation and the real system will degrade the control performance of classically designed controllers. Therefore, the design of data-driven intelligent controllers, which are robust, adaptive and reliable in the presence of various uncertainties and unknown dynamics, is of primary importance.

This PhD thesis first proposes intelligent controllers based on ADP of linear systems. In classical ADP-based control algorithms, it is commonly assumed that the state variables are measurable. However, this assumption is generally restrictive since, in many engineering applications, complete state information is often lacking, and only the output is available for measurement. In conditions of unknown system dynamics and unmeasurable states, using control techniques based on input/output measurements, also called output feedback techniques, is of great interest. Also, the stability of dynamic systems is achieved through output feedback, whereby input and output measurements are used to reconstruct unknown state variables.

To meet the persistent excitation requirements, an exploration noise needs to be added to the control input. By applying the theory of experiment design, this PhD thesis applies optimal exploration noise in the form of a sum of sinusoids, which will allow the system output to carry the maximum amount of information about the system. This will shorten the learning time, i.e. speed up the process of designing intelligent controllers. Implementing an event-triggered control strategy reduces the number of switching of the proposed intelligent controllers, thus saving energy computational and communication resources. The proposed control strategy will improve the fault-tolerant control algorithms, thus improving the performance of the control system in case of faults.

The practical application of the proposed intelligent controllers is increased by the fact that linear models with real-time dynamics estimation can faithfully represent nonlinear systems. The designed controllers are capable of controlling linear and nonlinear systems with completely unknown system matrices, and exponential convergence is achieved while the persistent excitation condition is satisfied. In addition to the theoretical framework, the practical aspect of the obtained results was verified on complex engineering problems, such as the control of a hydraulic servo actuator and a self-balancing robot. The validity and effectiveness of the proposed intelligent controllers based on ADP are evaluated through intensive simulations.

**Keywords:** intelligent controllers, adaptive dynamic programming, unknown dynamics, event-triggered control, fault-tolerant control

# Садржај

<b>1</b>	<b>Увод</b>	<b>1</b>
1.1	Преглед литературе и досадашњих истраживања . . . . .	1
1.1.1	Предмет, циљеви и хипотезе . . . . .	3
<b>2</b>	<b>Регулатори засновани на адаптивном динамичком програмирању</b>	<b>6</b>
2.1	Динамичко програмирање . . . . .	9
2.2	Итеративно решавање проблема динамичког програмирања . . . . .	12
2.3	Управљање засновано на АДП-у и повратној спрези по стањима . . . . .	13
2.4	Реконструкција стања . . . . .	18
2.5	Управљање засновано на АДП-у и повратној спрези по излазу . . . . .	19
2.6	Избор истраживачког шума . . . . .	24
2.6.1	Перзистентна побуда . . . . .	30
2.6.2	Оптимални истраживачки шум . . . . .	31
2.6.3	Минималан број синусоида . . . . .	32
2.7	Пример . . . . .	35
<b>3</b>	<b>Управљање засновано на догађајима</b>	<b>43</b>
3.1	Линеарни дискретни систем . . . . .	44
3.2	Реконструкција стања . . . . .	45
3.3	Алгоритам управљања заснован на АДП-у покренут догађајима . . . . .	46
3.4	Пример . . . . .	50
<b>4</b>	<b>Системи управљања са толерисаним отказима</b>	<b>55</b>
4.1	Линеарни систем са отказом актуатора . . . . .	57
4.2	Реконструкција стања . . . . .	59
4.3	Алгоритам УТО заснован на АДП-у . . . . .	60
4.4	Пример . . . . .	63
<b>5</b>	<b>Практични примери</b>	<b>68</b>
5.1	Хидраулички серво актуатор . . . . .	68
5.1.1	Управљање засновано на АДП-у . . . . .	73
5.1.2	Управљање засновано на АДП-у покренуто догађајима . . . . .	77
5.1.3	Управљање засновано на АДП-у толерантно на отказе . . . . .	79
5.2	Самобалансирајући робот . . . . .	82
5.2.1	Управљање засновано на АДП-у . . . . .	85
5.2.2	Управљање засновано на АДП-у покренуто догађајима . . . . .	91
5.2.3	Управљање засновано на АДП-у толерантно на отказе . . . . .	93
<b>6</b>	<b>Закључак и даља истраживања</b>	<b>97</b>



# Глава 1

## Увод

### 1.1 Преглед литературе и досадашњих истраживања

Проблем оптималног управљања у системима са непознатом динамиком је један од водећих изазова научне заједнице која се бави аутоматским управљањем. У савременој теорији управљања, оптимално управљање игра веома битну улогу у пројектовању регулатора. Наиме, главни изазов је пројектовање оптималних алгоритама управљања који ће обезбедити минималну потрошњу енергије [1, 2]. Међутим, оптимално управљање представља технику управљања, која обично зависи од савреног познавања модела система, који није могуће добити у већини практичних случајева. Чак и ако се може добити апроксимирани модел система, динамичка неодређеност, проузрокована неслагањем апроксимације и реалног система, ће деградирати перформансе управљања традиционално пројектованог оптималног регулатора [3, 4]. Стога су даља истраживања пројектовања оптималних регулатора веома важна и примарни циљ ове дисертације.

Адаптивно динамичко програмирање (АДП) представља ефикасан начин за постизање високих перформанси оптималног регулатора који се заснива на адаптивном управљању, оптималном управљању и учењу са поткрепљивањем (енг. reinforcement learning) [5–8]. Последњих година је поље примене АДП-а веома распрострањено у различитим областима истраживања укључујући роботске системе, ваздухопловне системе, навођене ракете, свемирске летелице, итд [9–11].

Класични алгоритми управљања засновани на АДП-у имају повратну спрегу по стању. Међутим, директно мерење вектора стања није изводљиво у практичним имплементацијама и осим тога захтевало би веома скупу мерну опрему. Зато је погодније користити алгоритме управљања који примењују методе засноване на реконструкцији стања, а не на њиховом директном мерењу. АДП представља технику управљања засновану на мерењима (енг. data-driven) која може гарантовати стабилност затвореног система управљања. Наиме, у условима непознате динамике система и немерљивих стања, од великог интереса је коришћење техника управљања заснованих на измереним улазно/излазним сигнаlima, које се још називају и технике са повратном спрегом по излазу (енг. output feedback). Главна корист ових техника управљања је да за њихову примену није потребно познавање динамике система. За непознат модел система, ова индиректна техника генерише низ субоптималних регулатора који конвергирају ка оптималном управљању са повећањем броја итерација.

Надаље, да би се обезбедила конзистентност апроксимација и добила јединственост решења у сваком кораку итерације, потребно је додати неки истраживачки шум (енг. exploration noise) управљачком улазу, да би се испунили захтеви перзистентне побуде. За истраживачки шум обично се примењује нека перзистентна побуа, као што је бели шум, поворка псеудо случајних бинарних сигнала (енг. PRBS) или сума синусоида [12–14]. Избор истраживачког шума је у општем смислу нетривијалан задатак за већину алгоритама управљања заснованих на учењу [15].

Спровођење алгоритама АДП-а обично се заснива на периодичном одабирању [16]. У циљу чувања ограничених комуникационих и рачунских ресурса, одавно су почеле да се примењују стратегије покренуте догађајима (енг. event-triggered) у алгоритмима управљања који су засновани на АДП-у [17–20]. Штавише, број ажурирања управљачких улаза на овај начин ће бити смањен у односу на периодично ажурирање регулатора, јер се ажурира само када је то неопходно (нпр. када се погоршају перформансе система). Спровођење алгоритама покренутих догађајима се заснива на аperiodичном одабирању. У литератури је предложено неколико регулатора заснованих на догађајима, од којих су већина регулатори са повратном спрегом по стању [21–25]. За разлику од тога, у тези је разматран проблем управљања заснованог на АДП-у покренутог догађајима у случају када је доступна само повратна информација о излазу.

Због све већих захтева за безбедношћу и стабилношћу у савременом индустријском применама, проблем детекције отказа динамичких система је веома актуелан последњих година [26–30]. У постојећој литератури детекција отказа и дијагностика система засноване на моделу (енг. model-based) су веома заступљене [31–34]. Са друге стране, могућности интелигентних регулатора заснованих на АДП-у, због предности које пружају методе засноване на учењу у реалном времену (енг. online learning methods), представљају добру основу за унапређење детекције и дијагностике отказа у системима аутоматског управљања, односно побољшање перформанси система управљања толерантних на отказе (енг. fault tolerant control).

Практична применљивост алгоритама управљања заснована на АДП-у је побољшана чињеницом да се нелинеарни системи могу веома прецизно представити линеарним моделима са естимацијом динамике у реалном времену [35, 36]. Притом, многе савремене инжењерске апликације као што су аутономна возила [37, 38], микромреже [39], високо осетљиви микрофони [40], предвиђање напрезања код замора материјала [41], одржавање безбедности сајбер-физичких система [42], управљање робота [43], користе пројектовање регулатора у реалном времену, заснованих на линеарним моделима. Пројектовани регулатори су способни да управљају линеарним и нелинеарним системима са потпуно непознатим матрицама система, а експоненцијална конвергенција се постиже док је услов перзистентне побуде задовољен. Поред теоријских оквира, практични аспект добијених резултата илустрован је интензивним симулацијама. Такође, конкретна примена предложене методологије приказана је за проблем управљања хидрауличког серво актуатора (ХСА) и самобалансирајућег робота.

Добро је познато да је ХСА веома сложен нелинеарни систем, и да се параметри система не могу тачно одредити због различитих неодређености, немогућности мерења неких параметара и поремећаја. У овој дисертацији се разматра проблем управљања ХСА-ом са непознатом динамиком. У последње две деценије, пројектовање регулатора високих перформанси ХСА привлачи све већу пажњу због интензивног повећања перформанси техничких система у индустрији [44–47]. Велики

број хидрауличких машина, често ради са великим оптерећењима у тешким условима рада. Као резултат утицаја спољашње температуре, прашине, влажности, хабања, различитих оптерећења и поремећаја, ХСА је обично подвргнут великим неодређеностима током рада. Дакле, управљање ХСА-ом са високом тачношћу увек је представљало велики изазов за истраживаче због немоделоване динамике, великих нелинеарности, параметарске неодређености, немерљивих стања у реалним условима рада, итд. Познат је проблем немогућности одређивања већине физичких параметара компоненти ХСА. Док су неки параметри доступни само са одређеном тачношћу, остали параметри су потпуно непознати. Ове чињенице представљају непремостиве проблеме у постизању управљања ХСА-ом високе тачности, које се не може постићи без познавања тачног ХСА модела [48–51].

Са друге стране, самобалансирајући робот је типичан робот који има потенцијалне изгледе за примену у многим областима, као што су транспорт и истраживање. Управљање самобалансирајућег робота привлачи велику пажњу, како у академским круговима, тако и у индустрији. Наиме, самобалансирајући робот је инхерентно нестабилан, високог реда, мултиваријабилан, нелинеаран и јако спрегнут систем, и притом представља недовољно активирани (енг. *underactuated system*) механички систем. За такав недовољно активирани систем, који има мање управљачких улаза од генералисаних координата, неопходно је индиректно управљати неактивираним генералисаним координатама, преко динамичке спреге. Недовољно активирање, иако резултира мањим бројем актуатора и на тај начин помаже у смањењу трошкова производње и стопе кварова, представља изазове за пројектовање регулатора. Штавише, за разлику од једноставнијих система, као што је инверзно клатно на колицима, који су ограничени на вођену путању, самобалансирајући робот се креће сопственом путањом док балансира клатно. Једна од потешкоћа при управљању самобалансирајућег робота је да се истовремено управља његовим линеарним кретањем, нагибним кретањем, као и скретањем.

Стога је развој интелигентних регулатора заснованих на АДП-у један од примарних задатака ове тезе, у циљу остваривања оптималног управљања у одсуству тачног познавања модела система. Притом се постиже и висока тачност система управљања.

### 1.1.1 Предмет, циљеви и хипотезе

У тези се разматра управљање засновано на учењу у реалном времену (енг. *online learning*), где током рада система, из измерених улаза и излаза, регулатор учи да компензује непознату динамику система, различите поремећаје и грешке моделовања обезбеђујући жељене перформансе система управљања. Оптимални закон управљања се итеративно остварује на основу повратне спреге по излазу, реконструкције стања и АДП-а. Непознати модел система се прво идентификује, након чега се алгебарска Рикатијева једначина итеративно решава. Да би се обезбедила конзистентност апроксимација и добила јединствена решења у сваком кораку итерације, улазу се додаје истраживачки шум у виду суме синусоида. Тако добијени улазни и излазни сигнали се користе за реконструкцију вектора стања модела, што има велики практични значај у односу на технике управљања са директним мерењем стања, које захтевају употребу великог броја сензора. Коришћењем повратне спреге по излазу и методе реконструкције стања, примењена техника управљања заснована на АДП-у представља користан алат за дигиталну имплементацију у ре-

алним системима. Закон управљања се итеративно учи и у реалном времену веома ефикасно обезбеђује решења проблема оптималног управљања заснованог само на мерењима.

Применом управљачке стратегије засноване на догађајима, број ажурирања управљачких улаза ће бити смањен у односу на периодично ажурирање регулатора, јер се регулатор ажурира само када се испуне одређени услови (нпр. када се погоршају перформансе система). На тај начин ће се значајно очувати енергетски, рачунски и комуникациони ресурси.

Могућности интелигентних регулатора заснованих на АДП-у, због предности које пружају методе засноване на учењу у реалном времену (енг. online learning methods), представљају добру основу за унапређење детекције и дијагностике отказа у системима аутоматског управљања заснованих на моделу (енг. model-based), односно побољшање перформанси система управљања толерантног на отказе (енг. fault-tolerant control).

Доминантне нелинеарности у ХСА-у, са којима је веома тешко руковати са великом тачношћу, су немогућност тачног одређивања параметара настале као резултат заштите власничких података појединих произвођача, индиректног мерења, губицима притиска, различитим условима струјања, коефицијентима трења и цурења, генерисањем дисконтинуалних управљачких сигнала услед ефеката засићења и промене смера серво вентила, итд. Штавише, променљиви услови рада током извршавања задатака, као што су температура уља, коефицијент стишљивости, флукутирајући напојни притисак и запремина у цеви, довешће до промена параметара модела које погоршавају постојеће перформансе управљања. Такође, неке од традиционалних методе управљања, укључујући ПИД регулаторе, фази регулаторе и клизни режим управљања, у стању су да управљају самобалансирајућим роботом у одсуству тачног познавања параметара система, али не постижу оптимално управљање [52–56]. У практичним применама, осим просте стабилизације, често је пожељно испунити и критеријум оптималности. Традиционални линеарно квадратни регулатор (енг. LQR), заснован на решењу Рикатијеве једначине, постиже циљ оптималног управљања [57, 58]. Међутим, линеарно квадратни регулатор захтева тачно познавање модела система. Основна предност предложене методологије управљања је избегавање познавања динамике система, што је веома важно у практичним условима.

Дакле, ова теза има за циљ да унапреди постојеће алгоритме управљања засноване на АДП-у, развије одговарајуће АДП регулаторе покренуте догађајима и адаптира их за имплементацију у системе управљања толерантних на отказе. Валидност и ефективност предложеног приступа управљања демонстрирана је кроз интензивне симулације.

Главни резултати истраживања, који су проистекли из рада на тези, су верификовани кроз објављене радове [59–64].

Теза је у наставку организована на следећи начин:

- У другом поглављу описани су принципи учења поткрепљивањем и дат је преглед развоја АДП-а, са акцентом на примену у управљању. Управљање засновано на АДП-у даје добре резултате у случају када је динамика система непозната. Када је управљање системима са непознатом динамиком у питању, неопходно је прикупити довољну количину улазних и излазних података потребних за фазу учења. У ту сврху, применом метода планирања експеримента, пројектује се побуди сигнал који максимизира количину информације

из мерења која су на располагању, будући да је прави систем априорно непознат. На основу прикупљених мерења, применом адаптивног динамичког програмирања итеративно се одређује оптимално појачање интелигентног регулатора. Притом се због потешкоћа директног израчунавања алгебарске Рикатијеве једначине користе ефикасни алгоритми нумеричких апроксимација. Конвергенција разматраних алгоритама је доказана у форми теореме.

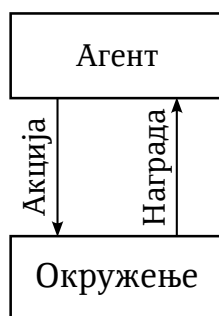
- У трећем поглављу разматрано је управљање системом са непознатом динамиком коришћењем регулатора заснованог на АДП-у покренутог догађајима. Увођењем стратегије покренуте догађајима у затворено коло система значајно се смањује број ажурирања управљања. На овај начин смањује се комуникација између регулатора и система и на тај начин чувају се рачунарски и комуникацијски ресурси.
- У четвртном поглављу разматрано је управљање толерантно на отказе актуатора засновано на АДП-у. Овај задатак је изазован због неизвесних поремећаја, непознате динамике и немерљивих стања система. Циљ је да се постигне асимптотско праћење и компензација отказа актуатора минимизирањем унапред дефинисаног индекса перформансе. Дискретна алгебарска Рикатијева једначина се итеративно решава применом АДП-а. Из практичних разлога, технике АДП-а и компензације отказа су интегрисане да итеративно одређују апроксимирано оптимално управљање толерантно на отказе користећи улазно/излазне податке, без а priori познавања динамике управљаног система и мерених стања. Да би се ово постигло примењено је АДП засновано на повратној спрези по излазу.
- У петом поглављу су перформансе предложених техника управљана показане на примерима хидрауличког серво актуатора и самобалансирајућег робота кроз интензивне симулације у софтверском пакету Матлаб.
- У шестом поглављу су дата закључна разматрања.
- У седмом поглављу дат је списак коришћење литературе.

## Глава 2

# Регулатори засновани на адаптивном динамичком програмирању

Учење поткрепљивањем (енг. reinforcement learning), поред надгледаног и не-надгледаног учења, представља једну од три главне парадигме машинског учења. Развијено је шездесетих година двадесетог века као начин проучавања вештачке интелигенције у заједници која се бави рачунарским наукама, али има доста блиских веза са оптималним и адаптивним управљањем. Пошто методе учења поткрепљивањем представљају добар начин за решавање проблема секвенцијалног одлучивања, последњих година нашле су примену у многим областима.

Код учења поткрепљивањем, агент учи оптимални низ акција, која се назива полиса или политика (енг. policy). Кроз интеракцију са околином, агент покушава да максимизује награду која представља меру за оцену ваљаности решења. На слици 2.1 дата је илустрација учења поткрепљивањем.



Слика 2.1: Учење поткрепљивањем.

Вредносна функција представља предикцију будућих награда на основу које агент одређује акције које ће предузети како би максимизирао награду. Најчешћа су два типа вредносне функције, то јест, вредносна функција заснована на стању  $V(s)$  и вредносна функција заснована на стању и акцији  $Q(s, a)$ . Старије методе учења поткрепљивањем засноване су на табеларном приказу вредносне функције. Наиме, вредносна функција је представљана и процењивана у табеларној форми за свако стање или пар стање-акција. Проблем код табеларних метода јесте спора конвергенција. Савремене методе су засноване на апроксимацији вредносне функције.

Агент практично учи методом покушаја и грешке, користећи две стратегије, истраживање (енг. exploration) и експлоатација (енг. exploitation). Истраживање представља стратегију где агент предузима акције које нису нужно оптималне са циљем да се пронађу нова стања и акције која ће резултирати већом оствареном наградом. Са друге стране, експлоатација подразумева употребу стеченог знања за које се зна да води већој награди на основу тренутне процене вредносне функције. Приликом пројектовања агента потребно је пронаћи одговарајући компромис између ове две стратегије.

Термини окружење, агент и политика се углавном користе у рачунарским наукама, док се у теорији управљања користе термини систем, регулатор и закон управљања, респективно. Такође, у теорији управљања уместо награде углавном се користи функција циља која се минимизира да би се постигле жељене перформансе система.

Постоје три групе алгорита које су засноване на:

- динамичком програмирању,
- Монте Карло методама,
- привременим разликама (енг. temporal difference).

Динамичко програмирање је засновано на раду Ричарда Белмана (енг. Richard Bellman) и његовом принципу оптималности [65]. Динамичко програмирање представља групу метода која тачно одређује оптимални закон управљања у случају кад је модел система у потпуности познат. Међутим, у пракси то често није случај. Такође, још једно ограничење представља и проблем димензионалности (енг. curse of dimensionality), који се јавља код система са великим или континуалним скуповима стања и управљања. Две најпопуларније методе динамичког програмирања су итерација вредности (енг. value iteration, VI) коју је развио Белман [65, 66] и итерација политике (енг. policy iteration, PI) коју је развио Хауард (енг. Howard) [67].

Алгоритмима учења поткрепљивањем заснованим на Монте Карло методама нису потребни модели [68]. Међутим, ове методе је тешко имплементирати онлајн јер се вредносна функција не оцењује у сваком кораку већ се оцењивање врши на крају епизоде. Карактеристика ових метода је споро учење из разлога што је потребно извршити епизоду до краја пре евалуације. Из истог разлога, овај тип алгоритама је немогуће применити за оптимизацију у проблемима са бесконачним временским хоризонтом.

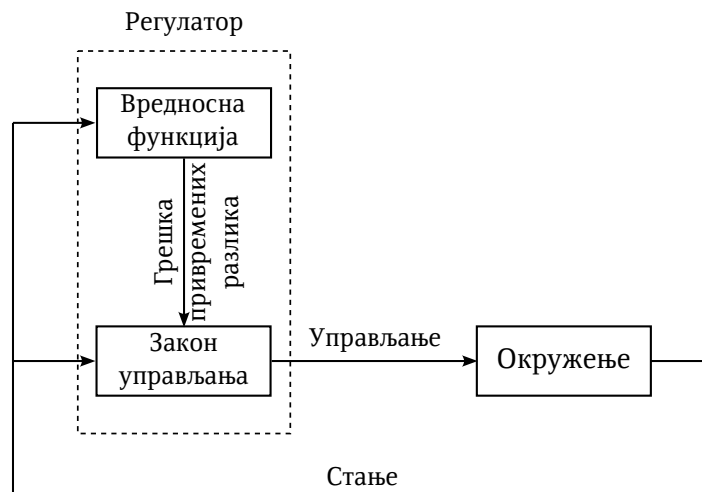
Алгоритми засновани на привременим разликама комбинују принципе динамичког управљања и Монте Карло метода. Код ових алгоритама није потребан модел за разлику од динамичког програмирања. За разлику од алгоритама заснованих на Монте Карло методама ове алгоритме можемо имплементирати онлајн јер се вредносна функција ажурира у сваком кораку.

Значајна класа алгоритама учења поткрепљивањем заснована је на структури актант-критичар. Критичар представља компоненту која оцењује вредност тренутног стања, док актант примењује политику на окружење.

Надгледано учење захтева постојање означеног скупа података, што је за потребе управљања сложених система немогуће обезбедити. Такође, алгоритми засновани на надгледаном учењу само имитирају одлуке на основу означеног скупа и из тог разлога не могу да надмаше експерта (оператера). За разлику од надгледаног

учења, учењу поткрепљивањем није потребан означен скуп већ само критеријум за оцену ваљаности решења, и због тога је привукло доста пажње у теорији управљања [3, 69–72].

Развијене су многе методе учења поткрепљивањем у реалном времену за одређивање онлајн оптималних закона управљања. Ове методе у теорији управљања су познате под називима апроксимативно/адаптивно динамичко програмирање (АДП) [4, 6, 73, 73–80] или неуро-динамичко програмирање [2, 81]. Адаптивно динамичко програмирање представља комбинацију динамичког програмирања, апроксиматора функција и структуре актант-критичар, слика 2.2.



Слика 2.2: Управљање засновано на адаптивном динамичком програмирању.

Развој теорије АДП-а може се поделити у три фазе. У првој фази, АДП је опсежно истражен у оквиру заједница рачунарских наука и оперативних истраживања. Обично се користе два основна алгорита: итерација политике или политике (енг. policy iteration) [67] и итерација вредности (енг. value iteration) [65]. Сатон (енг. Sutton) је увео методу привремене разлике у [70]. Воткинс (енг. Watkins) је предложио добро познату методу Q-учења [69]. Други сродни истраживачки рад у оквиру Марковљевог процеса одлучивања у дискретном времену и Марковљевом процесу одлучивања у дискретном простору-стања се може наћи у [2, 74, 81–86].

У другој фази, стабилност се доводи у контекст АДП-а док се проучавају проблеми управљања динамичких система у реалном времену. Према сазнањима аутора, Левис (енг. Lewis) први даје допринос интеграцији теорије стабилности и теорије АДП-а [71]. Суштинска предност теорије АДП-а је да се оптимална политика управљања може добити путем рекурзивног нумеричког алгорита користећи онлајн информације без решавања Хамилтон-Јакоби-Белманове једначине (ХЈБ) (за нелинеарне системе) и алгебарске Рикатијеве једначине (АРЕ) (за линеарне системе), чак и када динамика система није прецизно позната. Оптимално пројектовање управљања са повратном спрегом за линеарне и нелинеарне динамичке системе је предложило неколико истраживача у последњих неколико година [5, 87, 88]. Док је већина претходног рада о АДП-у била посвећена системима са дискретним временом, било је релативно мање истраживања АДП-а у континуалном времену. То је углавном зато што је АДП знатно тежи за континуалне системе него за дискретне системе. Заиста, многи резултати развијени за дискретне системе не могу се директно проширити на континуалне системе. Као што је раније речено, Левис је



предложио прво решење за анализу стабилности и доказе конвергенције за системе управљања заснованих на АДП-у помоћу теорије линеарно квадратних регулатора (LQR) [89].

Трећа фаза у развоју теорије АДП-а односи се на проширења претходних АДП резултата на нелинеарне системе са неодређеностима. Неуронске мреже и теорија игре се користе за решавање присуства неизвесности и нелинеарности у системима управљања [4, 90–92]. Имплицитна претпоставка у овим радовима је да је ред система познат и да је неодређеност статичка, а не динамичка. Присуство динамичке неодређености није систематски обрађено у литератури АДП-а. Под динамичком неодређеношћу подразумевамо неусклађеност између номиналног модела и реалног система, када је ред номиналног модела нижи од реда реалног система. Сасвим је уобичајено да често недостају потпуне информације стања у многим инжењерским апликацијама, већ су само доступна мерења на излазу или само делимична мерења стања. Адаптација постојеће АДП теорије за овај практични сценарио је важна, али није тривијална. Међутим, анализа стабилности проширеног система естиматора/регулатора није нимало лака, јер је укупан систем међусобно веома повезан.

АДП се може посматрати као адаптивно оптимално управљање. Класичне методе оптималног управљања захтевају добро познавање динамике система. Оптимални регулатор се пројектује офлајн, а затим се имплементира. Међутим, уколико дође до промене у динамици система, регулатор губи своје перформансе, и да би се постигло оптимално управљање потребно је поново пројектовати регулатор.

Са друге стране, за адаптивно управљање није неопходно познавање динамике система, а ако дође до промене у динамици, адаптивни регулатор се може прилагодити тој промени. Адаптивно управљање најчешће није оптимално.

Већина алгоритама управљања заснованих на АДП-у користи повратну спрегу по стањима. Међутим, у пракси често није могуће мерити сва стања. Уместо тога, доступна су нам мерења улазних и излазних сигнала.

На слици 2.3 је приказан регулатор са повратном спрегом по излазу заснован на АДП-у за дискретни систем који се састоји од три дела: реконструкција стања, критичар и актант. Реконструкција стања обезбеђује зависност између улазно-излазних мерења и стања система, која омогућава решавање проблема оптималног управљања система са непознатом динамиком. Критичар на основу улазно-излазних података оцењује закон управљања. Регулатор (актант) током рада учи како да максимизира перформансе система, и побољшава тренутни закон управљања.

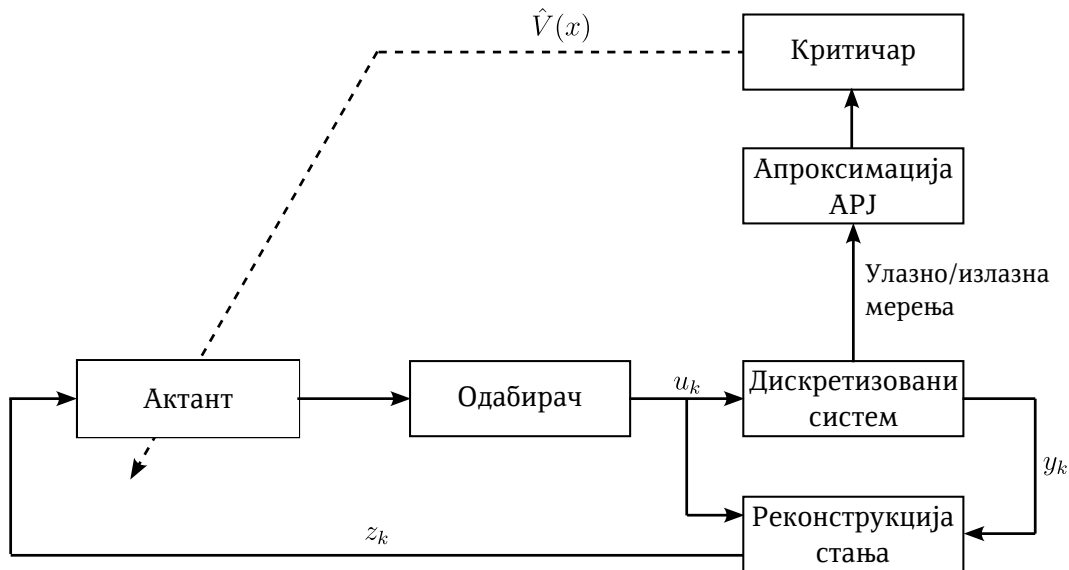
## 2.1 Динамичко програмирање

Посматраћемо дискретни линеарни систем у простору стања описан једначинама стања и излаза

$$x_{k+1} = Ax_k + Bu_k, \quad (2.1)$$

$$y_k = Cx_k, \quad (2.2)$$

где су  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  и  $y(t) \in \mathbb{R}^r$  вектори стања, улази и излаза, респективно, а  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$  и  $C \in \mathbb{R}^{r \times n}$  представљају непознате матрице система. Претпоставимо да је пар  $(A, B)$  управљив, а пар  $(A, C)$  осмотрив.



Слика 2.3: Управљање са повратном спрегом по стању засновано на адаптивном динамичком програмирању.

Квалитет понашања система оцењиваћемо на основу индекса перформансе који је дат у облику линеарне квадратне функције

$$J(x_k) = \sum_{j=k}^{\infty} (y_j^T Q y_j + u_j^T R u_j), \quad (2.3)$$

где је  $Q \in \mathbb{R}^{n \times n}$  реална симетрична позитивно семидефинитна тежинска матрица, а  $R \in \mathbb{R}^{m \times m}$  реална симетрична позитивно дефинитна тежинска матрица. Претпоставићемо да је пар  $(A, C)$  осмотрив.

Индекс перформансе можемо написати као

$$J(x_k) = y_k^T Q y_k + u_k^T R u_k + \sum_{j=k+1}^{\infty} (y_j^T Q y_j + u_j^T R u_j), \quad (2.4)$$

где последњи члан представља индекс перформансе у наредном кораку. Ово нас доводи до Белманове једначине облика

$$J(x_k) = y_k^T Q y_k + u_k^T R u_k + J(x_{k+1}), \quad (2.5)$$

која представља основу динамичког програмирања и учења подстицањем. Она успоставља рекурзивну зависност између вредности тренутног и следећег стања. Тражење оптималне вредносне функције се своди на решавање низа мањих потпроблема.

Оптимална вредност индекса перформансе може се одредити помоћу Хамилтон-Јакоби-Белманове једначине

$$J^*(x_k) = \min_{u_k} (y_k^T Q y_k + u_k^T R u_k + J(x_{k+1})), \quad (2.6)$$

док је оптимални закон управљања одређен изразом

$$u^*(x_k) = \arg \min_{u_k} (y_k^T Q y_k + u_k^T R u_k + J(x_{k+1})). \quad (2.7)$$

За случај линеарног квадратног регулатора, минимална вредност индекса перформансе може се представити у облику квадратне форме

$$J^*(x_k) = x_k^T P^* x_k, \quad (2.8)$$

где је  $P \in R^{n \times n}$  реална симетрична позитивно дефинитна матрица. Ако овај израз заменимо у Белманову једначину добијамо

$$x_k^T P x_k = y_k^T Q y_k + u_k^T R u_k + x_{k+1}^T P x_{k+1}. \quad (2.9)$$

Оптимальни закон управљања тражимо у облику

$$u_k = -K x_k, \quad (2.10)$$

где је  $K$  оптимальна матрица појачања.

Ако изразе (2.1), (2.2) и (2.10) заменимо у (2.9) и средимо, добићемо Љапуновљеву једначину

$$(A - BK)^T P (A - BK) - P + C^T Q C + K^T R K = 0, \quad (2.11)$$

која има позитивно дефинитно решење у случају када матрица појачања  $K$  обезбеђује стабилност система, а пар  $(A, C)$  је осмотрив. Решавањем Љапуновљеве једначине добијамо вредност матрице  $P$ , након чега можемо извршити евалуацију тренутног закона управљања.

Заменом (2.1) у (2.9), добићемо

$$x_k^T P x_k = y_k^T Q y_k + u_k^T R u_k + (A x_k + B u_k)^T P (A x_k + B u_k). \quad (2.12)$$

Хамилтонова функција у овом случају је дефинисана као

$$\mathcal{H}(x_k, u_k) = y_k^T Q y_k + u_k^T R u_k + (A x_k + B u_k)^T P (A x_k + B u_k) - x_k^T P x_k. \quad (2.13)$$

Ако знамо да оптимальни закон управљања  $u_k^*$  треба да задовољи услов

$$\frac{\partial \mathcal{H}(x_k, u_k)}{\partial u_k} = 0, \quad (2.14)$$

следи да је оптимальни закон управљања једнак

$$u^*(x_k) = - (R + B^T P^* B)^{-1} B^T P^* A x_k. \quad (2.15)$$

Из израза оптимальног закона управљања 2.15 видимо да је оптимальна матрица појачања једнака

$$K^* = (R + B^T P^* B)^{-1} B^T P^* A. \quad (2.16)$$

Ако у Љапуновљеву једначину (2.11) заменимо (2.16), сређивањем добићемо дискретну алгебарску Рикатијеву једначину

$$A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A - P + C^T Q C = 0. \quad (2.17)$$

а  $P^*$  представља јединствено симетрично позитивно дефинитно решење једначине

$$A^T P^* A - P^* + C^T Q C - A^T P^* B K^* = 0. \quad (2.18)$$

Пошто је једначина (2.18) нелинеарна за  $P^*$ , тешко се може директно одредити  $P^*$  за комплексне системе вишег реда. Ипак, развијени су многи ефикасни алгоритми који нумерички апроксимирају решење (2.18).

## 2.2 Итеративно решавање проблема динамичког програмирања

Алгоритам итерације закона управљања који је предложио Хјуар (енг. Hewer) [93] представљен је у Лема 2.1.

**Лема 2.1.** Нека је  $K_0 \in \mathbb{R}^{m \times n}$  стабилнишућа матрица појачања и нека је  $P_j$  симетрично позитивно дефинирано решење Љапуновљеве једначине

$$(A - BK_j)^T P_j (A - BK_j) + C^T Q C + K_j^T R K_j = 0, \quad (2.19)$$

код које се  $K_j$ ,  $j = 1, 2, \dots$  ажурира следећим изразом

$$K_{j+1} = (R + B^T P_j B)^{-1} B^T P_j A. \quad (2.20)$$

У овом случају важи да је

1.  $A - BK_j$  стабилна матрица;
2.  $P^* \leq P_{j+1} \leq P_j$ ;
3.  $\lim_{j \rightarrow \infty} K_j = K^*$ ,  $\lim_{j \rightarrow \infty} P_j = P^*$ .

Итеративним решавањем Љапуновљеве једначине 2.11, која је линеарна по  $P_j$ , и рекурзивним ажурирањем закона управљања  $K_j$  користећи израз 2.20, решење нелинеарне једначине 2.18 може се нумерички апроксимирати [93]. Доказано је да секвенце  $\{P_i\}_{i=0}^{\infty}$  и  $\{K_i\}_{i=0}^{\infty}$ , израчунате овим алгоритмом, конвергирају ка оптималним вредностима  $P^*$  и  $K^*$ , респективно. Такође, за  $j = 0, 1, \dots$  матрица  $A - BK_j$  је Шурова матрица. Треба напоменути да је Хјуаров метод заправо алгоритам итерација управљања (итеративно одређивање оптималног закона управљања) заснован на моделу. Наиме, ово је офлајн алгоритам који зависи од параметара система.

Псеудокод ове методе је дат у облику Алгоритма 1. Улазни параметри су матрице система, тежинске матрице, иницијална стабилнишућа матрица појачања и жељена тачност решења. Излазни параметри су приближне вредности матрица  $K$  и  $P$ .

---

**Алгоритам 1** Итеративно одређивање оптималног закона управљања.

---

1. **Иницијализација:** Изабрати стабилнишућу матрицу појачања  $K_0$  и довољно малу константу  $\varepsilon > 0$ . Подесити  $P_0 \leftarrow 0^{n \times n}$  и  $j \leftarrow 1$ .
2. **Евалуација закона управљања:** Решавањем Љапуновљеве једначине

$$(A - BK_j)^T P_j (A - BK_j) - P_j + C^T Q C + K_j^T R K_j = 0,$$

одређујемо непознату матрицу  $P_j$ .

3. **Унапређење закона управљања:** Ажурирамо закон управљања изразом

$$K_{j+1} = (R + B^T P_j B)^{-1} B^T P_j A.$$

4. **Итерације:** Постављамо  $j \leftarrow j + 1$  и понављамо корак 2 све док је испуњен услов  $\|P_j - P_{j-1}\| \geq \varepsilon$ .
-

Такође, Рикатијеву једначину 2.18 могуће је решити и итерацијом вредносне функције коришћењем Љапуновљеве рекурзије [3]. За разлику од Хјуарове методе, овој методи није потребно да иницијална матрица појачања буде стабилнишућа. Псеудокод ове методе је дат у облику Алгоритма 2. Улазни параметри су матрице система, тежинске матрице и жељена тачност решења, а излазни параметри су приближне вредности матрица  $K$  и  $P$ .

---

**Алгоритам 2** Итеративно одређивање оптималне вредносне функције.
 

---

1. **Иницијализација:** Изабрати довољно малу константу  $\varepsilon > 0$ . Подесити  $K_0 \leftarrow 0^{m \times n}$ ,  $P_0 \leftarrow 0^{n \times n}$  и  $j \leftarrow 1$ .

2. **Ажурирање вредносне функције:** Коришћењем Љапуновљеве рекурзије

$$P_j = (A - BK_j)^T P_{j-1} (A - BK_j) + C^T Q C + K_j^T R K_j,$$

ажурирамо матрицу  $P_j$ .

3. **Унапређење закона управљања:** Ажурирамо закон управљања изразом

$$K_{j+1} = (R + B^T P_j B)^{-1} B^T P_j A.$$

4. **Итерације:** Постављамо  $j \leftarrow j + 1$  и понављамо корак 2 све док је испуњен услов  $\|P_j - P_{j-1}\| \geq \varepsilon$ .

---

У наставку тезе биће коришћена метода итеративног одређивања закона управљања. Да бисмо применили овај алгоритам у реалном времену за дискретни систем (2.1)-(2.2), потребно је развити алгоритам управљања заснован на АДП-у користећи повратну спрегу по стању, која не зависи од познавања матрица система.

## 2.3 Управљање засновано на АДП-у и повратној спрези по стањима

Један од најважних принципа на којима се заснива АДП јесте грешка привремених разлика или Белманова грешка. Грешка привремених разлика произилази из Белманове једначине (2.9) и има облик:

$$\varepsilon_k = -x_k^T P x_k + y_k^T Q y_k + u_k^T R u_k + x_{k+1}^T P x_{k+1}, \quad (2.21)$$

који одговара Хамилтоновој функцији.

Што смо ближи оптималном решењу грешка ће бити мања. Такође, у изразу не конфигуришу матрице система, тако да немамо потребу за познавањем динамике система.

**Напомена 2.1.** У наставку тезе биће коришћена следећа нотација:

1.  $\mathbb{R}$  и  $\mathbb{Z}_+$  означавају скупи реалних и позитивних целих бројева, респективно.
-

2. Зајраде  $\| \cdot \|$  означавају Еуклидску норму вектора или матричну норму у случају матрице.
3. Символ  $\otimes$  се користи да значи Кронекеров производ.
4. Функција векторизације матрице

$$\text{vec}(A) = [a_1^T \ a_2^T \ \dots \ a_m^T]^T,$$

враћа  $m$ -вектор који се добија спајањем колоне  $A \in \mathbb{R}^{n \times m}$  једне истог групе, где  $a_i \in \mathbb{R}^n$  представљају колоне матрице  $A$ .

5. За произвољну симетричну матрицу  $S \in \mathbb{R}^{n \times n}$  векторизација се врши на следећи начин

$$\text{vecs}(S) = [s_{11} \ s_{12} \ \dots \ s_{1n} \ s_{22} \ s_{23} \ \dots \ s_{n-1,n} \ s_{nn}]^T \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

6. За произвољни вектор колону  $v \in \mathbb{R}^n$ ,

$$\text{vecv}(v) = [v_1^2 \ 2v_1v_2 \ \dots \ 2v_1v_n \ v_2^2 \ 2v_2v_3 \ \dots \ 2v_{n-1}v_n \ v_n^2]^T \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

За Хурвицову матрицу повратне сиреће  $A - BK$ ,  $K \in \mathbb{R}^{m \times n}$  се назива стабилнишућа матрица појачања за линеарни систем  $\dot{x} = Ax + Bu$ .

Други важан принцип на коме се заснива АДП јесте апроксимација вредносне функције. У тези се разматра апроксимација вредносне функције линеарном функцијом. Такође, у зависности од сложености проблема могу се користити и остали параметарски и непараметарски модели [1].

На основу релације

$$\text{vec}(ABC) = (C^T \otimes A) \text{vec}(P), \quad (2.22)$$

вредносна функција у квадратној форми  $J(x_k) = x_k^T P x_k$  се може написати у облику

$$J(x_k) = (x_k^T \otimes x_k^T) \text{vec}(P). \quad (2.23)$$

Ако ову апроксимацију заменимо у Белманову једначину добићемо

$$(x_k^T \otimes x_k^T) \text{vec}(P) = y_k^T Q y_k + u_k^T R u_k + (x_{k+1}^T \otimes x_{k+1}^T) \text{vec}(P), \quad (2.24)$$

односно

$$(x_k^T \otimes x_k^T - x_{k+1}^T \otimes x_{k+1}^T) \text{vec}(P) = y_k^T Q y_k + u_k^T R u_k. \quad (2.25)$$

С обзиром да је матрица  $P$  симетрична, број непознатих параметара можемо смањити на  $l = \frac{n(n+1)}{2}$  на следећи начин

$$(\text{vecv}(x_k)^T - \text{vecv}(x_{k+1})^T) \text{vecs}(P) = y_k^T Q y_k + u_k^T R u_k. \quad (2.26)$$

Сада се добијање вектора  $\text{vecs}(P)$  своди на решавање добро познатог проблема система једначина облика

$$\begin{bmatrix} \text{vecv}(x_k)^T - \text{vecv}(x_{k+1})^T \\ \text{vecv}(x_{k-1})^T - \text{vecv}(x_k)^T \\ \vdots \\ \text{vecv}(x_{k-l+1})^T - \text{vecv}(x_{k-l+2})^T \end{bmatrix} \text{vecs}(P) = \begin{bmatrix} y_k^T Q y_k + u_k^T R u_k \\ y_{k-1}^T Q y_{k-1} + u_{k-1}^T R u_{k-1} \\ \vdots \\ y_{k-l+1}^T Q y_{k-l+1} + u_{k-l+1}^T R u_{k-l+1} \end{bmatrix}. \quad (2.27)$$

Да бисмо могли одредити претходну матрицу применићемо "шаржну" методу најмањих квадрата (енг. batch least square)

$$\text{vecs}(P) = (\Phi^T \Phi)^{-1} \Phi^T Y, \quad (2.28)$$

где је

$$\Psi = \begin{bmatrix} \text{vecv}(x_k)^T - \text{vecv}(x_{k+1})^T \\ \text{vecv}(x_{k-1})^T - \text{vecv}(x_k)^T \\ \vdots \\ \text{vecv}(x_{k-l+1})^T - \text{vecv}(x_{k-l+2})^T \end{bmatrix}, \quad (2.29)$$

и

$$\Phi = \begin{bmatrix} y_k^T Q y_k + u_k^T R u_k \\ y_{k-1}^T Q y_{k-1} + u_{k-1}^T R u_{k-1} \\ \vdots \\ y_{k-l+1}^T Q y_{k-l+1} + u_{k-l+1}^T R u_{k-l+1} \end{bmatrix}. \quad (2.30)$$

Када одредимо матрицу  $P_j$  можемо да ажурирамо матрицу појачања  $K_j$  помоћу израза 2.20 и поступак се понавља све док не постигнемо задату тачност  $\varepsilon$ , односно све док је испуњен услов  $\|P_j - P_{j-1}\| > \varepsilon$ . Следи алгоритам 3 пројектовања регулатора заснованог на АДП-у са повратном спрегом по стању.

---

**Алгоритам 3** Пројектовање регулатора заснованог на АДП-у са повратном спрегом по стању.

---

1. **Иницијализација:** Изабрати стабилишућу матрицу појачања  $K_0$  и довољно малу константу  $\varepsilon > 0$ . Подесити  $j \leftarrow 1$  и  $P_0 \leftarrow 0^{n \times n}$ .
2. **Прикупљање података:** Примењујемо тренутни закон управљања  $u_j = -K_j x_j$  и на основу података формирамо матрице  $\Psi_j$  и  $\Phi_j$ .
3. **Евалуација закона управљања:** На основу израза

$$\text{vecs}(P_j) = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j,$$

одређујемо непознату матрицу  $P_j$ .

4. **Унапређење закона управљања:** Ажурирамо закон управљања изразом

$$K_{j+1} \leftarrow (R + B^T P_j B)^{-1} B^T P_j A.$$

5. **Итерације:** Постављамо  $j \leftarrow j + 1$  и понављамо корак 2 све док је испуњен услов  $\|\bar{P}_j - \bar{P}_{j-1}\| \geq \varepsilon$ .
  6. **Експлоатација:** Примена апроксимираног оптималног закона управљања  $u_j = -K_j x_j$ .
- 

Ефикасност управљања заснована на АДП-у и повратној спрези по стањима

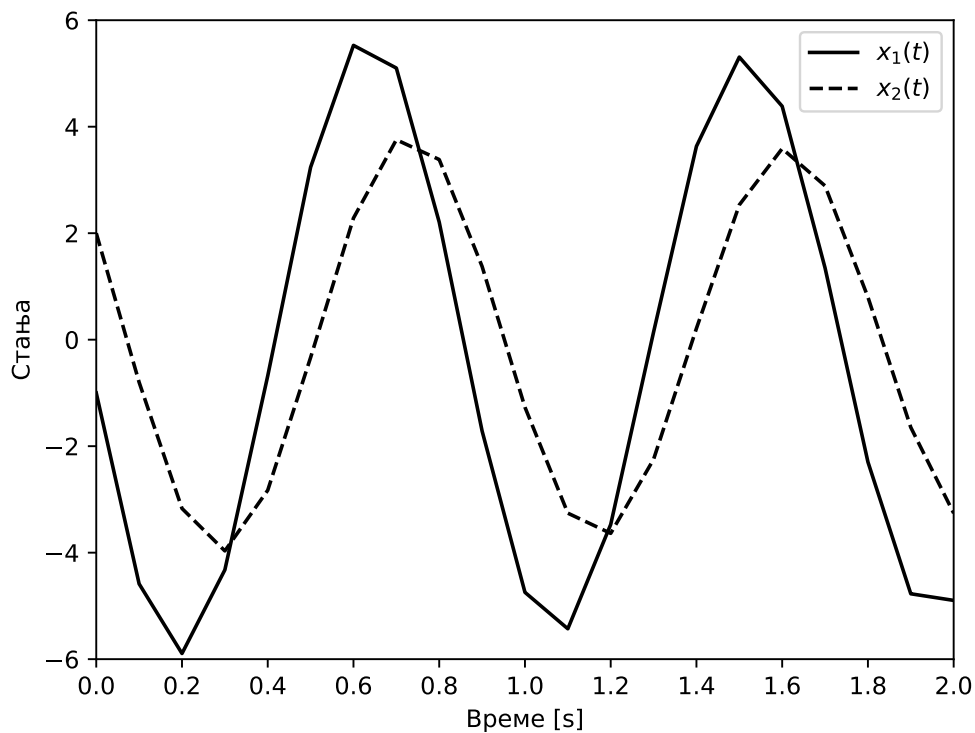
---

демонстрирана је на следећем систему

$$x_{k+1} = \begin{bmatrix} 1.5 & -0.6 \\ 0.7 & -0.05 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_k,$$

$$y_k = [1 \quad 0.3] x_k.$$

Сопствене вредности матрице  $A$  су  $\lambda_1 = 0.3$  и  $\lambda_2 = 1.15$ , на основу чега можемо да закључимо да је систем нестабилан. Да бисмо могли да применимо методу итерације управљања морамо изабрати стабилишућу матрицу појачања  $K_0 = [-0.05, 0.9179]$  чији одзив за почетне услове  $x_0 = [-1, 2]^T$  је приказан на слици 2.4.



Слика 2.4: Стања система за иницијалну матрицу појачања  $K_0$ .

Симулацијма су добијене следеће матрице  $P$  и  $K$

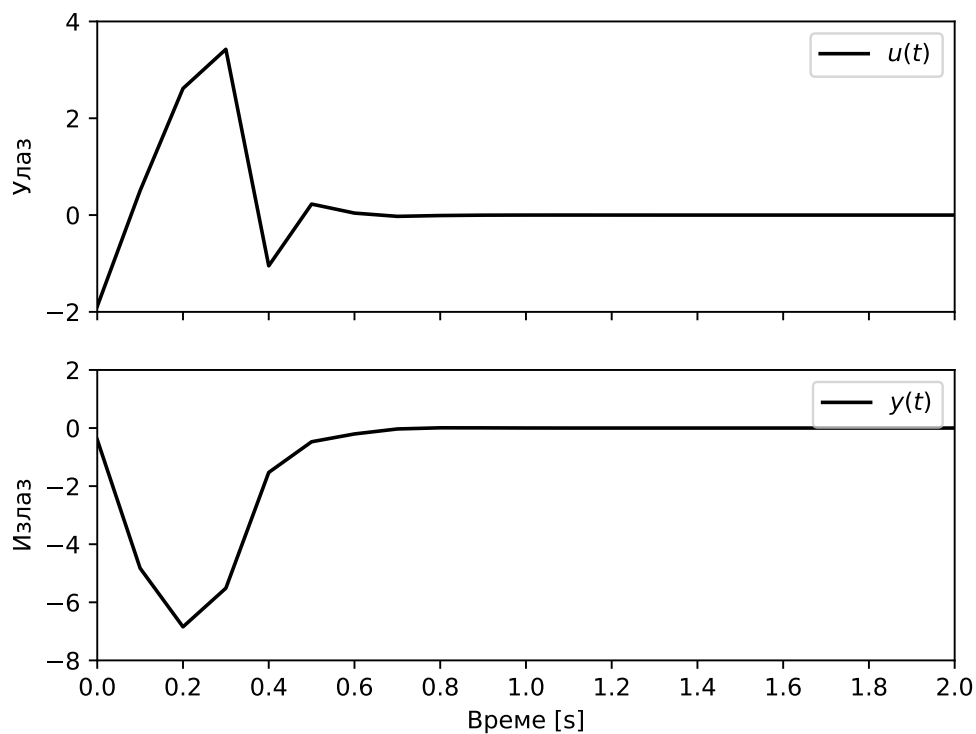
$$P = \begin{bmatrix} 2.5943 & -0.3170 \\ -0.3170 & 0.3453 \end{bmatrix}, \quad K = [1.0209 \quad -0.4287],$$

чије вредности су идентичне оптималним вредностима  $P^*$  и  $K^*$

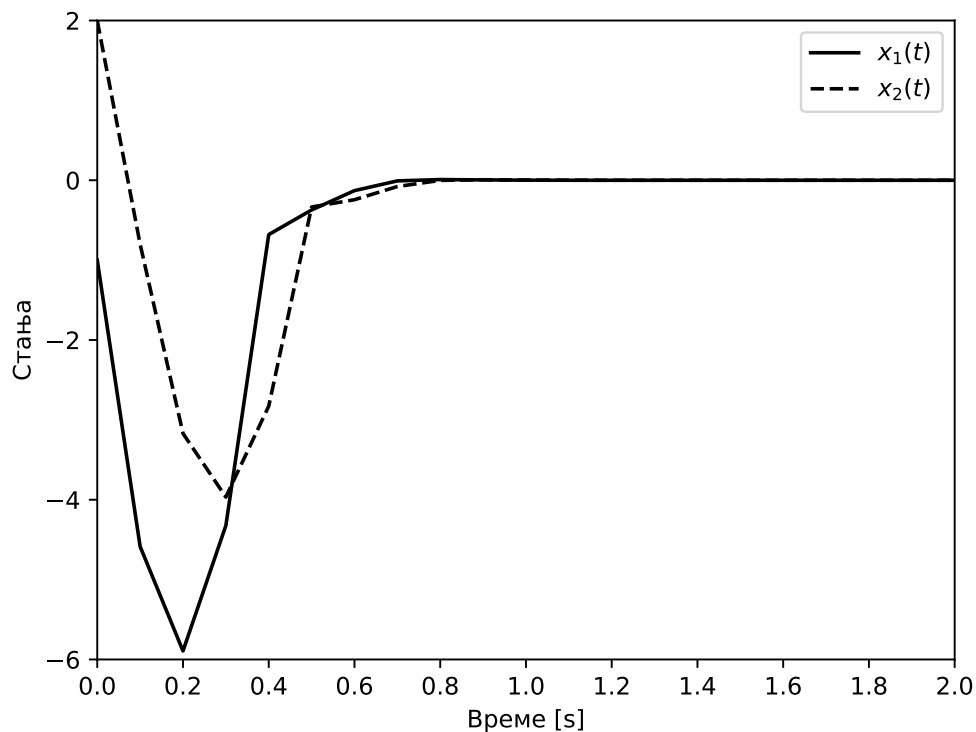
$$P^* = \begin{bmatrix} 2.5943 & -0.3170 \\ -0.3170 & 0.3453 \end{bmatrix}, \quad K^* = [1.0209 \quad -0.4287].$$

На слици 2.5 приказани су улаз и излаз, а слика 2.6 приказује стања, за почетно стање  $x_0 = [-1, 2]^T$ , система управљаног регулатором заснованог на АДП-у са повратном спрегом по стањима.





Слика 2.5: Улаз и излаз система управљаног регулатором заснованим на АДП-у и повратном спрегом по стањима.



Слика 2.6: Стања система управљаног регулатором заснованим на АДП-у и повратном спрегом по стањима.

У пракси, мерење стања система захтева додатну мерну опрему што повећава

цену управљачког система, а често је и немогуће измерити сва стања. Због тога је много практичније мерити улазне и излазне сигнале система и на основу њих извршити реконструкцију вектора стања.

## 2.4 Реконструкција стања

Дискретни систем (2.1)-(2.2) се може проширити улазно/излазном секвенцом на временском хоризонту  $[k - N, k - 1]$ , слично као што је урађено у [94, 95]. Проширимо прво једначину стања

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_{k-1} = A^2x_{k-2} + Bu_{k-1} + Bu_{k-2} = \dots \\ &= A^N x_{k-N} + Bu_{k-1} + ABu_{k-2} + \dots + A^{N-1}Bu_{k-N} = \\ &= A^N x_{k-N} + [B \ AB \ \dots \ A^{N-1}B] \begin{bmatrix} u_{k-1} \\ u_{k-2} \\ \vdots \\ u_{k-N} \end{bmatrix}. \end{aligned}$$

Сада употребом проширене једначине стања можемо проширити и једначине излаза за  $y_{k-1} \dots y_{k-N}$  на следећи начин

$$\begin{aligned} \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-N+1} \\ y_{k-N} \end{bmatrix} &= \begin{bmatrix} C(A^{N-1}x_{k-N} + Bu_{k-2} + ABu_{k-3} + \dots + A^{N-2}Bu_{k-N}) \\ C(A^{N-2}x_{k-N} + Bu_{k-3} + ABu_{k-4} + \dots + A^{N-3}Bu_{k-N}) \\ \vdots \\ C(Ax_{k-N+1} + Bu_{k-N+1}) \\ Cx_{k-N} \end{bmatrix} = \\ &= \begin{bmatrix} CA^{N-1} \\ CA^{N-2} \\ \vdots \\ CA \\ C \end{bmatrix} x_{k-N} + \begin{bmatrix} 0 & CB & CAB & \dots & CA^{N-2}B \\ 0 & 0 & CB & \dots & CA^{N-3}B \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & CB \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} u_{k-1} \\ u_{k-2} \\ \vdots \\ u_{k-N+1} \\ u_{k-N} \end{bmatrix}. \end{aligned}$$

Овако проширене једначине стања и излаза можемо записати у компактнијем облику

$$\begin{aligned} x_k &= A^N x_{k-N} + V(N)\bar{u}_{k-1,k-N}, \\ \bar{y}_{k-1,k-N} &= U(N)x_{k-N} + T(N)\bar{u}_{k-1,k-N}, \end{aligned} \tag{2.31}$$

у којем вектори

$$\begin{aligned} \bar{u}_{k-1,k-N} &= [u_{k-1}^T \ u_{k-2}^T \ \dots \ u_{k-N}^T]^T, \\ \bar{y}_{k-1,k-N} &= [y_{k-1}^T \ y_{k-2}^T \ \dots \ y_{k-N}^T]^T, \end{aligned}$$

представљају доступна мерења улазног и излазног сигнала, при чему матрице

$$\begin{aligned} V(N) &= [B \ AB \ \dots \ A^{N-1}B], \\ U(N) &= [(CA^{N-1})^T \ (CA)^T \ \dots \ C^T]^T, \end{aligned}$$

представљају матрицу управљивости и матрицу осмотривости, респективно, а матрица

$$T(N) = \begin{bmatrix} 0 & CB & CAB & \dots & CA^{N-2}B \\ 0 & 0 & CB & \dots & CA^{N-3}B \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & CB \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \quad (2.32)$$

представља матрицу Марковљевих параметара, који се у случају дискретног система поклапају са импулсним одзивом.

Индекс осмотривости је дефинисан као  $N = \max(\rho_u, \rho_v)$ , где је  $\rho_u$  најмањи цео број који гарантује да је матрица  $U(\rho_u)$  пуног ранга колона, док је  $\rho_v$  најмањи цео број који гарантује да је матрица  $V(\rho_v)$  пуног ранга врста [95]. Такође, постоји леви инверз матрице  $U(N)$ , дефинисан као

$$U^+(N) = [U^T(N)U(N)]^{-1} U^T(N).$$

На основу реконструисаног стања 2.31, можемо применити регулатор заснован на АДП-у са повратном спрегом по излазу да би се решио проблем оптималног управљања за систем са непознатом динамиком. Јединственост реконструкције стања исказана је у Лема 2.2 [96].

**Лема 2.2.** *Ако су задовољени услови осмотривости и управљивости за систем (2.1)-(2.2), онда се стања система могу јединствено добити на основу измерених улазно-излазних сигнала на следећи начин*

$$x_k = \Theta z_k, \quad (2.33)$$

где је  $\Theta = [M_u \ M_y]$  матрица пуног ранга врста,  $M_u = V(N) - M_y T(N)$ ,  $M_y = A_d^N U^+(N)$ ,  $z_k = [\bar{u}_{k-1, k-N}^T \ \bar{y}_{k-1, k-N}^T]^T \in \mathbb{R}^n$ ,  $q = N [\dim(u) + \dim(y)]$

## 2.5 Управљање засновано на АДП-у и повратној спрези по излазу

Апроксимација вредносне функције има доста сличности са идентификацијом система. Слично као у случају идентификације система потребан нам је сигнал који ће обезбедити перзистентну побуду. То можемо постићи додавањем истраживачког шума улазном сигналу током фазе учења на следећи начин

$$u_k = -K_j x_k + e_k, \quad (2.34)$$

где  $e_k$  представља истраживачку компоненту управљачког сигнала. Избор истраживачког шума представља нетривијални проблем који ће детаљније бити изложен у следећем одељку.

Додавањем истраживачког шума јавља се додатни члан поред алгебарске Рикатијеве једначине, који смањује тачност матрица  $P$  и  $K$ . Да бисмо умањили дејство

истраживачког шума преформулисаћемо једначину стања дискретног модела на следећи начин

$$x_{k+1} = Ax_k + B(-K_j x_k + e_k) = (A - BK_j)x_k + Be_k = A_j x_k + B(K_j x_k + u_k), \quad (2.35)$$

где је  $A_j = A - BK_j$ .

Сада се разлика вредносне функције између два суседна стања може написати на следећи начин

$$\begin{aligned} x_{k+1}^T P_j x_{k+1} - x_k^T P_j x_k &= [A_j x_k + B(K_j x_k + u_k)]^T P_j [A_j x_k + B(K_j x_k + u_k)] - \\ &\quad - x_k^T P_j x_k = x_k^T A_j^T P_j A_j x_k + \\ &\quad + 2(K_j x_k + u_k)^T B^T P_j A_j x_k + (K_j x_k + u_k)^T B^T P_j B(K_j x_k + u_k) - x_k^T P_j x_k = \\ &= 2(K_j x_k + u_k)^T B^T P_j (A - BK_j) x_k + (K_j x_k + u_k)^T B^T P_j B(K_j x_k + u_k) + \\ &\quad + A_j^T P_j A_j x_k - x_k^T P_j x_k. \end{aligned} \quad (2.36)$$

На основу Љапуновљеве једначине следи да је

$$\begin{aligned} x_k^T A_j^T P_j A_j x_k - x_k^T P_j x_k &= - (x_k^T C^T Q C x_k + x_k^T K_j^T R K_j x_k) = \\ &= - (y_k^T Q y_k + x_k^T K_j^T R K_j x_k), \end{aligned} \quad (2.37)$$

па ће израз 2.36 након замене и сређивања имати следећи облик

$$\begin{aligned} x_{k+1}^T P_j x_{k+1} - x_k^T P_j x_k &= 2(K_j x_k + u_k)^T B^T P_j A x_k + \\ &\quad + (K_j x_k + u_k)^T B^T P_j B(-K_j x_k + u_k) - (y_k^T Q y_k + x_k^T K_j^T R K_j x_k). \end{aligned} \quad (2.38)$$

Као што можемо видети израз и даље зависи од стања. Сада можемо применити реконструкцију стања  $x_k = \Theta z_k$  како би израз 2.38 зависио само до улазно-излазних података, након чега имамо да је

$$\begin{aligned} z_{k+1}^T \Theta^T P_j \Theta z_{k+1} - z_k^T \Theta^T P_j \Theta z_k &= 2(K_j \Theta z_k + u_k)^T B^T P_j A \Theta z_k + \\ &\quad + (K_j \Theta z_k + u_k)^T B^T P_j B(-K_j \Theta z_k + u_k) - (y_k^T Q y_k + z_k^T \Theta^T K_j^T R K_j \Theta z_k). \end{aligned} \quad (2.39)$$

Ако уведемо следеће ознаке

$$\begin{aligned} \bar{P}_j &= \Theta^T P_j \Theta, \\ \bar{K}_j &= K_j \Theta, \end{aligned}$$

израз 2.39 се може написати као

$$\begin{aligned} z_{k+1}^T \bar{P}_j z_{k+1} - z_k^T \bar{P}_j z_k &= 2(\bar{K}_j z_k + u_k)^T B^T P_j A \Theta z_k + \\ &\quad + (\bar{K}_j z_k + u_k)^T B^T P_j B(-\bar{K}_j z_k + u_k) - (y_k^T Q y_k + z_k^T \bar{K}_j^T R \bar{K}_j z_k). \end{aligned} \quad (2.40)$$

Извршићемо векторизацију употребом Кронекеровог производа

$$\begin{aligned} (z_{k+1}^T \otimes z_{k+1}^T) \text{vec}(\bar{P}_j) - (z_k^T \otimes z_k^T) \text{vec}(\bar{P}_j) &= \\ &= 2 [(z_k^T \otimes z_k^T)(I_m \otimes \bar{K}_j) + z_k^T \otimes u_k^T] \text{vec}(B^T P_j A \Theta) + \\ &\quad + [u_k^T \otimes u_k^T - (z_k^T \otimes z_k^T) (\bar{K}_j^T \otimes \bar{K}_j^T)] \text{vec}(B^T P_j B) - (y_k^T Q y_k + z_k^T \bar{K}_j^T R \bar{K}_j z_k). \end{aligned} \quad (2.41)$$

С обзиром да су матрице  $\bar{P}_j$  и  $B^T P_j B$  симетричне, претходни израз можемо преформулисати на следећи начин

$$\begin{aligned} & [\text{vecv}(z_{k+1})^T - \text{vecv}(z_k)^T] \text{vecs}(\bar{P}_j) = \\ & = 2 [(z_k^T \otimes z_k^T)(I_m \otimes \bar{K}_j) + z_k^T \otimes u_k^T] \text{vec}(B^T P_j A \Theta) + \\ & + [\text{vecv}(u_k)^T - \text{vecv}(z_k)^T \text{vecv}(K_j)^T] \text{vecs}(B^T P_j B) - (y_k^T Q y_k + z_k^T \bar{K}_j^T R \bar{K}_j z_k). \end{aligned} \quad (2.42)$$

Ако уведемо следећу матрицу

$$\Psi_j = [\Gamma_{\bar{u}} - \Gamma_{\bar{k}z}, \quad 2(\Gamma_{zu} + \Gamma_{zz}(I_m \otimes \bar{K}_j^T)), \quad -\delta_{zz}], \quad (2.43)$$

где је

$$\begin{aligned} \delta_{zz} &= [\text{vecv}(z_{k_1}) - \text{vecv}(z_{k_0}), \text{vecv}(z_{k_2}) - \text{vecv}(z_{k_1}), \dots, \text{vecv}(z_{k_{l+1}}) - \text{vecv}(z_{k_l})]^T, \\ \Gamma_{zz} &= [z_{k_1} \otimes z_{k_0}, z_{k_2} \otimes z_{k_1}, \dots, z_{k_{l+1}} \otimes z_{k_l}]^T, \\ \Gamma_{zu} &= [z_{k_1} \otimes u_{k_0}, z_{k_2} \otimes u_{k_1}, \dots, z_{k_{l+1}} \otimes u_{k_l}]^T, \\ \Gamma_{\bar{u}} &= [\text{vecv}(u_{k_0}), \text{vecv}(u_{k_1}), \dots, \text{vecv}(u_{k_l})]^T, \\ \Gamma_{\bar{k}z} &= [\text{vecv}(\bar{K}_j z_{k_0}), \bar{K}_j z_{k_1}, \dots, \bar{K}_j z_{k_l}]^T, \end{aligned} \quad (2.44)$$

као и матрицу

$$\Phi_j = \begin{bmatrix} y_{k_0}^T Q y_{k_0} + z_{k_0}^T \bar{K}_j^T R \bar{K}_j z_{k_0} \\ y_{k_1}^T Q y_{k_1} + z_{k_1}^T \bar{K}_j^T R \bar{K}_j z_{k_1} \\ \vdots \\ y_{k_l}^T Q y_{k_l} + z_{k_l}^T \bar{K}_j^T R \bar{K}_j z_{k_l} \end{bmatrix}, \quad (2.45)$$

онда израз 2.42 можемо написати у облику матричне једначине

$$\Psi_j \begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = \Phi_j. \quad (2.46)$$

Непознате матрице се могу одредити помоћу методе најмањих квадрата

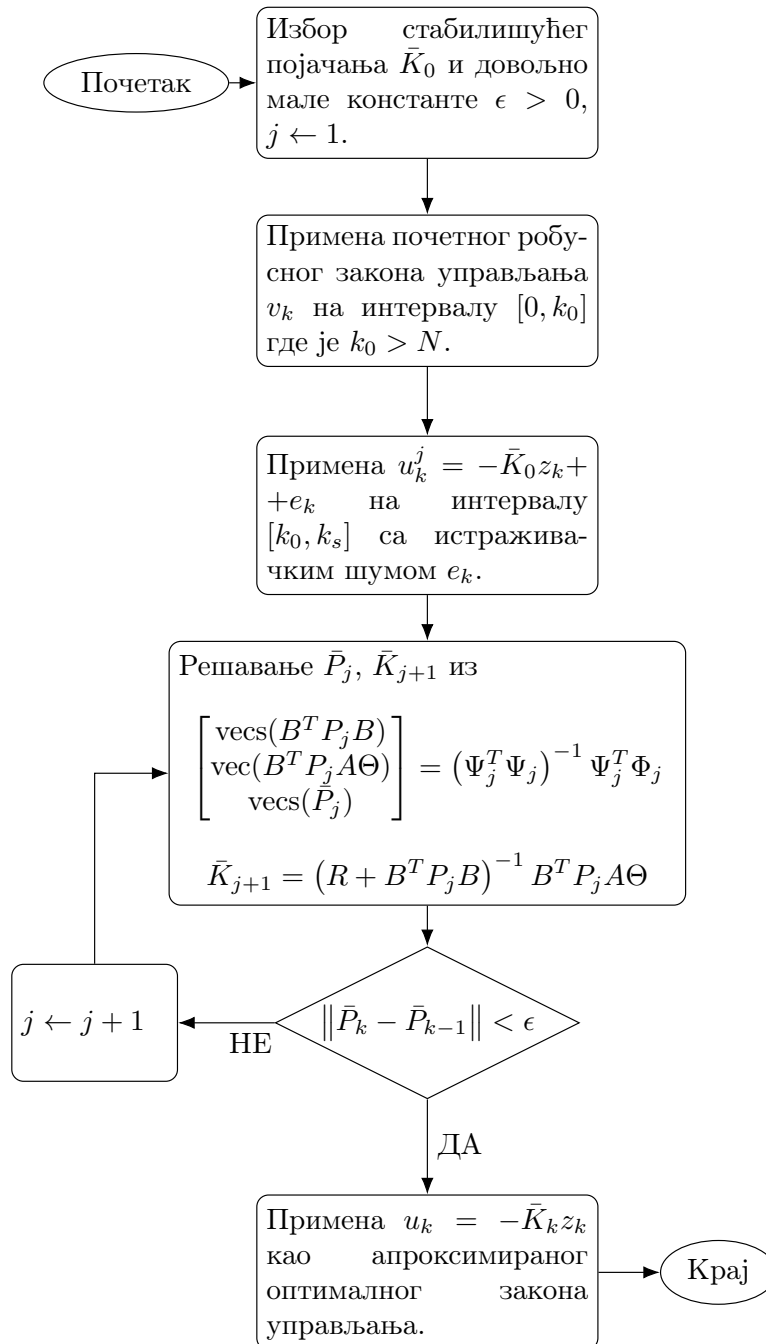
$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j. \quad (2.47)$$

Након што се одреде непознате матрице можемо ажурирати закон управљања користећи израз

$$K_{j+1} = (R + B^T P_j B)^{-1} B^T P_j A \Theta. \quad (2.48)$$

Овај поступак се понавља све док се не постигне одређена тачност решења, односно  $\|\bar{P}_j - \bar{P}_{j-1}\| < \varepsilon$ .

Треба напоменути да се (2.47) назива евалуација закона управљања, и њена сврха је да јединствено одреди  $\bar{P}_j$ , а (2.48) представља побољшање закона управљања које се користи за ажурирање појачања  $\bar{K}_{j+1}$ . На слици 2.7 је представљен управљачки алгоритам заснован на АДП-у.



Слика 2.7: Алгоритам пројектовања регулатора заснованог на АДП-у.

---

**Алгоритам 4** Пројектовње регулатора заснованог на АДП-у са повратном спрегом по излазу.

---

1. **Иницијализација:** Изабрати стабилишућу матрицу појачања  $K_0$  и довољно малу константу  $\varepsilon > 0$ . Подесити  $j \leftarrow 1$ .
2. **Прикупљање података:** Применити  $u_k^j = -\bar{K}_0 z_k + e_k$  на интервалу  $[k_0, k_l]$  где је  $e_k$  истраживачки шум. Формирати матрице  $\Psi_j$  и  $\Phi_j$  на основу улазно-излазних података тако да (2.47) има јединствено решење.
3. **Евалуација закона управљања:** На основу израза

$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j,$$

одређујемо непознате матрице  $B^T P_j B$ ,  $B^T P_j A \Theta$  и  $\bar{P}_j$ .

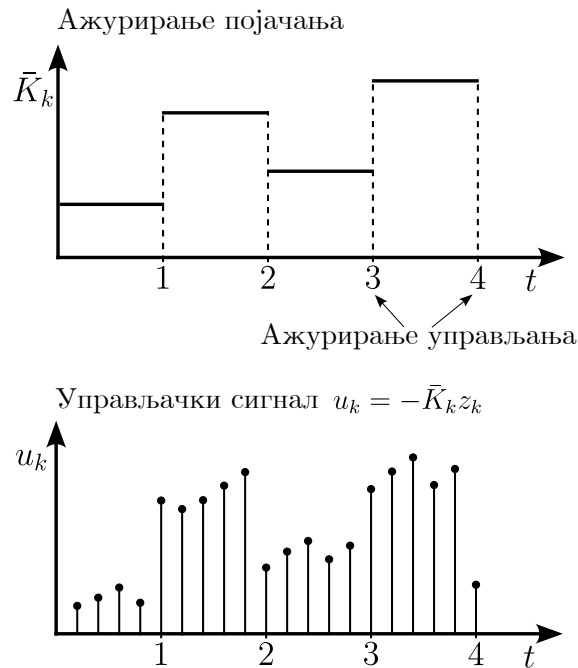
4. **Унапређење закона управљања:** Ажурирамо закон управљања изразом

$$K_{j+1} = (R + B^T P_j B)^{-1} B^T P_j A \Theta.$$

5. **Итерације:** Постављамо  $j \leftarrow j + 1$  и понављамо корак 2 све док је испуњен услов  $\|P_j - \bar{P}_{j-1}\| \geq \varepsilon$ .
  6. **Експлоатација:** Примена апроксимираног оптималног закона управљања  $u_k = -\bar{K}_k z_k$ .
- 

Треба приметити да решавање (2.47) уместо (2.11) у потпуности елиминише првобитну потребу за познавање тачне динамике система. Уместо тога, потребно је само мерити улаз  $u_k$  и излаз  $y_k$ . Наиме, ако посматрамо израз за  $z_k$  можемо уочити да закон управљања  $u_k = -\bar{K}_k^* z_k$  садржи само претходно измерене улазно-излазне податке.

Хибридна природа регулатора приказана је на слици 2.8. Може се видети да се појачање или закон управљања ажурирају у дискретним тренуцима на основу (2.48) након што се реши (2.47). Поред тога, управљање је дискретни сигнал који зависи од стања  $z_k$  у сваком тренутку  $k$ . Са слике 2.8 се може видети да се појачања ажурирају у дискретним тренуцима, па је управљачки сигнал на делове дискретан.



Слика 2.8: Хибридна природа управљачког сигнала.

## 2.6 Избор истраживачког шума

Посматрајмо линеарно дискретни систем

$$x_{k+1} = Ax_k + Bu_k. \quad (2.49)$$

где је  $x \in \mathbb{R}^n$  стање система које је у потпуности доступно за пројектовање управљања преко повратне спреге (енг. feedback control design),  $u \in \mathbb{R}^m$  је управљање, док су  $A \in \mathbb{R}^{n \times n}$  и  $B \in \mathbb{R}^{n \times m}$  непознате константне матрице. Притом се претпоставља да је систем стабилан.

Циљ пројектовања је наћи оптимални закон управљања у форми

$$u_k = -Kx_k, \quad (2.50)$$

који минимизира следећи индекс перформанси

$$J(x_k) = \sum_{j=k}^{\infty} (y_j^T Q y_j + u_j^T R u_j), \quad (2.51)$$

где је  $Q = Q^T \geq 0$ ,  $R = R^T > 0$ , са осмотривим паром  $(A, Q^{1/2})$ . Међутим, због нелинеатности алгебарске Рикатијеве једначине по  $P_k$ , оптимално решење је обично тешко директно добити. Притом, као што се може видети из

$$K_k = R^{-1} B^T P_{k-1}, \quad (2.52)$$

тачно познавање матрице система  $B$  је и даље потребно за итерације.

Пошто се  $x_k$  и  $u_k$  могу мерити у реалном времену, симетрично решење  $P_k$  се може јединствено решити под условом перзистентне побуде. Да би се обезбедио услов перзистентне побуде, стање је потребно да се ресетује при сваком кораку



итерације, али то може изазвати техничке проблеме за анализу стабилности система затворене петље [89]. Алтернативни начин је додавање истраживачког шума [15, 89, 97] тако да је сада улазни сигнал дат у форми

$$u_k = -Kx_k + e_k, \quad (2.53)$$

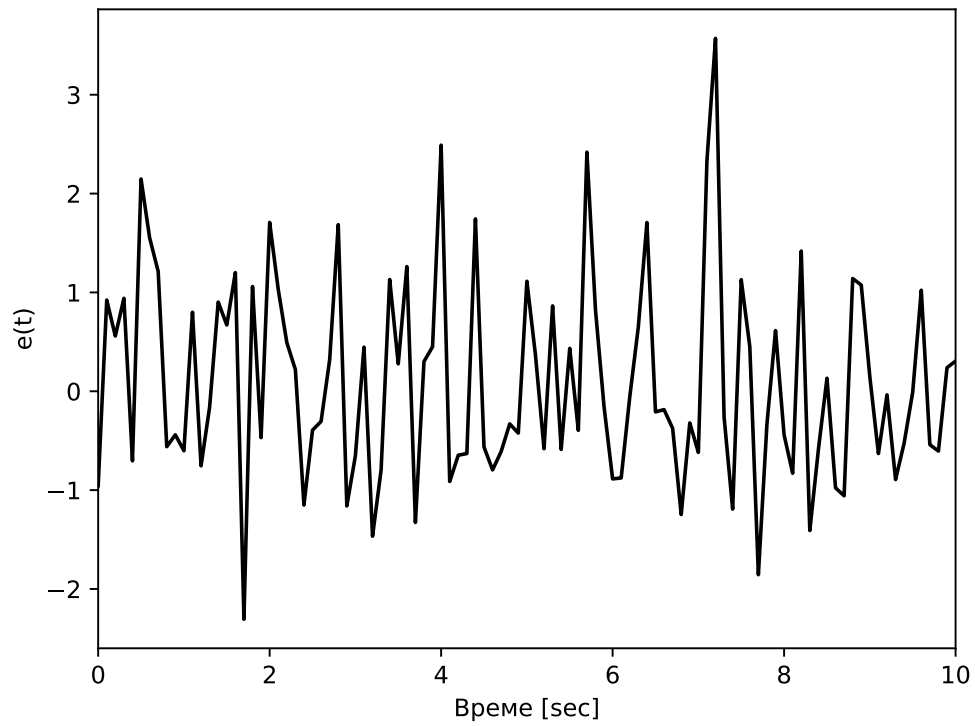
где је истраживачки шум  $e_k$ , без утицаја на конвергенцију у процесу учења. Након сваког ажурирања закона управљања, информације о стању и улазу морају бити прикупљене за следећу итерацију. Ово може успорити процес учења, посебно за високодимензионалне системе.

Треба имати на уму да нам знање о матрицама система,  $A$  и  $B$ , као и матрицама  $Q$  и  $R$  није потребно током процеса учења. Упркос сличности нотација,  $u_k$  се потпуно разликује од  $u$ . Стварни улаз система  $u$  се користи за генерисање података у реалном времену (енг. online) и стога одговара понашању закона управљања. С друге стране,  $u_k$  (укључујући  $u_0$ ) је регулатор који се ажурира кроз побољшање закона управљања, и стога је у ствари циљни закон управљања. С обзиром на то,  $u_k$  је функција од стања, док је  $u$  само управљачки улаз (функција времена) који чак теоретски може бити сигнал који је независан од стања у отвореној петљи. Ово раздвајање  $u$  од  $u_k$  сврстава изложене алгоритме управљања засноване на АДП-у у алгоритме независне од иницијалног закона управљања (енг. off-policy). Узимајући ово у обзир, додавање шума истраживања у оптимално управљање не угрожава тачност процене стварног управљања  $u$ , све док је услов перзистентне ексцитације (ПЕ) задовољен [82, 98, 99]. Дакле, користи се закон управљања  $u_k = -Kx_k + e_k$  да се генерише понашање од  $k_0$  до  $k_l$ . Након тога се процењује и побољшава закон управљања  $\hat{u}_k = -K_j x_k + e_k$  користећи прикупљене информације у интервалу  $[k_0, k_l]$ .

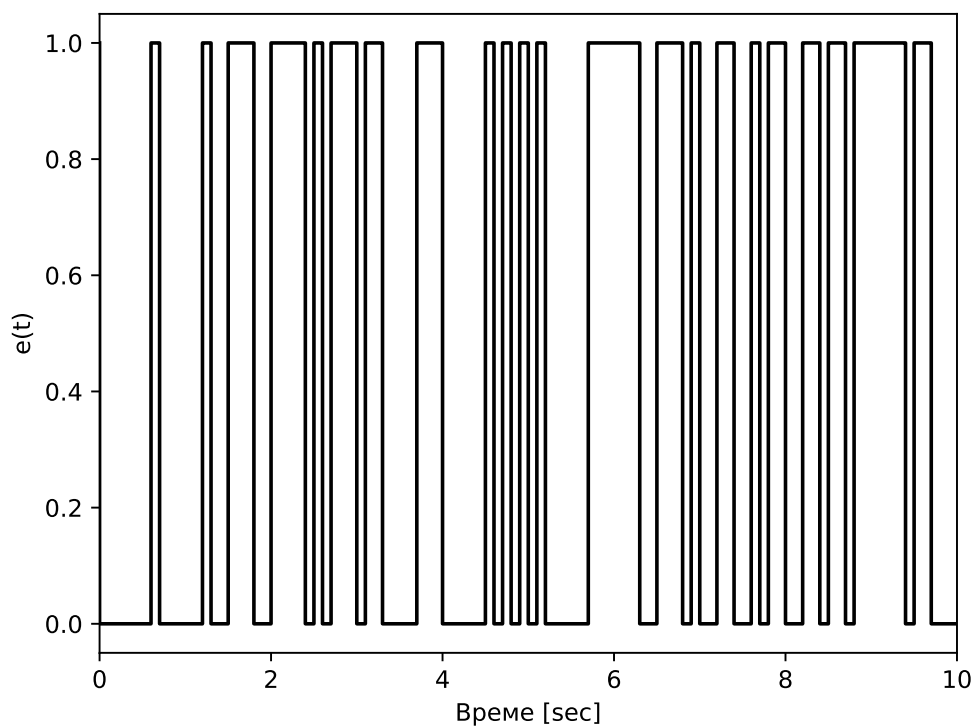
Техника додавања истраживачког шума је широко коришћена у методама машинског учења за побољшање перформанси учења [100] и адаптивног управљања како би се обезбедио услов перзистентне побуде. Као што ћемо видети у наставку, претпоставка да једначине АДП алгоритма, које генеришу  $P_k$  и  $K_k$  (2.47)-(2.48), нису дегенеративне за довољно велики индекс перформанси 2.51, је неопходна, да би се обезбедила конвергенција алгоритама, односно добила јединствена решења  $P_k$  и  $K_k$ . Након добијања  $K_k$  управљачки улаз 2.53 је једнозначно одређен.

Избор шума истраживања није тривијалан задатак за опште проблеме учења са поткрепљивањем и друге повезане проблеме машинског учења, посебно за високодимензионалне системе.

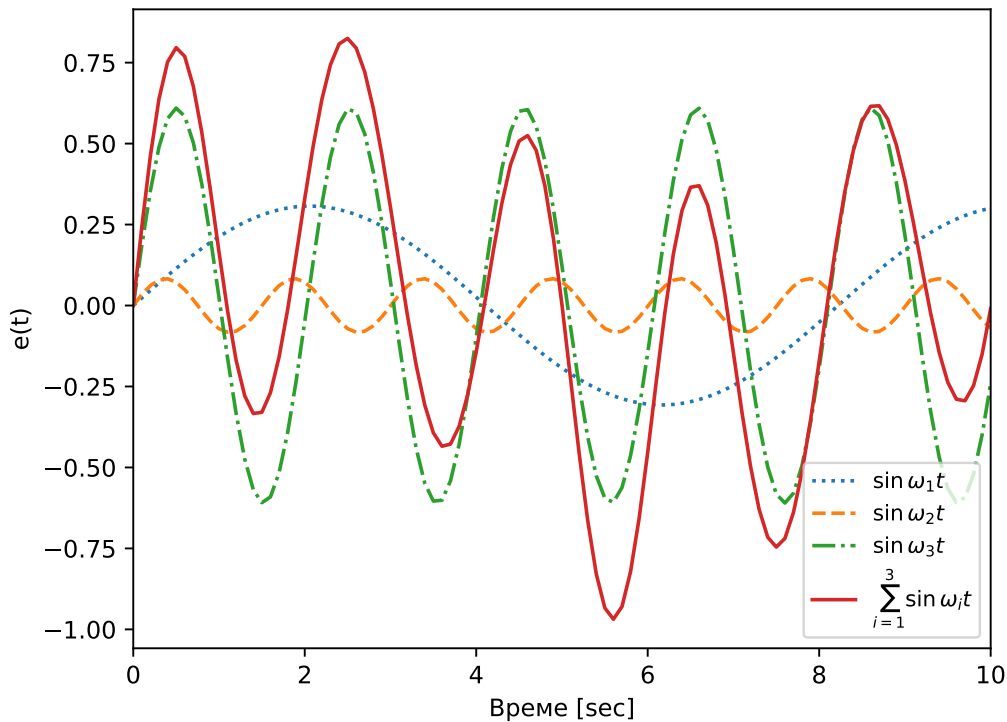
Примери истраживачког шума су бели шум (енг. white noise) [15, 97, 98], псеудо-случајна поворка бинарних сигнала (енг. Pseudo Binary Random Signal), као и сума синусних функција различитих фреквенција [101–103]. Примере ових истраживачких шума можемо видети на сликама 2.9, 2.10 и 2.11.



Слика 2.9: Бели шум.



Слика 2.10: Псеудо-случајни бинарни сигнал.



Слика 2.11: Сума синусоида.

У наставку тезе биће коришћен истраживачки шум облика суме синусоида

$$e_k = \sum_{i=1}^n \sin(\omega_i t),$$

који је најпогоднији за практичну примену.

Шум истраживања представља ограничену временски променљиву функцију, која је потребна у фази учења да би се гарантовала конвергенција АДП алгорита за системе са непознатом динамиком. Тачније, потребан нам је шум истраживања да би се испунио услов ранга дат у следећој Лема 2.3.

**Лема 2.3.** Ако постоји неки позитиван цео број  $l_0 \in Z^+$  тако да  $\forall l > l_0$  важи

$$\text{rank}([\Gamma_{\tilde{z}} \quad \Gamma_{zu} \quad \Gamma_{\tilde{u}}]) = \frac{q(q+1)}{2} + \frac{m(m+1)}{2} + qm, \quad (2.54)$$

тада

$$\Psi_j = [\Gamma_{\tilde{u}} - \Gamma_{kz}, \quad 2(\Gamma_{zu} + \Gamma_{zz}(I_m \otimes \bar{K}_j^T)), \quad -\delta_{zz}], \quad (2.55)$$

има јун ранг колона за свако  $j \in Z^+$ , при чему је  $q = n(m+r)$ .

Другим речима, ако је испуњен услов ранга 2.54, матрична једначина

$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j, \quad (2.56)$$

има јединствено решење, при чему је

$$\Phi_j = \begin{bmatrix} y_{k_0}^T Q y_{k_0} + z_{k_0}^T \bar{K}_j^T R \bar{K}_j z_{k_0} \\ y_{k_1}^T Q y_{k_1} + z_{k_1}^T \bar{K}_j^T R \bar{K}_j z_{k_1} \\ \vdots \\ y_{k_l}^T Q y_{k_l} + z_{k_l}^T \bar{K}_j^T R \bar{K}_j z_{k_l} \end{bmatrix}, \quad (2.57)$$

где су  $k_0 < k_1 < \dots < k_{l+1}$  позитивни цели бројеви.

*Доказ.* Претходна Лема се може доказати контрадикцијом. Претпоставимо да је  $W_v = [X_v \ Y_v \ Z_v]^T$  ненула решење једначине 2.56. Тада постоји  $Y_m \in \mathbb{R}^{m \times q}$  и симетричне матрице  $X_m \in \mathbb{R}^{m \times m}$  и  $Z_m \in \mathbb{R}^{n \times n}$  такве да важи

$$\begin{aligned} X_v &= \text{vecs}(X_m), \\ Y_v &= \text{vec}(Y_m), \\ Z_v &= \text{vecs}(\Theta^T Z_m \Theta). \end{aligned} \quad (2.58)$$

Користећи релације 2.49 и 2.42 важи

$$0 = \Psi_j W_v, \quad (2.59)$$

што имплицира

$$0 = \Gamma_{zz} \text{vec}(\Omega_1) - 2\Gamma_{zu} \text{vec}(\Omega_2) - \Gamma_{\tilde{u}} \text{vecs}(\Omega_3), \quad (2.60)$$

при чему је

$$\begin{aligned} \Omega_1 &= \Theta^T (A_j^T Z_m A_j - Z_m + C^T Q C + K_j^T R K_j) \Theta + \\ &+ \bar{K}_j^T (X_m - B^T Z_m B) \bar{K}_j - \bar{K}_j^T (Y_m - B^T Z_m A \Theta) - (Y_m - B^T Z_m A_d \Theta) \bar{K}_j, \end{aligned} \quad (2.61)$$

$$\Omega_2 = Y_m - B^T Z_m A \Theta, \quad (2.62)$$

$$\Omega_3 = X_m - B^T Z_m B. \quad (2.63)$$

Ако је  $\Omega_1$  симетрична матрица налази се да је

$$\Gamma_{zz} \text{vec}(\Omega_1) = \Gamma_{\tilde{z}} \text{vecs}(\Omega_1). \quad (2.64)$$

Користећи релацију 2.60 може се показати да важи

$$[\Gamma_{\tilde{z}} \quad -2\Gamma_{zu} \quad -\Gamma_{\tilde{u}}] \begin{bmatrix} \text{vecs}(\Omega_1) \\ \text{vec}(\Omega_2) \\ \text{vecs}(\Omega_3) \end{bmatrix} = 0. \quad (2.65)$$

Под претпоставком услова пуног ранга 2.54, јединствено решење једначине 2.65 је

$$\begin{bmatrix} \text{vecs}(\Omega_1) \\ \text{vec}(\Omega_2) \\ \text{vecs}(\Omega_3) \end{bmatrix} = 0. \quad (2.66)$$

На основу 2.66 релације 2.62 и 2.63 постају

$$Y_m = B^T Z_m A \Theta, \quad (2.67)$$

$$X_m = B^T Z_m B. \quad (2.68)$$

Пошто  $\Theta$  има пун ранг врста,  $Z_m$  је јединствено решење једначине

$$A_j^T Z_m A_j - Z_m + C^T Q C + K_j^T R K_j = 0, \quad (2.69)$$

што представља Љапуновљеву једначину. Коначно, имамо  $W_v = 0$  што је у контрадикторности са претпоставком. Тиме је доказ завршен.  $\square$

Дакле, шум истраживања је потребан у фази учења да би се могла добити апроксимација непознатог оптималног решења, тј. да се задовољи услов пуног ранга 2.54.

Да би се обезбедила јединственост решења матричне једначине 2.56,  $\forall k = 0, 1, 2, \dots$  постоји довољно велики цео број  $l_k > 0, l_k \in \mathbb{Z}^+$ , тако да  $\Psi_j$  има пун ранг колона 2.54.

Сваки интервал  $[t_{k,j}, t_{k,j+1}]$  се назива интервал узорковања. Морамо да прикупимо довољно узоркованих података (што значи довољно велики  $l_k > 0, l_k \in \mathbb{Z}^+$  за сваки корак  $k$  итерације. Избор шума истраживања игра кључну улогу.

Да би се задовољио услов ранга 2.54, добра пракса је да се обезбеди да сваки корак итерације може да користи податке из најмање двоструко већег броја интервала узорковања као непознатих, тј.  $l_k \geq q(q+1) + m(m+1) + 2qm$  за  $\forall k = 0, 1, 2, \dots$ . Поред тога, ако је истраживачки шум периодична функција, онда период одабирања сигнала мора бити довољно већи од периода истраживачког шума.

Приметимо да је услов ранга 2.54 у духу перзистенције побудних сигнала (ПЕ) у адаптивном управљању [104], што је неопходан услов за конвергенцију параметара.

**Напомена 2.2.** У пракси може доћи до нумеричке грешке приликом израчунавања матрица 2.44, услед чега решење 2.56 можда неће постојати. У том случају, решење матричне једначине 2.56 се може посматрати као решење најмањих квадрата 2.58.

Може се видети да се алгоритам заснован на АДП-у састоји из две одвојене фазе:

- Прво, почетни стабилишући закон управљања се примењује заједно са додатим шумом истраживања, у циљу прикупљања улазно/излазних информација које се складиште у матрицама 2.44, све док је задовољен услов ранга 2.54.
- Друго, без захтева за додатним информацијама о систему, матрице 2.44 се користе за имплементацију итеративног процеса. На тај начин се добија секвенца регулатора који конвергирају ка оптималном закону управљања.

Конвергенција алгоритма управљања заснованог на АДП-у дата је у форми Теореме 2.1.

**Теорема 2.1.** Ако је услов из Леме 2.3 задовољен, са иницијалним стабилишућом матрицом појачања  $\bar{K}_0$ , секвенце  $\{\bar{P}_j\}_{j=0}^{\infty}$  и  $\{\bar{K}_j\}_{j=0}^{\infty}$  добијене на основу овог алгоритма конвертирају ка њиховим оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ , респективно [16, 93]

*Доказ.* Ако  $P_j = P_j^T$  представља решење (2.11), за стабилну матрицу појачања  $\bar{K}_j$ , онда  $K_{j+1}$  се јединствено добија из (2.20). Лако се може показати да  $\bar{P}_j$  и  $\bar{K}_{j+1}$  задовољавају (2.47) и (2.48). Постављамо  $\bar{P}$  и  $\bar{K}$  као решења (2.47) и (2.48), а Лема 2.3 обезбеђује да су  $\bar{P}_j = \bar{P}$  и  $\bar{K}_{j+1} = \bar{K}$  јединствено изражене. Такође, из Леме 2.1, имамо да је  $\lim_{j \rightarrow \infty} \bar{K}_j = \bar{K}^*$  и  $\lim_{j \rightarrow \infty} \bar{P}_j = \bar{P}^*$ . Овим је конвергенција доказана.  $\square$

### 2.6.1 Перзистентна побуда

**Дефиниција 2.1** (Перзистентна побуда). *Стохастички вектор  $\beta_k \in \mathbb{R}^p$  је перзистентна побуда, ако постоје позитивне константе  $\delta$ ,  $\alpha$  и  $k_0 \geq 1$ , тако да је*

$$\sum_{k=k_0}^{k_0+\delta} E [\beta_k \beta_k^T] \geq \alpha I, \quad (2.70)$$

где је  $I$  јединична матрица, а  $E \{\cdot\}$  представља математичко очекивање.

**Лема 2.4.** *Вектор управљања са истраживачким шумом*

$$u_k = -Kx_k + e_k, \quad (2.71)$$

задовољава услов перзистентне побуде.

*Доказ.* Вектор истраживачког (тј. побудног) шума који се додаје вектору управљања у фази учења је:

$$e_k^{pe} = [e_k^{pe}(1) \ e_k^{pe}(2) \ \dots \ e_k^{pe}(m)]^T \in \mathbb{R}^{m \times 1}, \quad (2.72)$$

где је  $m$  број улаза система, док је  $l$ -ти скаларни елемент вектора колоне означен са  $e_k^{pe}(l)$ , и представља суму синусних функција различитих фреквенција

$$e_k^{pe}(l) = \sum_{j=1}^n \sin(\omega_{lj}k), \quad l = \overline{1, m}, \quad j = \overline{1, n}. \quad (2.73)$$

За дату позитивну константну  $\delta$  и  $k_0 \geq 1$ , вектор истраживачког шума се може додати вектору управљања система  $-Kx_k$  као што следи:

$$u_k = -Kx_k + e_k^{pe}, \quad (2.74)$$

где је  $k = k_0, \dots, k_0 + \delta, \dots$

Сагласно дефиницији 2.74,  $E [u_k u_k^T] \ \forall k = k_0, \dots, k_0 + \delta, \dots$ , задовољава неједнакост

$$\sum_{k=k_0}^{k_0+\delta} \begin{bmatrix} E[u_k(1)u_k^T(1)] & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & E[u_k(m)u_k^T(m)] \end{bmatrix} \geq \sum_{k=k_0}^{k_0+\delta} \begin{bmatrix} \sigma_k^2(1)I & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_k^2(m)I \end{bmatrix} \geq \sum_{k=k_0}^{k_0+\delta} \alpha_k^2 I \geq \alpha I, \quad (2.75)$$

где је  $\alpha_k = \min \{\sigma_k(1), \sigma_k(2), \dots, \sigma_k(m)\}$ , а  $\sigma_k$  стандардна девијација истраживачког шума. Другим речима,

$$\sum_{k=k_0}^{k_0+\delta} E [u_k u_k^T] \geq \alpha I, \quad (2.76)$$

при чему је  $\alpha = \sum_{k=k_0}^{k_0+\delta} \alpha_k^2$ . Дакле, када се истраживачки шум дода вектору управљања  $-Kx_k$ , задовољен је услов перзистентне побуде 2.70.  $\square$

## 2.6.2 Оптимални истраживачки шум

Да би смо дошли до оптималног истраживачког шума можемо поћи од процедуре планирања експеримента, која се користи у идентификацији динамичких система.

У претходном делу је показано да би фаза учења могла бити спроведена, мора се задовољити услов да истраживачки шум буде перзистентна побуда. Овај део је посвећен пројектовању истраживачког шума тако да излаз система у фази учења буде максимално информативан.

Први бенефит пројектовања оптималног улаза је смањење времена извршавања експеримента у поређењу са стандардним побудним сигналимa као и у редуковању побудних сигнала.

Пошто процес учења може бити дуготрајан, поготово за системе великог реда, употреба оптималних улазних сигнала ће омогућити да се то време смањи.

Такође, за управљање механичких система, у многим случајевима није проблем трајање експеримента. Уместо тога, оптимални улазни сигнали могу остварити уштеде у нивоу улазне побуде.

Општа формулација проблема пројектовања улаза

$$\min_{\Phi_u} \text{функција циља} \quad (2.77)$$

при датим ограничењима. Дакле, сви проблеми пројектовања улаза могу бити формулисани као оптимизациони проблем који укључује нека ограничења квалитета модела као и ограничења сигнала. Квалитативна ограничења су обично функција асимптотске коваријационе матрице  $\mathcal{P}$ , па је природније користити улазни спектар  $\Phi_u(\omega)$  за променљиве пројектовања улаза. Ограничења сигнала морају бити укључена да би се добио добро постављен проблем, тј. да би се избегло коришћење бесконачне улазне снаге.

Класични прилаз би био:

1. Минимизирати неку скаларну функцију циља при датим ограничењима. Примери обично коришћених критеријума су:

$$\begin{aligned} A - \text{оптимални критеријум} &: \min \text{tr}(\mathcal{P}), \\ E - \text{оптимални критеријум} &: \lambda_{\max}(\mathcal{P}), \\ D - \text{оптимални критеријум} &: \det(\mathcal{P}), \\ L - \text{оптимални критеријум} &: \text{tr}(W\mathcal{P}). \end{aligned} \quad (2.78)$$

2. Ограничења улазног сигнала

- Позитиван улазни спектар

$$\Phi_u(\omega) \geq 0. \quad (2.79)$$

- Ограничење улазне енергије је  $\alpha = \text{const}$ .

$$r_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \leq \alpha. \quad (2.80)$$

Општа формулација пројектовања улаза у свом првобитном облику је нерешива из више разлога:

- Ограничења су обично неконвексна и такви оптимизациони проблеми могу бити тешки за решавање
- Ограничења у многим случајевима су бесконачно димензионална, што позива на посебну пажњу при преузимању поступка оптимизације
- Такође се јавља проблем проналажења реализације улазног сигнала који има жељене спектралне особине. Овај проблем се назива спектрална факторизација. Дакле, употребљен алгоритам пројектовања улаза мора садржати и други корак који извршава факторизацију улазног спектра.
- Асимптотска коваријациона матрица обично зависи од тачних вредности параметара  $\mathcal{P} = \mathcal{P}(\theta)$ , који су непознати.

Прве три потешкоће, које су горе излистане, могу бити решене увођењем коначно димензионалне параметризације улазног спектра и евентуално крос-спектра  $\Phi_{u\varepsilon}(\omega)$ . Кључни алат у пројектовању улазног сигнала је увођење линеарне параметризације спектра. Са оваквом параметризацијом један број проблема оптималног пројектовања улаза може се формулисати као коначно димензионални програм конвексне оптимизације. Претпостављено је да параметри тачног система у датом коначном скупу параметара, док се најгори случај перформанси оптимизира на овом скупу. Пошто се улазни спектар састоји од реалног и комплексног дела, посматра се само реални позитивни део. Последња потешкоћа да оптимално решење зависи од карактера система који се идентификује је својствена за све проблеме пројектовања улазног сигнала. То је Ахилова пета у пројектовању улаза, јер решење зависи од тачног система, па се овај проблем може решити уз помоћ итеративне процедуре где је пројектовање улаза засновано на моделу система. Користи од оптималног пројектовања улаза су квантификоване у општим индустријским применама. Резултати показују да се време трајање експеримента може значајно скратити као и да се улазна снага може смањити. У реалним применама, ова чињеница се мора превазићи. У овом тренутку постоји веома мало систематских начина за то.

У литератури се често користи истраживачки шум следеће форме

$$e = \sum_{j=1}^n \sin(w_j t), \quad (2.81)$$

где су  $w_i$  случајно изабране фреквенције на неком посматраном интервалу времена. У наставку ће бити одређен минимални број синусоида који је довољан за генерисање оптималног истраживачког шума.

### 2.6.3 Минималан број синусоида

Посматрајмо линеарни дискретни систем у форми

$$y(k) = G_1(q)u(k) + G_2(q)e(k), \quad (2.82)$$

са стабилним рационалним преносним функцијама

$$G_1(z) = \frac{B(q)}{A(q)} = \frac{b_1 q^{-1} + \dots + b_n q^{-n}}{1 + a_1 q^{-1} + \dots + a_n q^{-n}}, \quad (2.83)$$

$$G_2(q) = \frac{D(q)}{C(q)} = \frac{d_1 q^{-1} + \dots + d_p q^{-p}}{1 + c_1 q^{-1} + \dots + c_p q^{-p}}, \quad (2.84)$$



при чему су полиномски парови  $(A, B)$  и  $(C, D)$  узајамно прости. Овде су  $\{u(k)\}$  и  $\{y(k)\}$  низови улаза и излаза, респективно, док је  $\{e(k)\}$  низ међусобно независних идентично расподељених случајних величина (енг. Independent identically distributed random variables) са Гаусовом расподелом  $\mathcal{N} \sim (0, \Sigma)$ , док је  $q^{-i}$  оператор помераја тј.  $q^{-i}s(k) = s(k - i)$ .

Комплексност генерисања оптималних улазних сигнала у временском домену је један од разлога који је мотивисао истраживаче да дефинишу овај проблем у фреквентном домену [67]. Зато ћемо увести Фишерову матрицу информација, која представља инверзну коваријациону матрицу тачних вредности параметара:

$$\mathcal{M} = \begin{bmatrix} \mathcal{M}_u & 0 \\ 0 & \mathcal{M}_c \end{bmatrix}, \quad (2.85)$$

где члан  $\mathcal{M}_u$  зависи од улаза и тачних вредности параметара система  $\theta$ , док је други члан Фишерове матрице  $\mathcal{M}_c$  независан од улаза. Треба нагласити да је за већи број одбирака  $l \in \mathbb{Z}^+$ , погодније радити са усредњеном матрицом информација, која се дефинише на следећи начин:

$$\bar{\mathcal{M}} = \lim_{N \rightarrow \infty} \frac{1}{l} \mathcal{M} = \begin{bmatrix} \bar{\mathcal{M}}_u & 0 \\ 0 & \bar{\mathcal{M}}_c \end{bmatrix}, \quad (2.86)$$

где је

$$\bar{\mathcal{M}}_u(\theta, \Phi_u) = \frac{1}{\pi} \int_0^\pi \tilde{\mathcal{M}}(\theta, \omega) \Phi_u(e^{j\omega}) d\omega, \quad (2.87)$$

са

$$\tilde{\mathcal{M}}(\theta, \omega) \doteq \text{Re} \left\{ \frac{\partial G_1(e^{j\omega})}{\partial \beta} |G_2(e^{j\omega})|^{-2} \left[ \frac{\partial G_1(e^{j\omega})}{\partial \beta} \right]^H \right\}. \quad (2.88)$$

Овде је  $\Phi_u$  функција спектралне густине улаза, док је  $(\cdot)^H$  оператор коњугованог транспоновања. Други члан усредњене Фишерове матрице информација је независан од улаза:

$$\bar{\mathcal{M}}_c = \frac{1}{\pi} \int_0^\pi \left[ G_2^{-1}(e^{j\omega}) \frac{\partial G_2(e^{j\omega})}{\partial \beta} \right] \left[ G_2^{-1}(e^{j\omega}) \frac{\partial G_2(e^{j\omega})}{\partial \beta} \right]^H d\omega + \frac{1}{2\Sigma^2} \left( \frac{\partial \Sigma}{\partial \beta} \right)^T \frac{\partial \Sigma}{\partial \beta}, \quad (2.89)$$

при чему се вектор  $\beta$  може дефинисати на следећи начин:

$$\beta = [\theta^T \quad \Sigma]^T. \quad (2.90)$$

За генерисање оптималног улаза важан нам је само члан  $\bar{\mathcal{M}}_u(\theta, \Phi_u)$ , јер зависи од улазног спектра  $\Phi_u(\omega)$  и тачних вредности параметара  $\theta$ .

Посматрајући релације горе наведене релације можемо запазити следеће:

- Члан  $\bar{\mathcal{M}}_u$  је афина функција улазног спектра  $\Phi_u(\omega)$ , па можемо користити  $\Phi_u(\omega)$  да обликујемо Фишерову матрицу информација  $(\Phi_u(\omega), \theta_0)$ .
- Са повећањем улазног спектра  $\Phi_u(\omega)$  биће већа Фишерова матрица информација  $\mathcal{M}$ , што значи да коришћење веће улазне снаге условљава да ће излаз система носити више информација о систему.

- Фишера матрица информација  $\mathcal{M}$  зависи од тачних вредности параметара система. Ова чињеница ће бити Ахилова пета у синтези оптималног истраживачког шума.

Чињеница да се уз помоћ улазног сигнала  $u$ , односно улазног спектра  $\Phi_u(\omega)$ , може обликовати Фишера матрица информација је веома важна из перспективе пројектовања оптималних улаза. Улазни сигнали су дакле, ограничене снаге и оптимални су у смислу максимизирања информације о систему.

**Теорема 2.2.** *За систем описан релацијом 2.82, довољан услов да средња матрица информација  $\bar{\mathcal{M}}$  буде регуларна је да улаз садржи  $n$  различитих синусоидних компоненти на интервалу  $(0, \pi)$ .*

*Доказ.* Претпоставимо да су  $n$  различитих фреквенција улаза  $\omega_1, \omega_2, \dots, \omega_n$  и да  $\lambda_i, i = \overline{1, n}$  представља део укупне улазне снаге за  $i$ -ту фреквенцију. Према претпоставци имамо

$$0 < \omega_1 < \omega_2 < \dots < \omega_n < \pi, \quad \lambda_i \neq 0, \quad i = \overline{1, n}. \quad (2.91)$$

Напоменимо да за посматрани систем 2.82 информациона матрица има структуру 2.86, при чему је  $\bar{\mathcal{M}}_c$  несингуларна константа матрица (независна од улаза), а  $\bar{\mathcal{M}}_u$  се може представити као:

$$\bar{\mathcal{M}}_u = \frac{1}{\Sigma} S \left[ \sum_{i=1}^n \lambda_i \Gamma(\omega_i) \right] S^T, \quad (2.92)$$

где је

$$S = \begin{bmatrix} 1 & a_1 & \dots & a_n & 0 & \dots & 0 \\ 0 & 1 & a_1 & \dots & a_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & 0 \\ 0 & \dots & 0 & 1 & a_1 & \dots & a_n \\ 0 & -b_1 & \dots & -b_n & 0 & \dots & 0 \\ 0 & & -b_1 & \dots & -b_n & \ddots & \vdots \\ \vdots & \ddots & & \ddots & & \ddots & 0 \\ 0 & \dots & 0 & & -b_1 & & b_n \end{bmatrix}, \quad (2.93)$$

$$\Gamma(\omega_i) = H(\omega_i) \begin{bmatrix} 1 & \cos(\omega_i) & \dots & \cos((2n-1)\omega_i) \\ \cos(\omega_i) & 1 & & \\ \vdots & & \ddots & \vdots \\ \cos((2n-1)\omega_i) & \dots & & 1 \end{bmatrix}, \quad (2.94)$$

при чему је  $H(\omega_i)$  дато као

$$H(\omega_i) = \frac{C(e^{j\omega_i})C(e^{-j\omega_i})}{D(e^{j\omega_i})A^2(e^{j\omega_i})A^2(e^{-j\omega_i})D(e^{-j\omega_i})}. \quad (2.95)$$

Помоћу индукције се може показати да је детерминанта матрице информација дата као:

$$\det \bar{\mathcal{M}} = \frac{4^{n(n-1)}}{\Sigma^{2n}} \det \bar{\mathcal{M}}_c (\det S)^2 \prod_{i=1}^n [\lambda_i H(\omega_i) \sin(\omega_i)]^2 \prod_{1 \leq s < t \leq n} [\cos(\omega_s) - \cos(\omega_t)]^4. \quad (2.96)$$

Напоменимо да је  $S$  Силвестерова матрица за полиномски пар  $(A(q), B(q))$ . Пошто је према претпоставци  $(A(q), B(q))$  узајамно прост полиномски пар,  $S$  је несингуларна матрица. Дакле, може се закључити да је  $\det \bar{M}$  несингуларно пошто је  $\lambda_i \neq 0$ ,  $\sin(\omega_i) \neq 0$ ,  $\cos(\omega_i) \neq \cos(\omega_j)$ ,  $i \neq j$  за  $i = \overline{1, n}$  и  $j = \overline{1, n}$ .  $\square$

Претходна теорема показује да је могуће смањити број захтеваних синусоида на  $n$  (једнак реду система), што представља најмањи број синусоида, јер је то најмањи могући број да информативна матрица буде регуларна [17].

## 2.7 Пример

Посматрајмо континуални систем у простору стања описан једначинама

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ -4 & -13 & -4 \end{bmatrix} x(t) + \begin{bmatrix} 12 \\ 1 \\ 1 \end{bmatrix} u(t), \quad (2.97)$$

$$y(t) = [1 \quad 0.8 \quad 1] x(t). \quad (2.98)$$

За потребе симулација дискретизоваћемо систем са периодом одабирања  $h = 1$  s. Потребно је решити следећи проблем:

$$\min_{\Phi_u} (-\log \det \mathcal{M}), \quad (2.99)$$

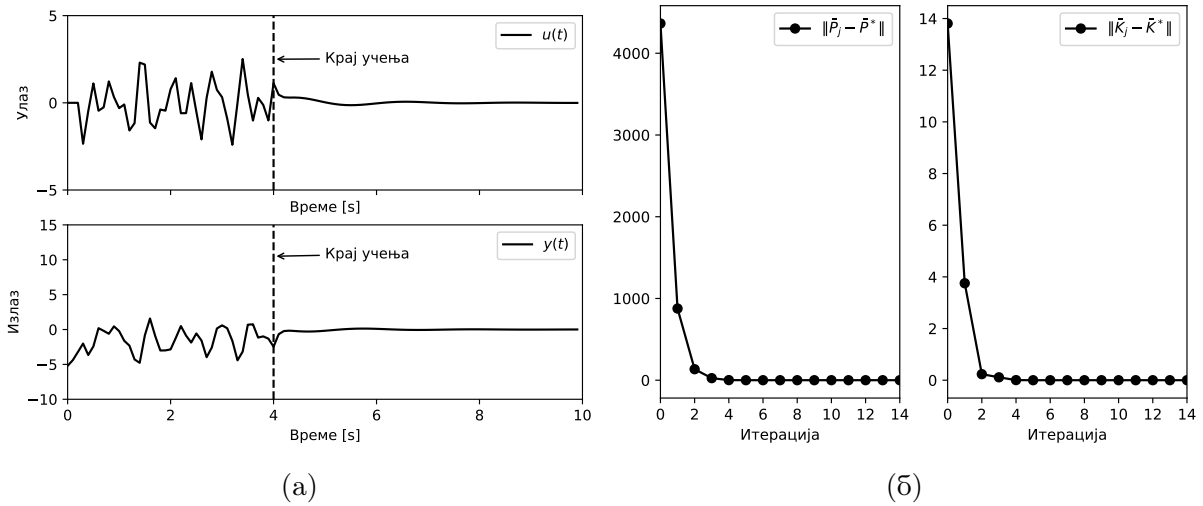
под ограничењима

$$\Phi_u(\omega) \geq 0, \quad (2.100)$$

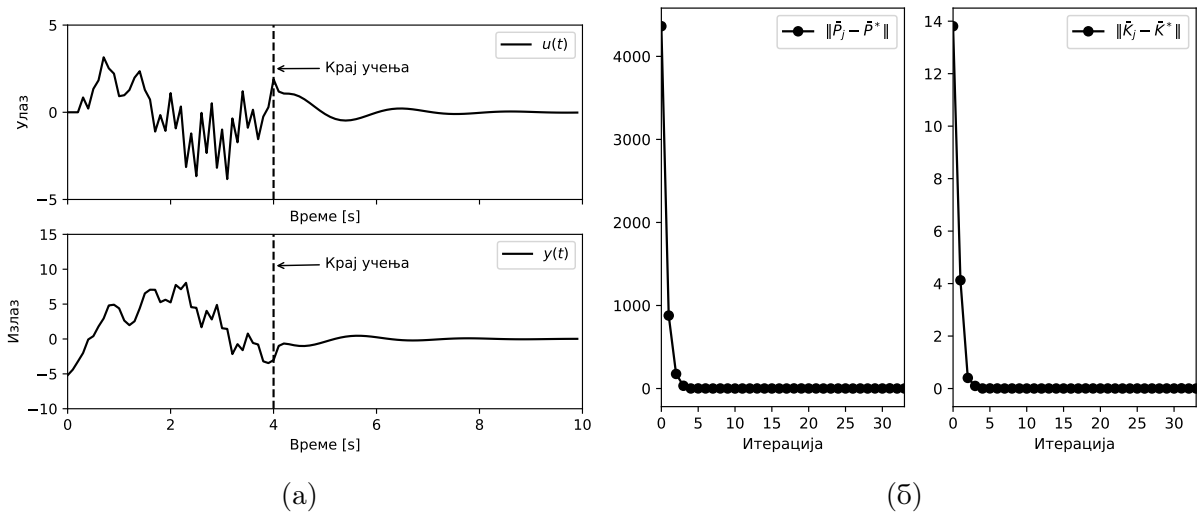
$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \leq 1. \quad (2.101)$$

Важно је напоменути да за решавање постављеног проблема оптимизације постоје врло ефикасни солвери, као што су CVX или YALMIP (за MATLAB окружење). За постављени проблем добијене су следеће фреквенције оптималног истраживачког шума  $\omega_i = [52.0083, 79.2211, 81.3458]$ .

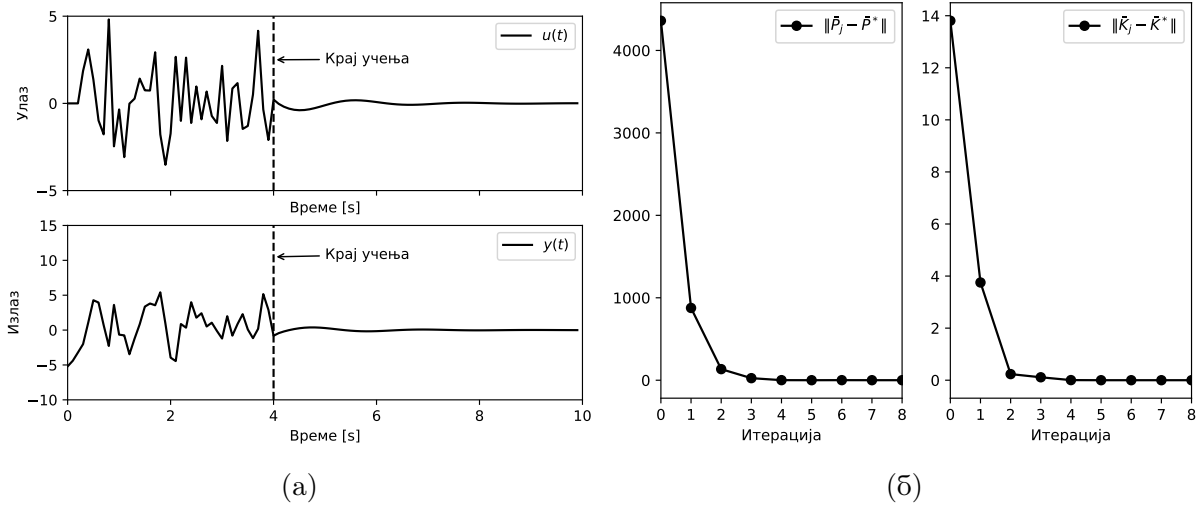
Слике 2.12, 2.13, 2.14, 2.15, 2.16, 2.17 и 2.18 приказују различите врсте истраживачког шума који се додаје управљачком сигналу током фазе учења. Може се приметити да су побудни шумови различитих фреквенција и величина.



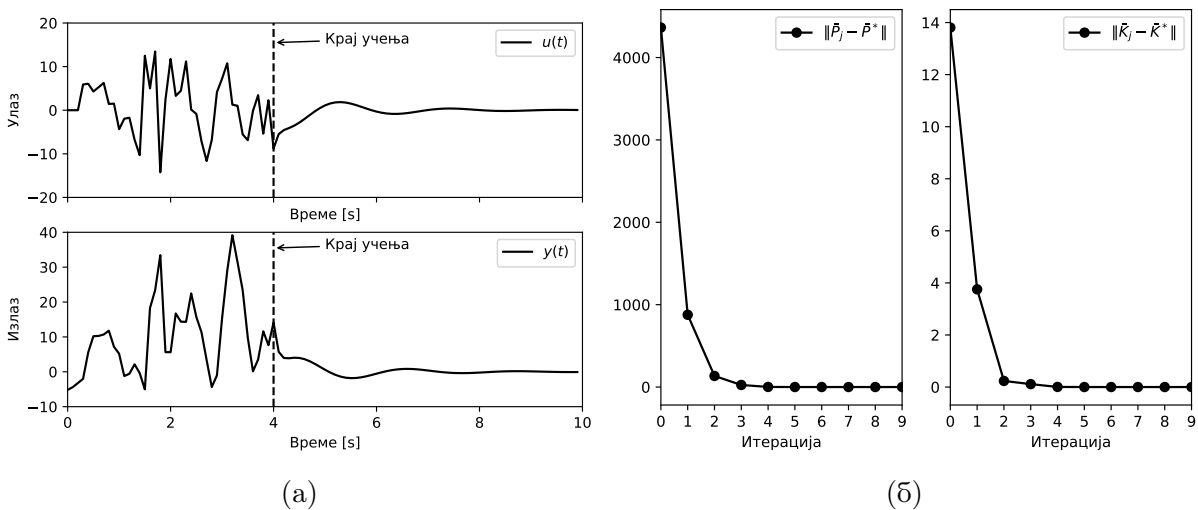
Слика 2.12: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум сума три насумичне синусоиде са учесталостима  $\omega_i \in \mathcal{U}(0, 100)$ .



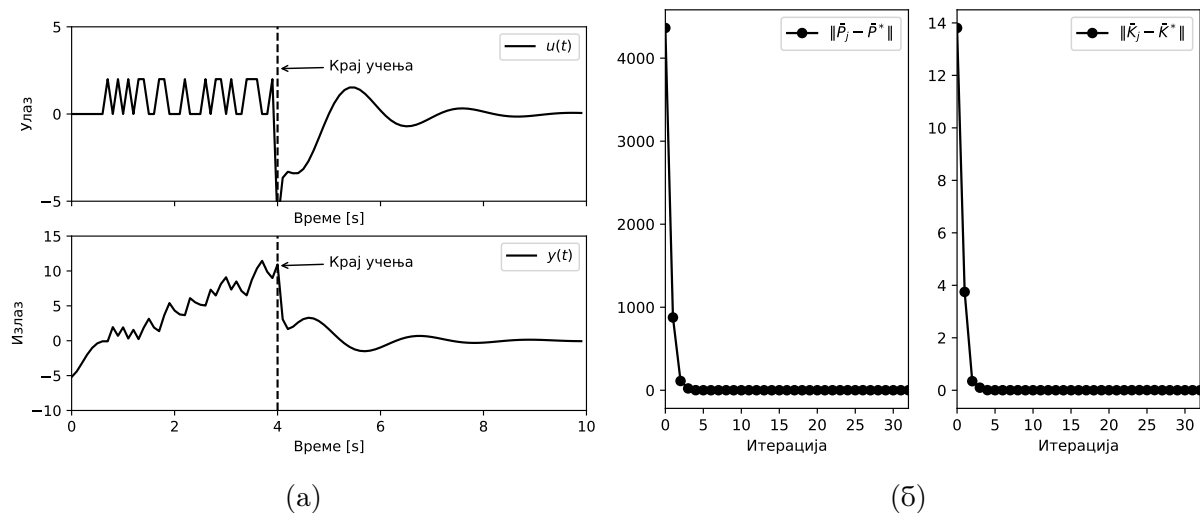
Слика 2.13: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум сума пет насумичних синусоида са учесталостима  $\omega_i \in \mathcal{U}(0, 100)$ .



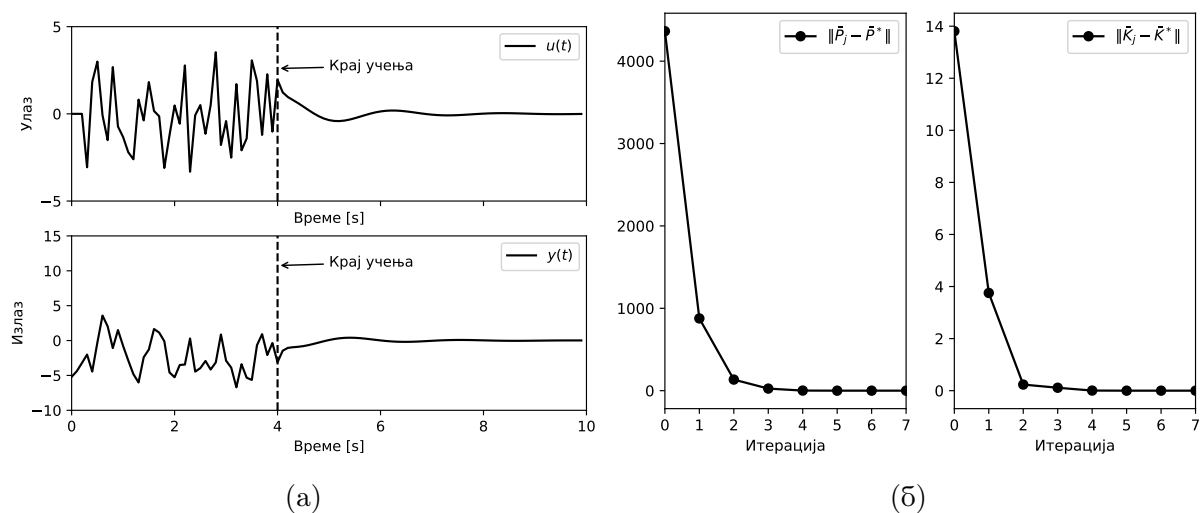
Слика 2.14: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум сума десет насумичних синусоида са учесталостима  $\omega_i \in \mathcal{U}(0, 100)$ .



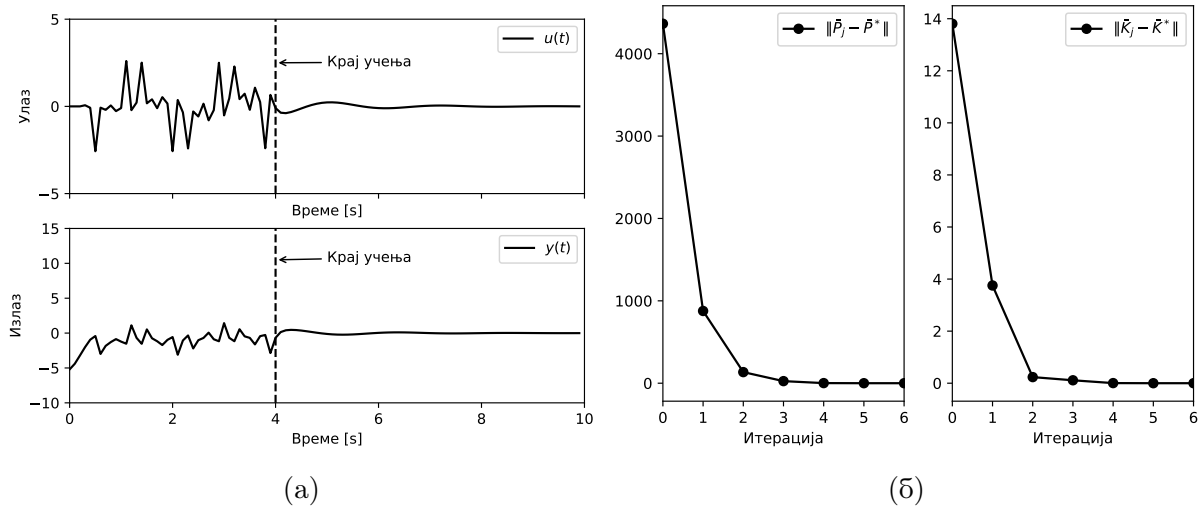
Слика 2.15: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум сума сто насумичних синусоида са учесталостима  $\omega_i \in \mathcal{U}(0, 100)$ .



Слика 2.16: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум псеудо-случајни бинарни сигнал  $e_k \in \{0, 2\}$ .



Слика 2.17: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум бели шум  $\mathcal{N}(0, 10)$ .



Слика 2.18: (а) Улазни и излазни сигнал; (б) Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  у случају када је истраживачки шум оптимална сума синусоида са учесталостима  $\omega_i = [52.0083, 79.2211, 81.3458]$ .

Показује се да, иако различити побудни шумови резултирају различитим временом конвергенције, научно појачање регулатора може конвергирати до теоријске оптималне вредности у свим случајевима.

У случају када је истраживачки шум оптимална сума синусоида са учесталостима  $\omega_i = [52.0083, 79.2211, 81.3458]$  добијене су апроксимирани вредности оптималних матрица  $\bar{P}$  и  $\bar{K}$

$$\bar{P} = \begin{bmatrix} 63.9811 & -90.9007 & 39.1155 & 4.5478 & -68.0606 & 63.9769 \\ -90.9007 & 132.3579 & -57.8435 & -4.9060 & 97.2702 & -94.6070 \\ 39.1155 & -57.8435 & 25.6239 & 1.9158 & -41.7981 & 41.9096 \\ 4.5478 & -4.9060 & 1.9158 & 1.5999 & -4.0759 & 3.1333 \\ -68.0606 & 97.2702 & -41.7981 & -4.0759 & 72.9491 & -68.3635 \\ 63.9769 & -94.6070 & 41.9096 & 3.1333 & -68.3635 & 68.5457 \end{bmatrix}, \quad (2.102)$$

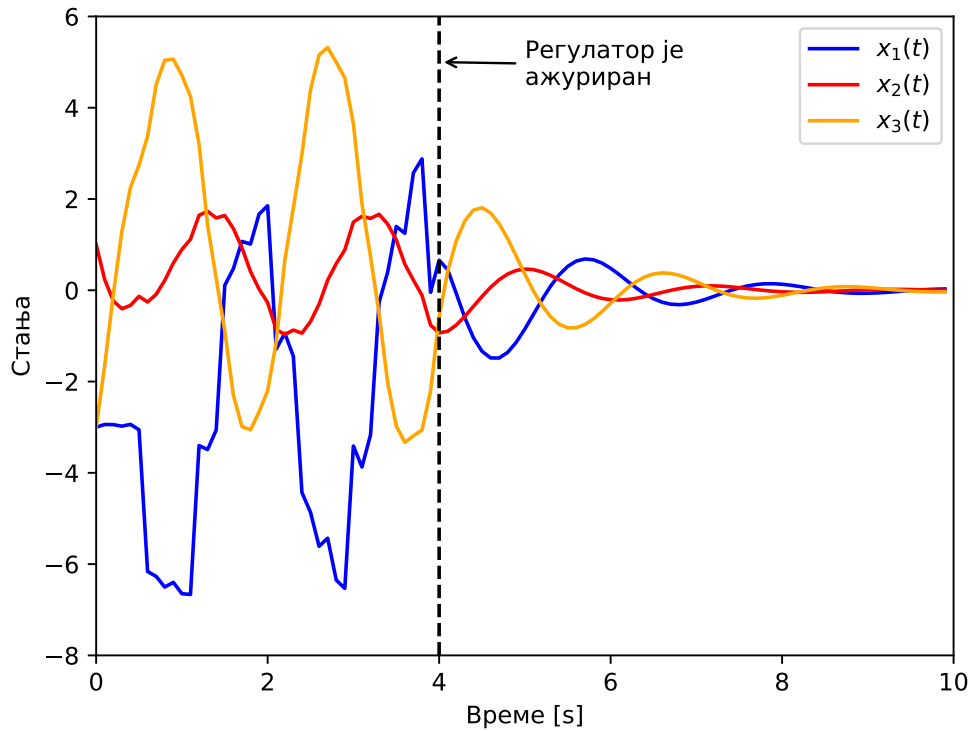
$$\bar{K} = [2.4279 \quad -3.0070 \quad 1.1725 \quad 0.3889 \quad -2.5017 \quad 1.9177], \quad (2.103)$$

док су њихове тачне вредности

$$\bar{P}^* = \begin{bmatrix} 63.9811 & -90.9004 & 39.1153 & 4.5478 & -68.0602 & 63.9760 \\ -90.9004 & 132.3571 & -57.8429 & -4.9060 & 97.2696 & -94.6063 \\ 39.1153 & -57.8429 & 25.6236 & 1.9157 & -41.7978 & 41.9093 \\ 4.5478 & -4.9060 & 1.9157 & 1.6000 & -4.0760 & 3.1332 \\ -68.0602 & 97.2696 & -41.7978 & -4.0760 & 72.9490 & -68.3634 \\ 63.9760 & -94.6063 & 41.9093 & 3.1333 & -68.3634 & 68.5457 \end{bmatrix}, \quad (2.104)$$

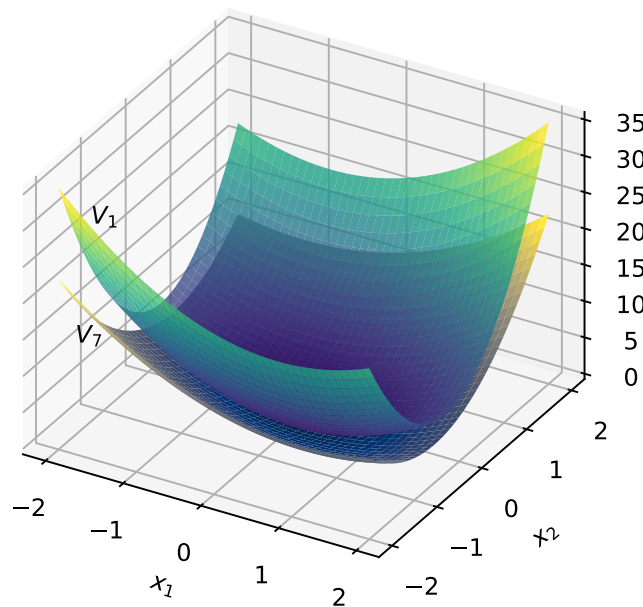
$$\bar{K}^* = [2.4279 \quad -3.0070 \quad 1.1725 \quad 0.3889 \quad -2.5017 \quad 1.9177]. \quad (2.105)$$

На слици 2.19 приказана су стања система управљаног регулатором заснованим на АДП-у са повратном спрегом по излазу.



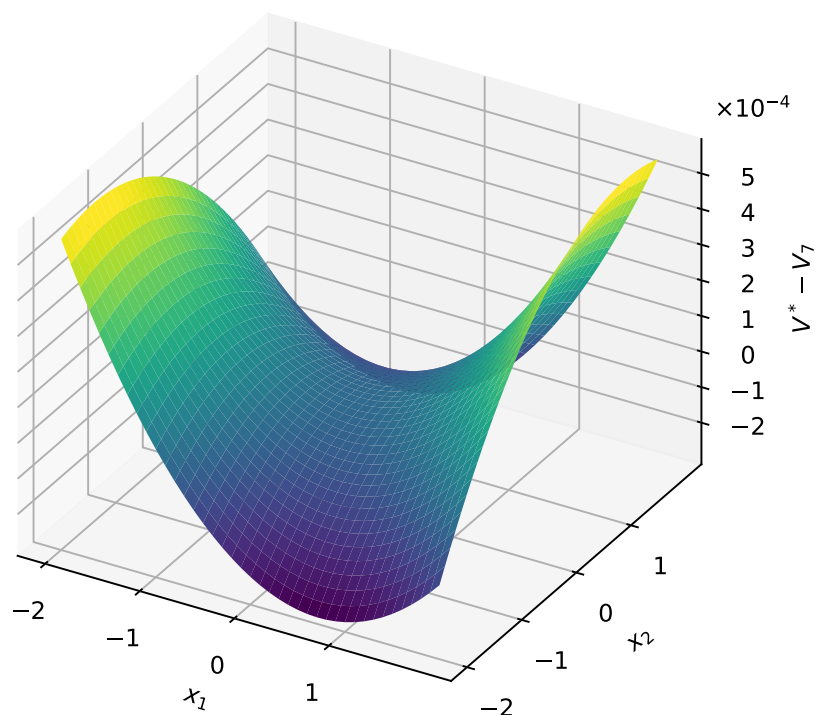
Слика 2.19: Стања система управљаног регулатором заснованим на АДП-у.

Поређење апроксимираних вредносних функција након седме итерације,  $V_7$ , и вредносне функције након прве итерације,  $V_1$ , је приказано на слици 2.20. Разлика оптималне вредносне функције,  $V^*$ , и субоптималне вредносне функције,  $V_7$ , се може видети на слици 2.21.



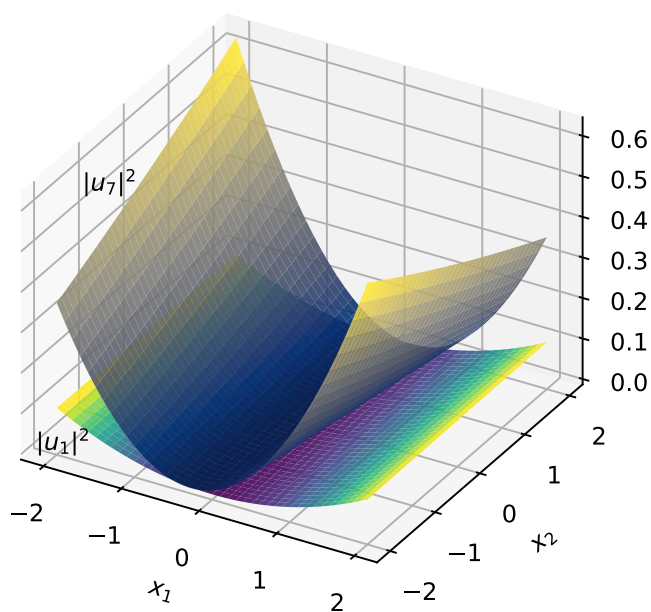
Слика 2.20: Вредносна функција за прву и последњу итерацију фазе учења.



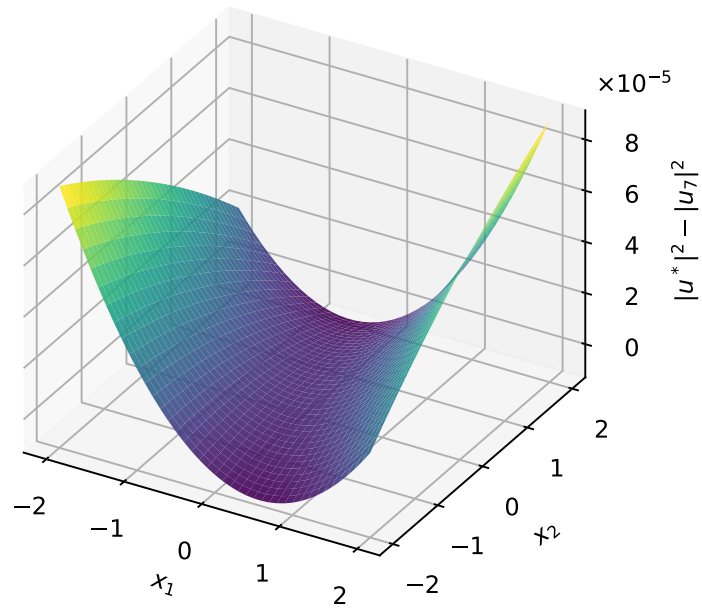


Слика 2.21: Разлика између оптималне и апроксимираних вредносних функција.

Слика 2.22 приказује поређење апроксимираних закона управљања након седме итерације,  $u_7$ , и апроксимираних закона управљања након прве итерације  $u_1$ . На слици 2.23 је приказана разлика оптималног закона управљања,  $u^*$ , и субоптималног закона управљања,  $u_7$ .



Слика 2.22: Закон управљања за прву и последњу итерацију фазе учења.



Слика 2.23: Разлика између оптималног и апроксимираног закона управљања.

## Глава 3

# Управљање засновано на догађајима

Имплементација АДП алгоритама обично се заснива на периодичном одабирању [16]. У циљу чувања ограничених комуникационих и рачунарских ресурса, код алгоритама заснованих на АДП-у се у задње време користе стратегије покренуте догађајима [17–20]. С обзиром да се ажурирање врши само када је то неопходно (на пример, када се перформансе система погоршају) број ажурирања сигнала управљања је мањи у односу на регулатор са периодичним ажурирањем. Имплементација алгоритама заснованог на догађајима заснована је на аperiodичном одабирању. Неколико регулатора заснованих на догађајима је предложено у литератури, од којих су већина регулатори са управљањем по стању [25, 105–108]. Овде ће се разматрати случај регулатора заснованог на АДП-у који је покренут догађајима када је само управљање по излазу могуће [62].

Разматра се техника учења у реалном времену, код које регулатор у току рада на основу измерених улазно-излазних података учи како да компензује непознату динамику објекта управљања, разне поремећаје и грешке моделирања, при том обезбеђујући жељене перформансе управљаног система. Оптимални закон управљања се остварује итеративно на основу излаза, реконструкције стања и АДП-а. Најпре се идентификује непознати модел објекта управљања а затим се алгебарска Рикатијева једначина решава итеративно. Да би се обезбедила конзистентност апроксимације и добила јединствена решења у свакој итерацији, потребно је увести истраживачки шум улазном сигналу, чиме се постиже перзистентна побуда [12–14]. Као истраживачки шум, обично се користи нека перзистентна побуда као што је бели шум, песудо случајни бинарни сигнал или сума синусоида. Избор истраживачког шума је нетривијалан задатак за већину проблема учења, јер утиче на прецизност решења, нарочито код сложених система [15]. Прикупљени улазни и излазни сигнали се користе за реконструкцију вектора стања модела.

Имплементација управљања заснованог на АДП-у једноставнија је у случају временски-дискретног система због лакшег прикупљања потребних податка. Методологија управљања заснована АДП-у за временски-дискретне системе предложена је у [95].

На основу измерених улазно-излазних података можемо реконструисати вектор стања дискретног модела, након чега се може имплементирати управљање засновано на АДП-у. Закон управљања се учи итеративно и врло ефикасно обезбеђује решења за оптимално управљање само на основу мерења добијених у реалном

времену. Главна предност предложене методе управљања је то што није потребно познавање динамике објекта којим се управља, што је врло важно из практичних разлога.

Ако применимо стратегију засновану на догађајима, број ажурирања сигнала управљања се смањује, с обзиром да се ажурирање врши само када је то неопходно. На тај начин се значајно чувају енергетски, рачунарски и комуникациони ресурси.

### 3.1 Линеарни дискретни систем

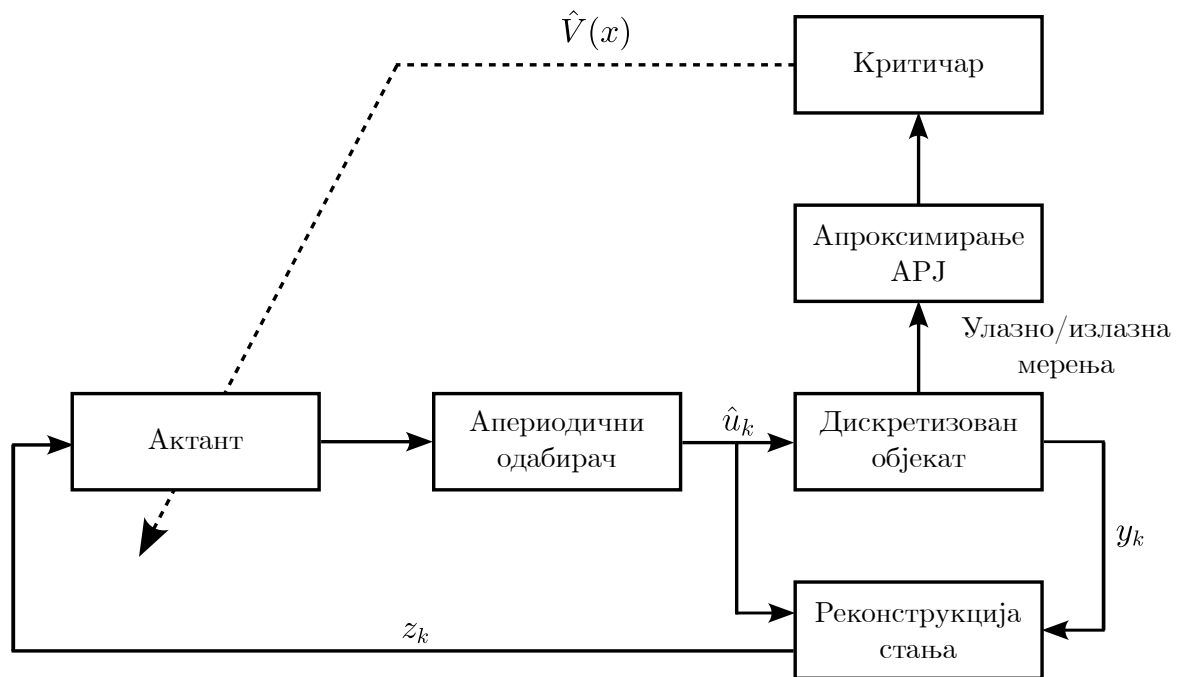
Посматрамо линеарни дискретни систем у простору стања

$$x_{k+1} = Ax_k + Bu_k, \quad (3.1)$$

$$y_k = Cx_k, \quad (3.2)$$

код кога се узорковање вектора стања  $x_k$ , улаза  $u_k$  и излаза  $y_k$  врши тренутно. Дати парови  $(A, C)$  и  $(A, Q^{\frac{1}{2}}C)$  су осмотриви, а пар  $(A, B)$  је управљив.

Као што се може видети на слици 3.1, регулатор заснован на АДП-у у случају дискретног система састоји се од три дела: реконструкција стања, критичар и актант. Реконструкција стања обезбеђује везу између улазно/излазних података и стања објекта, што нам даје могућност решавања проблема оптималног управљања објекта са непознатом динамиком. Критичар на основу улазно/излазних података оцењује перформансе тренутног закона управљања. Регулатор током рада учи како да максимизује перформансе система. На крају, актант примењује побољшан закон управљања на објекат. Број ажурирања закона управљања се услед дејства механизма покретања догађајима смањује, и самим тим се смањује и количина података која се преноси од регулатора ка објекту.



Слика 3.1: Алгоритам управљања заснован на АДП-у покренутим догађајем.

Стратегије покренуте догађајима се заснивају на апериодичном одабирању. Наравно, одабирање мора бити такво да не резултира појавом лажних учесталости.

Користићемо  $\hat{u}_k$  да означимо одабрану вредност  $u_k$ , то јест

$$\hat{u}_k = u_{k_j}, \quad k \in [k_j, k_{j+1}), \quad (3.3)$$

где је  $\{k_j\}_0^\infty$  монотono растућа секвенца одбирака, а управљачки сигнал се само ажурира у дискретним тренуцима:  $k_0, k_1, k_2, \dots$

Дефинисаћемо грешку одабирања улазних података као

$$\Delta_k = \hat{u}_k - u_k. \quad (3.4)$$

Сада једначине (3.1)-(3.2) можемо написати као

$$x_{k+1} = Ax_k + B(u_k + \Delta_k), \quad (3.5)$$

$$y_k = Cx_k. \quad (3.6)$$

Такође, индекс перформансе за дискретни систем (3.1)-(3.2) је једнак:

$$J(x_0) = \sum_{j=0}^{\infty} y_j^T Q y_j + u_j^T R u_j, \quad (3.7)$$

где су  $Q$  и  $R$  тежинске матрице. Оптимални закон управљања који минимизира индекс перформансе 3.7) има облик

$$u_k = -K^* x_k, \quad (3.8)$$

у коме оптимална матрица појачања је једнака

$$K^* = (R + B^T P^* B)^{-1} B^T P^* A, \quad (3.9)$$

а  $P^*$  представља јединствено симетрично позитивно дефинитно решење једначине

$$A^T P^* A - P^* + C^T Q C - A^T P^* B K^* = 0. \quad (3.10)$$

Пошто је једначина (3.10) нелинеарна за  $P^*$ , тешко се из ње може одредити  $P^*$  у случају комплексних система вишег реда. Ипак, до решења једначине (3.10) можемо доћи нумерички, употребом Хјуаровог алгоритма дефинисаног у претходном поглављу у облику Лему 2.1.

Пошто је Хјуарова метода заснована на моделу, не може се директно применити у случају када није позната динамика система. Да бисмо могли да применимо овај алгоритма на дискретни систем описан једначинама (3.1)-(3.2), потребно је развити алгоритам управљања по излазу заснован на АДП-у, који неће зависити од познавања матрица система.

## 3.2 Реконструкција стања

Инспирисано [94, 95], једначине (3.5)-(3.6) се могу проширити коришћењем улазно/излазних секвенци на временском хоризонту  $[k - N, k - 1]$  на следећи начин

$$\begin{aligned} x_k &= A^N x_{k-N} + V(N) \bar{u}_{k-1, k-N}, \\ \bar{y}_{k-1, k-N} &= U(N) x_{k-N} + T(N) \bar{u}_{k-1, k-N}, \end{aligned} \quad (3.11)$$

где су

$$\begin{aligned}\bar{u}_{k-1,k-N} &= [\hat{u}_{k-1}^T \quad \hat{u}_{k-2}^T \quad \cdots \quad \hat{u}_{k-N}^T]^T, \\ \bar{y}_{k-1,k-N} &= [y_{k-1}^T \quad y_{k-2}^T \quad \cdots \quad y_{k-N}^T]^T,\end{aligned}$$

вектори доступних мерења улазног и излазног сигнала, док је

$$\bar{\Delta}_{k-1,k-N} = [\Delta_{k-1}^T \quad \Delta_{k-2}^T \quad \cdots \quad \Delta_{k-N}^T]^T,$$

вектор грешака одабирања улазног сигнала. Матрице

$$\begin{aligned}V(N) &= [B \quad AB \quad \cdots \quad A^{N-1}B], \\ U(N) &= [(CA^{N-1})^T \quad (CA)^T \quad \cdots \quad C^T]^T,\end{aligned}$$

представљају матрице управљивости и осмотирвости, док је матрица Марковљевих параметара

$$T(N) = \begin{bmatrix} 0 & CB & CAB & \cdots & CA^{N-2}B \\ 0 & 0 & CB & \cdots & CA^{N-3}B \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & CB \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Број  $N$  представља индекс осмотирвости  $N = \max(\rho_u, \rho_v)$  где је  $\rho_u$  најмањи цео број који обезбеђује да је  $U(\rho_u)$  пуног ранга колона, а  $\rho_v$  је најмањи цео број који обезбеђује да је  $V(\rho_v)$  пуног ранга врста [95]. Овај индекс гарантује постојање левог инверза матрице  $U(N)$ , дефинисан као

$$U^+(N) = [U^T(N)U(N)]^{-1}U^T(N). \quad (3.12)$$

Са реконструкцијом стања у (3.11), алгоритам управљања по излазу заснован на АДП-у може се применити на решавање проблема оптималног управљања дискретног система са непознатом динамиком. Јединственост реконструкције стања је дата у форми Леме 3.1.

**Лема 3.1.** *Ако су задовољени услови осмотирвости и управљивости система описаној (3.1)-(3.2), онда се стања дискретној систему могу јединствено одредити на основу измерених улазних и излазних сигнала на следећи начин*

$$x_k = \Theta z_k \quad (3.13)$$

где је  $\Theta = [M_u \quad M_y]$  пуног ранга врсца,  $M_u = V(N) - M_y T(N)$ ,  $M_y = A^N U^+(N)$ ,  $z_k = [\bar{u}_{k-1,k-N}^T \quad \bar{y}_{k-1,k-N}^T]^T \in \mathbb{R}^q$ ,  $q = N[\dim(u) + \dim(y)]$ .

### 3.3 Алгоритам управљања заснован на АДП-у покренут догађајима

Сада, на основу (3.9)-(3.10), може се применити онлајн стратегија учења на основу повратне спреге по излазу облика  $u_k^* = -\bar{K}z_k$ , која обезбеђује субоптималну карактеристику затвореног система. Дискретни модел (3.5) се може изразити као

$$x_{k+1} = A_j x_k + B(K_j x_k + \hat{u}_k), \quad (3.14)$$

где је  $A_j = A - BK_j$ . Ако поставимо да је  $\bar{K}_j = K_j\Theta$  и  $\bar{P}_j = \Theta^T P_j \Theta$ , на основу (3.10) и (3.14) се добија једначина

$$\begin{aligned} z_{k+1}^T \bar{P}_j z_{k+1} - z_k^T \bar{P}_j z_k &= [\hat{u}_k^T \otimes \hat{u}_k^T - (z_k^T \otimes z_k^T)(\bar{K}_j^T \otimes \bar{K}_j^T)] \text{vec}(B^T \bar{P}_j B) + \\ &2 [(z_k^T \otimes z_k^T)(I_q \otimes \bar{K}_j^T) + (z_k^T \otimes u_k^T)] \text{vec}(B^T \bar{P}_j A \Theta) - \\ &- (y_k^T Q y_k + z_k^T \bar{K}_j^T R \bar{K}_j z_k), \end{aligned} \quad (3.15)$$

коју, уколико је задовољен услов (2.54), можемо решити на следећи начин

$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j, \quad (3.16)$$

где је

$$\begin{aligned} \Phi_j &= \begin{bmatrix} y_{k_0}^T Q y_{k_0} + z_{k_0}^T \bar{K}_j^T R \bar{K}_j z_{k_0} \\ y_{k_1}^T Q y_{k_1} + z_{k_1}^T \bar{K}_j^T R \bar{K}_j z_{k_1} \\ \vdots \\ y_{k_l}^T Q y_{k_l} + z_{k_l}^T \bar{K}_j^T R \bar{K}_j z_{k_l} \end{bmatrix}, \\ \Psi_j &= [\Gamma_{\bar{u}} - \Gamma_{\bar{k}z}, \quad 2(\Gamma_{zu} + \Gamma_{zz}(I_m \otimes \bar{K}_j^T)), \quad -\delta_{zz}], \\ \delta_{zz} &= [\text{vecv}(z_{k_1}) - \text{vecv}(z_{k_0}), \text{vecv}(z_{k_2}) - \text{vecv}(z_{k_1}), \dots, \text{vecv}(z_{k_{l+1}}) - \text{vecv}(z_{k_l})]^T, \\ \Gamma_{zz} &= [z_{k_1} \otimes z_{k_0}, z_{k_2} \otimes z_{k_1}, \dots, z_{k_{l+1}} \otimes z_{k_l}]^T, \\ \Gamma_{zu} &= [z_{k_1} \otimes u_{k_0}, z_{k_2} \otimes u_{k_1}, \dots, z_{k_{l+1}} \otimes u_{k_l}]^T, \\ \Gamma_{\bar{u}} &= [\text{vecv}(u_{k_0}), \text{vecv}(u_{k_1}), \dots, \text{vecv}(u_{k_l})]^T, \\ \Gamma_{\bar{k}z} &= [\text{vecv}(\bar{K}_j z_{k_0}), \bar{K}_j z_{k_1}, \dots, \bar{K}_j z_{k_l}]^T. \end{aligned}$$

Конвергенција онлајн учења закона управљања по излазу је загарантована условом који је дефинисан у форми Леме 2.3 [94]. Лема 2.3 се односи на услов перзистентне побуде у теорији адаптивног управљања који је описан у [82, 104].

Да би се задовољило услов у погледу перзистентне побуде, потребно је додати истраживачки шум  $e_k$  улазном сигналу током фазе онлајн учења. Овим ће бити задовољен услов ранга (2.54), притом не утичући на конвергенцију приликом фазе учења [15, 89, 109]. Треба напоменути да (3.15) представља евалуацију закона управљања, која се користи да би се нашло јединствено решење  $\bar{P}_j$ , док (2.48) побољшања закон управљања, који се користи да се ажурира матрица појчања  $\bar{K}_{j+1}$ . На слици 3.2, представљен је онлајн алгоритам управљања заснован на АДП-у.

Треба приметити да решавање (3.15) уместо (2.19), у потпуности елиминише почетну потребу да прецизно познавање динамике система којим се управља. Уместо тога, потребно је само мерити улазни сигнал  $u_k$  и излазни сигнал  $y_k$ . Ако погледамо израз за  $z_k$ , видећемо да закон управљања  $\hat{u}_k = -\bar{K}_k^* z_k + \Delta_k$  садржи само претходно измерене улазно-излазне податке.

**Теорема 3.1.** *Са законом управљања покренутиим дојаћајем  $\hat{u}_k$ , систем (3.14) је глобално асимптотски стабилан ако и само ако*

$$\|\Delta_k\|^2 \leq \frac{\alpha \gamma \|y_k\|^2 + \lambda_{\min}(R) \|\hat{u}_k\|^2}{\eta}. \quad (3.17)$$

где је  $\alpha \in (0, 1)$ ,  $\gamma$  позитивна константа, а  $\eta$  је позитивна константа која задовољава услов  $\eta \geq \lambda_{\max}(R + B^T \bar{P} B)$ .

*Доказ.* Дефинишимо  $A_j = A - BK$ . Сада имамо

$$x_{k+1} = A_j x_k + B \Delta_k. \quad (3.18)$$

Нека је  $V(x_k) = x_k^T \bar{P} x_k$ , тада имамо

$$\begin{aligned} V(x_{k+1}) - V(x_k) &= x_{k+1}^T \bar{P} x_{k+1} - x_k^T \bar{P} x_k = \\ &= x_k^T A_j^T \bar{P} A_j x_k + \Delta_k^T B^T \bar{P} B \Delta_k + 2 \Delta_k^T B^T \bar{P} A_j x_k - x_k^T \bar{P} x_k. \end{aligned} \quad (3.19)$$

Имајући на уму Љапуновљево једначину, важи

$$x_k^T A_j^T \bar{P} A_j x_k - x_k^T \bar{P} x_k = -y_k^T Q y_k - (K x_k)^T R K x_k. \quad (3.20)$$

Сада се (3.19), може записати као

$$\begin{aligned} V(x_{k+1}) - V(x_k) &= \\ &= -y_k^T Q y_k - (K x_k)^T R K x_k + \Delta_k^T B^T \bar{P} B \Delta_k + 2 \Delta_k^T B^T \bar{P} A_j x_k - 2 \Delta_k^T B^T \bar{P} B K x_k. \end{aligned} \quad (3.21)$$

На основу  $K = (R + B^T \bar{P} B)^{-1} B^T \bar{P} A$  важи да је  $B^T \bar{P} A = (R + B^T \bar{P} B) K$ . Сада се може добити

$$\begin{aligned} V(x_{k+1}) - V(x_k) &\leq \\ &= -y_k^T Q y_k - (K x_k)^T R K x_k + \Delta_k^T B^T \bar{P} B \Delta_k + 2 \Delta_k^T R K x_k \leq \\ &= -y_k^T Q y_k - \lambda_{\min}(R) \|\hat{u}_k\|^2 + \Delta_k^T (R + B^T \bar{P} B) \Delta_k \leq \\ &= -(1 - \alpha) \gamma \|y_k\|^2 - \alpha \gamma \|y_k\|^2 - \lambda_{\min}(R) \|\hat{u}_k\|^2 + \lambda_{\max}(R + B^T \bar{P} B) \|\Delta_k\|^2. \end{aligned} \quad (3.22)$$

Ако важи услов  $\|\Delta_k\|^2 \leq \frac{\alpha \gamma \|y_k\|^2 + \lambda_{\min}(R) \|\hat{u}_k\|^2}{\eta}$ , онда имамо

$$V(x_{k+1}) - V(x_k) \leq -(1 - \alpha) \gamma \|y_k\|^2. \quad (3.23)$$

Засновано на особини опсервабилности дискретног система (3.1)-(3.2), и директној примени Ласалеовог принципа инваријантности (енг. LaSalle's Invariance Principle) лако се може доказати глобална асимптотска стабилност решења система (3.14). Овим је доказ теореме завршен.  $\square$

Да би се обезбедило задовољење услова (3.17), дефинисан је следећи механизам покретања догађајем

$$k_{j+1} = \inf \left\{ k \in \mathbb{Z}_+ \mid k > k_j \wedge \|\Delta_k\|^2 \geq \frac{\alpha \gamma \|y_k\|^2 + \lambda_{\min}(R) \|\hat{u}_k\|^2}{\eta} \right\} \quad (3.24)$$

при чему је  $k_0 = 0$ . Као резултат тога, комуникација између регулатора и објекта управљања се редукује.

Конвергенција алгоритма управљања заснованог на АДП-у је презентована у форми Теореме 3.2. За Хурвицову матрицу  $A - BK$ ,  $K \in R^{m \times n}$  се назива стабилишућом матрицом појачања за линеарни систем  $\dot{x} = Ax + Bu$ .

**Теорема 3.2.** *Ако је услов из Леме 2.3 задовољен, са неком почетном стабилишућом матрицом појачања  $\bar{K}_0$ , секвенце  $\{\bar{P}_j\}_{j=0}^{\infty}$  и  $\{\bar{K}_j\}_{j=0}^{\infty}$  добијене помоћу овог алгоритма, конвертирају њиховим оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ , респективно [93, 94].*



*Доказ.* Ако  $P_j = P_j^T$  представља решење (2.19), за стабилишућу матрицу појачања  $\bar{K}_j$ , онда  $K_{j+1}$  се јединствено може одредити из (2.20). Лако се може показати да  $\bar{P}_j$  и  $\bar{K}_{j+1}$  задовољавају (3.15) и (2.48). Узимајући  $\bar{P}$  и  $\bar{K}$  као решења (3.15) и (2.48), Лема 2.3 обезбеђује да су  $\bar{P}_j = \bar{P}$  и  $\bar{K}_{j+1} = \bar{K}$  јединствено изражени. Такође, из Леме 2.1, имамо да  $\lim_{j \rightarrow \infty} \bar{K}_j = \bar{K}^*$ ,  $\lim_{j \rightarrow \infty} \bar{P}_j = \bar{P}^*$ . Овим је конвергенција доказана.  $\square$

---

### Алгоритам 5 Пројектовње регулатора заснованог на догађајима и АДП-у.

---

1. **Иницијализација:** Изабрати стабилишућу матрицу појачања  $K_0$  и довољно малу константу  $\varepsilon > 0$ . Подесити  $j \leftarrow 1$ .
2. **Прикупљање података:** Применити  $\hat{u}_k^j = -\bar{K}_0 z_k + \Delta_k + e_k$  на интервалу  $[k_0, k_l]$  где је  $e_k$  истраживачки шум. Формирати матрице  $\Psi_j$  и  $\Phi_j$  на основу улазно-излазних података тако да буде задовољен услов ранга.
3. **Евалуација закона управљања:** На основу израза

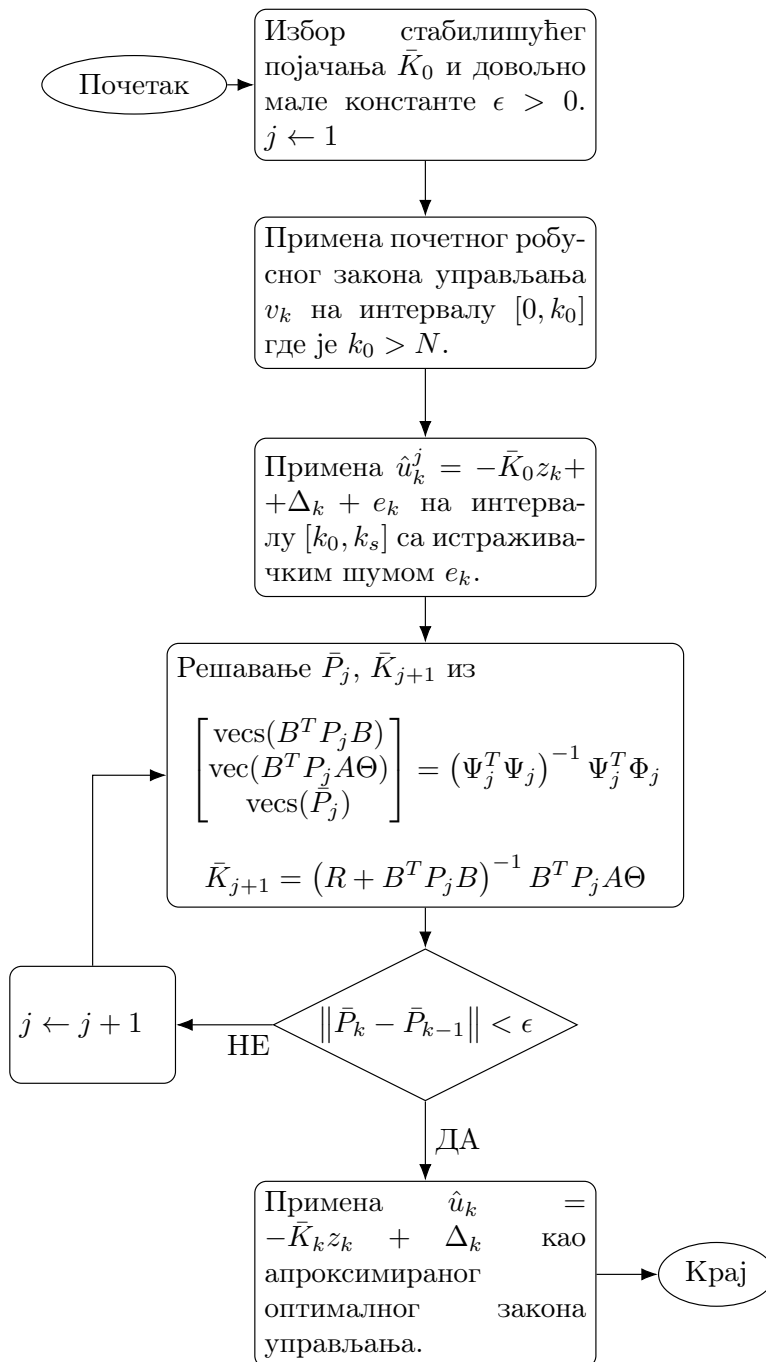
$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j,$$

одређујемо непознате матрице  $B^T P_j B$ ,  $B^T P_j A \Theta$  и  $\bar{P}_j$ .

4. **Унапређење закона управљања:** Ажурирамо закон управљања изразом

$$K_{j+1} = (R + B^T P_j B)^{-1} B^T P_j A \Theta.$$

5. **Итерације:** Постављамо  $j \leftarrow j + 1$  и понављамо корак 2 све док је задовољен услов  $\|\bar{P}_j - \bar{P}_{j-1}\| \geq \varepsilon$ .
  6. **Експлоатација:** Примена апроксимираног оптималног закона управљања  $\hat{u}_k = -\bar{K}_k z_k + \Delta_k$ .
-



Слика 3.2: Алгоритам пројектовања регулатора заснованог на догађајима и АДП-у.

### 3.4 Пример

Посматраћемо временски континуални систем описан следећим једначинама улаза и излаза

$$x_{k+1} = \begin{bmatrix} -3.5 & 1.1 & 3 \\ 1.1 & -4 & 1.5 \\ 2 & 1.3 & 3.3 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} u_k, \quad (3.25)$$

$$y_k = [8 \ 6 \ 3] x_k. \quad (3.26)$$

Дискретизацијом за период одабирања  $h = 0.1$  s добијамо временски дискретни модел

$$x_{k+1} = \begin{bmatrix} 0.7371 & 0.0943 & 0.3147 \\ 0.0903 & 0.6841 & 0.1650 \\ 0.2113 & 0.1401 & 1.4371 \end{bmatrix} x_k + \begin{bmatrix} 0.1360 \\ 0.3523 \\ 0.2774 \end{bmatrix} u_k, \quad (3.27)$$

$$y_k = [8 \quad 6 \quad 3] x_k. \quad (3.28)$$

који је коришћен у симулацијама. За потребе симулација узето је почетно стање  $x_0 = [5, -3, 3]^T$ , а за праг конвергенције усвојено је  $\varepsilon = 10^{-5}$ . За оптимални истраживачки шум је усвојена сума синусоида са фреквенцијама  $\omega_i \in [-0.138, -2.384, 4.894]$ .

Оптималне вредности матрица  $\bar{P}^*$  и  $\bar{K}^*$  израчунате на основу дискретног модела система имају вредности

$$\bar{P}^* = \begin{bmatrix} 18.8530 & -48.5527 & 22.1398 & 18.1325 & -19.2531 & 5.8661 \\ -48.5527 & 328.6091 & -179.1931 & -100.6211 & 138.7758 & -47.4784 \\ 22.1398 & -179.1931 & 99.3919 & 53.6544 & -76.1593 & 26.3346 \\ 18.1325 & -100.6211 & 53.6544 & 31.7237 & -42.1430 & 14.2161 \\ -19.2531 & 138.7758 & -76.1593 & -42.1430 & 58.7464 & -20.1789 \\ 5.8661 & -47.4784 & 26.3346 & 14.2161 & -20.1789 & 6.9775 \end{bmatrix}, \quad (3.29)$$

$$\bar{K}^* = [1.2391 \quad -4.8575 \quad 2.2649 \quad 1.6407 \quad -1.9406 \quad 0.6001]. \quad (3.30)$$

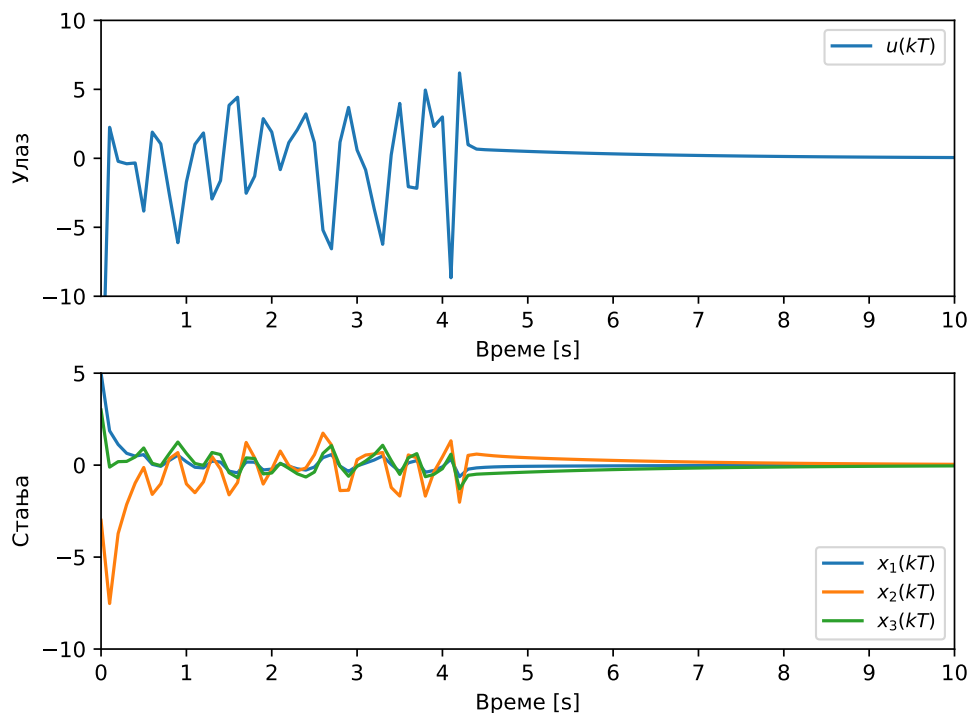
Коришћењем регулатора заснованог на АДП-у покренутог догађајима добијене су следеће вредности матрица  $\bar{P}$  и  $\bar{K}$

$$\bar{P} = \begin{bmatrix} 18.8530 & -48.5527 & 22.1398 & 18.1325 & -19.2531 & 5.8661 \\ -48.5527 & 328.6091 & -179.1931 & -100.6211 & 138.7758 & -47.4784 \\ 22.1398 & -179.1931 & 99.3919 & 53.6544 & -76.1593 & 26.3346 \\ 18.1325 & -100.6211 & 53.6544 & 31.7237 & -42.1430 & 14.2161 \\ -19.2531 & 138.7758 & -76.1593 & -42.1430 & 58.7464 & -20.1789 \\ 5.8661 & -47.4784 & 26.3346 & 14.2161 & -20.1789 & 6.9775 \end{bmatrix}, \quad (3.31)$$

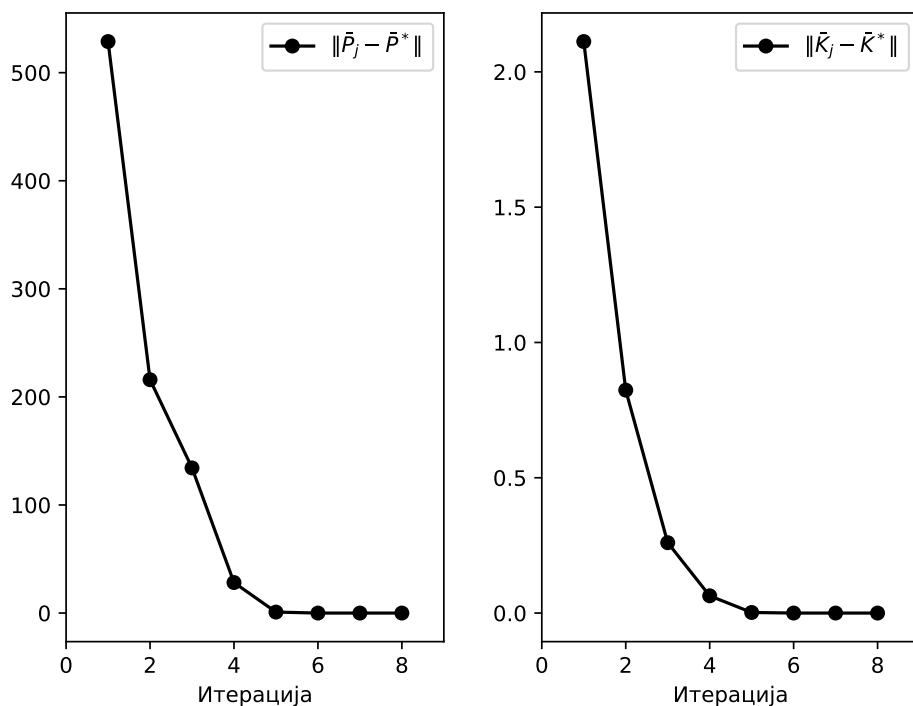
$$\bar{K} = [1.2391 \quad -4.8575 \quad 2.2649 \quad 1.6407 \quad -1.9406 \quad 0.6001]. \quad (3.32)$$

Као што се може видети, вредности оптималних и апроксимираних матрица су идентичне на тачности од четири децимале.

Слика 3.3 приказује улаз и стања система управљаног регулатором заснованим на АДП-у са периодичним одабирањем. На слици 3.4 приказана је конвергенција матрица  $\bar{P}$  и  $\bar{K}$  њиховом оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ .



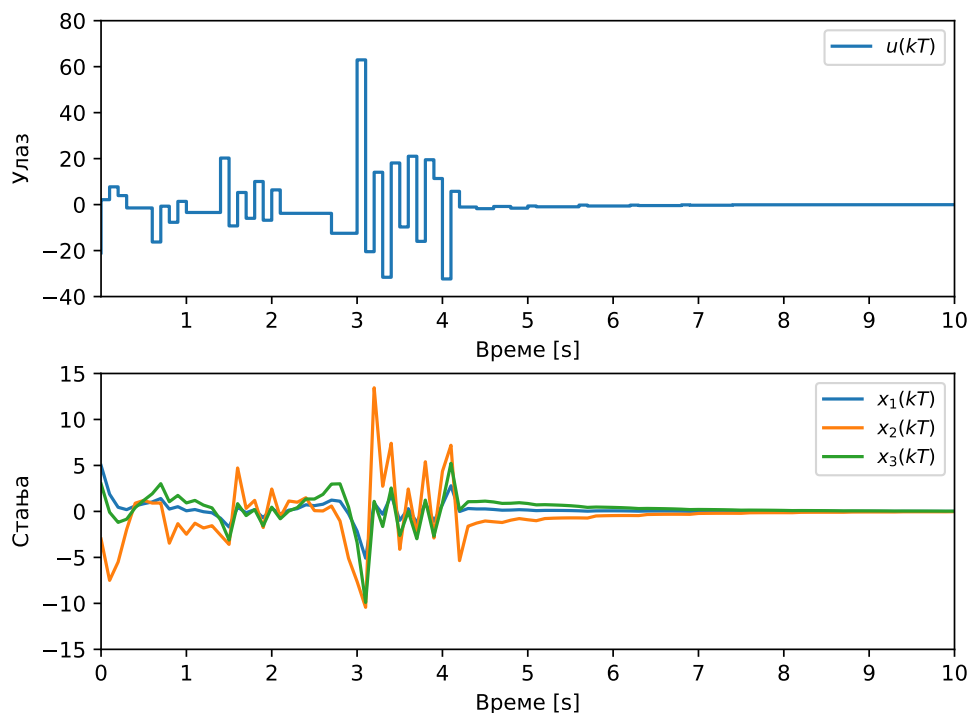
Слика 3.3: Улаз и стања система управљаног регулатором заснованим на АДП-у са периодични одабирањем.



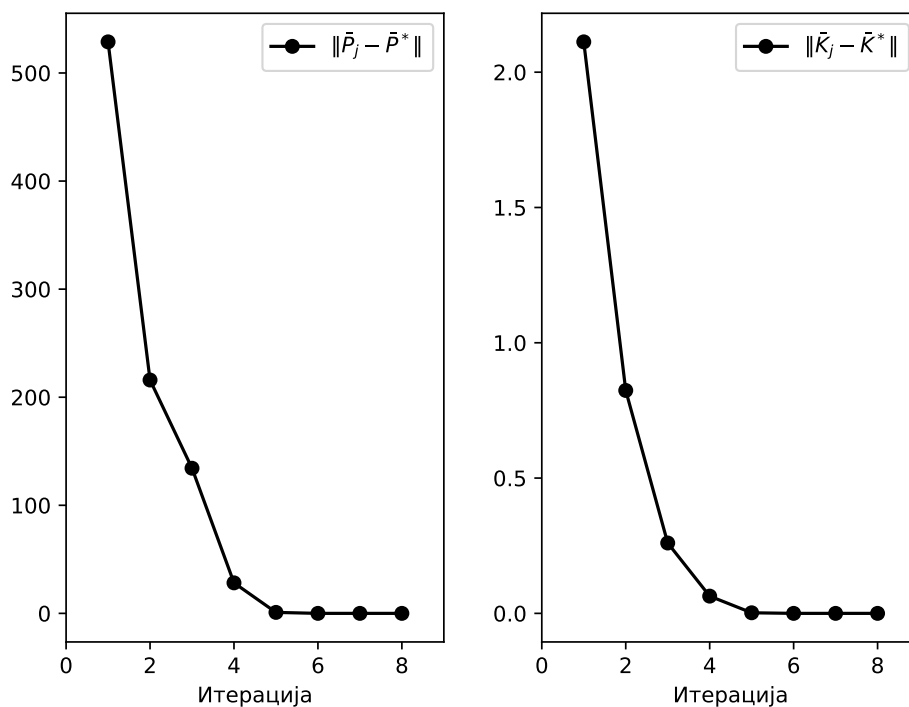
Слика 3.4: Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  њиховом оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  приликом пројектовања регулатора заснованог на АДП-у са периодичним одабирањем.

Управљање и стања система управљаним регулатором заснованим на АДП-у покренутог догађајима приказана је на слици 3.5, док је конвергенција апроксимираних вредности матрица  $\bar{P}$  и  $\bar{K}$  њиховим оптималним вредностима током фазе

учења приказана на слици 3.6. Са слика можемо видети да апериодично одабирање не утиче значајно на перформансе управљаног система и конвергенцију током фазе учења.

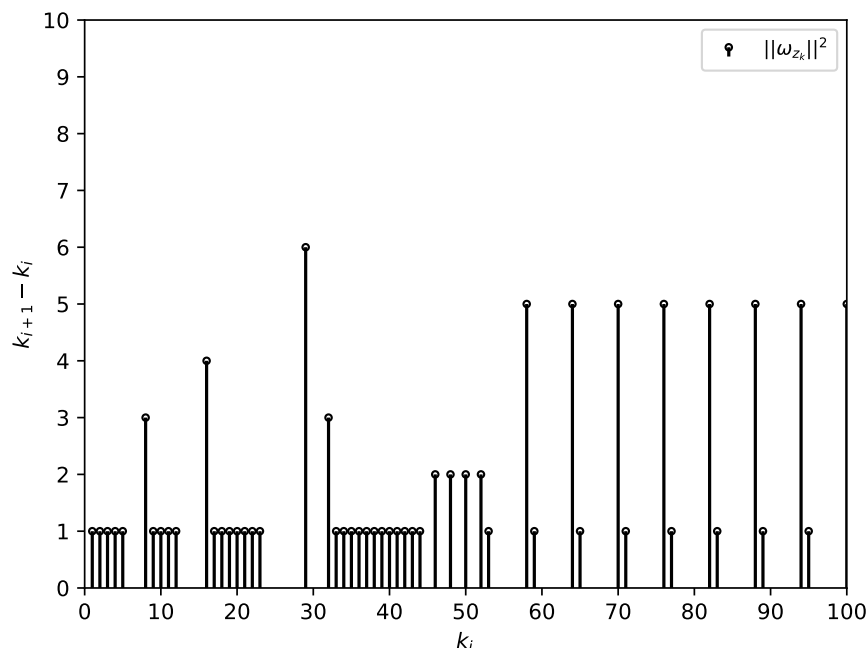


Слика 3.5: Улаз и стања система управљаног регулатором заснованим на АДП-у покренутог догађајима.



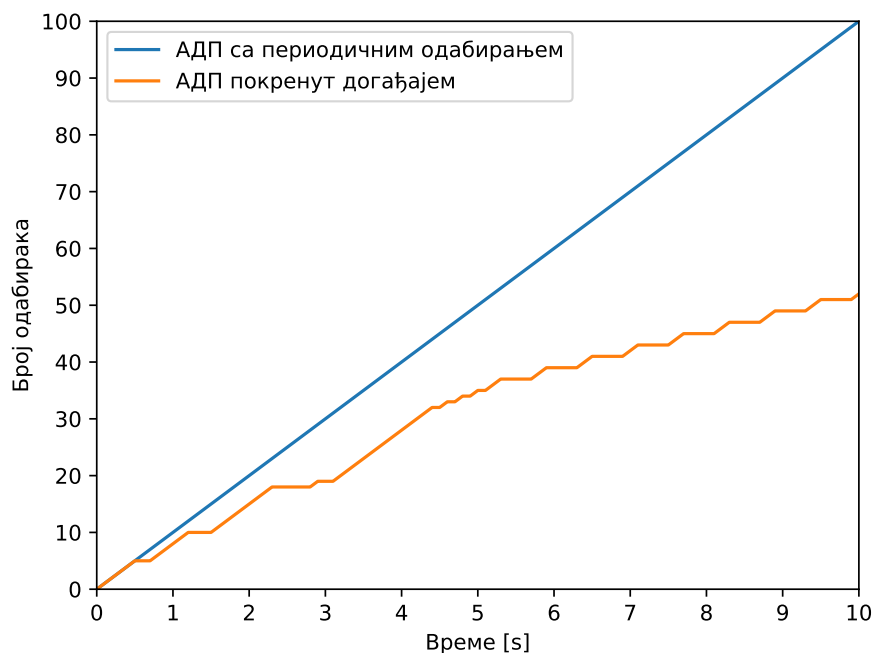
Слика 3.6: Конвергенција матрица  $\bar{P}$  и  $\bar{K}$  њиховом оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$  приликом пројектовања регулатора заснованог на АДП-у покренутиим догађајима.

Апериодично одабирање регулатора заснованог на АДП-у покренутог догађајима приказано је на слици 3.7.



Слика 3.7: Апериодично одабирање регулатора заснованог на АДП-у покренутог догађајима.

На слици 3.8 упоредно је приказано одабирање управљачких система заснованих на АДП-у са периодичним одабирањем и покренутог догађајима. Као што се може видети на слици, управљање засновано на АДП-у покренутог догађајима значајно смањује комуникацију како време пролази.



Слика 3.8: Упоредни приказ броја одабирања регулатора заснованих на АДП-у са периодичним одабирањем и покренутог догађајима.

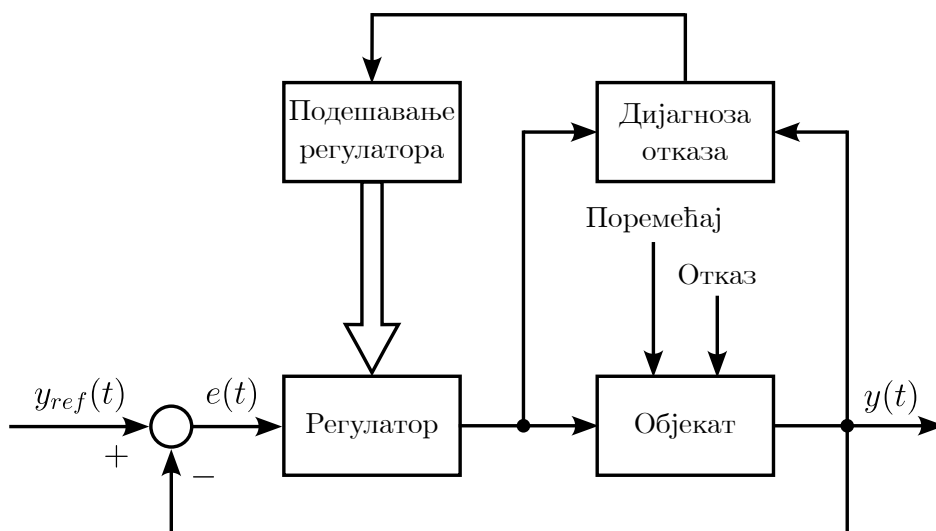
## Глава 4

# Системи управљања са толерисаним отказима

Са брзим развојем савремених технологија, системи постају све сложенији и обимнији. Као последица, јављају се неизбежни откази у системима који могу да утичу на квалитет производа, оштећење опреме или повреду људи. Добро је познато да у постојећим системима, системске компоненте као што су актуатори и сензори се могу оштетити услед изненадних отказа које се могу јавити појединачно или истовремено током рада, изазивајући озбиљне хаварије [110, 111]. Од свих типова отказа, откази актуатора највише доприносе смањењу перформанси управљачког система. Поред тога, откази на актуаторима могу се јавити током дуготрајног рада услед старења опреме и других фактора. Због тога је важно развити управљање толерантно на отказе који раде упркос таквим отказима и одржавати прихватљиве перформансе управљачког система [112].

Постоје два типа управљања толерантног на отказе (УТО): пасивно и активно. Пасивно УТО је робусна техника управљања, у смислу да је унапред познат скуп отказа. Пасивно УТО се састоји из два дела: управљање повратном спрегом, којом управља преко перформанси регулатора, док се у случају нетачног модела и спољашњих поремећаја користе робусни регулатори. Због тога, пасивно УТО је потпуно ограничено у случају појаве значајних отказа. Са друге стране, активно УТО има већу могућност да толерише отказе, одржавајући стабилност и захтеване перформансе техникама подешавања управљања на основу одлука јединице за дијагностику отказа. Постоје различите методе рада активног УТО, као што су компензација отказа или реконфигурација отказа. УТО има широку употребу и велику важност у различитим областима истраживања, нарочито у интелигентној производној индустрији. УТО је напредна метода управљања која обезбеђује безбедан рад система у случају отказа компоненти или параметара [113–116]. Општа структура УТО система заснованог на подацима приказана је на слици 4.1.

У последње време, УТО засновано на подацима постало је једно од најактивнијих области истраживања. Нова онлајн стратегија за детекцију и изолацију отказа заснована на концепту вишеструких модела авионских млазних мотора предложена је у [117]. У предложеној стратегији нелинеарна динамика система је линеаризована скупом линеарних модела [118]. Проблем детекције и изолације отказа код нелинеарних система употребом линеарних вишеструких модела разматран је у [119]. Интелигентно оптимално управљање у случају непознатих временски-дискретних нелинеарних система разматран је у [5, 120]. Ова метода је заснована на адаптивном



Слика 4.1: Систем управљања толерантан на отказе.

динамичком програмирању [75, 121] који представља методу машинског учења за решавање одговарајуће Хамилтон-Јакоби-Белманове (ХЈБ) једначине, које је врло компликовано решити директно.

Задњих година, развијено је доста УТО метода заснованих на АДП-у, код којих отказ актуатора представља најчешћи отказ. Развијен је алгоритам УТО заснован на клизном режиму који може радити са отказима промене појачања и померености актуатора и представљена у [122]. Стратегија УТО са гарантованим индексом перформансе заснован на актант-критичар структури за сложени нелинеарни систем у случају губитка ефикасности актуатора предложен је у [123]. Захваљујући опсерверу отказа, модификована функција циља је развијена у [124], док се отказ померености актуатора компензује преко оптималног пројектовања регулатора. Адаптивни фази УТО метод, заснован на опсерверу, за СИСО нелинеарне системе је развијен у [125].

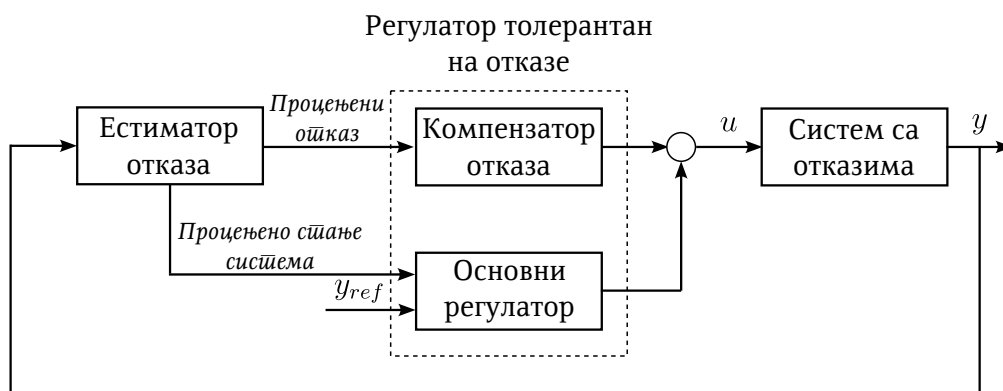
АДП методе обезбеђују ефикасан начин да се постигну добре перформансе оптималног регулатора, који је заснован на адаптивном и оптималном управљању као и учењу поткрепљивањем [4–8]. АДП представља технику управљања која је заснована на подацима која гарантује стабилност управљаног система [4]. У случајевима када је непозната динамика система и када стања система нису мерљива, од великог значаја је употреба метода АДП заснованих на измереним улазно-излазним подацима линеарног система, који је заснован на повратној спрези по излазу. Главна предност техника које су засноване на повратној спрези по излазу је да није потребно познавање динамике система за њихову примену. Ова техника за непознати модел система генерише секвенцу субоптималних регулатора који конвергирају оптималном закону управљања у свакој наредној итерацији. Практична применљивост алгоритма управљања је поткрепљена и чињеницом да се нелинеарни системи могу прецизно представити линеарним моделом са променљивим параметрима [35, 36]. Такође, мерење свих стања система није изводљиво у практичним применама и захтевало би скупу мерну опрему. Из овог разлога, погоднија је примена управљачких алгоритама који користе реконструкцију стања како би се избегла непотребна мерења [4].

На основу претходног, предложена је метода УТО засновано на АДП-у. На



основу измерених улазно-излазних података, регулатор учи да компензује непознату динамику система, поремећаје и отказе обезбеђујући жељене перформансе. УТО се постиже итеративно на основу повратне спреге по излазу, реконструкције стања и АДП. Непознати модел система се прво идентификује, након чега се итеративно решава алгебарска Рикатијева једначина (АРЈ). Да би се обезбедила конзистентност апроксимације и добила јединствена решења у свакој итерацији, потребно је додати истраживачки шум да би се задовољио услов перзистентне побуде [14]. Из практичних разлога, код управљања заснованог на АДП-у лакше је користити временски-дискретни модел код кога је једноставно извршити аквизицију података у односу на временски-континуални модел. Због тога, користимо измерене улазно-излазне податке да реконструирамо вектор стања дискретног модела, након чега ће бити имплементирано УТО засновано на АДП-у.

На слици 4.2 приказана је шема активног управљачког система толерантног на отказе. Пројектовање овог управљачког система може се поделити у три фазе. Прво се врши пројектовање основног регулатора, за случај када нема отказа у систему којим управљамо. Након тога, пројектује се естиматор отказа који ће дијагностиковати отказ и одредити динамику система са отказима. На крају, врши се реконфигурација управљачког сигнала да би се обезбедило да се систем са отказима понаша као исправан систем [126, 127].



Слика 4.2: Шема система управљања толерантног на отказе.

## 4.1 Линеарни систем са отказом актуатора

Посматрајмо линеарни модел система са отказима актуатора описан једначинама

$$\dot{x}(t) = Ax(t) + B[u(t) - f_a(t)], \quad (4.1)$$

$$y(t) = Cx(t), \quad (4.2)$$

где  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$ ,  $y(t) \in \mathbb{R}^r$ ,  $f_a(t) \in \mathbb{R}^m$ , представљају стање система, управљачки сигнал, излаз система, непознати отказ актуатора, респективно.  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{r \times n}$  су непознате матрице система, уз претпоставку да је пар  $(A, B)$  управљив, а пар  $(A, C)$  осмотрив.

За систем описан једначинама (4.1)-(4.2), за случај када је  $f_a(t) = 0$ , односно

када нема отказа у систему, индекс перформансе се може дефинисати као

$$J(x_0) = \int_0^{\infty} [y^T(\tau)Qy(\tau) + u^T(\tau)Ru(\tau)] d\tau, \quad (4.3)$$

где  $x_0 \in \mathbb{R}^n$  представља почетно стање система,  $Q = Q^T \geq 0$ ,  $R = R^T > 0$ , а пар  $(A, Q^{1/2}C)$  је осмотрив.

Циљ је пројектовати линеарни оптимални закон управљања облика

$$u = -K^*x, \quad (4.4)$$

који минимизира индекс перформансе (4.3). Оптимална матрица појачања  $K^*$  се може одредити помоћу израза

$$K^* = R^{-1}B^T P^*, \quad (4.5)$$

где је  $P^* = (P^*)^T > 0$  јединствено симетрично позитивно дефинитно решење следеће добро познате алгебарске Рикатијеве једначине (АРЈ)

$$A^T P^* + P^* A + C^T Q C - P^* B R^{-1} B^T P^* = 0, \quad (4.6)$$

под условом да су матрице система у потпуности познате, као и да је пар  $(A, B)$  управљив а пар  $(A, Q^{1/2}C)$  осмотрив [3]. Треба приметити да је пројектовање оптималног управљања на овај начин углавном применљиво на системе нижег реда. Наиме, у случају комплексних система високог реда је обично тешко директно решити  $P^*$  на основу једначине (4.6), која је нелинеарна за  $P$ .

Такође, из практичних разлога имплементације управљачког система, лакше је реализовати аквизицију података у случају временски-дискретног система у односу на временски-континуални систем. Због тога, дискретизоваћемо временски-континуални модел у одговарајући временски-дискретни модел следећег облика

$$x_{k+1} = Ax_k + Bu_k, \quad (4.7)$$

$$y_k = Cx_k, \quad (4.8)$$

где је  $A_d = e^{Ah}$ ,  $B_d = \int_0^h (e^{A\tau} d\tau) B$ , а  $h > 0$  је одговарајући период одабирања, који задовољава услов  $\omega_h = 2\pi/h$  да приликом одабирања не дође до појаве непостојећих учесталости, чије је постајање добро познато [128]. Другим речима, управљивост и осмотривост оригиналног временски-континуалног система треба остати и након дискретизације. Наиме, ако су  $x_k$ ,  $u_k$ ,  $y_k$  вектори стања, улаза и излаза узорковани у тренутку  $kh$ , респективно, онда парови  $(A_d, C)$  и  $(A_d, Q^{1/2}C)$  су осмотриви, кад је пар  $(A_d, B_d)$  управљив.

У случају када у систему нема отказа, то јест када је  $f_a = 0$ , минимизацијом критеријума (4.3) се може доћи до решења коришћењем технике итерације закона управљања. Циљ је пројектовање управљања које може компензовати утицај отказа актуатора за систем описан једначинама (4.1)-(4.2) и обезбедити стабилности затвореног система.

Алгоритам итерације закона управљања се састоји од евалуације тренутног закона управљања на основу израза (2.19), а затим се тренутни закон управљања унапређује коришћењем израза (2.20).

Такође, индекс перформансе за дискретни систем (4.7)-(4.8) :

$$J_d(x_0) = \sum_{j=0}^{\infty} y_j^T Q_d y_j + u_j^T R_d u_j, \quad (4.9)$$

где је  $Q_d = Qh$  и  $R_d = Rh$ . Оптимални закон управљања који минимизира критеријум (4.9) има облик

$$u_k = -K_d^* x_k, \quad (4.10)$$

где се оптимална матрица појачања у дискретном случају може одредити изразом

$$K_d^* = (R + B_d^T P_d^* B_d)^{-1} B_d^T P_d^* A_d, \quad (4.11)$$

у коме  $P_d^*$  представља јединствено симетрично позитивно дефинитно решење дискретне алгебарске Рикатијеве једначине

$$A_d^T P_d^* A_d - P_d^* + C^T Q C - A_d^T P_d^* B_d K_d^* = 0. \quad (4.12)$$

Пошто је једначина (4.12) нелинеарна по  $P_d^*$ , не може се једноставно директно решити  $P_d^*$  у случају сложених система вишег реда. Ипак, развијени су ефикасни алгоритми који нумерички апроксимирају решење једначине (4.12). Један од таквих алгорита развио је Хјуар [93], који се уводи у облику Леме 2.1.

Итеративним решавањем Љапуновљевих једначина (2.19), која је линеарна по  $P_j$ , и рекурзивним ажурирањем закона управљања  $K_j$  употребом израза (2.20), решава се нелинеарна једначина (4.12) нумеричком апроксимацијом [93]. Може се закључити да секвенце  $\{P_j\}_{j=0}^{\infty}$  и  $\{K_j\}_{j=0}^{\infty}$ , израчунате помоћу овог алгорита, конвергирају вредностима  $P_d^*$  и  $K_d^*$ , респективно. Такође, за  $j = 0, 1, \dots$ ,  $A_d - B_d K_j$  је Шурова матрица (модуо сопствених вредности матрице су мањи од 1). Треба приметити да Хјуарова метода представља алгоритам итерације закона управљања заснованог на моделу, који се не може директно применити на овај проблем, јер су за његов рад неопходни параметри система. Да би се овај алгоритам применио на онлајн дискретни систем (4.7)-(4.8), пројектоваћемо алгоритам управљања заснован на АДП-у помоћу повратне спреге по излазу, који не захтева познавање матрица система.

## 4.2 Реконструкција стања

На основу [94, 95], временски-дискретни систем (4.7)-(4.8) може се проширити употребом улазно/излазне секвенце на временском интервалу  $[k - N, k - 1]$  на следећи начин

$$\begin{aligned} x_k &= A_d^N x_{k-N} + V(N) \bar{u}_{k-1, k-N}, \\ \bar{y}_{k-1, k-N} &= U(N) x_{k-N} + T(N) \bar{u}_{k-1, k-N}, \end{aligned} \quad (4.13)$$

где је

$$\begin{aligned}\bar{u}_{k-1,k-N} &= [u_{k-1}^T \quad u_{k-2}^T \quad \cdots \quad u_{k-N}^T]^T, \\ \bar{y}_{k-1,k-N} &= [y_{k-1}^T \quad y_{k-2}^T \quad \cdots \quad y_{k-N}^T]^T, \\ V(N) &= [B_d \quad A_d B_d \quad \cdots \quad A_d^{N-1} B_d], \\ U(N) &= [(C A_d^{N-1})^T \quad (C A_d)^T \quad \cdots \quad C^T]^T, \\ T(N) &= \begin{bmatrix} 0 & C B_d & C A_d B_d & \cdots & C A_d^{N-2} B_d \\ 0 & 0 & C B_d & \cdots & C A_d^{N-3} B_d \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & C B_d \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},\end{aligned}$$

а  $N = \max(\rho_u, \rho_v)$  представља индекс осмотривости, где је  $\rho_u$  најмањи цео број који обезбеђује да матрица  $U(\rho_u)$  има пун ранг колона, док је  $\rho_v$  најмањи цео број који обезбеђује да матрица  $V(\rho_v)$  има пун ранг врста [95]. Зато постоји леви инверз матрице  $U(N)$ , облика

$$U^+(N) = [U^T(N)U(N)]^{-1} U^T(N).$$

Помоћу реконструкције стања (4.13), регулатор заснован на АДП-у са повратном спрегом по излазу се може применити за решавање оптималног проблема управљања у случају система са непознатом динамиком. Јединственост реконструкције стања је већ изражена у форми Леме 2.3.

### 4.3 Алгоритам УТО заснован на АДП-у

Сада, на основу израза (2.19)-(2.20), онлајн стратегија учења која користи повратне информације о излазима може се изразити у облику  $u_k^* = -\bar{K}_d z_k$ , обезбеђујући субоптимално својство затвореном систему управљања. Временски-дискретни модел (4.7) може се изразити на следећи начин

$$x_{k+1} = A_j x_k + B_d (K_j x_k + u_k), \quad (4.14)$$

где је  $A_j = A_d - B_d K_j$ . Заменом  $\bar{K}_j = K_j \Theta$  и  $\bar{P}_j = \Theta^T P_j \Theta$ , из израза (2.19) и (4.14) се добија једначина

$$\begin{aligned}z_{k+1}^T \bar{P}_j z_{k+1} - z_k^T \bar{P}_j z_k &= [u_k^T \otimes u_k^T - (z_k^T \otimes z_k^T)(\bar{K}_j^T \otimes \bar{K}_j^T)] \text{vec}(B_d^T P_j B_d) + \\ &+ 2 [(z_k^T \otimes z_k^T)(I_q \otimes \bar{K}_j^T) + (z_k^T \otimes u_k^T)] \text{vec}(B_d^T P_j A_d \Theta) - (y_k^T Q y_k + z_k^T \bar{K}_j^T R \bar{K}_j z_k),\end{aligned} \quad (4.15)$$

коју можемо лако решити уколико је задовољен услов (2.54), на следећи начин

$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j, \quad (4.16)$$

где је

$$\Phi_j = \begin{bmatrix} y_{k_0}^T Q y_{k_0} + z_{k_0}^T \bar{K}_j^T R \bar{K}_j z_{k_0} \\ y_{k_1}^T Q y_{k_1} + z_{k_1}^T \bar{K}_j^T R \bar{K}_j z_{k_1} \\ \vdots \\ y_{k_l}^T Q y_{k_l} + z_{k_l}^T \bar{K}_j^T R \bar{K}_j z_{k_l} \end{bmatrix},$$

$$\Psi_j = [\Gamma_{\bar{u}} - \Gamma_{\bar{k}z}, \quad 2(\Gamma_{zu} + \Gamma_{zz}(I_m \otimes \bar{K}_j^T)), \quad -\delta_{zz}],$$

$$\delta_{zz} = [\text{vecv}(z_{k_1}) - \text{vecv}(z_{k_0}), \text{vecv}(z_{k_2}) - \text{vecv}(z_{k_1}), \dots, \text{vecv}(z_{k_{l+1}}) - \text{vecv}(z_{k_l})]^T,$$

$$\Gamma_{zz} = [z_{k_1} \otimes z_{k_0}, z_{k_2} \otimes z_{k_1}, \dots, z_{k_{l+1}} \otimes z_{k_l}]^T,$$

$$\Gamma_{zu} = [z_{k_1} \otimes u_{k_0}, z_{k_2} \otimes u_{k_1}, \dots, z_{k_{l+1}} \otimes u_{k_l}]^T,$$

$$\Gamma_{\bar{u}} = [\text{vecv}(u_{k_0}), \text{vecv}(u_{k_1}), \dots, \text{vecv}(u_{k_l})]^T,$$

$$\Gamma_{\bar{k}z} = [\text{vecv}(\bar{K}_j z_{k_0}), \bar{K}_j z_{k_1}, \dots, \bar{K}_j z_{k_l}]^T.$$

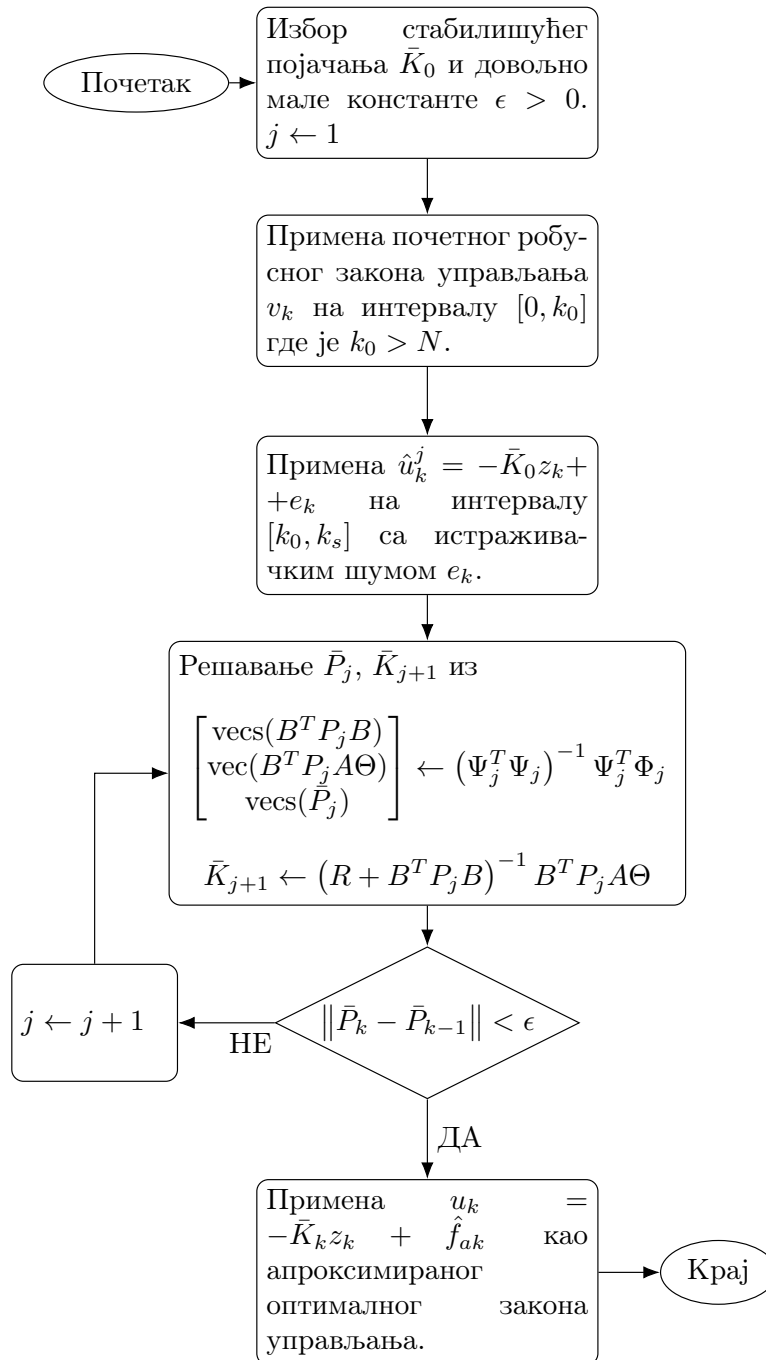
Конвергенција онлајн учења закона управљања са повратном спрегом по излазу је загарантована условом ранга који је изражен у Лемми 2.3 [94]. Лема 2.3 се односи на услов перзистентне побуде из теорије адаптивног управљања [82].

Истраживачки шум  $e_k$ , који задовољава услов перзистентне побуде, је потребно додати улазном сигналу током фазе учења, да би се задовољио услов ранга (2.54), а при том да не утиче на конвергенцију током фазе учења [15, 89, 109]. Помоћу израза (4.15) се врши евалуација закона управљања, која се користи да се јединствено реши  $\bar{P}_j$ , док се изразом (2.48) побољшава закон управљања, и на тај начин се ажурира матрица појачања  $\bar{K}_{j+1}$ . Треба приметити да решавањем (4.15) уместо (2.19), се у потпуности елиминише почетни захтев за потпуним познавањем динамике система. На овај начин, потребна су нам само мерења улаза  $u_k$  и излаза  $y_k$ . Наиме, имајући у виду израз за  $z_k$ , можемо видети да у изразу који дефинише закон управљања  $\hat{u}_k = -\bar{K}_k^* z_k$  фигуришу само измерени улазно-излазни подаци.

Пројектовани закон управљања заснован на компензацији отказа је  $u_k = \hat{u}_k + \hat{f}_a$  где је  $\hat{f}_a$  члан који компензује отказ да не би дошло до погоршања перформанси система, и може се одредити из израза  $\hat{f}_{a,k+1} = l_2 (2\hat{u}_k R - x^T B)^T$ , где је  $l_2 > 0$ .

**Напомена 4.1.** Матрица  $B$  која је неопходна за процену отказа може се лако одредити користећи израз  $B = (K_{k+1} P_k^{-1})^T$ , који следи из  $K_{k+1} = B^T P_k$ . Матрице  $P_k$  и  $K_{k+1}$  могу се одредити на основу израза 4.15 и 2.48

Алгоритам управљања толерантан на отказе заснован на АДП-у приказан је на слици 4.3.



Слика 4.3: Алгоритам пројектовања управљања толерантног на отказе заснованог на АДП-у

---

**Алгоритам 6** Пројектовање регулатора заснованог на АДП-у толерантног на отказе
 

---

1. **Иницијализација:** Изабрати стабилишућу матрицу појачања  $K_0$  и довољно малу константу  $\varepsilon > 0$ . Подесити  $j \leftarrow 1$ .
2. **Прикупљање података:** Применити  $\hat{u}_k^j = -\bar{K}_0 z_k + e_k$  на интервалу  $[k_0, k_l]$  где је  $e_k$  истраживачки шум. Формирати матрице  $\Psi_j$  и  $\Phi_j$  на основу улазно-излазних података тако да буде задовољен услов ранга.
3. **Евалуација закона управљања:** На основу израза

$$\begin{bmatrix} \text{vecs}(B^T P_j B) \\ \text{vec}(B^T P_j A \Theta) \\ \text{vecs}(\bar{P}_j) \end{bmatrix} = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T \Phi_j,$$

одређујемо непознате матрице  $B^T P_j B$ ,  $B^T P_j A \Theta$  и  $\bar{P}_j$ .

4. **Унапређење закона управљања:** Ажурирамо закон управљања изразом

$$K_j = (R + B^T P_j B)^{-1} B^T P_j A \Theta.$$

5. **Итерације:** Постављамо  $j \leftarrow j + 1$  и понављамо корак 2 све док је испуњен услов  $\|\bar{P}_j - \bar{P}_{j-1}\| \geq \varepsilon$ .
  6. **Експлоатација:** Примена апроксимираног оптималног закона управљања  $u_k = -\bar{K}_k z_k + \hat{f}_{ak}$ .
- 

## 4.4 Пример

Посматраћемо временски континуални систем описан следећим једначинама стања и излаза

$$\dot{x}(t) = \begin{bmatrix} -3.5 & 1.1 & 3 \\ 1.1 & -4 & 1.5 \\ 2 & 1.3 & 3.3 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} (u(t) - f_a(t)), \quad (4.17)$$

$$y(t) = [8 \ 6 \ 3] x(t). \quad (4.18)$$

Дискретизацијом за период одабирања  $h = 0.1$  s добијамо временски дискретни модел

$$x_{k+1} = \begin{bmatrix} 0.7371 & 0.0943 & 0.3147 \\ 0.0903 & 0.6841 & 0.1650 \\ 0.2113 & 0.1401 & 1.4371 \end{bmatrix} x_k + \begin{bmatrix} 0.1360 \\ 0.3523 \\ 0.2774 \end{bmatrix} u_k, \quad (4.19)$$

$$y_k = [8 \ 6 \ 3] x_k, \quad (4.20)$$

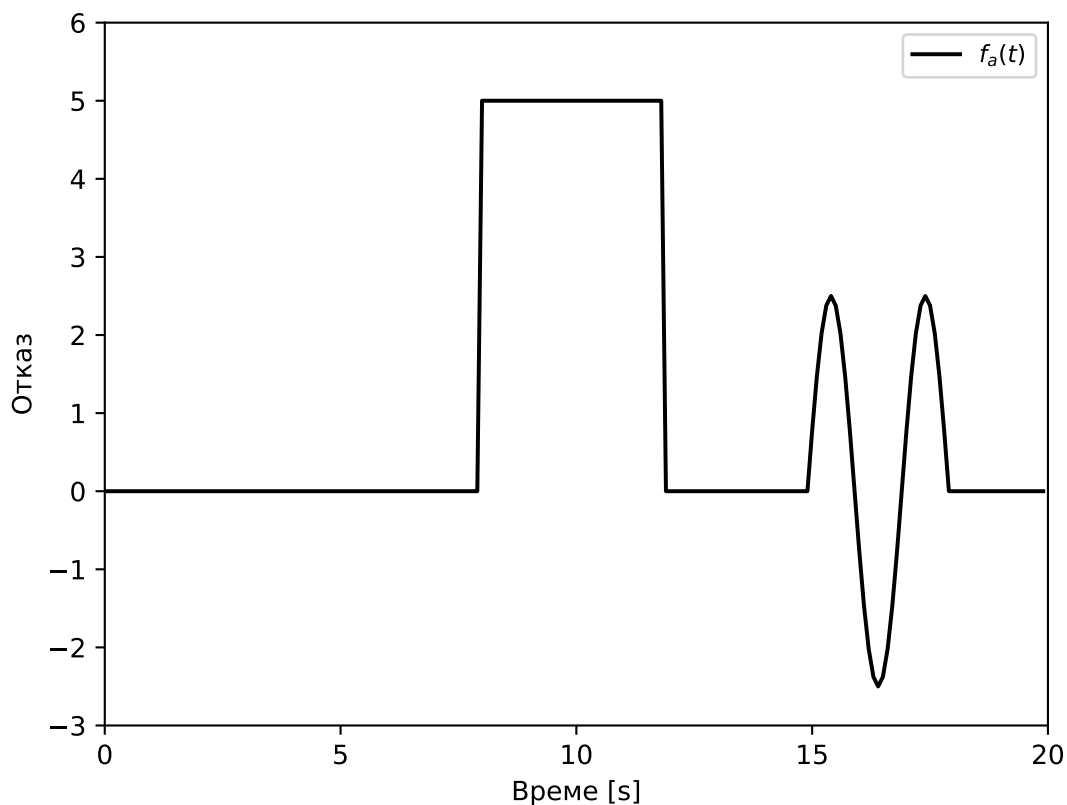
који је коришћен у симулацијама. За потребе симулација узето је почетно стање  $x_0 = [1, 0.1, 0.1]^T$ . За оптимални истраживачки шум је примењена сума синусоида са фреквенцијама  $\omega_i \in [-1.918, -2.124, 0.595]$ .

---

Отказ актуатора је представљен функцијом

$$f_a(t) = \begin{cases} 0, & 0 \leq t < 8 \text{ s} \\ 5, & 8 \text{ s} \leq t < 12 \text{ s} \\ 0, & 12 \text{ s} \leq t < 15 \text{ s} \\ 2.5 \sin(t - 15) & 15 \text{ s} \leq t < 18 \text{ s} \\ 0, & 18 \text{ s} \leq t < 20 \text{ s} \end{cases}$$

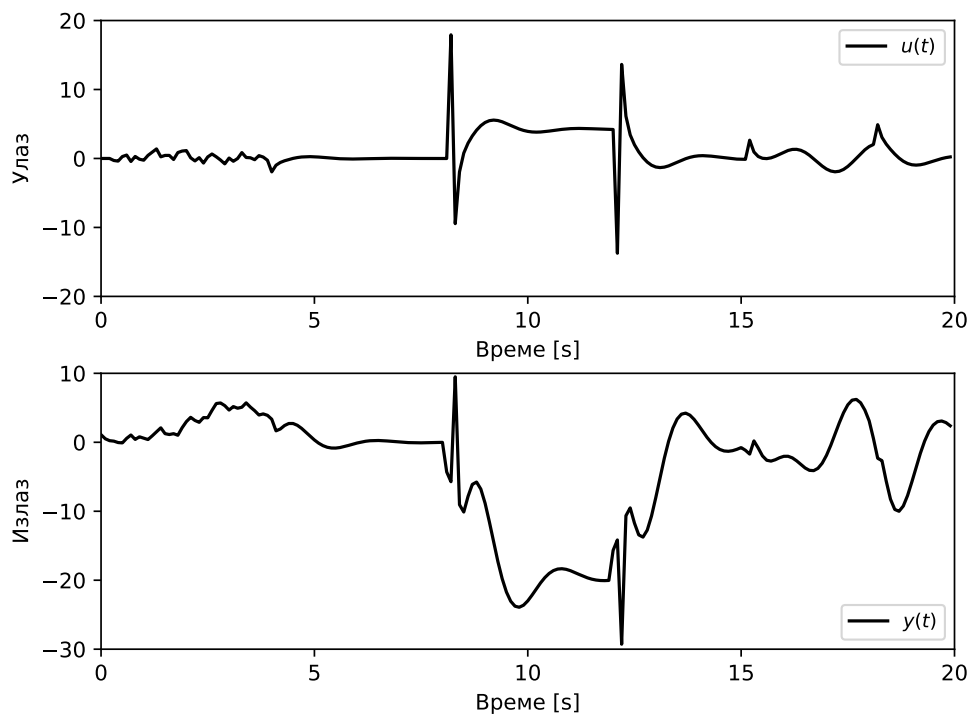
чији облик је приказан на слици 4.4.



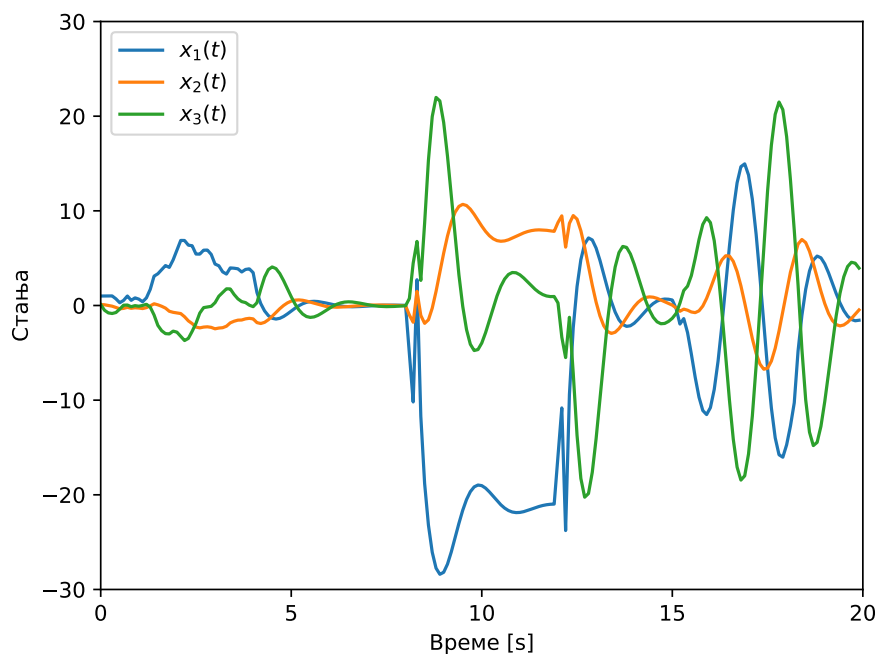
Слика 4.4: Отказ актуатора.

На сликама 4.5 и 4.6 приказани су улаз, излаз и стања система у случају када регулатор заснована на АДП-у није толерантан на отказе. Са слика се јасно види деградација перформанси када се појави отказ актуатора.



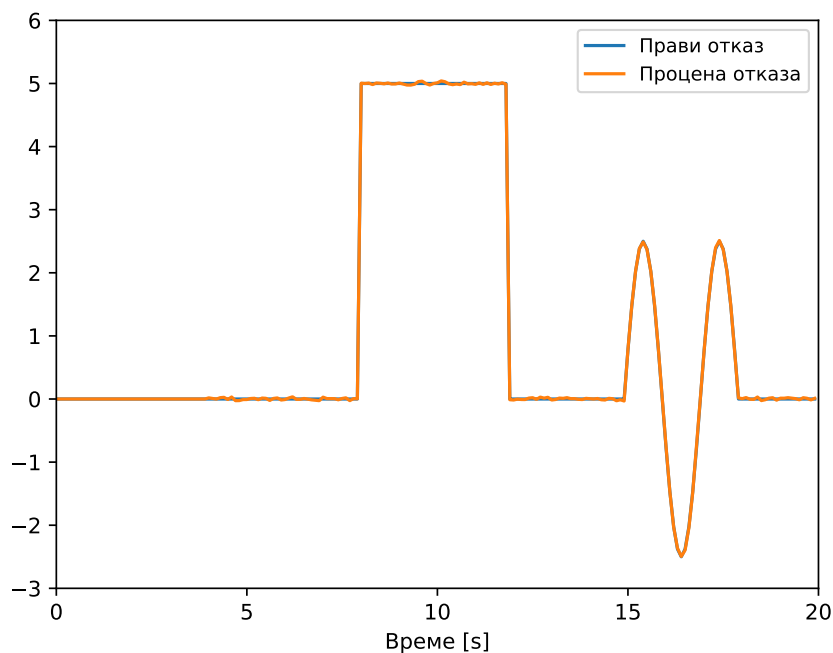


Слика 4.5: Улаз и излаз система управљаног регулатором заснованим на АДП-у без компензације отказа актуатора.



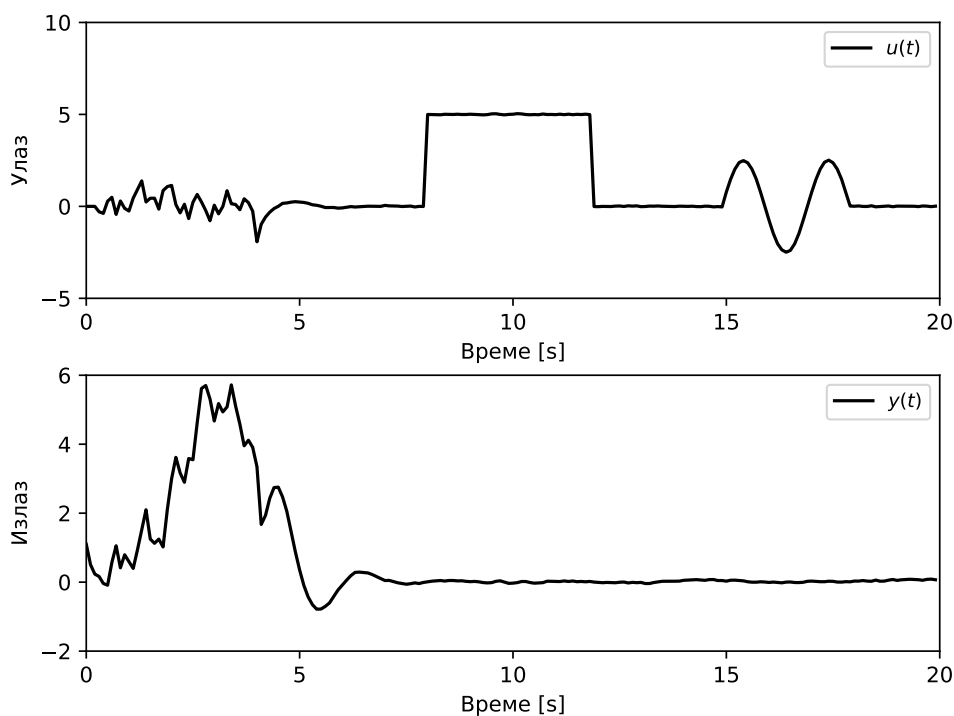
Слика 4.6: Стања система управљаног регулатором заснованим на АДП-у без компензације отказа актуатора.

Упоредни приказ процене отказа и његове праве вредности дат је на слици 4.7. Можемо уочити добро поклапање процене са стварним отказом.

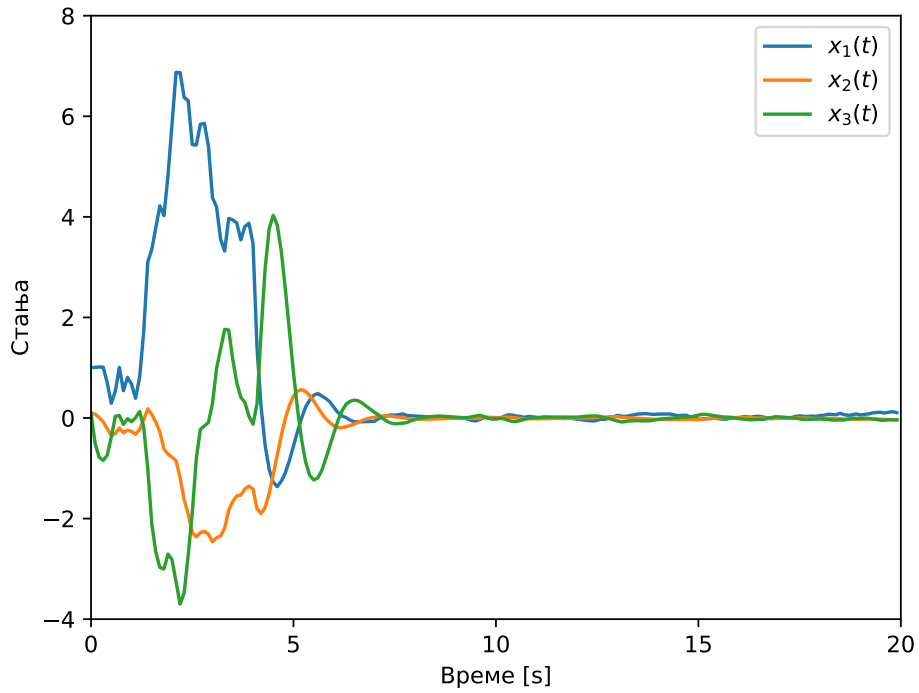


Слика 4.7: Процена отказа актуатора.

Улаз, излаз и стања система управљаног регулатором заснованим на АДП-у толерантим на отказе прикази су на сликама 4.8 и 4.9. На основу слика можемо закључити да перформансе управљаног система остају задовољавајуће након појаве отказа актуатора.

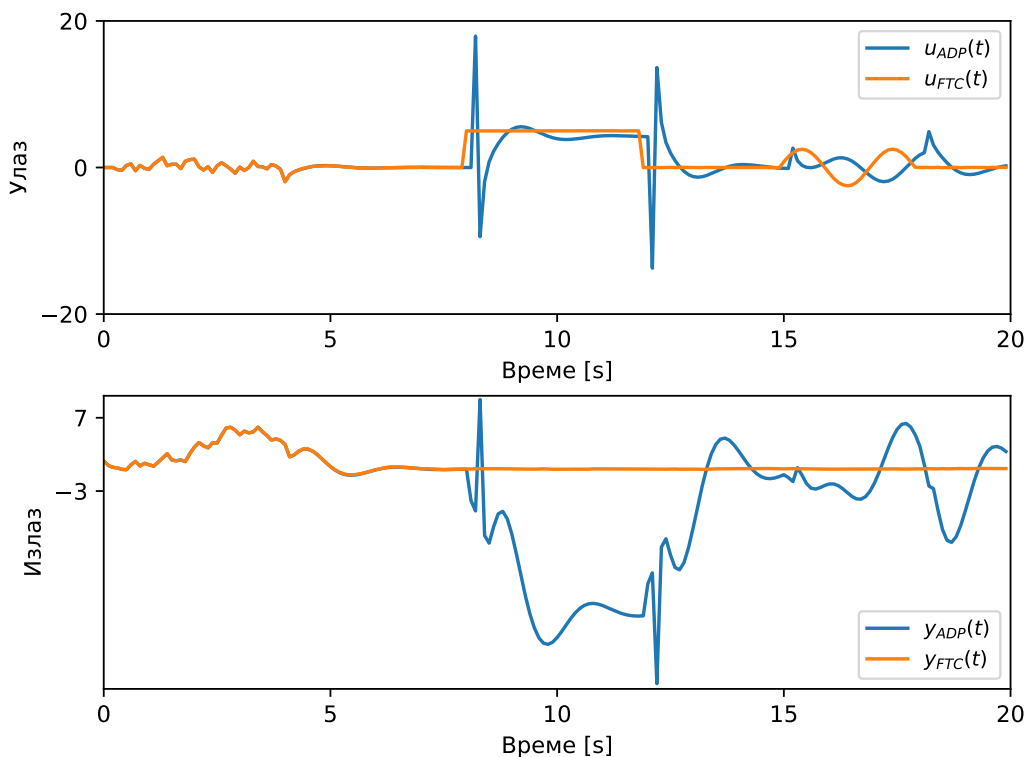


Слика 4.8: Улаз и излаз система управљаног регулатором заснованим на АДП-у толерантан на отказе.



Слика 4.9: Стања система управљаног регулатором заснованим на АДП-у толерантан на отказе.

Коначно, на слици 4.10 дат је упоредни приказ улаза и излаза система управљаног регулатором заснованим на АДП-у са и без компензације поремећаја.



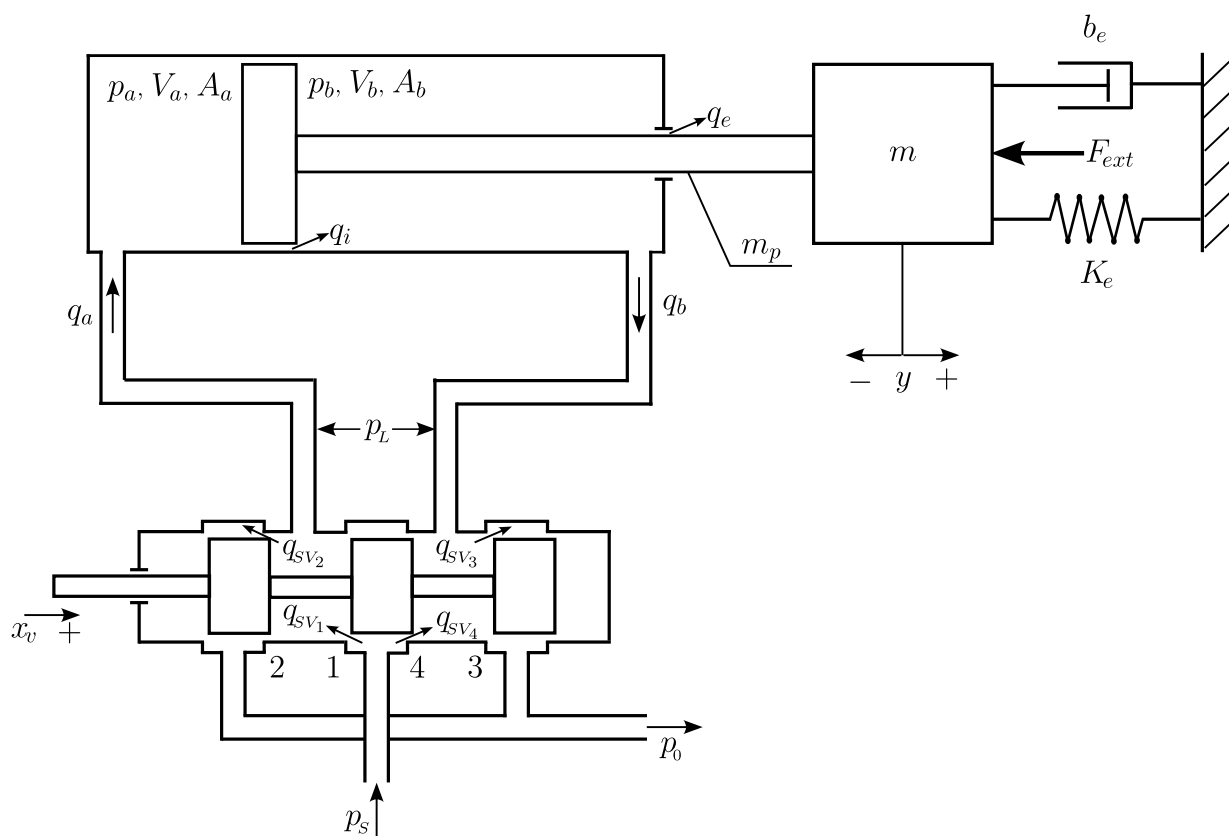
Слика 4.10: Упоредни приказ улаза и излаза у случају управљања заснованим на АДП-у са и без толеранције на отказе.

# Глава 5

## Практични примери

### 5.1 Хидраулички серво актуатор

Посматра се хидраулички серво актуатор (ХСА) приказан на слици 5.1. Са слике можемо видети да се хидраулички серво актуатор састоји од серво вентила и хидрауличког цилиндра. Динамика хидрауличког серво актуатора је дефинисана системом сложених нелинеарних једначина које зависе од многих параметара који се не могу тачно одредити [50, 51]. У табели 5.1 је дат опис параметара.



Слика 5.1: Шематски приказ хидрауличког серво актуатора.

На основу ознака са слике 5.1, можемо дефинисати однос површина клипа ци-

Симбол	Опис
$x_v$	померање клипа разводника
$p_a, p_b$	притисци у коморама цилиндра
$q_a, q_b$	протоци у коморама цилиндра
$y$	померање клипа цилиндра
$L$	ход клипа
$K_e$	крутост опруге
$b_e$	коэффицијент пригушења
$p_s, p_0$	притисак напајања и резервоара
$m_t, m_p, m$	укупна маса, маса клипа, маса оптерећења
$F_f$	силе трења
$F_{ext}$	силе поремећаја
$A_a, A_b$	ефективне површине клипа
$V_a, V_b, V_{a0}, V_{b0}$	запремине обе коморе цилиндра као и њихове почетне запремине
$q_{Li}, q_{Le}$	унутрашње и спољашње цурење
$\beta_e$	коэффицијент стишљивости

Табела 5.1: Параметри хидрауличког серво актуатора

линдра као

$$\alpha = \frac{A_b}{A_a}.$$

Запремине обе коморе зависе од померања клипа цилиндра, и дефинисане су функцијама

$$\begin{aligned} V_a(y) &= V_{a0} + yA_a, \\ V_b(y) &= V_{b0} + (L - y)\alpha A_a. \end{aligned}$$

Унутрашње цурење у цилиндру износи

$$q_{Li} = c_{Li} (p_a - p_b),$$

где је  $c_{Li}$  коэффициент унутрашњег цурења, а  $c_{vi}$  су коэффициенти пражњења. Претпоставићемо да је спољашње цурење занемарљиво.

Динамика хидрауличког серво актуатора је дефинисана једначином равнотежа сила на клипу цилиндра

$$m_t \ddot{y} = A_a p_a - A_b p_b - F_f(\dot{y}) - K_e y - F_{ext}, \quad (5.1)$$

једначинама промена притиска у коморама цилиндра

$$\dot{p}_a = \frac{\beta_e}{V_a(y)} (q_a - A_a \dot{y} - q_{Li} - q_{Lea}), \quad (5.2)$$

$$\dot{p}_b = \frac{\beta_e}{V_b(y)} (q_b + \alpha A_a \dot{y} + q_{Li} - q_{Leb}), \quad (5.3)$$

и једначинама протока кроз вентил

$$\begin{aligned} q_a &= q_{sv1} - q_{sv2} = \\ &= c_{v1} \text{sg}(x_v) \text{sign}(p_s - p_a) \sqrt{|p_s - p_a|} - c_{v2} \text{sg}(-x_v) \text{sign}(p_a - p_0) \sqrt{|p_a - p_0|}, \end{aligned} \quad (5.4)$$

$$\begin{aligned} q_b &= q_{sv3} - q_{sv4} = \\ &= c_{v3} \text{sg}(x_v) \text{sign}(p_s - p_b) \sqrt{|p_s - p_b|} - c_{v4} \text{sg}(-x_v) \text{sign}(p_b - p_0) \sqrt{|p_b - p_0|}, \end{aligned} \quad (5.5)$$

где је  $\text{sg}$  функција која је дефинисана као

$$\text{sg}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Ако дефинишемо променљиве стања и улаза

$$x(t) = [x_1(t) \ x_2(t) \ x_3(t) \ x_4(t)]^T = [y(t) \ \dot{y}(t) \ p_a(t) \ p_b(t)]^T, \quad (5.6)$$

$$u(t) = x_v(t), \quad (5.7)$$

једначине могу се представити као нелинеарни динамички модел у простору стања

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t), u(t)) + h(t), \\ y(t) &= \eta(x(t)), \end{aligned} \quad (5.8)$$

где су  $f(x(t))$  и  $g(x(t), u(t))$  векторске функције представљају динамику стања и функцију улаза, респективно,

$$f(x(t)) = \begin{bmatrix} x_2 \\ \frac{1}{m_t} (A_a x_3 - \alpha A_a x_4 - F_f(x_2) - K_e x_1) \\ -\frac{\beta_e}{A_a x_1 + V_{a0}} (A_a x_2 + c_{Li}(x_3 - x_4)) \\ \frac{\beta_e}{\alpha A_a (L - x_1) + V_{b0}} (\alpha A_a x_2 + c_{Li}(x_3 - x_4)) \end{bmatrix}, \quad (5.9)$$

$$g(x(t), u(t)) = \begin{bmatrix} 0 \\ 0 \\ \frac{\beta_e}{A_a x_1 + V_{a0}} \left( c_{v1} \text{sg}(u) \text{sign}(p_s - x_3) \sqrt{|p_s - x_3|} - c_{v4} \text{sg}(-u) \text{sign}(x_3 - p_0) \sqrt{|x_3 - p_0|} \right) \\ \frac{\beta_e}{\alpha A_a (L - x_1) + V_{b0}} \left( c_{v3} \text{sg}(-u) \text{sign}(p_s - x_4) \sqrt{|p_s - x_4|} - c_{v4} \text{sg}(u) \text{sign}(x_4 - p_0) \sqrt{|x_4 - p_0|} \right) \end{bmatrix}, \quad (5.10)$$

функција поремећаја  $h(t)$  укључује и оптерећења, немоделовану динамику и нетачност параметара

$$h(t) = \left[ h_1(t) \quad -\frac{F_{ext}}{m_t} + h_2(t) \quad h_3(t) \quad h_4(t) \right]^T,$$

док је  $\eta(x(t))$  функција излаза  $\eta(x(t)) = x_1(t)$ .

Једну од главних нелинеарности у моделу цилиндра представља нелинеарна сила трења  $F_f$  која се састоји од статичког трења, Кулоновог трења и Стрибековог ефекта брзине. Детаљнија истраживања која се односе на дејста сила трења на хидрауличке серво актуаторе може се пронаћи у [50].

Динамичке једначине модела се могу поједноставити увођењем променљиве притисака оптерећења дефинисане као

$$p_L = p_a - \alpha p_b. \quad (5.11)$$

Сада је нови вектор стања

$$[x_1(t) \ x_2(t) \ x_3(t)]^T = [y(t) \ \dot{y}(t) \ p_L(t)]^T,$$

што омогућује да модел ХСА изразимо у компактнијем облику. Ако узмемо радну тачку

$$x_0 = [y_0 \ \dot{y}_0 \ p_{L0}]^T,$$

и линеаризујемо систем око задате радне тачке, добићемо линеарни модел у простору стања

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (5.12)$$

$$y(t) = Cx(t), \quad (5.13)$$

са следећим матрицама система

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{B_C}{m_t} & \frac{A_a}{m_t} \\ 0 & -K_d & K_p \end{bmatrix},$$

$$B = [0 \ 0 \ K_x]^T,$$

$$C = [1 \ 0 \ 0].$$

Коефицијенти осетљивости у матрици стања система и матрици улаза су дефинисани следећим изразима

$$K_d = A_a \left( \frac{\beta_e}{V_A} + \alpha^2 \frac{\beta_e}{V_B} \right)^{-1}, \quad (5.14)$$

$$K_p = \frac{\beta_e (K_{pA} - C_{Li} (1 + \alpha^2))}{V_A (1 + \alpha^2)} - \frac{\alpha \beta_e (K_{pB} \alpha^2 + C_{Li} (1 + \alpha^2))}{V_B (1 + \alpha^2)}, \quad (5.15)$$

$$K_x = \frac{\beta_e}{V_A} K_{xA} - \alpha \frac{\beta_e}{V_B} K_{xB}, \quad (5.16)$$

у којима је појачање протока у зависности од притиска у коморама цилиндра одређено изразима

$$K_{pA} = \begin{cases} \frac{-c_v x_{v0}}{\sqrt{p_s - p_{A0}}} & \text{за } x_v > 0 \\ \frac{-c_v x_{v0}}{\sqrt{p_{A0} - p_0}} & \text{за } x_v < 0 \end{cases}, \quad (5.17)$$

$$K_{pB} = \begin{cases} \frac{-c_v x_{v0}}{\sqrt{p_{B0} - p_0}} & \text{за } x_v > 0 \\ \frac{-c_v x_{v0}}{\sqrt{p_s - p_{B0}}} & \text{за } x_v < 0 \end{cases}, \quad (5.18)$$

док је појачање протока у зависности од померања вентила је одређено изразима

$$K_{xA} = \begin{cases} c_v \sqrt{p_s - p_{A0}} & \text{за } x_v > 0 \\ -c_v \sqrt{p_{A0} - p_0} & \text{за } x_v < 0 \end{cases}, \quad (5.19)$$

$$K_{xB} = \begin{cases} -c_v \sqrt{p_{B0} - p_0} & \text{за } x_v > 0 \\ c_v \sqrt{p_s - p_{B0}} & \text{за } x_v < 0 \end{cases}. \quad (5.20)$$

Поменути појачања имају велики утицај на стабилност и остале карактеристике система [51]. Наиме, појачање протока  $K_x$  има директни утицај на стабилност ХСА, зато што директно утиче на појачање отвореног кола ХСА. Такође, директан утицај на степен пригушења има појачање  $K_p$ . Због тога, коефицијент осетљивости на притисак  $K_{px} = \frac{K_x}{K_p}$  је изузетно велик, што објашњава способност ХСА да савлада велика оптерећења изазвана трећем уз малу грешку.

Основни услов за енергетску ефикасност у многим хидрауличким индустријским системима је добро пројектован оптимални регулатор. Да би се приказала ефикасност алгоритма управљања заснованог на АДП-у у случају непознате динамике и немерљивих стања система у овом делу биће приказани резултати симулација на моделу хидрауличног серво актуатора описаног једначинама (5.12)-(5.13).

За потребе симулација, модел хидрауличног серво актуатора је дискретизован за период одабирања  $h = 0.1$  s, уз задржавање нултог реда (енг. zero-order holder). Апроксимирана оптимална матрица појачања и индекс перформансе за дискретизован модел ХСА добијају се итеративно.

Ефикасност алгоритма управљања заснованог на АДП-у биће разматрана на моделу ХСА (5.12)-(5.13) са вредностима параметара датим у табели 5.2.

Параметар	Вредност
$B_C$	200 N s/m
$p_S$	45 bar
$p_0$	1.6 bar
$\beta_e$	$2 \cdot 10^8$ Pa
$m$	25 kg
$V_{a0}, V_{b0}$	$8.2 \cdot 10^{-6}$ m <sup>3</sup>
$K_e$	$10^{-1}$
$A_a$	$4.91 \cdot 10^{-4}$ m <sup>2</sup>
$A_b$	$2.43 \cdot 10^{-4}$ m <sup>2</sup>
$c_{Li}$	$5 \cdot 10^{-14}$
$c_{vi}$	1.15
$L$	1 m

Табела 5.2: Вредности параметара хидрауличног серво актуатора коришћених у симулацијама

Да би се демонстрирале перформансе предложеног управљања ХСА-ом заснованог на АДП-у, за вредности тежинских матрица,  $Q$  и  $R$ , узете су јединичне матрице, индекс осмотривости  $N = 3$ , почетни вектор стања  $x_0 = [3 \ 1 \ 3]$  а за праг конвергенције  $\varepsilon$  узета је вредност  $10^{-10}$ .

Решавањем једначине (2.18) добијене су следеће оптималне вредности

$$\bar{P}_d^* = \begin{bmatrix} 269.1657 & -377.7014 & 139.5220 & 7.7063 & -276.6863 & 293.3140 \\ -377.7014 & 535.9737 & -199.3177 & -10.0756 & 387.6731 & -419.0086 \\ 139.5220 & -199.3177 & 74.4521 & 3.7667 & 3.7667 & 156.5180 \\ 7.7063 & -10.0756 & 3.7667 & 1.5862 & -7.2146 & 7.9186 \\ -276.6830 & 387.6731 & -142.9420 & -7.2146 & 284.9547 & -300.5013 \\ 293.3140 & -419.0086 & 156.5180 & 7.9186 & -300.5013 & 329.0404 \end{bmatrix}, \quad (5.21)$$

$$\bar{K}_d^* = [3.3615 \ -4.4696 \ 1.6358 \ 0.3700 \ -3.3307 \ 3.4388]. \quad (5.22)$$



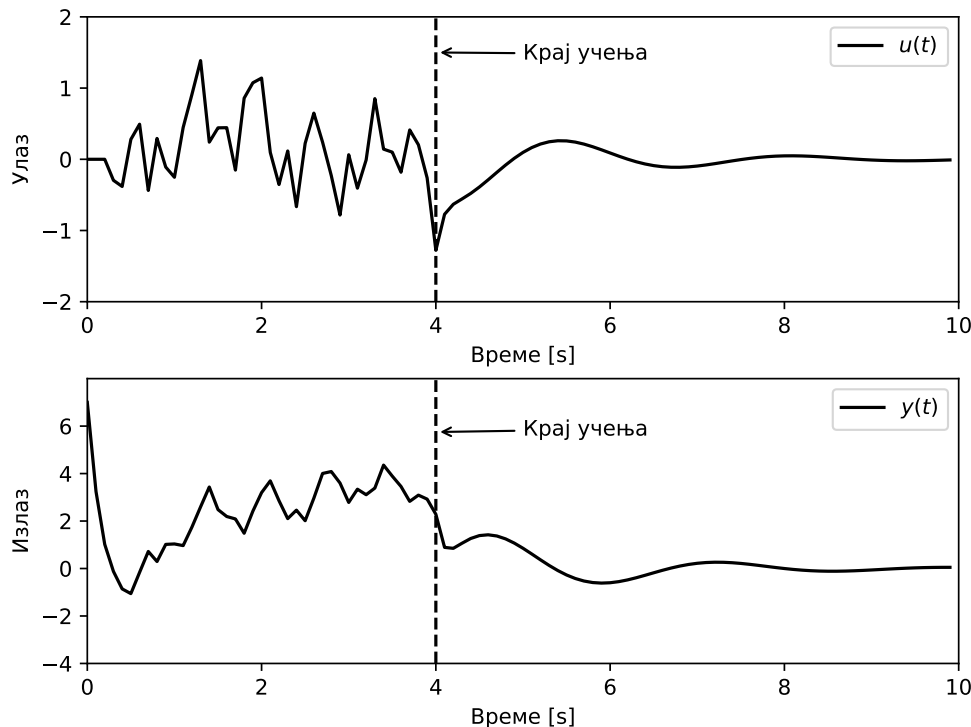
### 5.1.1 Управљање засновано на АДП-у

Фаза учења током рада траје од 0.8 до 4 секунде, у коме се улазно-излазни подаци прикупљају, при чему се конвергенција постиже након 7 итерација. На основу структуре вектора  $z_k$ , може се приметити да није могуће скупити измерене податке од  $k = 0$  до  $k = N - 1$ , а последица овога је да фаза учења не може почети од 0 секунде, већ се робусни закон управљања примењује у првих 0.8 секунди. Након 4 секунде, почиње итеративно рачунање оптималних вредности матрице појачања и индекса перформансе, при чему се конвергенција постиже после 7 итерација. Из следећих израза, може се видети да су добијене апроксимирани вредности  $\bar{P}_7^*$  и  $\bar{K}_7^*$  приближне њиховим оптималним вредностима (5.21)-(5.22).

$$\bar{P}_7^* = \begin{bmatrix} 269.1659 & -377.7057 & 139.5221 & 7.7082 & -276.6863 & 293.3106 \\ -377.7057 & 535.9646 & -199.3158 & -10.0757 & 387.6722 & -419.0119 \\ 139.5221 & -199.3158 & 74.4526 & 3.7667 & -142.9427 & 156.5181 \\ 7.7082 & -10.0757 & 3.7667 & 1.5861 & -7.2146 & 7.9186 \\ -276.6863 & 387.6722 & -142.9427 & -7.2146 & 284.9548 & -300.5016 \\ 293.3106 & -419.0119 & 156.5181 & 7.9186 & -300.5016 & 329.0404 \end{bmatrix}, \quad (5.23)$$

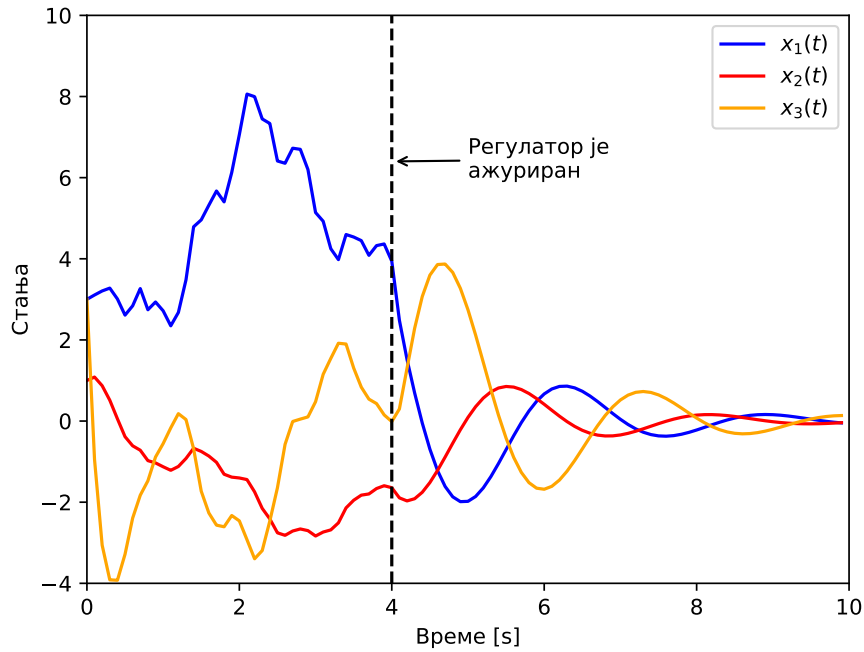
$$\bar{K}_7^* = [3.3615 \quad -4.4695 \quad 1.6358 \quad 0.3700 \quad -3.3306 \quad 3.4387]. \quad (5.24)$$

На слици 5.2 приказан је управљачки сигнал и излаз оригиналног временски-континуалног линеарног модела ХСА. Након 4 секунде, закон управљања је ажуриран и научно управљање се може применити. Као оптимални истраживачки шум се користи сума синусоида, са фреквенцијама  $\omega_i \in [-0.944, 2.116, -0.349]$ .



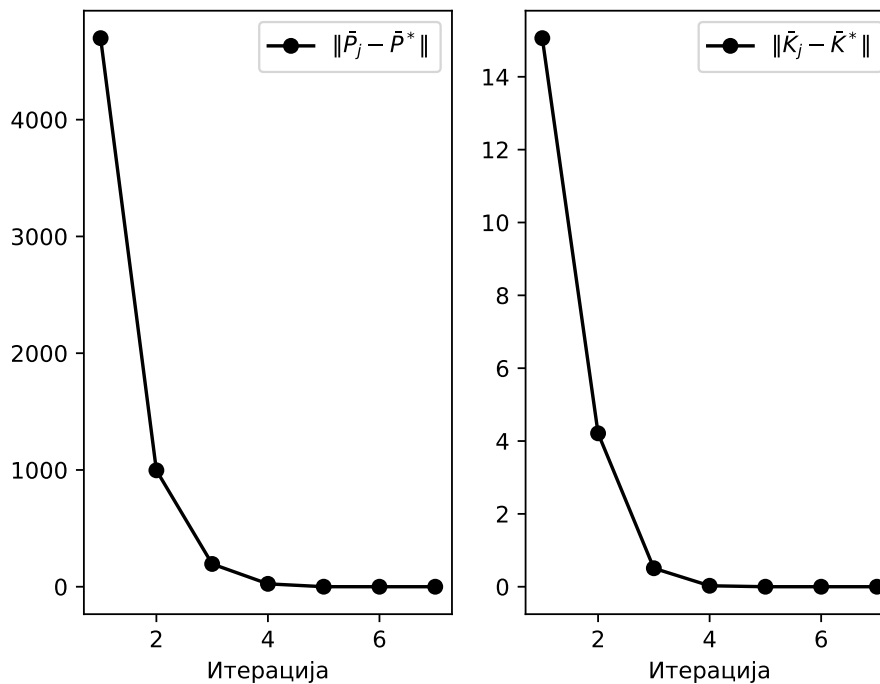
Слика 5.2: Улазни и излазни сигнал ХСА.

Кретање система управљаног регулатором заснованим на АДП-у приказано је на слици 5.3.



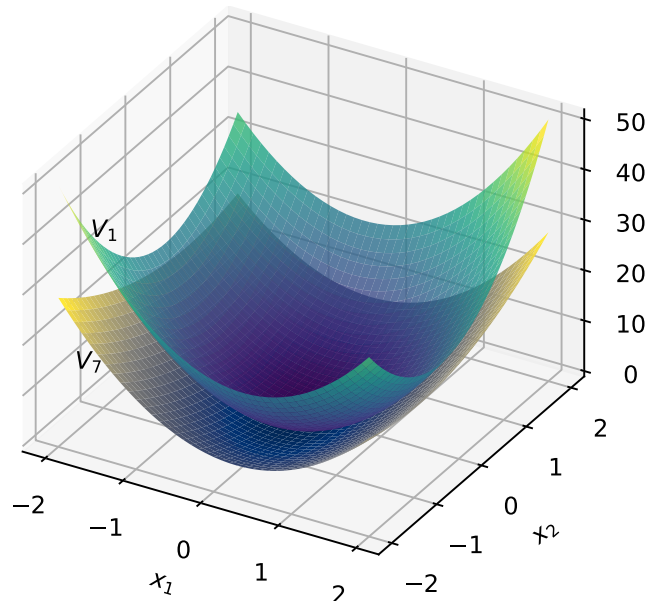
Слика 5.3: Трајекторије стања ХСА.

На слици 5.4 приказана је грешка између  $\bar{P}_j$  и  $\bar{P}^*$ , као и грешка између  $\bar{K}_j$  и  $\bar{K}^*$ . На овој слици може се уочити конвергенција ка оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ .

Слика 5.4: Конвергенција вредности апроксимираних матрица  $\bar{P}$  и  $\bar{K}$  ка оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ .

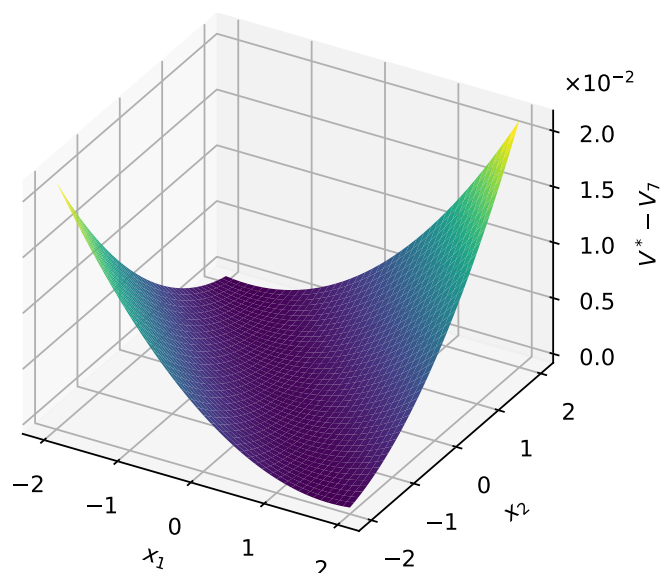
Промена максималне вредности критеријума у случају ХСА приказан је на слици 5.5, где  $V_1$  представља максималну вредност критеријума за почетни закон

управљања, док  $V_7$  представља максималну вредност критеријума за ажуриран закон управљања након 7 итерација. Са слике се може видети да се апроксимирана функција циља  $V_7$  значајно смањила у односу на почетну вредност  $V_1$ .



Слика 5.5: Упоредни приказ вредносне функције за прву и последњу итерацију фазе учења.

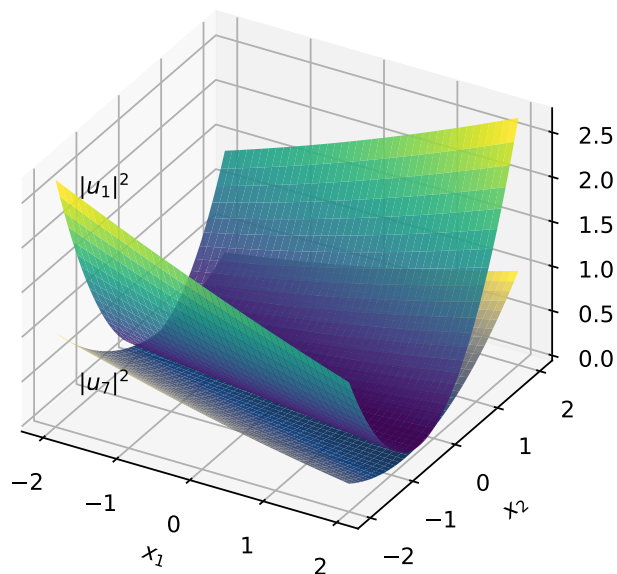
Слика 5.6 приказује грешку апроксимације функције циља. Ова грешка је близу нуле, што потврђује да се током процеса учења постиже добра апроксимација оптималне функције циља.



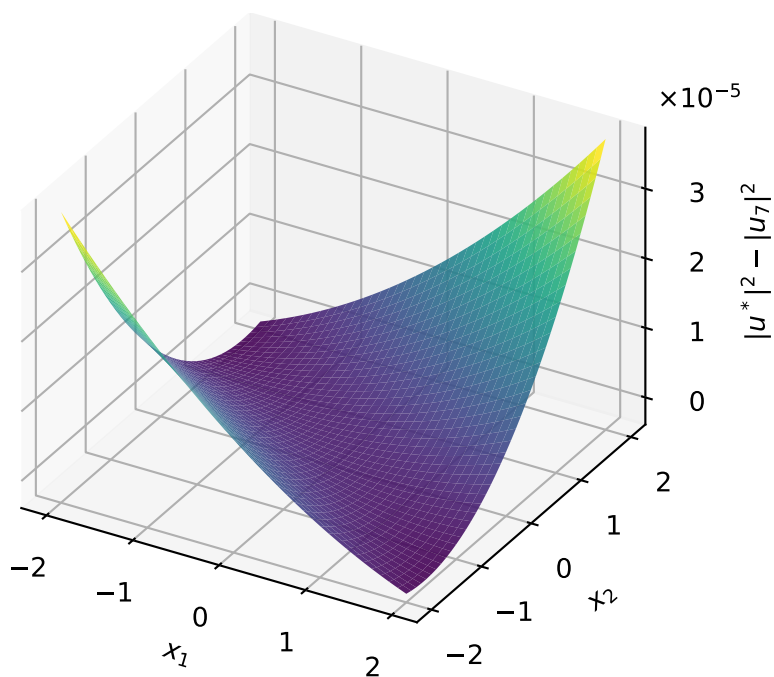
Слика 5.6: Грешка између оптималне и апроксимиране вредносне функције.

Почетни и унапређењи закон управљања након 7. итерације су упоређени на слици 5.7. Такође, на слици 5.8 је приказана и разлика између апроксимираног и

оптималног управљања. Са слика се може видети да су вредности грешака близу нуле, што потврђује да је постигнута добра апроксимација током процеса учења.



Слика 5.7: Упоредни приказ закона управљања за прву и последњу итерацију фазе учења.



Слика 5.8: Грешка између оптималног и апроксимираног управљачког сигнала.

### 5.1.2 Управљање засновано на АДП-у покренуто догађајима

У овом делу биће представљени резултати за предложени алгоритам управљања заснованог на АДП-у покренуто догађајима примењеног на ХСА-у. У случају непознате динамике и немерљивих стања ХСА, можемо користити методу засновану на АДП-у. Због тога, урађене су симулације ХСА описаног линеаризованим временски-континуалним моделом у простору стања (5.12)-(5.13) да би се показала ефикасност управљачког алгоритма заснованог на АДП-у. Основни услов за уштеду енергије у многим хидрауличким индустријским системима је добро пројектовано управљање ХСА-ом засновано на АДП-у покренуто догађајем.

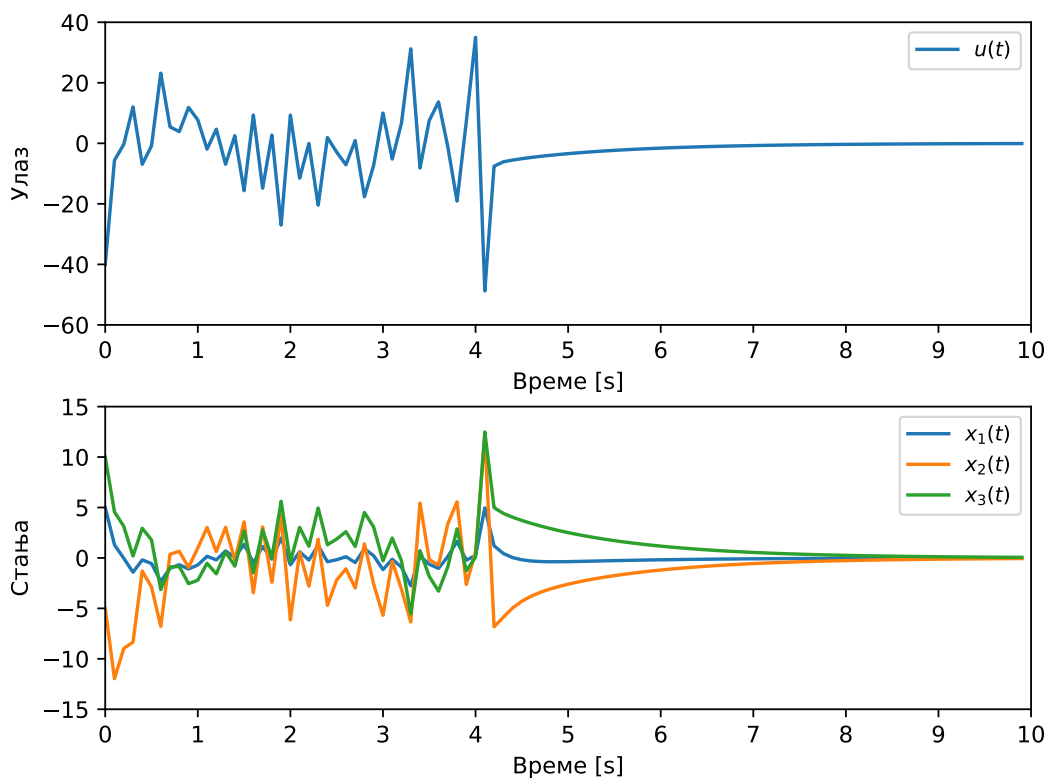
За потребе симулација, модел ХСА је дискретизован периодичним одабирањем  $h = 0.1$  s уз задржавање нултог реда. Апроксимирано оптимално појачање и индекс перформансе за дискретни модел ХСА су итеративно одређени.

Ефикасност алгоритма управљања заснованог на АДП-у приказана је на моделу ХСА описаног (5.12)-(5.13) за вредности параметара који су дати у Табели 5.2.

За потребе симулације методе АДП-а покренуто догађајима на ХСА-у, за вредности тежинских матрица  $Q$  и  $R$  узете су јединичне матрице. Индекс осмотривости је  $N = 3$ , почетно стање  $x_0 = [5, -5, 10]$ , док је праг конвергенције  $\varepsilon = 10^{-10}$ .

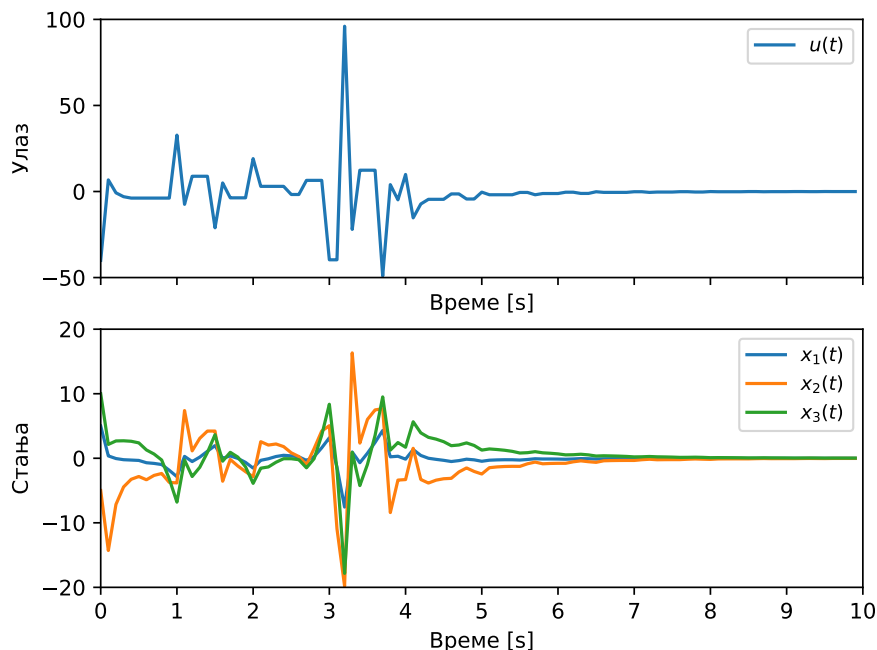
Сама метода управљања заснована на АДП-у покренуто догађајима не захтева тачно познавање матрица система. Међутим, због потреба верификације резултата преко симулација, претпостављено је да су матрице у једначинама (5.12) и (5.13) познате.

На слици 5.9 приказан је управљачки сигнал и стања ХСА која су добијена применом регулатора заснованог на АДП-у са периодичним узорковањем.



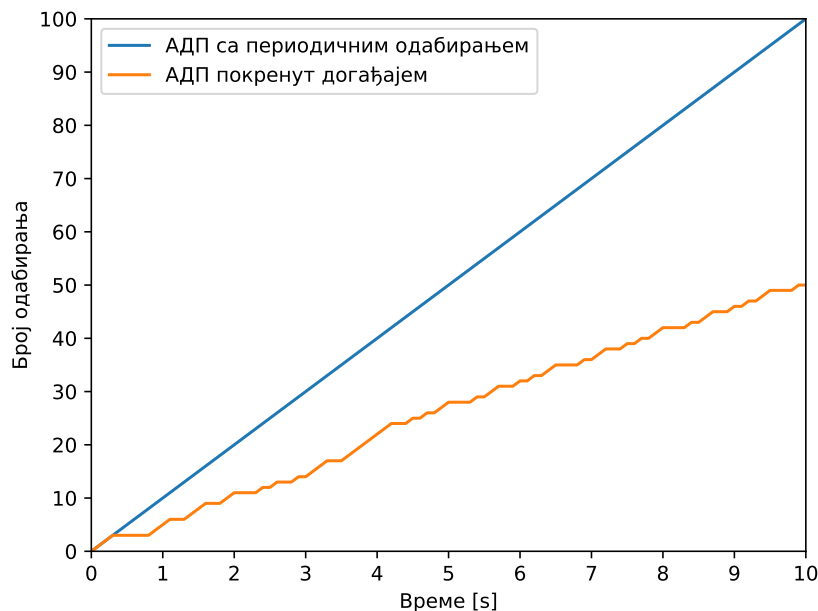
Слика 5.9: Управљање и стања модела ХСА са регулатором заснованим на АДП-у.

На слици 5.10 приказано је управљање и стања ХСА за случај управљања заснованог на АДП-у покренутог догађајима.



Слика 5.10: Управљање и стања модела ХСА са регулатором заснованим на АДП-у покренутог догађајима.

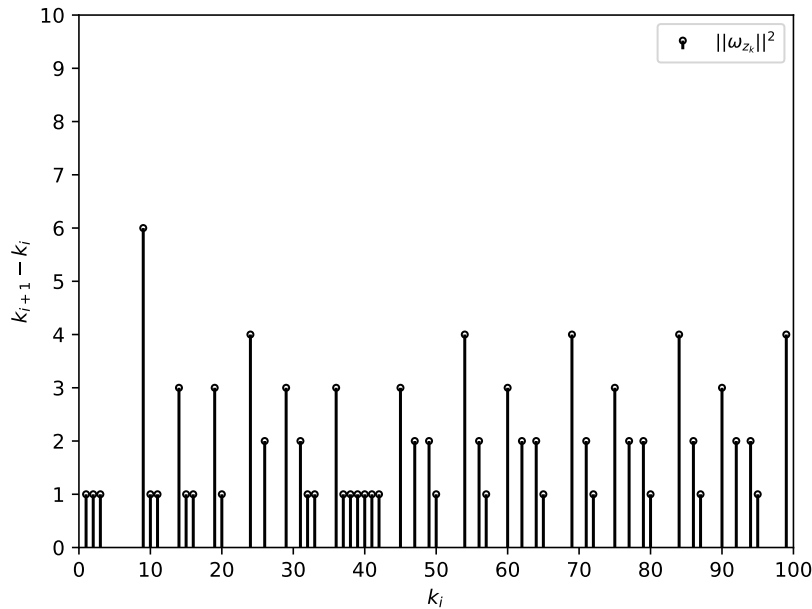
Упоредни приказ бројева узорковања у случају регулатора заснованог на АДП-у покренутог догађајима у односу на регулатор заснованог на АДП-у са периодичним одабирањем приказан је на слици 5.11.



Слика 5.11: Упоредни приказ укупног броја одбирака.

Обе методе дају сличне управљачке ефекте, међутим, у случају методе АДП-а покренутог догађајима, управљачки сигнал се ажурира само када квадратна норма

грешке покретања достигне праг. У супротном се одржава константна вредност. Такође, показано је да се комуникација између регулатора и ХСА смањила за 54% употребном методе АДП-а покренутог догађајима у односу на методу АДП-а са периодичним одабирањем. Секвенца одабирања регулатора покренутог догађајима приказана је на слици 5.12.



Слика 5.12: Секвенца одабирања покренутог догађајима.

### 5.1.3 Управљање засновано на АДП-у толерантно на отказе

Посматра се линеарни ХСА са отказом актуатора  $f_a(t)$ :

$$\dot{x}(t) = Ax(t) + B(t)(u(t) - f_a(t)), \quad (5.25)$$

$$y(t) = Cx(t), \quad (5.26)$$

где су вредности параметара ХСА исте као и у претходним случајевима.

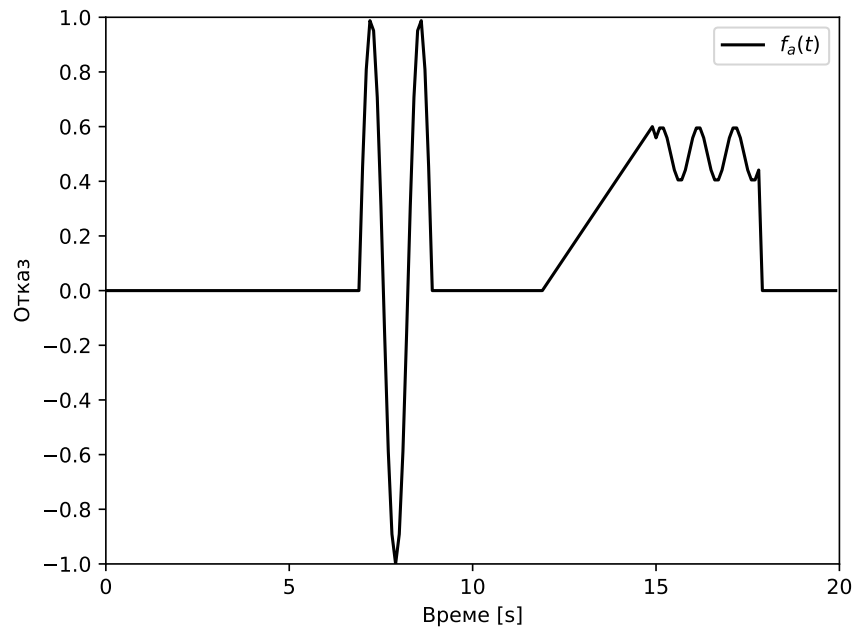
Члан  $f_a(t)$  означава непознати адитивни отказ актуатора описан изразом:

$$f_a(t) = \begin{cases} 0 & 0 \leq t \leq 6 \text{ s} \\ \sin(0.15\pi \cdot (t - 6)) & 6 \text{ s} < t \leq 10 \text{ s} \\ 0 & 10 \text{ s} < t \leq 12 \text{ s} \\ 0.02 \cdot (t - 12) & 12 \text{ s} < t \leq 15 \text{ s} \\ 0.5 + 0.1 \cdot \sin(0.2\pi \cdot (t - 15)) & 15 \text{ s} < t \leq 18 \text{ s} \\ 0 & 18 \text{ s} < t \leq 20 \text{ s} \end{cases}, \quad (5.27)$$

који је коришћен за потребе симулација, слика 5.13. Као што се може видети са слике, отказ који је коришћен за потребе симулација комбинује типове отказа различитих временских карактеристика: периодичне, постепене и нагле.

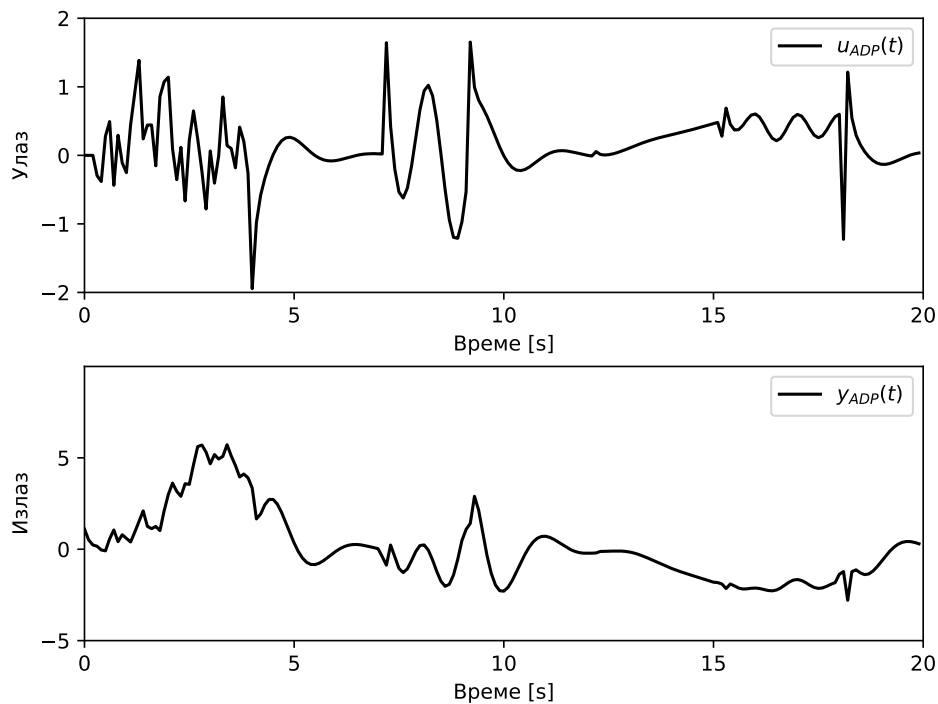
Да би се показале перформансе предложеног управљања заснованог на АДП-у, за тежинске матрице  $Q$  и  $R$  изабране су јединичне матрице. Индекс осмотривости

$N = 3$ , почетни вектор стања који је коришћен у симулацијама је  $x_0 = [3, 1, 3]$ , док је праг конвергенције  $\varepsilon = 10^{-10}$ .



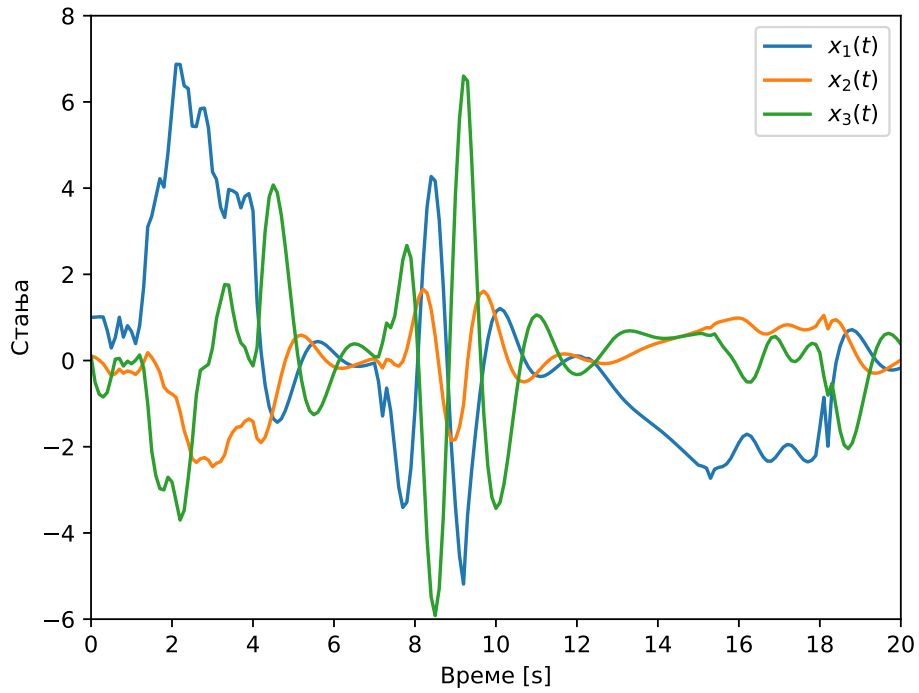
Слика 5.13: Приказ отказа  $f_a(t)$ .

На сликама 5.14 и 5.15 приказани су улаз и излаз, као и стања система са управљањем заснованим на АДП-у без компензације дејства отказа. Када дође до непознатог отказа актуатора, закон управљања заснован на АДП-у  $u_k$  не може да гарантује стабилост система. Да би се отказ компензовао у току рада, потребно је извршити адаптивну процену отказа.



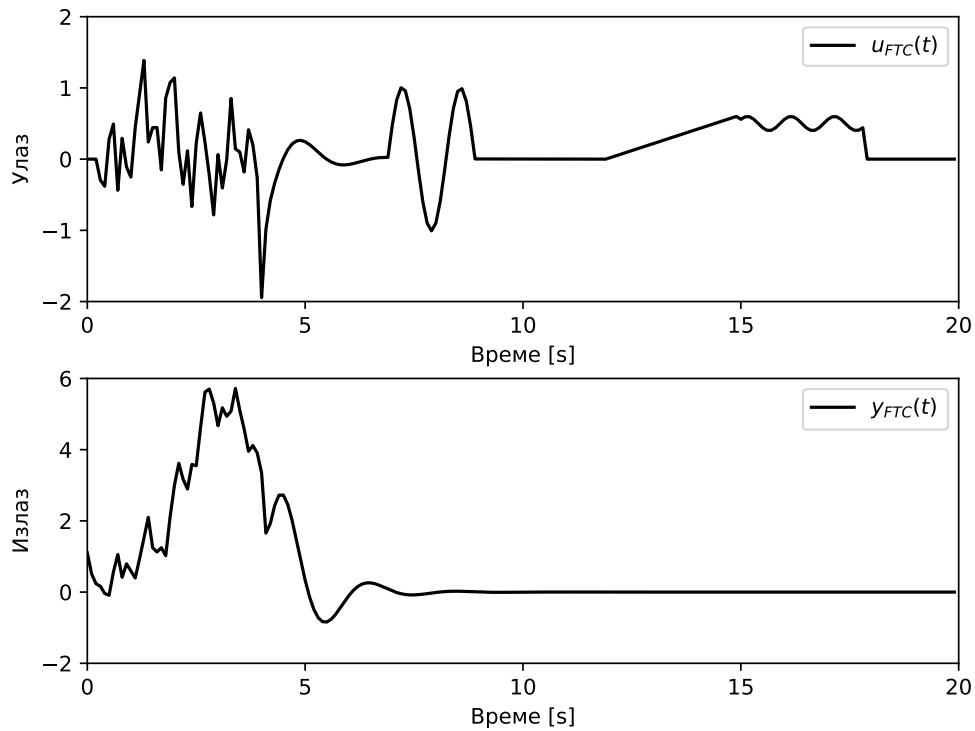
Слика 5.14: Улаз и излаз система без компензације отказа.



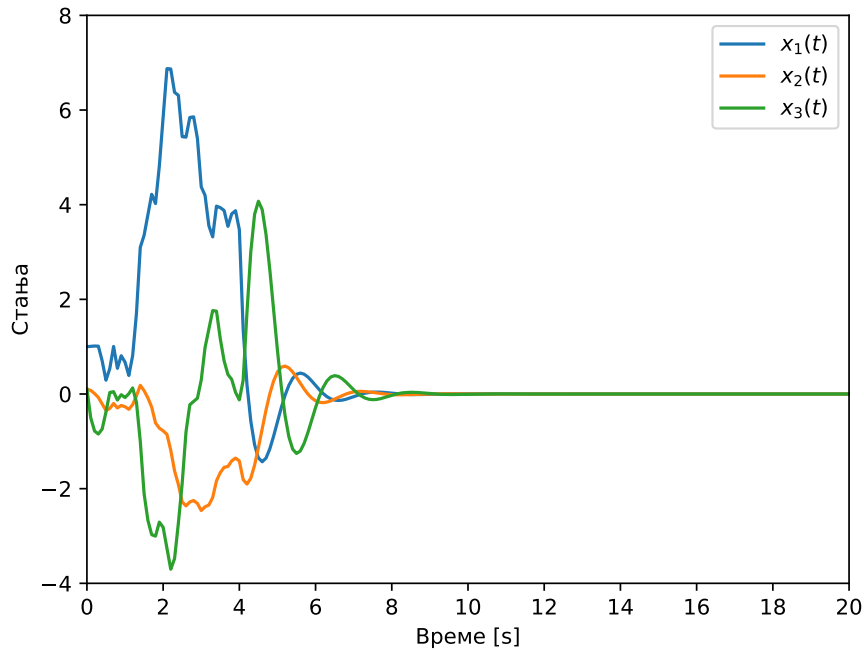


Слика 5.15: Стања система без компензације отказа.

Улаз и излаз ХСА са управљањем заснованим на АДП-у и компензацијом отказа приказани су на слици 5.16, док су стања приказана на слици 5.17. На слици се може јасно видети да стања конвергирају равнотежном стању у кратком периоду након што се отказ појави.



Слика 5.16: Улаз и излаз система са компензованим отказом.



Слика 5.17: Стања система са компензованим отказом.

Можемо видети да својства управљања заснованог на АДП-у и компензацији отказа су заснована на тачности процене отказа, али је временско кашњење узроковано проценом отказа и даље присутно. Након процене параметара, перформансе управљања се могу у великој мери повратити.

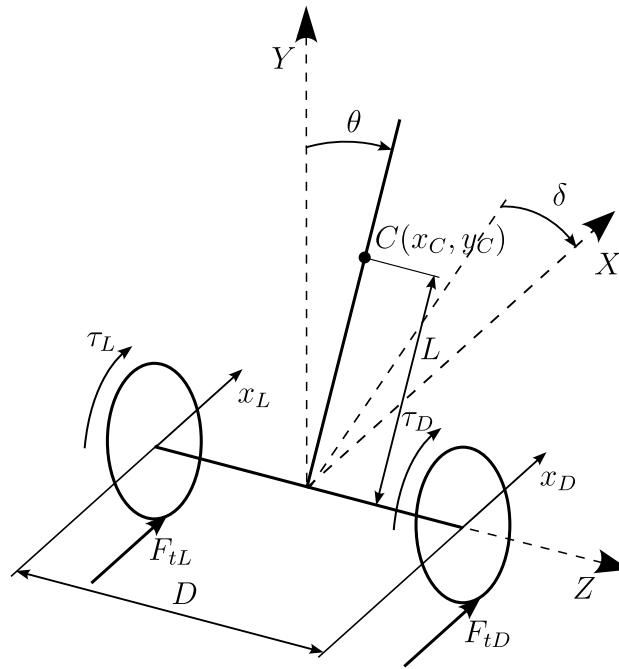
## 5.2 Самобалансирајући робот

Самобалансирајући робот представља мултиваријабилни непотпуно активирани (енг. *underactuated*) механички систем [129, 130]. Непотпуно активирани механички системи имају већи број генерисаних координата од броја актуатора. Код ових система, генерисаним координатама се управља индиректно, преко њихове међусобне повезаности. Такође, самобалансирајући робот је нелинеаран и нестабилан систем вишег реда.

Робот се састоји од шасије за коју су причвршћени точкови које покрећу електромотори, слика 5.18. Робот има три степена слободе (генерисане координате):

1. угао нагиба  $\theta$
2. линеарно кретање  $x$
3. угао скретања  $\delta$

Ове три генерисане координате и њихове одговарајуће брзине у потпуности описују динамику робота и представљају елементе вектора стања.



Слика 5.18: Самобалансирајући робот.

Роботом се управља помоћу два електромотора који покрећу одговарајуће тачкове моментима  $\theta_L$  и  $\theta_D$ . У табели 5.3 дати су параметри самобалансирајућег робота.

Симбол	Опис
$M$	маса шасије
$m$	маса тачка
$R$	пречник тачка
$D$	растојање између тачкова
$L$	растојање између центра масе робота и осе $z$
$J_y$	момент инерције шасије у односу на $y$ осу
$J_z$	момент инерције шасије у односу на $z$ осу
$J_T$	момент инерције тачка у односу на $z$ осу
$\theta$	угао нагиба робота
$\delta$	угао закретања робота
$x_L, x_D$	померање левог и десног тачка
$x_C, y_C$	координате центра масе робота
$X_L, X_D$	унутрашње силе између тачкова и шасије пројектоване на $x$ осу
$Y_L, Y_D$	унутрашње силе између тачкова и шасије пројектоване на $y$ осу
$F_{tL}, F_{tD}$	силе трења између тачкова и подлоге
$\tau_L, \tau_D$	моменти мотора левог и десног тачка
$\theta_L, \theta_D$	углови закретања левог и десног тачка

Табела 5.3: Параметри самобалансирајућег робота.

Динамика тачкова дефинисана је сумама сила

$$m\ddot{x}_L = F_{tL} - X_L, \quad (5.28)$$

$$m\ddot{x}_D = F_{tD} - X_D, \quad (5.29)$$

и сумама момената

$$J_T \ddot{\theta}_L = \tau_L - F_{tL} R, \quad (5.30)$$

$$J_T \ddot{\theta}_D = \tau_D - F_{tD} R. \quad (5.31)$$

Динамика шасије описана је сумама сила за осе  $x$  и  $y$ , као и сумама момената за осе  $z$  и  $y$  на следећи начин

$$M \ddot{x}_C = X_L + X_D, \quad (5.32)$$

$$M \ddot{y}_C = Y_L + Y_D - Mg, \quad (5.33)$$

$$J_z \ddot{\theta} = (Y_L + Y_D) L \sin \theta - (X_L + X_D) L \cos \theta - (\tau_L + \tau_D), \quad (5.34)$$

$$J_y \ddot{\delta} = \frac{D}{2} (X_L - X_D). \quad (5.35)$$

Ако претпоставимо да нема проклизавања тачкова, онда важи да је  $x_L = R\theta_L$  и  $x_D = R\theta_D$ . На основу овога, можемо средити једначине (5.28)-(5.31) и из њих елиминисати углове закретања тачкова  $\theta_L$  и  $\theta_D$ , као и силе трења на тачковима  $F_{tL}$  и  $F_{tD}$ . Као резултат добијамо следеће једначине које описују динамику левог и десног тачка

$$\left(m + \frac{J_T}{R^2}\right) \ddot{x}_L = \frac{1}{R} \tau_L - X_L, \quad (5.36)$$

$$\left(m + \frac{J_T}{R^2}\right) \ddot{x}_D = \frac{1}{R} \tau_D - X_D. \quad (5.37)$$

Однос између угла закретања  $\delta$  и померања тачкова  $x_L$  и  $x_D$  дат је изразом

$$D\delta = x_L - x_D. \quad (5.38)$$

Центар масе робота одређен је координатама

$$x_C = x + L \sin \theta, \quad (5.39)$$

$$y_C = L \cos \theta, \quad (5.40)$$

где је

$$x = \frac{1}{2} (x_L + x_D). \quad (5.41)$$

На основу горњих једначина добијамо нелинеарне једначине система

$$\ddot{x} \left(M + 2m + \frac{2J_T}{R^2}\right) + ML \left(\ddot{\theta} \cos \theta - \dot{\theta}^2 \sin \theta\right) = \frac{1}{R} (\tau_L + \tau_D) \quad (5.42)$$

$$\left(\frac{2J_y}{D} + \frac{DJ_T}{R^2} + Dm\right) \ddot{\delta} = \frac{1}{R} (\tau_L - \tau_D), \quad (5.43)$$

$$J_z \ddot{\theta} = 2\ddot{x}L \left(m + \frac{J_T}{R^2}\right) \cos \theta + MgL \sin \theta - ML^2 \ddot{\theta} \sin^2 \theta - ML^2 \dot{\theta}^2 \sin \theta \cos \theta - \left(1 + \frac{L \cos \theta}{R}\right) (\tau_L + \tau_D). \quad (5.44)$$

Линеаризацијом ових нелинеарних једначина око радне тачке  $\theta = 0$  добијамо следећи модел система у простору стања

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \\ \dot{\theta} \\ \ddot{\theta} \\ \dot{\delta} \\ \ddot{\delta} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & a_{43} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \theta \\ \dot{\theta} \\ \delta \\ \dot{\delta} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ b_{21} & b_{22} \\ 0 & 0 \\ b_{41} & b_{42} \\ 0 & 0 \\ b_{61} & b_{62} \end{bmatrix} \begin{bmatrix} \tau_L \\ \tau_D \end{bmatrix}, \quad (5.45)$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \theta \\ \dot{\theta} \\ \delta \\ \dot{\delta} \end{bmatrix}, \quad (5.46)$$

где је  $[x \ \dot{x} \ \theta \ \dot{\theta} \ \delta \ \dot{\delta}]^T$  вектор стања,  $[\tau_L \ \tau_D]^T$  вектор улаза,  $[x \ \theta \ \delta]^T$  вектор излаза, а параметри модела су дефинисани као

$$\begin{aligned} a_{23} &= \frac{-M^2 L^2 g}{M J_z + 2(J_z + M L^2)(m + J_T/R^2)}, \\ a_{43} &= \frac{M^2 g L + 2M g L(m + J_T/R^2)}{M J_z + 2(J_z + M L^2)(m + J_T/R^2)}, \\ b_{21} = b_{22} &= \frac{(J_z + M L^2)/R + M L}{M J_z + 2(J_z + M L^2)(m + J_T/R^2)}, \\ b_{41} = b_{42} &= \frac{-(R + L)M/R - 2(m + J_T/R^2)}{M J_z + 2(J_z + M L^2)(m + J_T/R^2)}, \\ b_{61} = -b_{62} &= \frac{D/2R}{J_y + \frac{D^2}{2R}(mR + J_T/R)}. \end{aligned}$$

### 5.2.1 Управљање засновано на АДП-у

За потребе симулација уместо реалног система коришћен је линеаризовани временски-дискретни модел у простору стања са два улаза и три излаза. Дискретизација је извршена за период одабирања  $h = 0.1$  s уз задржавање нултог реда. Вектор стања система се састоји од шест елемената. Параметри робота који су коришћени у симулацијама дати су табели 5.4.

Тежинске матрице  $Q$  и  $R$  су јединичне матрице, док је за потребе учења, у периоду трајању од 4 s, коришћен је оптимални истраживачки шум у облику сума синусоида

$$e_i(t) = \sum_{j=1}^6 \sin \omega_{ij} t,$$

где су оптималне фреквенције за оба улаза

$$\omega = \begin{bmatrix} -4.65 & 15.44 & 17.90 & 44.51 & 20.92 & -38.06 \\ 39.09 & -16.58 & 19.87 & -30.21 & -46.94 & 24.40 \end{bmatrix}.$$

Параметар	Вредност
$M$	21 kg
$m$	0.42 kg
$R$	0.106 m
$D$	0.44 m
$L$	0.3 m
$J_y$	0.3388 kg m <sup>2</sup>
$J_z$	0.63 kg m <sup>2</sup>
$J_T$	$2.4 \cdot 10^{-3}$ kg m <sup>2</sup>

Табела 5.4: Параметри самобалансирајућег робота коришћени у симулацијама.

Такође, у симулацијама је као иницијална вредност вектора стања узета  $x_0 = [0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1 \ 0.1]^T$ , док је за праг конвергенције  $\varepsilon = 10^{-10}$ . У овом случају индекс осматривости износи  $N = 6$ .

Коришћењем регулатора заснованог на АДП-у добијене су следеће апроксимирани вредности матрица

$$\bar{P}_7 = \begin{bmatrix} 1.6026 \cdot 10^2 & 5.3492 \cdot 10^{-1} & 2.0597 \cdot 10^1 & 6.3398 & -6.1216 & 3.6217 \cdot 10^{-2} \\ 5.3492 \cdot 10^{-1} & 7.5673 \cdot 10^{-2} & 1.2319 \cdot 10^{-1} & 1.3730 \cdot 10^{-1} & 3.4535 \cdot 10^{-1} & 4.6660 \cdot 10^{-4} \\ 2.0597 \cdot 10^1 & 1.2319 \cdot 10^{-1} & 3.1688 \cdot 10^1 & -1.1052 \cdot 10^{-1} & 9.6155 & 3.1731 \cdot 10^{-3} \\ 6.3398 & 1.3730 \cdot 10^{-1} & -1.1052 \cdot 10^{-1} & 4.6055 & -6.7851 \cdot 10^{-1} & -2.4241 \cdot 10^{-3} \\ -6.1216 & 3.4535 \cdot 10^{-1} & 9.6155 & -6.7851 \cdot 10^{-1} & 2.4401 \cdot 10^1 & 2.2179 \cdot 10^{-3} \\ 3.6217 \cdot 10^{-2} & 4.6660 \cdot 10^{-4} & 3.1731 \cdot 10^{-3} & -2.4241 \cdot 10^{-3} & 2.2179 \cdot 10^{-3} & 5.7339 \cdot 10^{-3} \end{bmatrix}, \quad (5.47)$$

$$\bar{K}_7 = \begin{bmatrix} -3.4134 \cdot 10^{-1} & -6.2813 \cdot 10^{-2} & -9.4974 \cdot 10^{-2} & 4.1142 \cdot 10^{-1} & -3.7725 \cdot 10^{-1} & -9.2785 \cdot 10^{-4} \\ 1.9763 & 1.1756 \cdot 10^{-1} & 1.6552 \cdot 10^{-1} & 2.6958 \cdot 10^{-1} & 9.8224 \cdot 10^{-2} & 9.3404 \cdot 10^{-4} \end{bmatrix}. \quad (5.48)$$

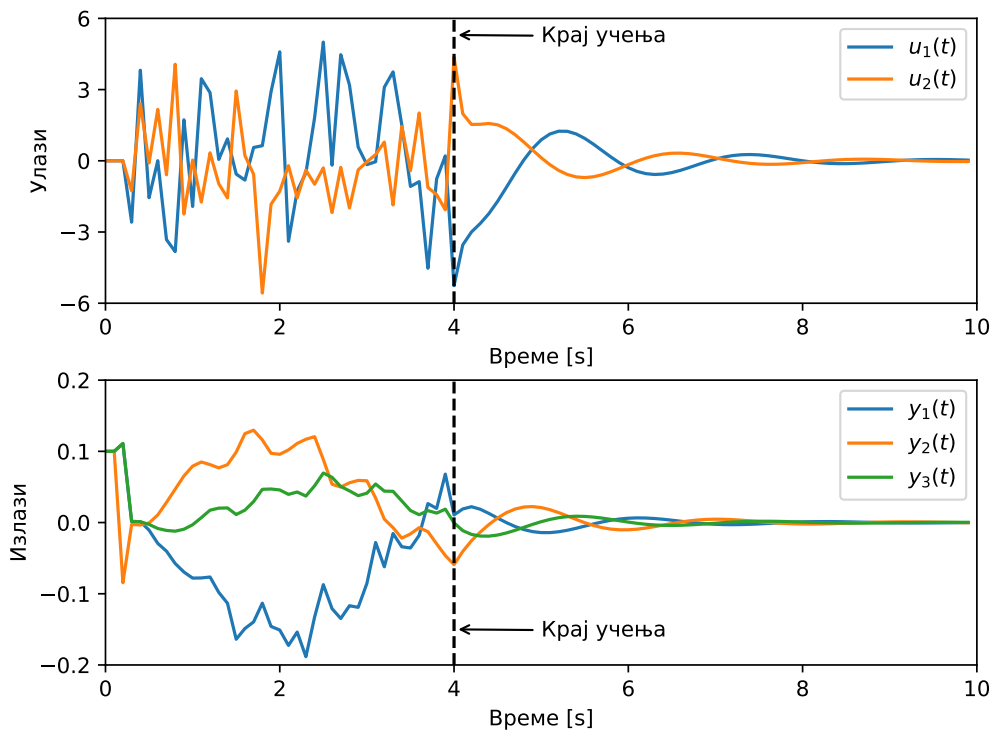
Оптималне вредности матрица  $\bar{P}^*$  и  $\bar{K}^*$  су дате следећим изразима

$$\bar{P}^* = \begin{bmatrix} 1.6026 \cdot 10^2 & 5.3492 \cdot 10^{-1} & 2.0597 \cdot 10^1 & 6.3398 & -6.1216 & 3.6217 \cdot 10^{-2} \\ 5.3492 \cdot 10^{-1} & 7.5673 \cdot 10^{-2} & 1.2319 \cdot 10^{-1} & 1.3730 \cdot 10^{-1} & 3.4535 \cdot 10^{-1} & 4.6660 \cdot 10^{-4} \\ 2.0597 \cdot 10^1 & 1.2319 \cdot 10^{-1} & 3.1688 \cdot 10^1 & -1.1052 \cdot 10^{-1} & 9.6155 & 3.1731 \cdot 10^{-3} \\ 6.3398 & 1.3730 \cdot 10^{-1} & -1.1052 \cdot 10^{-1} & 4.6055 & -6.7851 \cdot 10^{-1} & -2.4241 \cdot 10^{-3} \\ -6.1216 & 3.4535 \cdot 10^{-1} & 9.6155 & -6.7851 \cdot 10^{-1} & 2.4401 \cdot 10^1 & 2.2179 \cdot 10^{-3} \\ 3.6217 \cdot 10^{-2} & 4.6660 \cdot 10^{-4} & 3.1731 \cdot 10^{-3} & -2.4241 \cdot 10^{-3} & 2.2179 \cdot 10^{-3} & 5.7339 \cdot 10^{-3} \end{bmatrix}, \quad (5.49)$$

$$\bar{K}^* = \begin{bmatrix} -3.4134 \cdot 10^{-1} & -6.2813 \cdot 10^{-2} & -9.4974 \cdot 10^{-2} & 4.1142 \cdot 10^{-1} & -3.7725 \cdot 10^{-1} & -9.2785 \cdot 10^{-4} \\ 1.9763 & 1.1756 \cdot 10^{-1} & 1.6552 \cdot 10^{-1} & 2.6958 \cdot 10^{-1} & 9.8224 \cdot 10^{-2} & 9.3404 \cdot 10^{-4} \end{bmatrix}. \quad (5.50)$$

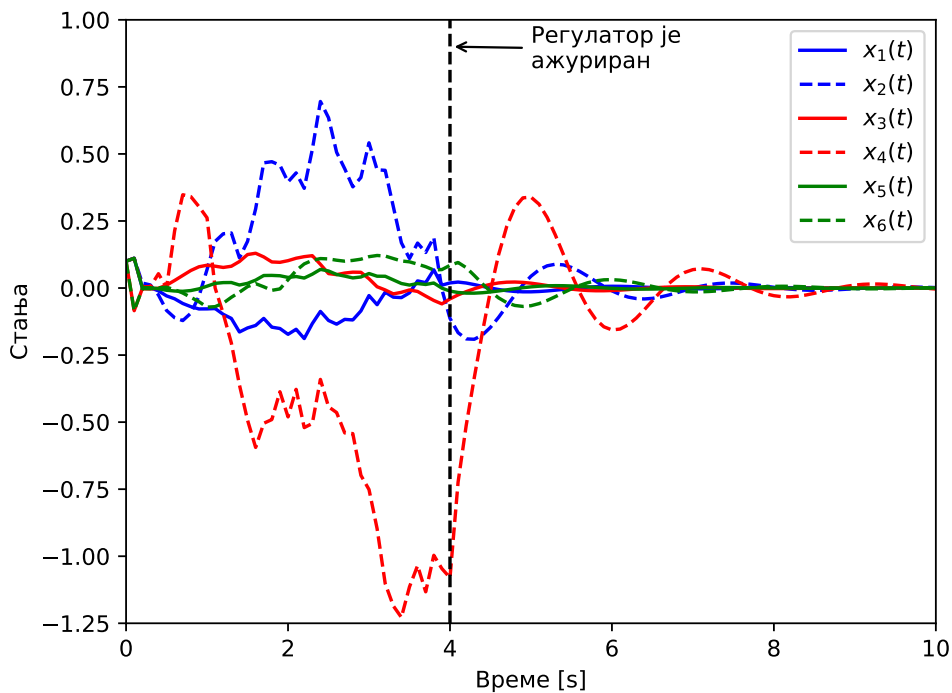
Из претходних израза се може видети да су добијене апроксимирани вредности матрица  $\bar{P}$  и  $\bar{K}$  идентичне њиховим оптималним вредностима на чак 4 децимале.

Улази и излази робота управљаног регулатором заснованим на АДП-у приказан и су на слици 5.19.

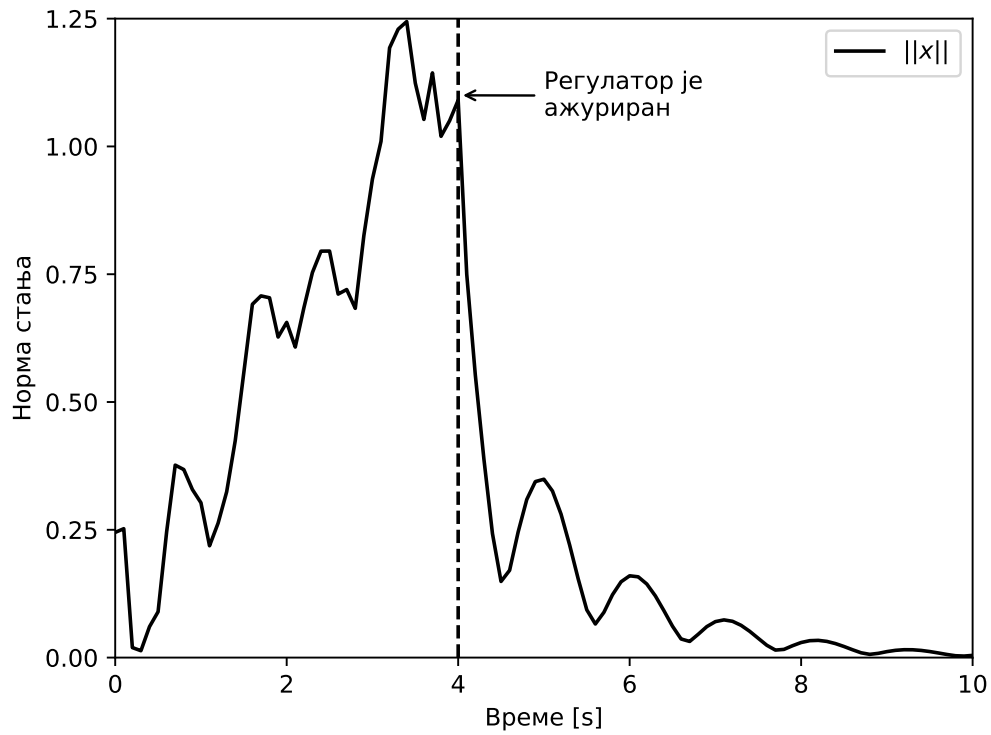


Слика 5.19: Улазни и излазни сигнали робота.

На слици 5.20 приказане су трајекторије стања робота, док је на слици 5.21 приказана промена норме вектора стања током времена. Са слика да може видети да стања конвергирају равнотежном стању брзо након што се закон управљања ажурира.

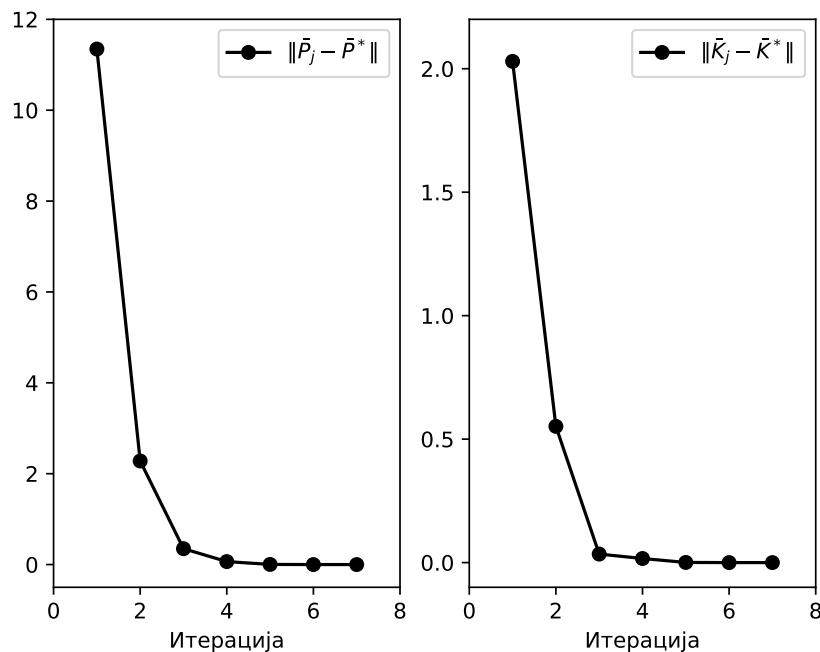


Слика 5.20: Стања робота.



Слика 5.21: Норма вектора стања робота.

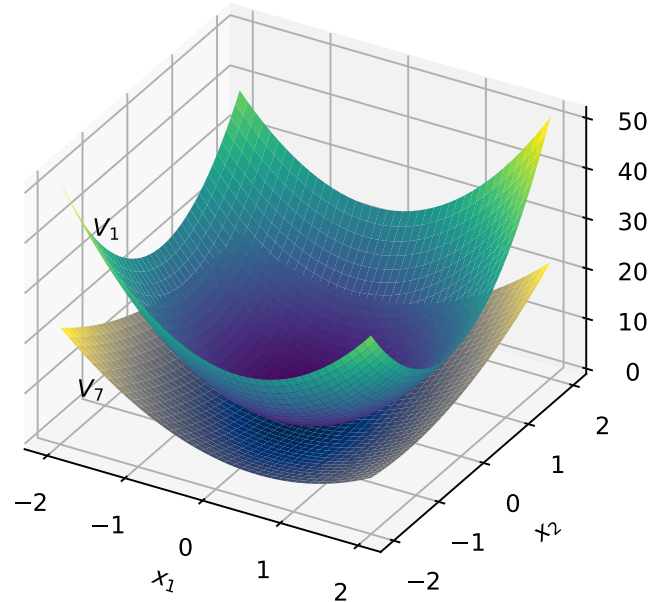
Слика 5.22 приказује конвергенцију апроксимираних вредности матрица  $\bar{P}_j$  и  $\bar{K}_j$  ка оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ .

Слика 5.22: Конвергенција вредности апроксимираних матрица  $\bar{P}$  и  $\bar{K}$  ка оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ .

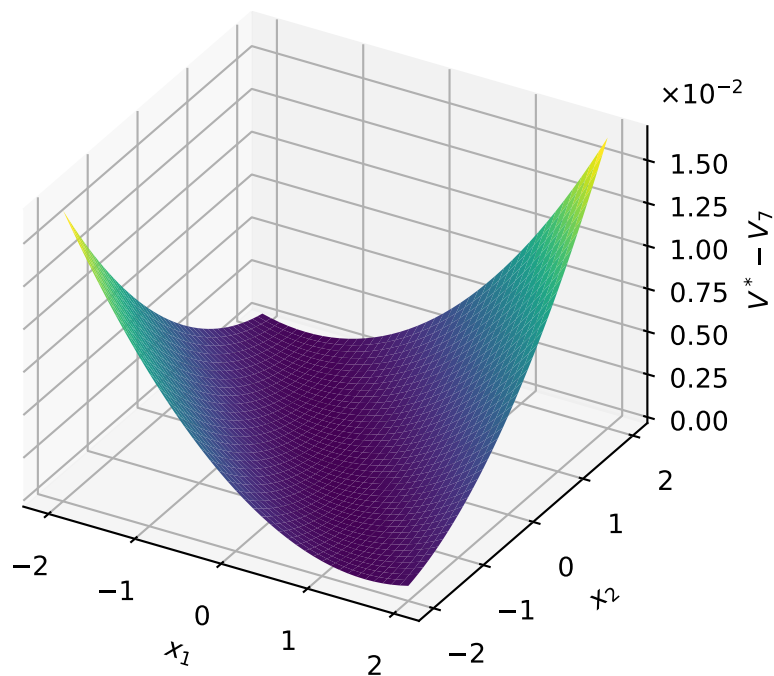
На слици 5.23 дато је поређење апроксимираних вредности функције након сед-



ме итерације,  $V_7$ , и апроксимирани вредносне функције након прве итерације,  $V_1$ . Слика 5.24 приказује разлику између оптималне и апроксимирани вредносне функције. Мале вредности са слике 5.24 имплицирају добро поклапање апроксимирани и оптималне вредносне функције.



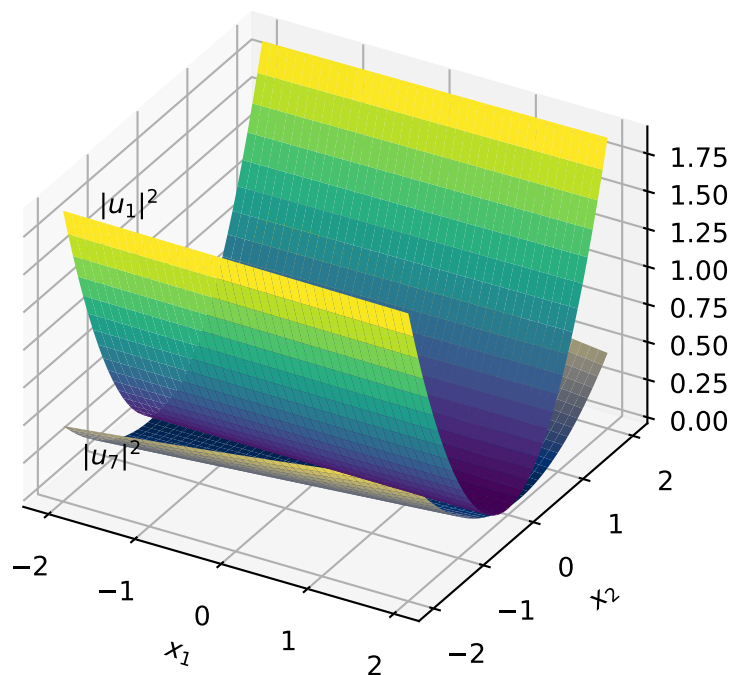
Слика 5.23: Упоредни приказ вредносне функције за прву и последњу итерацију фазе након учења.



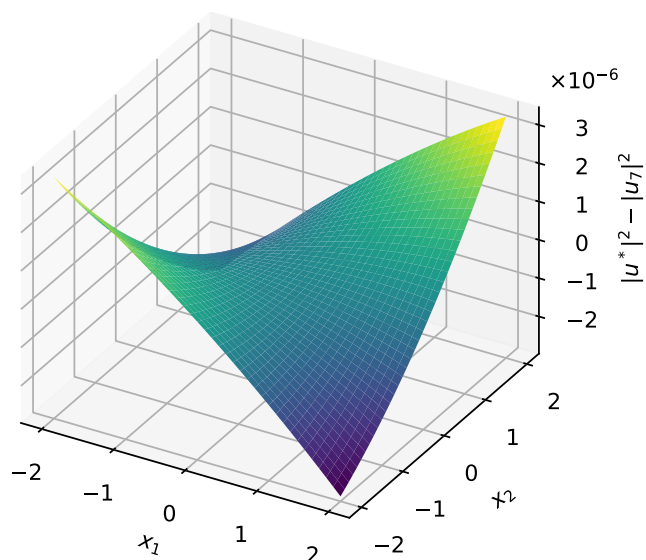
Слика 5.24: Грешка између оптималне и апроксимирани вредносне функције.

Слика 5.25 приказује упоредне апроксимирани законе управљања након седме

итерације,  $u_7$ , и након прве итерације  $u_1$ . Разлика између оптималног и апроксимираног закона управљања приказана је на слици 5.26. На основу слике 5.26 можемо закључити да је добро поклапање апроксимираног и оптималног закона управљања.



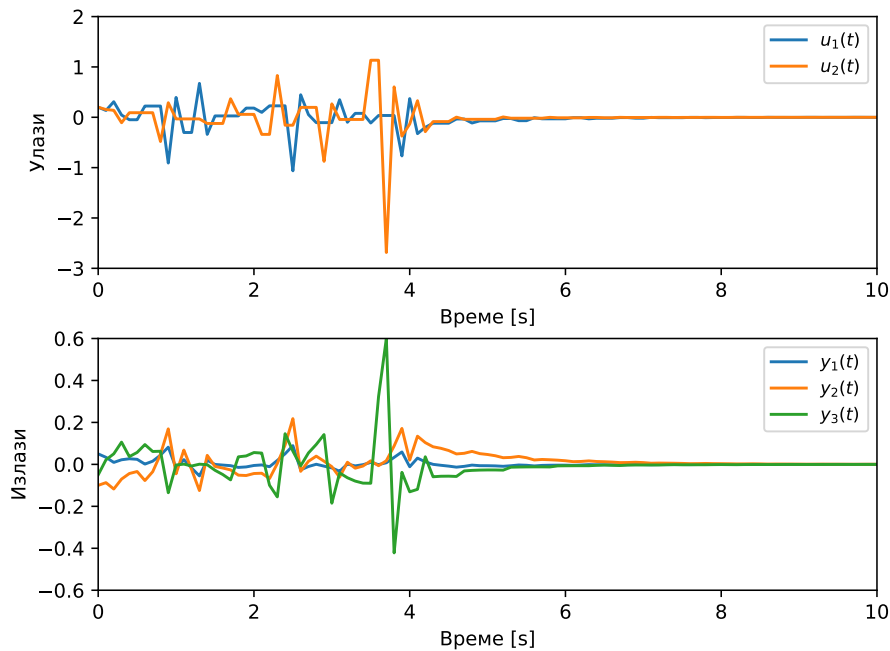
Слика 5.25: Упоредни приказ закона управљања за прву и последњу итерацију фазе учења.



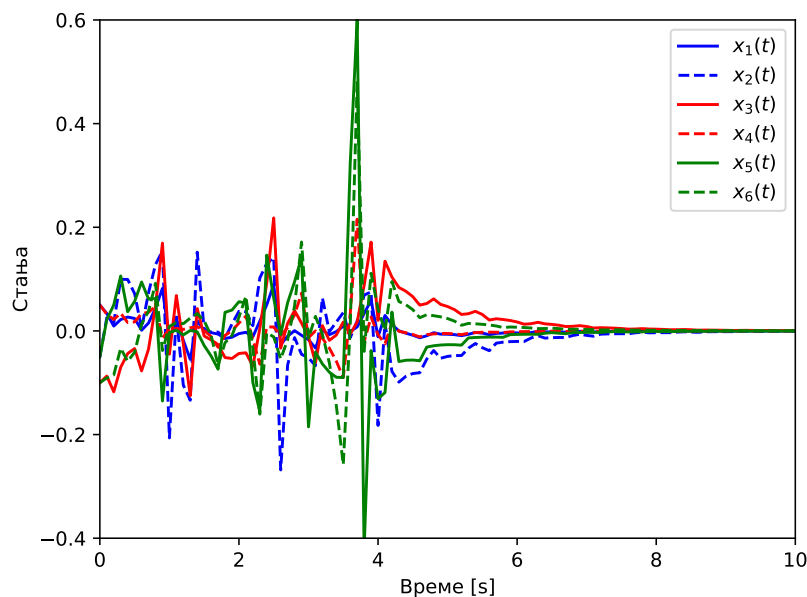
Слика 5.26: Грешка између оптималног и апроксимираног управљачког сигнала.

## 5.2.2 Управљање засновано на АДП-у покренуто догађајима

Слике 5.27 и 5.28 приказују улазе, излазе и стања самобалансирајућег робота управљаног регулатором заснованим на АДП-у покренутиим догађајима. На основу слика можемо видети да се квалитет управљања не смањује упркос мањем броју одбирака. У симулацијама је као иницијална вредност вектора стања узета  $x_0 = [0.05 \quad -0.05 \quad -0.1 \quad 0.05 \quad -0.05 \quad -0.1]^T$ .

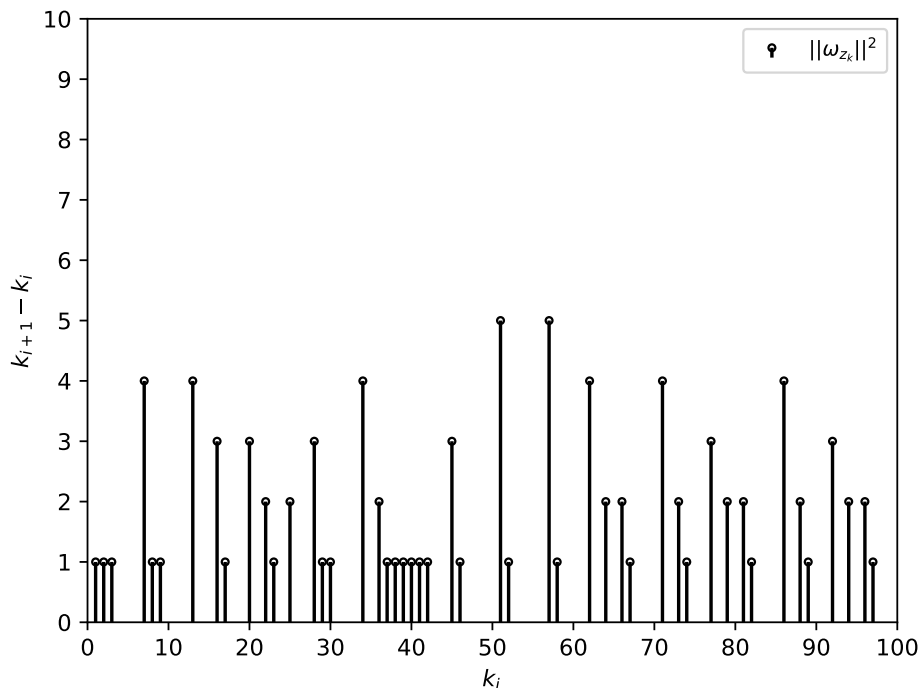


Слика 5.27: Управљање и излази модела робота са регулатором заснованим на АДП-у покренутог догађајима.



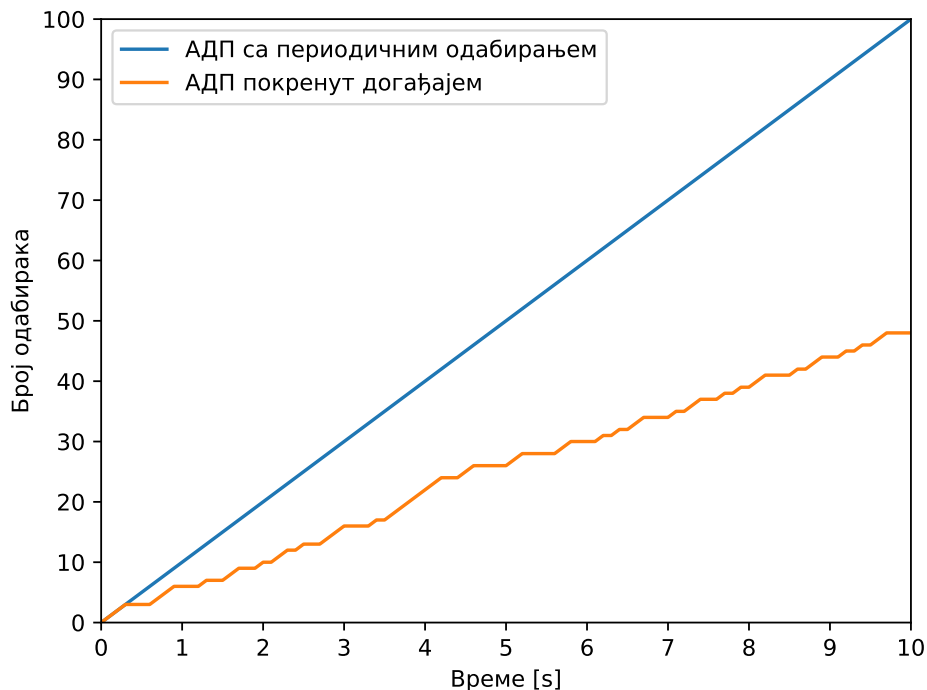
Слика 5.28: Стања модела робота са регулатором заснованим на АДП-у покренутог догађајима.

На слици 5.29 је приказана секвенца одбирака. Број одбирака је смањен за 51% у односу на управљање регулатором заснованим на АДП-у са периодичним одабирањем.



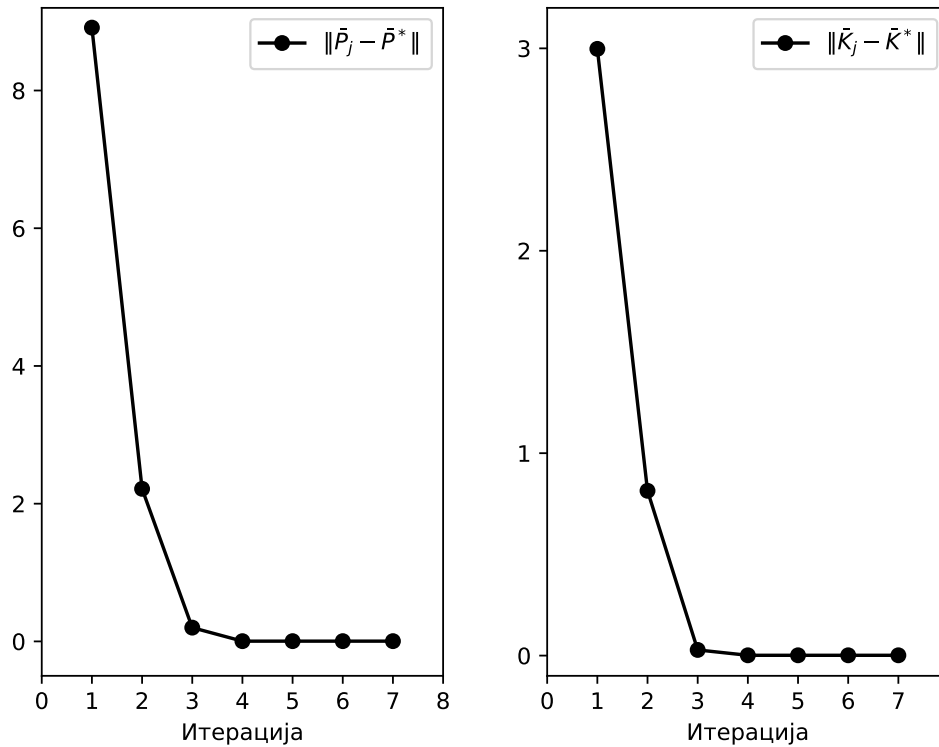
Слика 5.29: Секвенца одбирака у случају аperiодичног одабирања.

Упоредни приказ броја одбирака у случају управљања заснованим на АДП-у са периодичним и аperiодичним одабирањем дат је на слици 5.30.



Слика 5.30: Упоредни приказ укупног броја одбирака у случају периодичног и аperiодичног одабирања.

На слици 5.31 приказана је конвергенција у случају управљања самобалансирајућег робота регулатором заснованим на АДП-у покренутим догађајима.



Слика 5.31: Конвергенција вредности апроксимираних матрица  $\bar{P}$  и  $\bar{K}$  ка оптималним вредностима  $\bar{P}^*$  и  $\bar{K}^*$ .

### 5.2.3 Управљање засновано на АДП-у толерантно на отказе

Посматра се линеарни моделе самобалансирајућег робота са отказима актуатора описан једначинама

$$\dot{x}(t) = Ax(t) + B(u(t) - f_a(t)), \quad (5.51)$$

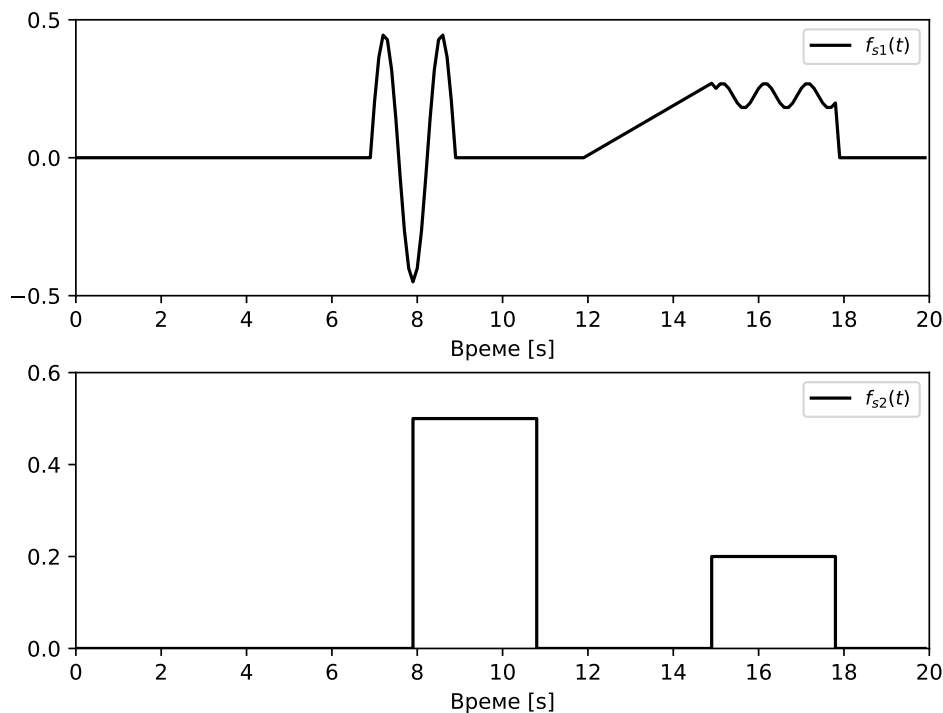
$$y(t) = Cx(t), \quad (5.52)$$

где члан  $f_a(t)$  представља непознате адитивне отказе актуаторе. У симулацијама коришћене су следеће функције отказа

$$f_{a1}(t) = \begin{cases} 0 & 0 \leq t \leq 7\text{s} \\ 0.4 \cdot \sin(0.15\pi \cdot (t - 6)) & 7\text{s} < t \leq 9\text{s} \\ 0 & 9\text{s} < t \leq 12\text{s} \\ 0.2 \cdot (t - 12) & 12\text{s} < t \leq 15\text{s} \\ 0.25 + 0.05 \cdot \sin(0.2\pi \cdot (t - 15)) & 15\text{s} < t \leq 18\text{s} \\ 0 & 18\text{s} < t \leq 20\text{s} \end{cases}, \quad (5.53)$$

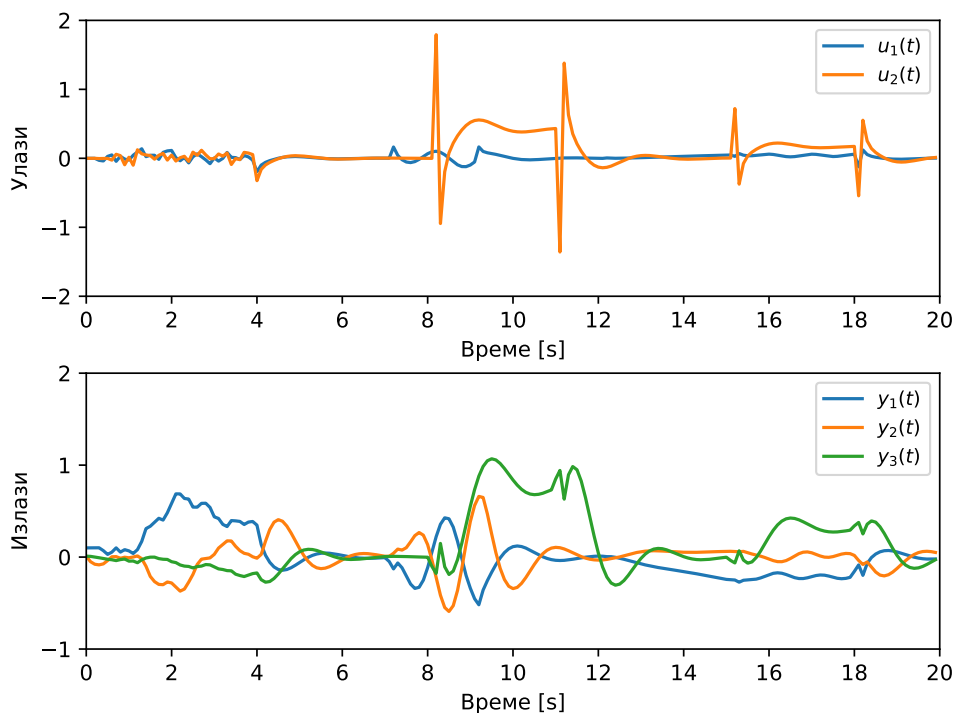
$$f_{a2}(t) = \begin{cases} 0 & 0 \leq t \leq 8 \text{ s} \\ 0.5 & 8 \text{ s} < t \leq 11 \text{ s} \\ 0 & 11 \text{ s} < t \leq 15 \text{ s} , \\ 0.2 & 15 \text{ s} < t \leq 18 \text{ s} \\ 0 & 18 \text{ s} < t \leq 20 \text{ s} \end{cases} \quad (5.54)$$

чији су облици приказани на слици 5.32. Такође, у симулацијама је као иницијална вредност вектора стања узета  $x_0 = [0.1 \ 0.01 \ 0.01 \ 0.1 \ 0.01 \ 0.01]^T$ .

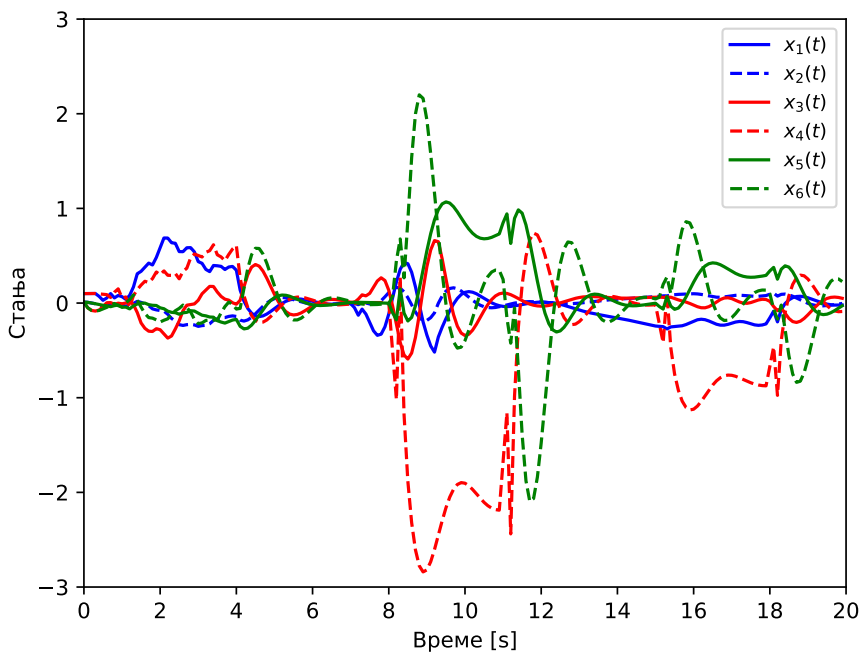


Слика 5.32: Приказ отказа  $f_{s1}$  и  $f_{s2}$ .

Излази, улази и стања робота у случају управљања заснованог на АДП-у без компензације отказа приказани су на сликама 5.33 и 5.34. На сликама можемо видети да откази актуатора значајно деградирају преформансе система.

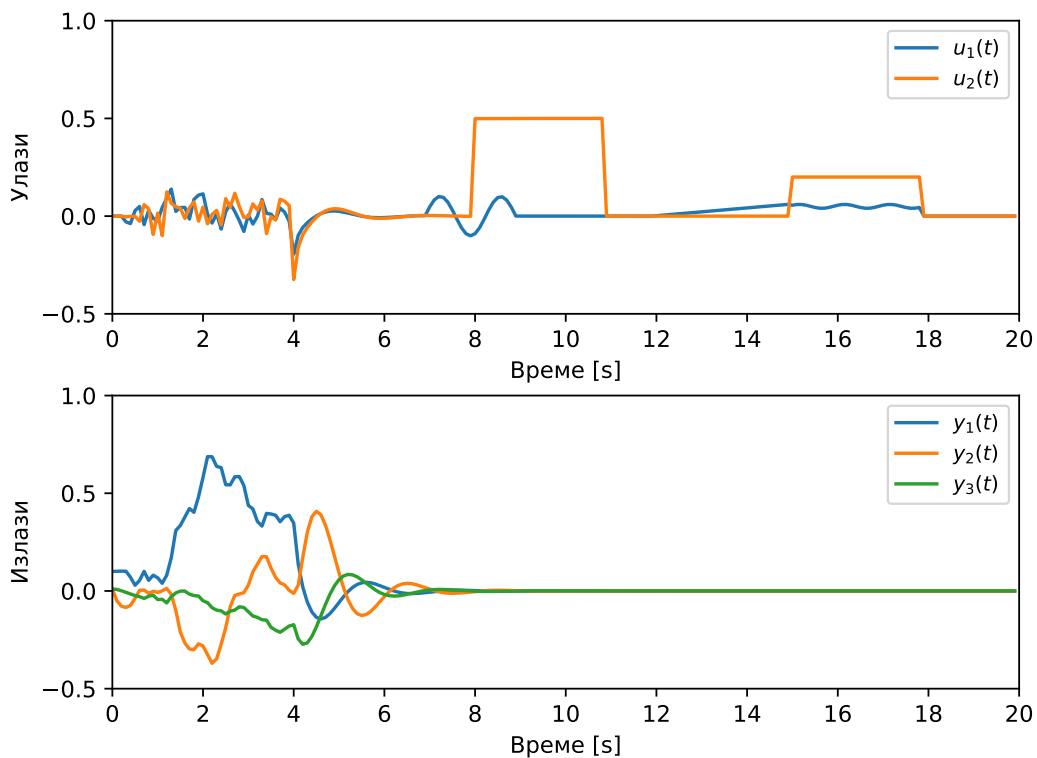


Слика 5.33: Улази и излази система без компензованим отказом.

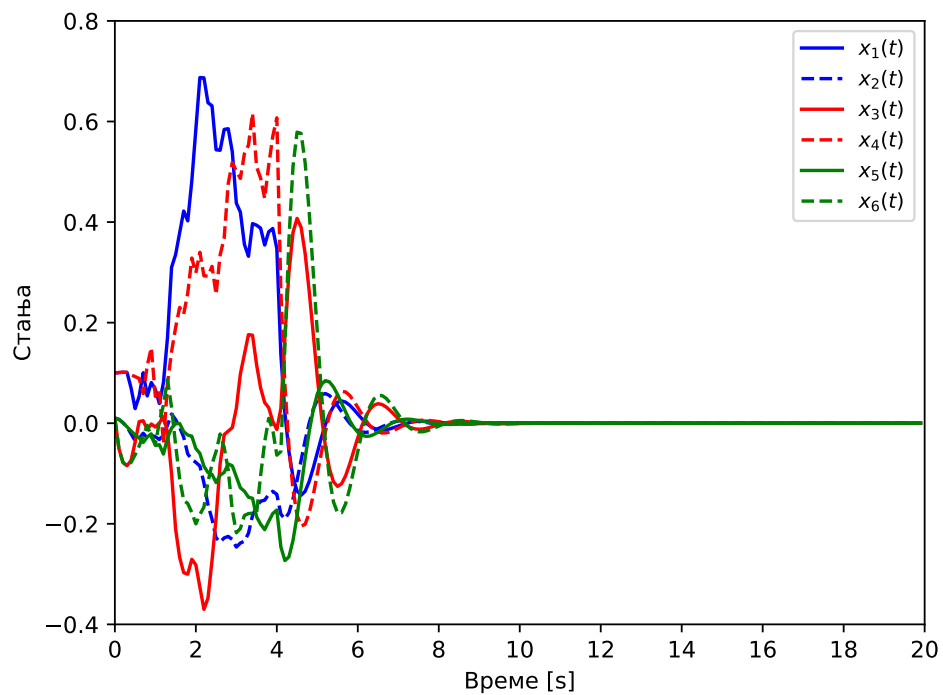


Слика 5.34: Стања система без компензованим отказом.

Са друге стране, на сликама 5.35 и 5.36 приказани су излази, улази и стања работа управљаног регулатором заснованим на АДП-у толерантаним на отказе. Са слика се може видети да се отказ актуатора у потпуности неутралише и не утиче на перформансе систем.



Слика 5.35: Улази и излази система са компензованим отказом.



Слика 5.36: Стања система са компензованим отказом.



## Глава 6

# Закључак и даља истраживања

У овом поглављу су сумирани резултати до којих се дошло кроз рад на развоју интелигентних система управљања заснованих на АДП-у. Такође су изложени предлози за будуће правце истраживања. Главни доприноси ове дисертације се огледају у неколико аспекта.

- У тези је представљен нови оквир за оптимално управљање засновано на учењу за класу линеарних и временски дискретних динамичких система. Предности учења са поткрепљивањем и адаптивног динамичког програмирања се систематски користе у комбинацији са алатима из теорије управљања, у циљу пројектовања адаптивних оптималних регулатора заснованих на учењу са гарантованом стабилношћу затвореног система.
- Развијени су АДП регулатори за дискретне линеарне системе са комплетно непознатом динамиком, унапређењем постојећих АДП алгоритама, у циљу асимптотског праћења у оптималном смислу. ПИИ методе, које су показале велики успех у решавању традиционалних проблема динамичког програмирања и проблема АДП-а, су уопштене за решавање проблема временски дискретних система са непознатом динамиком у простору стања. Да би предложени регулатори могли да се баве решавањем проблема праћења линеарних система у дискретном времену, у дисертацији је предложен нови ПИИ приступ за итеративну апроксимацију решења Рикатијеве једначине. Наиме, уобичајено се претпоставља да су променљиве величине стања мерљиве. Међутим, овај услов је генерално рестриктиван, јер у многим инжењерским применама често недостаје потпуна информација о стању, и само је излаз доступан за мерење. Овом приликом се стабилност динамичких система постиже путем повратне спреге по излазу, при чему се за реконструкцију непознатих величина стања користе узоркована улазна и излазна мерења. Урађена је анализа конвергенције алгоритама адаптивног учења, док је стабилност затвореног система управљања загарантована према теорији система са узоркованим мерењима (енг. sampled-data systems).
- Применом теорије планирања експеримента, у тези се користи оптимални сигнал, у виду суме синусоида, који је омогућио да излаз система носи максимум информације о систему. Тиме је скраћено време учења, односно убрзан је процес пројектовања интелигентних регулатора.

- Имплементацијом стратегије управљања засноване на догађајима развијени су одговарајући АДП регулатори. Тиме се смањио број укључивања регулатора, чиме су се уштедели енергетски, рачунски и комуникациони ресурси.
- Имплементирана стратегија управљања унапредила је детекцију и дијагностику отказа, чиме су побољшане перформансе система управљања у случају појаве отказа.
- Уз помоћ предложених алгоритмима заснованих на подацима основни проблеми оптималног управљања (линеарних квадратних регулатора LQR) и проблеми регулације по излазу су проширени, из перспективе теорије управљања засноване на учењу, на класу нелинеарних система. Користећи чињеницу да се нелинеарни системи могу верно представити сумом линеарних система са променљивим параметрима, предложена методологија је успешно примењена на континуалне нелинеарне системе.
- Предложени приступ управљању заснован на учењу има широку примену у практичним инжењерским системима, у ери роботике и вештачке интелигенције. Практична примењивост предложених АДП регулатора је успешно верификована на инжењерским проблемима, као што је управљање хидрауличког серво актуатора и самобалансирајућег робота. Валидност и ефективност предложених АДП регулатора су евалуирани кроз интензивне симулације.

Будући рад биће усмерен на уопштавање предложених метода и алата управљања заснованих на учењу у неколико праваца:

#### **Управљање стохастичких нелинеарних система засновано на учењу.**

С обзиром на богату литературу о савременом нелинеарном управљању, постоји много могућности да се представљени резултати генерализују на много општије форме нелинеарних система, као што су неафини нелинеарни системи, изван референтне класе линеарних система која се разматрала у овој тези. Будући рад на овој теми ће се концентрисати на пројектовање оптималних АДП регулатора, користећи информације улазних сигнала и повратне спреге по излазу, за системе са нелинеарним динамичким неодређеностима и стохастичким поремећајима. Будући рад у овој области укључује развој АДП регулатора за парцијално линеарне системе са нелинеарним динамичким неодређеностима.

#### **Управљање хибридниh и прекидачких система засновано на учењу.**

Хибридни системи [131–134] и прекидачки системи [135–137] су две важне класе динамичких система који показују и континуалну и дискретну динамику у времену. Како генерализовати приступ управљања заснован на учењу, који је овде развијен за системе у дискретном времену на важне класе хибридниh и прекидачких система је занимљива тема која заслужује детаљно истраживање.

#### **Управљање бесконачно димензионалних система засновано на учењу.**

Системи представљени функционалним диференцијалним једначинама [138–140] и парцијалним диференцијалним једначинама [141–143] су две важне класе бесконачних система. Постоји свега неколико истраживања посвећених

управљању заснованом на учењу за ову врсту бесконачно димензионалних динамичких система, упркос њиховом јасном значају у теорији и применама. Потребно је развити нове алате и методе који превазилазе резултате који су овде предложени за управљање коначно-димензионалним динамичким системима заснованих на учењу.

**Примене у роботизи и индустрији.** На крају, али не и најмање важно, учење са поткрепљивањем је пронашло многе примене у роботизи [68, 144, 145], која као веома актуелна област истраживања има пуно отворених питања за системе управљања заснованих на учењу. Поред даљег развоја и усавршавања теорије АДП-а, такође је важно проучавати питања практичне имплементације за примене у реалном свету. Са становишта практичне примене важно је обезбедити софтверску подршку у форми пакета за предложене интелигентне регулаторе засноване на учењу.

# Библиографија

- [1] D. Bertsekas. *Reinforcement Learning and Optimal Control*. Athena Scientific, Belmont, Massachusetts (USA), 2019. ISBN 9781886529397.
- [2] D. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, Massachusetts (USA), 2012. ISBN 9781886529083.
- [3] F.L. Lewis, D. Vrabie, and V.L. Syrmos. *Optimal Control*. John Wiley & Sons, Inc., Hoboken, New Jersey (USA), 3rd edition, 2012. ISBN 9780470633496. doi:10.1002/9781118122631.
- [4] F.L. Lewis and D. Liu, editors. *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. John Wiley & Sons, Inc., Hoboken, New Jersey (USA), 2013. ISBN 978-1-118-45397-1.
- [5] J.J. Murray, C.J. Cox, G.G. Lendaris, and R. Saeks. Adaptive dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 32(2):140–153, 2002. doi:10.1109/TSMCC.2002.801727.
- [6] P. Werbos. *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University, 1974.
- [7] W. Gao and Z.P. Jiang. Learning-based adaptive optimal tracking control of strict-feedback nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6):2614–2624, 2018. doi:10.1109/TNNLS.2017.2761718.
- [8] T. Bian and Z.P. Jiang. Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design. *Automatica*, 71:348–360, 2016. doi:10.1016/j.automatica.2016.05.003.
- [9] M. Roozegar, M.J. Mahjoob, and M.M. Jahromi. Optimal motion planning and control of a nonholonomic spherical robot using dynamic programming approach: simulation and experimental results. *Mechatronics*, 39:174–184, 2016. doi:10.1016/j.mechatronics.2016.05.002.
- [10] J.L. Sun and C.S. Liu. An overview on the adaptive dynamic programming based missile guidance law. *Acta Automatica Sinica*, 43(7):1101–1113, 2017.
- [11] Q. Hu. Robust adaptive sliding mode attitude maneuvering and vibration damping of three-axis-stabilized flexible spacecraft with actuator saturation limits. *Nonlinear Dynamics*, 55(4):301–321, 2009. doi:10.1007/s11071-008-9363-1.

- [12] L Ljung. *System Identification: Theory for the User*. Information and System Sciences. Prentice-Hall, New Jersey (USA), 2nd edition, 1999. ISBN 9780136566953.
- [13] R. Pintelon and J. Schoukens. *System identification: A frequency domain approach*. Wiley-IEEE Press, New Jersey (USA), 2nd edition, 2012. ISBN 978-1-118-28739-2.
- [14] C.R. Rojas, J.C. Aguero, J.S. Welsh, G.C. Goodwin, and A. Feuer. Robustness in experiment design. *IEEE Transactions on Automatic Control*, 57(4):860–874, 2011. doi:10.1109/TAC.2011.2166294.
- [15] A. Al-Tamimi, F.L. Lewis, and M. Abu-Khalaf. Model-free q-learning designs for linear discrete-time zero-sum games with application to h-infinity control. *Automatica*, 43(3):473–481, 2007. doi:10.1016/j.automat.2006.09.019.
- [16] Weinan Gao, Yu Jiang, Zhong-Ping Jiang, and Tianyou Chai. Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming. *Automatica*, 72:37–45, 2016. doi:10.1016/j.automat.2016.05.008.
- [17] K.J. Aström. Event based control. In *Analysis and design of nonlinear control systems*, pages 127–147. Springer-Verlag, Berlin, Heidelberg (Germany), 2008. ISBN 978-3-540-74357-6. doi:10.1007/978-3-540-74358-3\_9.
- [18] W.P.M.H. Heemels, M.C.F. Donkers, and A.R. Teel. Periodic event-triggered control for linear systems. *IEEE Transactions on automatic control*, 58(4):847–861, 2013. doi:10.1109/TAC.2012.2220443.
- [19] B. Jiang, H.R. Karimi, Y Kao, and C. Gao. Takagi–sugeno model based event-triggered fuzzy sliding-mode control of networked control systems with semi-markovian switchings. *IEEE Transactions on Fuzzy Systems*, 28(4):673–683, 2020. doi:10.1109/TFUZZ.2019.2914005.
- [20] T. Liu and Z.P. Jiang. A small-gain approach to robust event-triggered control of nonlinear systems. *IEEE Transactions on Automatic Control*, 60(8):2072–2085, 2015. doi:10.1109/TAC.2015.2396645.
- [21] P. Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 52(9):1680–1685, 2007. doi:10.1109/TAC.2007.904277.
- [22] M. Lemmon. Event-triggered feedback in control, estimation, and optimization. In *Networked control systems*, pages 293–358. Springer, London (UK), 2010. ISBN 978-0-85729-032-8. doi:10.1007/978-0-85729-033-5\_9.
- [23] J. Lunze and D. Lehmann. A state-feedback approach to event-based control. *Automatica*, 46(1):211–215, 2010. doi:10.1016/j.automat.2009.10.035.
- [24] M. Mazo, A. Anta, and P. Tabuada. An iss self-triggered implementation of linear controllers. *Automatica*, 46(8):1310–1314, 2010. doi:10.1016/j.automat.2010.05.009.

- [25] A. Sahoo, H. Xu, and S. Jagannathan. Neural network-based event-triggered state feedback control of nonlinear continuous-time systems. *IEEE Transactions on Neural Networks and Learning Systems*, 27(3):497–509, 2015. doi:10.1109/TNNLS.2015.2416259.
- [26] S. Lu, Q. He, and J. Wang. A review of stochastic resonance in rotating machine fault detection. *Mechanical Systems and Signal Processing*, 116:230–260, 2019. doi:10.1016/j.ymsp.2018.06.032.
- [27] Y. Li, M.J. Zuo, J. Lin, and J. Liu. Fault detection method for railway wheel flat using an adaptive multiscale morphological filter. *Mechanical Systems and Signal Processing*, 84:642–658, 2017. doi:10.1016/j.ymsp.2016.07.009.
- [28] T.D. Popescu and D. Aiordachioaie. Fault detection of rolling element bearings using optimal segmentation of vibrating signals. *Mechanical Systems and Signal Processing*, 116:370–391, 2019. doi:10.1016/j.ymsp.2018.06.033.
- [29] L. Li, M. Chadli, S.X. Ding, J. Qiu, and Y. Yang. Diagnostic observer design for t–s fuzzy systems: Application to real-time-weighted fault-detection approach. *IEEE Transactions on Fuzzy Systems*, 26(2):805–816, 2018. doi:10.1109/TFUZZ.2017.2690627.
- [30] A. Chibani, M. Chadli, P. Shi, and N.B. Braiek. Fuzzy fault detection filter design for t–s fuzzy systems in the finite-frequency domain. *IEEE Transactions on Fuzzy Systems*, 25(5):1051–1061, 2016. doi:10.1109/TFUZZ.2016.2593921.
- [31] S. Cho, Z. Gao, and T. Moan. Model-based fault detection, fault isolation and fault-tolerant control of a blade pitch system in floating wind turbines. *Renewable Energy*, 120:306–321, 2018. doi:10.1016/j.renene.2017.12.102.
- [32] J. Lan and R.J. Patton. *Robust integration of model-based fault estimation and fault-tolerant control*. Advances in Industrial Control. Springer, Cham (Switzerland), 1st edition, 2020. ISBN 978-3-030-58759-8. doi:10.1007/978-3-030-58760-4.
- [33] R. Isermann. *Fault-diagnosis applications: model-based condition monitoring: actuators, drives, machinery, plants, sensors, and fault-tolerant systems*. Springer-Verlag, Berlin, Heidelberg (Germany), 1st edition, 2011. ISBN 978-3-642-12766-3. doi:10.1007/978-3-642-12767-0.
- [34] S.X. Ding. *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer-Verlag, Berlin, Heidelberg (Germany), 1st edition, 2008. ISBN 978-3-540-76304-8. doi:10.1007/978-3-540-76304-8.
- [35] M. Tomás-Rodríguez and S.P. Banks. *Linear, time-varying approximations to nonlinear dynamical systems: with applications in control and optimization*. Lecture Notes in Control and Information Sciences. Springer, London (UK), 1st edition, 2010. doi:10.1007/978-1-84996-101-1.
- [36] V. Stojanovic and D. Prsic. Robust identification for fault detection in the presence of non-gaussian noises: application to hydraulic servo drives. *Nonlinear Dynamics*, 100(3):2299–2313, 2020. doi:10.1007/s11071-020-05616-4.

- [37] M. Mynuddin and W. Gao. Distributed predictive cruise control based on reinforcement learning and validation on microscopic traffic simulation. *IET Intelligent Transport Systems*, 14(5):270–277, 2020. doi:10.1049/iet-its.2019.0404.
- [38] M. Mynuddin, W. Gao, and Z.P. Jiang. Reinforcement learning for multi-agent systems with an application to distributed predictive cruise control. In *2020 American Control Conference (ACC)*, pages 315–320. IEEE, July 2020. ISBN 978-1-5386-8267-8. doi:10.23919/ACC45564.2020.9147968.
- [39] M. Davari, W. Gao, Z.P. Jiang, and F.L. Lewis. An optimal primary frequency control based on adaptive dynamic programming for islanded modernized microgrids. *IEEE Transactions on Automation Science and Engineering*, 18(3): 1109–1121, 2021. doi:10.1109/TASE.2020.2996160.
- [40] A. van de Walle, F. Naets, and W. Desmet. Virtual microphone sensing through vibro-acoustic modelling and kalman filtering. *Mechanical Systems and Signal Processing*, 104:120–133, 2018. doi:10.1016/j.ymsp.2017.08.032.
- [41] K. Maes, A. Iliopoulos, W. Weijtjens, C. Devriendt, and G. Lombaert. Dynamic strain estimation for fatigue assessment of an offshore monopile wind turbine using filtering and modal expansion algorithms. *Mechanical Systems and Signal Processing*, 76:592–611, 2016. doi:10.1016/j.ymsp.2016.01.004.
- [42] Y.H. Chang, Q. Hu, and C.J. Tomlin. Secure estimation based kalman filter for cyber–physical systems against sensor attacks. *Automatica*, 95:399–412, 2018. doi:10.1016/j.automatica.2018.06.010.
- [43] A. Cavallo, G. De Maria, C. Natale, and S. Pirozzi. Slipping detection and avoidance based on kalman filter. *Mechatronics*, 24(5):489–499, 2014. doi:10.1016/j.mechatronics.2014.05.006.
- [44] A. Vacca and G. Franzoni. *Hydraulic fluid power: Fundamentals, applications, and circuit design*. John Wiley & Sons, Hoboken, New Jersey (USA), 2021. ISBN 978-1-119-56911-4.
- [45] N. Manring. *Fluid power pumps and motors: analysis, design, and control*. McGraw-Hill Education, New York (USA), 2013. ISBN 978-0-071-81220-7.
- [46] N. Nedic, V. Stojanovic, and V. Djordjevic. Optimal control of hydraulically driven parallel robot platform based on firefly algorithm. *Nonlinear Dynamics*, 82(3):1457–1473, 2015. doi:10.1007/s11071-015-2252-5.
- [47] J.J. Vyas, B. Gopalsamy, and H. Joshi. *Electro-hydraulic actuation systems: design, testing, identification and validation*. SpringerBriefs in Applied Sciences and Technology. Springer, Singapore, 2018. ISBN 978-981-13-2546-5. doi:10.1007/978-981-13-2547-2.
- [48] V. Stojanovic, N. Nedic, D. Prsic, Lj. Dubonjic, and V. Djordjevic. Application of cuckoo search algorithm to constrained control problem of a parallel robot platform. *The International Journal of Advanced Manufacturing Technology*, 87(9):2497–2507, 2016. doi:10.1007/s00170-016-8627-z.

- [49] V. Filipovic, N. Nedic, and V. Stojanovic. Robust identification of pneumatic servo actuators in the real situations. *Forschung im Ingenieurwesen*, 75(4):183–196, 2011. doi:10.1007/s10010-011-0144-5.
- [50] J.F. Blackburn, G. Reethof, and J.L. Shearer. *Fluid power control*. MIT Press, Cambridge, Massachusetts (USA), 1960.
- [51] M. Jelali and A. Kroll. *Hydraulic servo-systems: Modelling, identification and control*. Advances in Industrial Control. Springer, London (UK), 2002. ISBN 978-1-85233-692-9. doi:10.1007/978-1-4471-0099-7.
- [52] Z. Li, C. Yang, and L. Fan. *Advanced control of wheeled inverted pendulum systems*. Springer, London (UK), 1st edition, 2012. ISBN 978-1-4471-2962-2. doi:10.1007/978-1-4471-2963-9.
- [53] S.C. Lin, C.C. Tsai, and H.C. Huang. Adaptive robust self-balancing and steering of a two-wheeled human transportation vehicle. *Journal of Intelligent & Robotic Systems*, 62:103–123, 2011. doi:10.1007/s10846-010-9460-5.
- [54] Z. Li and C. Xu. Adaptive fuzzy logic control of dynamic balance and motion for wheeled inverted pendulums. *Fuzzy Sets and Systems*, 160(12):1787–1803, 2009. doi:10.1016/j.fss.2008.09.013.
- [55] J. Huang, Z.H. Guan, T. Matsuno, T. Fukuda, and K. Sekiyama. Sliding-mode velocity control of mobile-wheeled inverted-pendulum systems. *IEEE Transactions on robotics*, 26(4):750–758, 2010. doi:10.1109/TRO.2010.2053732.
- [56] S.C. Lin, C.C. Tsai, and H.C. Huang. Nonlinear adaptive sliding-mode control design for two-wheeled human transportation vehicle. In *2009 IEEE International Conference on Systems, Man and Cybernetics*, pages 1965–1970. IEEE, 2009. doi:10.1109/ICSMC.2009.5346583.
- [57] S. Jeong and T. Takahashi. Wheeled inverted pendulum type assistant robot: design concept and mobile control. *Intelligent Service Robotics*, 1(4):313–320, 2008. doi:10.1007/s11370-008-0024-5.
- [58] Changkai Xu, Ming Li, and Fangyu Pan. The system design and lqr control of a two-wheels self-balancing mobile robot. In *2011 International Conference on Electrical and Control Engineering*, pages 2786–2789. IEEE, 2011. doi:10.1109/ICECENG.2011.6057680.
- [59] V Djordjevic, V. Stojanovic, H. Tao, X. Song, S. He, and Gao W. Data-driven control of hydraulic servo actuator based on adaptive dynamic programming. *Discrete and Continuous Dynamical Systems - S*, 15(7):1633–1650, 2022. ISSN 1937-1632. doi:10.3934/dcdss.2021145.
- [60] V. Djordjevic, H. Tao, X. Song, S. He, W. Gao, and V. Stojanovic. Data-driven control of hydraulic servo actuator: An event-triggered adaptive dynamic programming approach. *Mathematical Bioscience and Engineering*, 20(5):8561–8582, 2023. ISSN 1551-0018. doi:10.3934/mbe.2023376.



- [61] V. Djordjevic, Lj. Dubonjic, M.M. Morato, D. Prsic, and V. Stojanovic. Sensor fault estimation for hydraulic servo actuator based on sliding mode observer. *Mathematical Modelling and Control*, 2(1):34–43, 2022. ISSN 2767-8946. doi:10.3934/mmc.2022005.
- [62] V. Djordjevic, V. Stojanovic, H. Tao, X. Song, S. He, and W. Gao. Event-triggered adaptive dynamic programming based optimal control for hydraulic servo actuator. In *HM 2023 Proceedings*, pages D.1–D.6. Faculty of Mechanical and Civil Engineering in Kraljevo, June 2023. ISBN 978-86-82434-01-6.
- [63] V. Djordjevic, M.M. Morato, and V. Stojanovic. Adaptive dynamic programming based optimal control for hydraulic servo actuator. In *HM 2021 Proceedings*, pages C.37–C.42. Faculty of Mechanical and Civil Engineering in Kraljevo, June 2021. ISBN 978-86-81412-09-1.
- [64] V. Djordjevic, V. Stojanovic, D. Prsic, Lj. Dubonjic, and M.M. Morato. Observer-based fault estimation in steer-by-wire vehicle. *Engineering Today*, 1(1):7–17, 2022. ISSN 2812-9474. doi:10.5937/engtoday2201007D.
- [65] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton (USA), 1957.
- [66] R. Bellman and S. Dreyfus. Functional approximations and dynamic programming. *Mathematical Tables and Other Aids to Computation*, 13(68):247–251, 1959.
- [67] R.A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, Cambridge (USA), 1960.
- [68] R.S. Sutton and A.G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018. ISBN 9780262039246.
- [69] C.J.C.H. Watkins. *Learning from delayed rewards*. PhD thesis, King’s College, Cambridge United Kingdom, 1989.
- [70] R.S. Sutton. Learning to predict by the methods of temporal differences. *Machine learning*, 3:9–44, 1988. doi:10.1007/BF00115009.
- [71] F.L. Lewis and D. Vrabie. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits and Systems Magazine*, 9(3):32–50, 2009. doi:10.1109/MCAS.2009.933854.
- [72] D. Vrabie, K.G. Vamvoudakis, and F.L. Lewis. *Optimal adaptive control and differential games by reinforcement learning principles*. The Institution of Engineering and Technology, London (UK), 2013. doi:10.1049/PBCE081E.
- [73] P.J. Werbos. A menu of designs for reinforcement learning over time. In W.T. Miller, R.S. Sutton, and P.J. Werbos, editors, *Neural networks for control*, pages 67–95. MIT press, Cambridge (USA), 1996. ISBN 0-262-13261-3.
- [74] W.B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. John Wiley & Sons, New Jersey (USA), 2nd edition, 2011. ISBN 978-1-118-02916-9.

- [75] F.Y. Wang, H. Zhang, and D. Liu. Adaptive dynamic programming: An introduction. *IEEE Computational intelligence magazine*, 4(2):39–47, 2009. doi:10.1109/MCI.2009.932261.
- [76] P.J. Werbos. Neural networks for control and system identification. In *Proceedings of the 28th IEEE Conference on Decision and Control*, pages 260–265, Tampa (USA), 1989. IEEE. doi:10.1109/CDC.1989.70114.
- [77] P. Werbos. Approximate dynamic programming for real-time control and neural modeling. In D. White and D. Sofge, editors, *Handbook of intelligent control: Neural, fuzzy and adaptive approaches*, chapter 13, pages 493–525. Van Nostrand, 1992.
- [78] P.J. Werbos. From adp to the brain: foundations, roadmap, challenges and research priorities. In *2014 International Joint Conference on Neural Networks (IJCNN)*, pages 107–111, Beijing (China), 2014. IEEE. doi:10.1109/IJCNN.2014.6889359.
- [79] J. Si, A.G. Barto, W.B. Powell, and D. Wunsch, editors. *Handbook of learning and approximate dynamic programming*, volume 2 of *Computational Intelligence*. John Wiley & Sons, New Jersey (USA), 2004. ISBN 978-0-471-66054-5.
- [80] H. Zhang, D. Liu, Y. Luo, and D. Wang. *Adaptive dynamic programming for control: algorithms and stability*. Communications and Control Engineering. Springer, London (UK), 2012. ISBN 978-1-4471-4756-5. doi:10.1007/978-1-4471-4757-2.
- [81] D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-dynamic programming*. Athena Scientific, Belmont (USA), 1996. ISBN 978-1-8865-2910-6.
- [82] Karl J Åström and Björn Wittenmark. *Adaptive control*. Dover Publications, New York (USA), 2nd edition, 2008.
- [83] V.S. Borkar. *Stochastic approximation: A dynamical systems viewpoint*, volume 48. Hindustan Book Agency, Gurgaon (India), 2009. ISBN 978-93-86279-38-5. doi:10.1007/978-93-86279-38-5.
- [84] L. Busoniu, R. Babuska, B. De Schutter, and D. Ernst. *Reinforcement learning and dynamic programming using function approximators*. CRC press, Boca Raton (USA), 2010. ISBN 978-1-4398-2108-4.
- [85] M.L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. Probability and Statistics. John Wiley & Sons, New Jersey (USA), 2014. ISBN 978-1-118-62587-3.
- [86] C. Szepesvári. Reinforcement learning algorithms for mdps. Technical Report TR09-13, Department of Computing Science, University of Alberta, Edmonton (Canada), October 2009.
- [87] S. Bhasin, N. Sharma, P. Patre, and W. Dixon. Asymptotic tracking by a reinforcement learning-based adaptive critic controller. *Journal of Control Theory and Applications*, 9:400–409, 2011. doi:10.1007/s11768-011-0170-8.

- [88] H. Modares, F.L. Lewis, and M.B. Naghibi-Sistani. Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks. *IEEE Transactions on neural networks and learning systems*, 24(10):1513–1525, 2013. doi:10.1109/TNNLS.2013.2276571.
- [89] K.G. Vamvoudakis and F.L. Lewis. Multi-player non-zero-sum games: Online adaptive learning solution of coupled hamilton–jacobi equations. *Automatica*, 47(8):1556–1569, 2011. doi:10.1016/j.automatica.2011.03.005.
- [90] H. He, Z. Ni, and J. Fu. A three-network architecture for on-line learning and optimization based on adaptive dynamic programming. *Neurocomputing*, 78(1):3–13, 2012. doi:10.1016/j.neucom.2011.05.031.
- [91] H. Zhang, Q. Wei, and D. Liu. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. *Automatica*, 47(1):207–214, 2011. doi:10.1016/j.automatica.2010.10.033.
- [92] X. Xu, C. Wang, and F.L. Lewis. Some recent advances in learning and adaptation for uncertain feedback control systems. *International journal of adaptive control and signal processing*, 28(3-5):201–204, 2014. doi:10.1002/acs.2475.
- [93] G. Hewer. An iterative technique for the computation of the steady state gains for the discrete optimal regulator. *IEEE Transactions on Automatic Control*, 16(4):382–384, 1971. doi:10.1109/TAC.1971.1099755.
- [94] W. Gao, Y. Jiang, Z.P. Jiang, and T. Chai. Adaptive and optimal output feedback control of linear systems: An adaptive dynamic programming approach. In *Proceeding of the 11th World Congress on Intelligent Control and Automation*, pages 2085–2090. IEEE, June-July 2014. ISBN 978-1-4799-5825-2. doi:10.1109/WCICA.2014.7053043.
- [95] F.L. Lewis and K.G. Vamvoudakis. Reinforcement learning for partially observable dynamic processes: Adaptive dynamic programming using measured output data. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41(1):14–25, 2011. doi:10.1109/TSMCB.2010.2043839.
- [96] W. Aangenent, D. Kostic, B. de Jager, R. van de Molengraft, and M. Steinbuch. Data-based optimal control. In *Proceedings of the 2005 American Control Conference*, pages 1460–1465. IEEE, June 2005. doi:10.1109/ACC.2005.1470171.
- [97] S.J. Bradtke, B.E. Ydstie, and A. G. Barto. Adaptive linear quadratic control using policy iteration. In *Proceedings of 1994 American Control Conference-ACC'94*, volume 3, pages 3475–3479. IEEE, 1994. doi:10.1109/ACC.1994.735224.
- [98] Gang Tao. *Adaptive control design and analysis*, volume 37 of *Adaptive and Learning Systems for Signal Processing, Communications, and Control*. John Wiley & Sons, New Jersey (USA), 2003. ISBN 9780471274520.
- [99] W. Gao and Z.P. Jiang. Adaptive dynamic programming and adaptive optimal output regulation of linear systems. *IEEE Transactions on Automatic Control*, 61(12):4164–4169, 2016. doi:10.1109/TAC.2016.2548662.

- [100] S. Thrun. Efficient exploration in reinforcement learning. Technical Report CMU-CS-92-102, Carnegie-Mellon University, School of Computer Science, Pittsburgh (USA), 1992.
- [101] Y. Jiang and Z.P. Jiang. Approximate dynamic programming for output feedback control. In *Proceedings of the 29th Chinese control conference*, pages 5815–5820, Beijing (China), 2010. IEEE.
- [102] Y. Jiang and Z.P. Jiang. Robust approximate dynamic programming and global stabilization with nonlinear dynamic uncertainties. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 115–120, Orlando (USA), 2011. IEEE. doi:10.1109/CDC.2011.6160279.
- [103] Y. Jiang and Z.P. Jiang. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. *Automatica*, 48(10): 2699–2704, 2012. doi:10.1016/j.automatica.2012.06.096.
- [104] P.A. Ioannou and J Sun. *Robust Adaptive Control*. Dover Publications, New York (USA), 2012.
- [105] Y.S. Ma, W.W. Che, C. Deng, and Z.G. Wu. Observer-based event-triggered containment control for mass under dos attacks. *IEEE Transactions on Cybernetics*, 52(12):13156–13167, 2022. doi:10.1109/TCYB.2021.3104178.
- [106] X. Wang, H.R. Karimi, M. Shen, D. Liu, L.W. Li, and J. Shi. Neural network-based event-triggered data-driven control of disturbed nonlinear systems with quantized input. *Neural Networks*, 156:152–159, 2022. doi:10.1016/j.neunet.2022.09.021.
- [107] M. Shen, Y. Gu, J.H. Park, Y. Yi, and W.W. Che. Composite control of linear systems with event-triggered inputs and outputs. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 69(3):1154–1158, 2022. doi:10.1109/TCSII.2021.3098820.
- [108] X. Wang, M. Shen, and J.H. Park. Event-triggered control of uncertain nonlinear discrete-time systems with extended state observer. (preprint on webpage at <https://doi.org/10.21203/rs.3.rs-644060/v1>), 2021.
- [109] H. Xu, S. Jagannathan, and F.L. Lewis. Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses. *Automatica*, 48(6):1017–1030, 2012. doi:10.1016/j.automatica.2012.03.007.
- [110] C. Chen, K. Xie, F.L. Lewis, S. Xie, and A. Davoudi. Fully distributed resilience for adaptive exponential synchronization of heterogeneous multiagent systems against actuator faults. *IEEE Transactions on Automatic Control*, 64(8):3347–3354, 2019. doi:10.1109/TAC.2018.2881148.
- [111] C.H. Xie and G.H. Yang. Approximate guaranteed cost fault-tolerant control of unknown nonlinear systems with time-varying actuator faults. *Nonlinear Dynamics*, 83:269–282, 2016. doi:10.1007/s11071-015-2324-6.

- [112] G. Dong, H. Li, H. Ma, and R. Lu. Finite-time consensus tracking neural network ftc of multi-agent systems. *IEEE Transactions on Neural Networks and Learning Systems*, 32(2):653–662, 2021. doi:10.1109/TNNLS.2020.2978898.
- [113] M. Benosman and K.Y. Lum. Passive actuators’ fault-tolerant control for affine nonlinear systems. *IEEE Transactions on Control Systems Technology*, 18(1): 152–163, 2010. doi:10.1109/TCST.2008.2009641.
- [114] L. Su, X. Zhu, J. Qiu, C. Zhou, Y. Zhao, and Y. Guo. Robust passive fault-tolerant control for uncertain non-linear stochastic systems with distributed delays. In *Chinese Control Conference 2010: Proceedings of the 29th Chinese Control Conference*, pages 1949–1953, 2010. URL <https://ieeexplore.ieee.org/abstract/document/5573628>.
- [115] Y. Li, F. Zhou, and B. Zhao. Dynamic output feedback based active decentralized fault-tolerant control for reconfigurable manipulator with concurrent failures. *Mathematical Problems in Engineering*, 2015. doi:10.1155/2015/528086.
- [116] W. Guo, A. Qiu, and C. Wen. Active actuator fault tolerant control based on generalized internal model control and performance compensation. *IEEE Access*, 7:175514–175521, 2019. doi:10.1109/ACCESS.2019.2957824.
- [117] Y. Fu and T. Chai. Nonlinear adaptive decoupling control based on neural networks and multiple models. *International Journal of Innovative Computing, Information and Control*, 8(3A):1867–1878, 2012. URL <http://www.ijicic.org/ijicic-10-12020.pdf>.
- [118] N. Meskin, E. Naderi, and K. Khorasani. A multiple model-based approach for fault diagnosis of jet engines. *IEEE Transactions on Control Systems Technology*, 21(1):254–262, 2011. doi:10.1109/TCST.2011.2177981.
- [119] Z. Abbasfard, A. Baniamerian, and K. Khorasani. Fault diagnosis of gas turbine engines: A symbolic multiple model approach. In *2014 European Control Conference (ECC)*, pages 944–951, 2014. doi:10.1109/ECC.2014.6862358.
- [120] D. Liu, D. Wang, and D. Zhao. Adaptive dynamic programming for optimal control of unknown nonlinear discrete-time systems. In *2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pages 242–249, 2011. doi:10.1109/ADPRL.2011.5967357.
- [121] Z.P. Jiang and Y. Jiang. Robust adaptive dynamic programming: recent results and applications. In *Chinese Control Conference 2013: Proceedings of the 32nd Chinese Control Conference*, pages 968–973, 2013. URL <https://ieeexplore.ieee.org/abstract/document/6639567>.
- [122] Q.Y. Fan and G.H. Yang. Adaptive fault-tolerant control for affine non-linear systems based on approximate dynamic programming. *IET Control Theory & Applications*, 10(6):655–663, 2016. doi:10.1049/iet-cta.2015.1081.
- [123] D. Ye and T. Song. Decentralized reliable guaranteed cost control for large-scale nonlinear systems using actor-critic network. *Neurocomputing*, 320:121–128, 2018. doi:10.1016/j.neucom.2018.09.029.

- [124] B. Zhao, D. Liu, and Y. Li. Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems. *Information Sciences*, 384: 21–33, 2017. doi:10.1016/j.ins.2016.12.016.
- [125] Y. Li, K. Sun, and S. Tong. Observer-based adaptive fuzzy fault-tolerant optimal control for siso nonlinear systems. *IEEE transactions on cybernetics*, 49(2):649–661, 2019. doi:10.1109/TCYB.2017.2785801.
- [126] V. Stojanovic. Fault-tolerant control of a hydraulic servo actuator via adaptive dynamic programming. *Mathematical Modelling and Control*, 3(3):181–191, 2023. ISSN 2767-8946. doi:10.3934/mmc.2023016.
- [127] V Stojanovic and D. Prsic. *Detekcija i dijagnostika otkaza u sistemima automatskog upravljanja*. Fakultet za mašinstvo i građevinarstvo, Kraljevo (Srbija), 2022. ISBN 978-86-81412-16-9.
- [128] T. Chen and B.A. Francis. *Optimal Sampled-Data Control Systems*. Communications and Control Engineering. Springer, London, 2011. ISBN 978-1-4471-3039-0. doi:10.1007/978-1-4471-3037-6.
- [129] F. Grasser, A. D’arrigo, S. Colombi, and A.C. Rufer. Joe: a mobile, inverted pendulum. *IEEE Transactions on industrial electronics*, 49(1):107–114, 2002. doi:10.1109/41.982254.
- [130] L. Guo, S.A.A. Rizvi, and Z. Lin. Optimal control of a two-wheeled self-balancing robot by reinforcement learning. *International Journal of Robust and Nonlinear Control*, 31(6):1885–1904, 2021. doi:10.1002/rnc.5058.
- [131] R. Goedel, R.G. Sanfelice, and A.R. Teel. *Hybrid dynamical systems: modeling stability, and robustness*. Princeton University Press, Princeton, New Jersey (USA), 2012. ISBN 9781400842636. URL <http://www.jstor.org/stable/j.ctt7s02z>.
- [132] A.J. Van Der Schaft and H. Schumacher. *An introduction to hybrid dynamical systems*. Lecture Notes in Control and Information Sciences. Springer, London (UK), 1st edition, 2013. ISBN 978-1-4471-3916-4. doi:10.1007/BFb0109998.
- [133] H. Lin and P.J. Antsaklis. *Hybrid Dynamical Systems: Fundamentals and Methods*. Advanced Textbooks in Control and Signal Processing. Springer, Cham (Switzerland), 1st edition, 2022. ISBN 978-3-030-78731-8. doi:10.1007/978-3-030-78731-8.
- [134] W.M. Haddad, V. Chellaboina, and S.G. Nersesov. *Impulsive and hybrid dynamical systems: stability, dissipativity, and control*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, New Jersey (USA), 2014. ISBN 9781400865246. doi:10.1515/9781400865246.
- [135] D. Liberzon. *Switching in systems and control*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, Massachusetts (USA), 2003. ISBN 978-0-8176-4297-6. doi:10.1007/978-1-4612-0017-8.

- [136] Y. Liu and Q.H. Wu. *Adaptive Switching Control of Large-Scale Complex Power Systems: Theory and Applications*. Power Systems. Springer, Singapore, 2023. ISBN 978-981-99-1038-0. doi:10.1007/978-981-99-1039-7.
- [137] E.K. Boukas. *Stochastic switching systems: analysis and design*. Control Engineering. Birkhäuser, Boston, Massachusetts (USA), 1st edition, 2007. ISBN 978-0-8176-4452-9. doi:10.1007/0-8176-4452-0.
- [138] J.K. Hale and S.M.V. Lunel. *Introduction to functional differential equations*. Applied Mathematical Sciences. Springer Science & Business Media, New York (USA), 1st edition, 2013. ISBN 978-1-4612-8741-4. doi:10.1007/978-1-4612-4342-7.
- [139] P.K. Sahoo and P. Kannappan. *Introduction to functional equations*. CRC press, Boca Raton (USA), 2011. ISBN 978-1-4398-4116-7.
- [140] I. Karafyllis and Z.P. Jiang. *Stability and stabilization of nonlinear systems*. Communications and Control Engineering. Springer, London (UK), 1st edition, 2011. ISBN 978-0-85729-512-5. doi:10.1007/978-0-85729-513-2.
- [141] P.D Christofides. *Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, Massachusetts (USA), 1st edition, 2001. ISBN 978-0-8176-4156-6. doi:10.1007/978-1-4612-0185-4.
- [142] I. Karafyllis and M. Krstic. *Input-to-state stability for PDEs*. Communications and Control Engineering. Springer, Cham (Switzerland), 2018. ISBN 978-3-319-91010-9. doi:10.1007/978-3-319-91011-6.
- [143] M. Krstic. Delay compensation for nonlinear, adaptive, and pde systems. 2009. doi:10.1007/978-0-8176-4877-0.
- [144] J. Kober, J.A. Bagnell, and J. Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11):1238–1274, 2013. doi:10.1177/0278364913495721.
- [145] J Li, F.L. Lewis, and J. Fan. *Reinforcement Learning: Optimal Feedback Control with Industrial Applications*. Advances in Industrial Control. Springer, Cham (Switzerland), 1st edition, 2023. ISBN 978-3-031-28393-2. doi:10.1007/978-3-031-28394-9.

## Биографија кандидата

Владимир Ђорђевић је рођен 18.02.1985. године у Крушевцу. Завршио је ОШ „Јован Поповић“ у Крушевцу. Средње образовање је стекао у периоду 1999.-2003. год. у Машинско-електротехничкој школи у Крушевцу – образовни профил машински техничар.

Након завршетка средње школе уписао је Машински факултет у Краљеву. Дипломирао је на смеру Конструисање и пројектовање у машиноградњи, група Аутоматско управљање и флуидна техника на тему Фази управљање инверзним клатном, 2010. године. Током студија остварио је просечну оцену 8,83 (осам и 83/100). Докторске академске студије на Факултету за машинство и грађевинарство у Краљеву уписао је школске године 2019/2020, и тренутно је студент треће године докторских студија. Положио је све испите предвиђене програмом студија са просечном оценом 9,4 (девет и 40/100).

Од јануара 2011. године је запослен на Факултету за машинство и грађевинарство у Краљеву, и био је укључен у реализацију пројекта код Министарства за науку:

1. ТР33026, финансиран од стране Министарства просвете, науке и технолошког развоја, 2011-2019, Повећање енергетске ефикасности постројења за производњу топлотне енергије помоћу аутоматског управљања. Руководилац пројекта проф. др Драган Пршић. Носилац истраживања Факултет за машинство и грађевинарство у Краљеву.
2. ТР33027, финансиран од стране Министарства просвете, науке и технолошког развоја, 2011-2019, Развој енергетске ефикасности постројења за гасификацију и когенерацију чврсте биомасе. Руководилац пројекта проф. др Раде Карамарковић. Носилац истраживања Факултет за машинство и грађевинарство у Краљеву.

Од јануара 2019. године изабран је у звање асистента за ужу научну област Аутоматско управљање и флуидна техника и ангажован на извођењу вежби предмета Катедре за аутоматско управљање и флуидна техника.



**ИЗЈАВА АУТОРА О ОРИГИНАЛНОСТИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ**

Изјављујем да докторска дисертација под насловом:

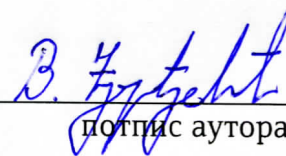
**Интелигентни регулатори засновани на адаптивном динамичком  
програмирању**

представља *оригинално ауторско дело* настало као резултат *сојствене истраживачкој рада*.

Овом Изјавом *такође потврђујем*:

- да сам *једини аутор* наведене докторске дисертације,
- да у наведеној докторској дисертацији *нисам извршио/ла повреду* ауторског нити другог права интелектуалне својине других лица,

У Краљеву, 17.11.2023. године,

  
\_\_\_\_\_  
потпис аутора

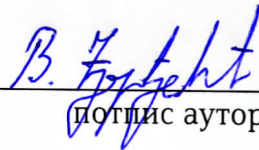
**ИЗЈАВА АУТОРА О ИСТОВЕТНОСТИ ШТАМПАНЕ И ЕЛЕКТРОНСКЕ ВЕРЗИЈЕ  
ДОКТОРСКЕ ДИСЕРТАЦИЈЕ**

Изјављујем да су штампана и електронска верзија докторске дисертације под  
насловом:

**Интелигентни регулатори засновани на адаптивном динамичком  
програмирању**

истоветне.

У Краљеву, 17.11.2023. године,



\_\_\_\_\_  
потпис аутора

**ИЗЈАВА АУТОРА О ИСКОРИШЋАВАЊУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ**

Ја, Владимир Ђорђевић,

дозвољавам

не дозвољавам

Универзитетској библиотеци у Крагујевцу да начини два трајна умножена примерка у електронској форми докторске дисертације под насловом:

**Интелигентни регулатори засновани на адаптивном динамичком програмирању**

и то у целини, као и да по један примерак тако умножене докторске дисертације учини трајно доступним јавности путем дигиталног репозиторијума Универзитета у Крагујевцу и централног репозиторијума надлежног министарства, тако да припадници јавности могу начинити трајне умножене примерке у електронској форми наведене докторске дисертације путем *преузимања*.

Овом Изјавом такође

дозвољавам

не дозвољавам<sup>1</sup>


---

<sup>1</sup> Уколико аутор изабере да не дозволи припадницима јавности да тако доступну докторску дисертацију користе под условима утврђеним једном од *Creative Commons* лиценци, то не искључује право припадника јавности да наведену докторску дисертацију користе у складу са одредбама Закона о ауторском и сродним правима.

припадницима јавности да тако доступну докторску дисертацију користе под условима утврђеним једном од следећих *Creative Commons* лиценци:

- 1) Ауторство
- 2) Ауторство - делити под истим условима
- ③ Ауторство - без прерада
- 4) Ауторство - некомерцијално
- 5) Ауторство - некомерцијално - делити под истим условима
- 6) Ауторство - некомерцијално - без прерада<sup>2</sup>

У Краљеву, 17.11.2023. године,

  
\_\_\_\_\_  
потпис аутора

---

<sup>2</sup> Молимо ауторе који су изабрали да дозволе припадницима јавности да тако доступну докторску дисертацију користе под условима утврђеним једном од *Creative Commons* лиценци да заокруже једну од понуђених лиценци. Детаљан садржај наведених лиценци доступан је на: <http://creativecommons.org.rs/>