Journal of Aerospace Mechanics/ 2024/ Vol.20/ No.1/ 141-161

Journal of Aerospace Mechanics

DOR: 20.1001.1.26455323.1403.20.1.9.1

Fault-Tolerant Optimal Attitude Tracking Control of Quadrotor Subject to State and Input Constraints Using Safe Reinforcement Learning

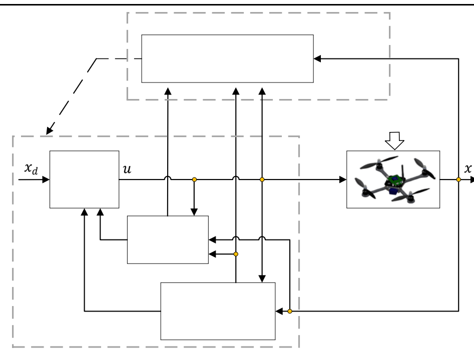
Sajad Roshanravan[1], Saeed Shamaghdari[2*]

[1] Ph.D. Student, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran
[2] Associate Professor, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

HIGHLIGHTS

- Model-free method
- improvement in the convergence properties of the identifier and critic NNs.
- Ensuring input and state constraints.
- Guaranteeing system stability at all times
- HJB-based fault detection without requiring any additional filter.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Article Type: Research paper
Received: 7 October 2023
Received in revised form: 29 October 2023
Accepted: 2 December 2023
Available online: 9 March 2024
*Correspondence:
shamaghdari@iust.ac.ir

How to cite this article:
S. Roshanravan, S. Shamaghdari. Fault-Tolerant Optimal Attitude Tracking Control of Quadrotor Subject to State and Input Constraints Using Safe Reinforcement Learning. Journal of Aerospace Mechanics. 2024; 20(1):141-161.

Keywords:
Quadrotor attitude control
Component and actuator faults
Fault-tolerant optimal control
Fault detection
Safe reinforcement learning

ABSTRACT

In this article, a method for designing a fault-tolerant optimal attitude tracking control (FTOATC) for a quadrotor UAV subject to component and actuator faults is presented. The proposed fault-tolerant method is based on safe reinforcement learning (SRL) and is capable of ensuring input and state constraints without the need for prior knowledge of the quadrotor dynamics. To this end, the proposed optimal method is presented with a dual neural network (NN) structure consisting of identifier-critic neural networks. In the identifier NN update law, in addition to considering the variable forgetting factor dependent on measurement noise, the experience response method is used, which increases convergence speed and robustness to measurement noise and reduces estimation error. In this method, solving the constrained FTOATC problem is equivalent to solving an unconstrained optimal stabilization problem for an augmented system, where control input constraints and states are guaranteed by selecting suitable cost functions on the input signal and appropriate control barrier functions (CBF)on the states, respectively. Furthermore, fault detection is performed without the need for any model or filter bank, simply by comparing the residual value of the Hamilton-Jacobi-Bellman (HJB) equation with a predetermined threshold. The Uniformly Ultimately Boundedness (UUB) of identifier and critic NN weight errors and, as a result, the convergence of the control input to the neighborhood of the optimal solution are all proved by Lyapunov theory and the performance of the method is validated through simulation results.

مکانیک هوافضا/ سال ۱۴۰۳/ دوره ۲۰/ شماره ۱/ صفحه ۱۴۱-۱۶۱

نشریه علمی مکانیک هوافضا

DOR: 20.1001.1.26455323.1403.20.1.9.1

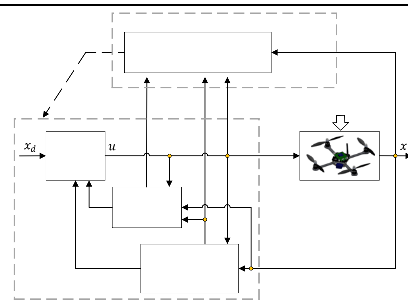
کنترل ردیاب وضعیت بهینه تحمل‌پذیر عیب کوادروتور در حضور قیود حالت و ورودی با استفاده از یادگیری تقویتی ایمن

سجاد روشن‌روان[1]، سعید شمقدری[2]*

[1] دانشجوی دکتری، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران

[2] دانشیار، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران

چکیده گرافیکی



برجسته‌ها

- عدم نیاز به شناخت دینامیک پرنده
- بهبود سرعت شبکه‌های شناساگر و نقاد
- تضمین قیود ورودی و حالت
- تضمین پایداری سیستم در تمامی زمان‌ها
- آشکارسازی وقوع عیب مبتنی بر خطای HJB بدون نیاز به مدل

چکیده

در این مقاله، به ارائه روشی جهت طراحی سیستم کنترل وضعیت ردیاب بهینه برای پرنده کوادروتور که در معرض عیوب اجزا و عملگر قرار دارد پرداخته‌شده است. روش کنترل تحمل‌پذیر عیب یکپارچه پیشنهادی مبتنی بر یادگیری تقویتی ایمن ارائه‌شده است و قادر است بدون نیاز به شناخت قبلی از دینامیک پرنده، قیود ورودی و حالات را تضمین نماید. به این منظور، روش بهینه پیشنهادی با ساختار شبکه عصبی دوگانه شامل شبکه‌های عصبی شناساگر-نقاد ارائه‌شده است. در قانون به‌روزرسانی وزن‌های شبکه شناساگر علاوه بر متغیر در نظر گرفتن ضریب فراموشی از روش پاسخ تجربه استفاده‌شده که باعث افزایش سرعت همگرایی و مقاومت نسبت به نویز اندازه‌گیری و کاهش خطای تخمین می‌شوند. در این روش، حل مسئله کنترل ردیاب وضعیت بهینه تحمل‌پذیر عیب در حالت مقید با حل مسئله پایدارسازی بهینه نامقید برای یک سیستم افزوده معادل می‌شود که در آن قیود ورودی کنترلی و حالات به ترتیب با انتخاب تابع هزینه مناسب بر سیگنال ورودی و توابع کنترل مانع مناسب بر حالات، تضمین داده می‌شوند. همچنین آشکارسازی وقوع عیب بدون نیاز به هیچ‌گونه بانکی از مدل یا فیلتر و صرفاً با مقایسه مقدار باقی‌مانده معادله همیلتون-ژاکوبی-بلمن با یک آستانه از پیش تعیین‌شده انجام می‌پذیرد. پایداری فراگیر یکنواخت وزن‌های هر دو شبکه و درنتیجه همگرایی قانون کنترل به پاسخ بهینه با استفاده از قضیه لیاپانوف اثبات و با استفاده از نتایج شبیه‌سازی صحت عملکرد آن نشان داده‌شده است.

مشخصات مقاله

تاریخچه مقاله:

نوع مقاله: علمی پژوهشی

دریافت: ۱۴۰۲/۰۷/۱۵

بازنگری: ۱۴۰۲/۰۸/۰۷

پذیرش: ۱۴۰۲/۰۹/۱۱

ارائه برخط: ۱۴۰۲/۱۲/۱۹

*نویسنده مسئول:

shamaghdari@iust.ac.ir

کلیدواژه‌ها:

کنترل وضعیت کوادروتور

عیوب اجزا و عملگر

کنترل بهینه تحمل‌پذیر عیب

آشکارسازی وقوع عیب

یادگیری تقویتی ایمن

۱- مقدمه

پیشرفت‌های چشمگیر به‌دست‌آمده در حوزه عملگرهای دوار پربازده و منابع تغذیه جریان بالا در دهه‌های اخیر، باعث گردید ساخت ربات‌های زمینی و هوایی بدون سرنشین بسیار موردتوجه قرار گیرد. دراین‌بین، پرنده‌های مولتی روتور با توجه به هزینه ساخت پایین، قابلیت مانور بالا و کم بودن هزینه‌های پروازی در حوزه‌های مختلف تجاری، صنعتی و نظامی جهت مأموریت‌های پروازی گوناگون به‌صورت فراگیر مورداستفاده قرار گرفتند [۱ و ۲].

این نوع پرنده‌های عمودپرواز با توجه به قابلیت شناور ماندن در یک مکان ثابت، نداشتن محدودیت حداقل سرعت جابه‌جایی و عدم نیاز به باند فرود، بیشتر از سایر پرنده‌های بدون سرنشین جهت انجام مأموریت در محیط‌های خطرناک و مانع‌دار مورداستفاده قرار می‌گیرند. لذا به دلیل حضور موانع محیطی، این دسته از پرنده‌ها بسیار در معرض عیوب عملگر و اجزا از قبیل آسیب‌های موتوری و شکستگی پره‌ها قرار دارند. ازاین‌رو لازم است سیستم کنترل در آن‌ها به نحوی طراحی گردد که بتواند عملکرد آن‌ها را در مقابل این دسته از عیوب احتمالی حفظ نماید [۳].

به این منظور روش‌های کنترلی بسیار وسیعی تحت عنوان کنترل تحمل‌پذیر عیب[1] (FTC) ارائه‌شده است که بر اساس نیاز به بازطراحی کنترل‌کننده به‌طورکلی به دو دسته فعال[2] و غیرفعال[3] تقسیم می‌شوند [۴]. در دسته غیرفعال یک قانون کنترل ثابت به نحوی طراحی می‌شود که در مقابل دسته مشخصی از عیوب پیش‌بینی‌شده مقاوم[4] باشد و بتواند پایداری و عملکرد سیستم معیوب را همانند حالت بدون عیب حفظ نماید. درحالی‌که در روش‌های فعال از اطلاعات بلادرنگ عیب که توسط قسمت آشکارسازی و تشخیص[5] (FDD) به دست می‌آید، جهت بازطراحی کنترل‌کننده استفاده می‌شود [۵]. لذا روش‌های غیرفعال اگرچه در مقایسه با فعال، از بار محاسباتی کمتر و سهولت پیاده‌سازی بیشتری برخوردار می‌باشند، ولی محافظه‌کاری بالا و عدم توانایی آن‌ها در جبران عیوب غیرقابل‌پیش‌بینی، استفاده از آن‌ها را در سیستم عملی با محدودیت‌های زیادی مواجه می‌کند [۶].

استفاده از اطلاعات عیب به‌دست‌آمده از واحد FDD در بازطراحی کنترل‌کننده، منجر به یک نامعینی دو سویه در ساختار سیستم FTC فعال می‌شود بدین معنی که عملکرد هر یک از آن‌ها عملکرد و طراحی بخش دیگر را تحت تأثیر قرار می‌دهد [۷]. بررسی‌های انجام‌شده نشان می‌دهد اگرچه مطالعات زیادی به تجمیع واحدهای FDD و FTC پرداخته‌اند، ولی تعداد بسیار کمی از مقالات موضوع یکپارچه‌سازی واحد FDD را در طراحی FTC موردبحث قرار داده‌اند [۸ و ۹].

در اکثر سیستم‌های واقعی ازجمله سیستم‌های پروازی، نه‌تنها پایداری بلکه بهینگی عملکرد سیستم بایستی در تمامی شرایط عملیاتی تضمین داده شود [۱۰-۱۲]. روش‌های مرسوم حل مسئله کنترل بهینه در سیستم‌های غیرخطی، عموماً با حل معادله غیرخطی همیلتون-ژاکوبی-بلمن[6] (HJB) روبه‌رو هستند که حل تحلیلی آن با چالش‌های زیادی همراه است [۱۳]. ازاین‌رو روش‌های تخمین به‌منظور حل این معادله بکار گرفته می‌شوند. در این میان، یادگیری تقویتی[7] (RL) یک ابزار شناخته‌شده قدرتمند جهت تخمین پاسخ معادله HJB به‌صورت زمان پیشرو است که به‌کارگیری آن را در روش‌های کنترل بهینه تحمل‌پذیر عیب فعال با استقبال مواجه نموده است [۱۴]. حل مسئله کنترل بهینه توسط روش RL در دو حالت مبتنی بر مدل و مستقل از آن انجام می‌شود که با توجه به در معرض عیب بودن پرنده، دسته مستقل از مدل در این مقاله مورد هدف قرارگرفته است [۱۵-۱۷]. قابل‌توجه است که روش‌های مستقل از مدل نیز خود با دو رویکرد مبتنی بر داده و شناساگر ارائه‌شده‌اند که نیاز روش‌های داده محور به فاز جمع‌آوری داده، استفاده از آن‌ها را در سیستم‌های در معرض عیب با چالش‌های جدی در تضمین پایداری مواجه می‌کند.

قیود حالت عضو جدایی‌ناپذیر سیستم‌های ایمنی بحرانی مانند هواپیماها، ربات‌های جراح و ... می‌باشند که تأثیر بسیار زیادی بر عملکرد سیستم دارند؛ به‌طوری‌که نقض شدن آن‌ها ممکن

[1] Fault-Tolerant Control
[2] Active
[3] Passive
[4] Robust
[5] Fault Detection and Diagnosis
[6] Hamilton-Jacobi-Bellman
[7] Reinforcement learning

سیستم را به سمت ناپایداری سوق دهد و عواقب فاجعه‌باری را به دنبال آورد. در سال‌های اخیر، روش RL ایمن[۱] (SRL) توجه بسیار زیادی را به‌عنوان یک رویکرد کارآمد در یادگیری قانون کنترل بهینه مقید به خود معطوف ساخته است [۱۸ و ۱۹]. در این روش قیود حالات با اضافه شدن یک جمله از توابع کنترل مانع[۲] (CBF) به تابع هدف موردنظر طراح برآورده می‌شوند.

پیرو نتایج به‌دست‌آمده از SRL در حل مسائل کنترل بهینه، در این مقاله یک روش جدید جهت طراحی کنترل ردیاب وضعیت بهینه تحمل‌پذیر عیب[۳] (FTOATC) یکپارچه برای پرنده کوادروتور[۴] با مدل غیرخطی ارائه می‌شود که در معرض عیوب عملگر و اجزا با دو رفتار زمانی ناگهانی و تکرارشونده قرار دارد. روش پیشنهادی، مستقل از مدل بوده و قادر است قیود ورودی و حالت را ضمن ردیابی وضعیت مرجع مطلوب تضمین نماید. در این رابطه مسئله FTOATC مقید برای پرنده کوادروتور با مسئله پایدارسازی بهینه نامقید برای یک سیستم افزوده معادل می‌شود که از دینامیک خطای ردیابی و مسیر مرجع تشکیل‌شده است. با استفاده از این رویکرد، FTOATC مقید از حل معادله HJB برای سیستم افزوده به دست می‌آید که تابع هدف آن شامل یک جمله CBF است.

حل معادله HJB به شناخت دقیق دینامیک سیستم احتیاج دارد [۲۰]. درحالی‌که این یک نیازمندی بسیار دشوار به شمار می‌آید و برای سیستم‌های در معرض عیب تقریباً ناممکن خواهد بود. در این مقاله، به‌منظور حذف نیازمندی حل معادله HJB به شناخت دینامیک سیستم و تخمین مقدار عیب، روش پیشنهادی مبتنی ساختار شبکه عصبی دوگانه شامل شبکه‌های شناساگر[۵] و نقاد[۶] ارائه‌شده است.

ساختار شناساگر ارائه‌شده به‌صورت عصبی تطبیقی در نظر گرفته‌شده و بر تخمین سیستم برحسب توابع پایه فیلتر شده مبتنی است. برخلاف شناساگرهای عصبی تطبیقی ارائه‌شده در [۲۱-۲۳]، ضریب فراموشی در قانون به‌روزرسانی مبتنی بر پاسخ تجربه[۷] (ER) ارائه‌شده متغیر است و تابعی از خطای تخمین حالات و تخمین نویز اندازه‌گیری می‌باشد. در مقایسه با ضریب فراموشی ثابت و بدون استفاده از روش ER، نشان داده می‌شود قانون به‌روزرسانی پیشنهادی، ضمن حفظ مقاومت نسبت به نویز اندازه‌گیری، باعث افزایش سرعت همگرایی و کاهش خطای تخمین می‌شود. همچنین با استفاده از قضیه لیاپانوف[۸] نشان داده می‌شود که علاوه بر همگرایی خروجی شناساگر به متغیرهای حالت، وزن‌های شبکه شناساگر نیز به مقادیر واقعی خود همگرا می‌شوند.

سپس به‌منظور تخمین برخط تابع ارزش کنترل‌کننده از یک شبکه نقاد استفاده‌شده است. در انتها قانون کنترل بهینه به‌صورت مستقیم از تخمین‌های به‌دست‌آمده از دینامیک سیستم و تابع ارزش بهینه که به ترتیب توسط شبکه‌های عصبی شناساگر و نقاد محاسبه‌شده‌اند حاصل می‌شود.

بعد از آشکارسازی عیب، لازم است ساختار کنترل‌کننده با وضعیت جدید انطباق پیدا کند. از سوی دیگر، قید کنترل‌کننده پایدارساز اولیه یک شرط ضروری در همگرایی فرآیندهای یادگیری مبتنی بر RL به قانون کنترل بهینه و تضمین پایداری سیستم محسوب می‌شود. لذا پیدا کردن یک کنترل‌کننده پایدارساز اولیه، یک نیازمندی کلیدی برای همگرایی فرآیند آموزش شمار می‌آید که با توجه ناشناخته بودن دینامیک سیستم معیوب، کاری بسیار دشوار محسوب می‌شود [۲۴]. اخیراً به‌منظور کاهش این محدودیت، یک قانون به‌روزرسانی اصلاح‌شده برای شبکه نقاد در [۲۵] ارائه‌شده که در آن از یک جمله پایدارساز استفاده گردیده است. این جمله زمانی فعال می‌شود که تابع لیاپانوف کاندید برای سیستم حلقه بسته بهینه در طول مسیر سیستم غیر کاهشی باشد. سپس در [۲۶]، به‌منظور مقابله با تغییرات در دینامیک سیستم، یک جمله مقاوم در قانون به‌روزرسانی شبکه نقاد اضافه‌شده است. اگرچه در هر دو روش ارائه‌شده در [۲۵ و ۲۶]، قانون کنترل بهینه مبتنی بر شناخت پیشین از دینامیک سیستم محاسبه می‌شود. با هدف رفع محدودیت شناخت کنترل‌کننده پایدارساز اولیه، در این مقاله نیز مشابه با [۲۵ و ۲۶]، از یک جمله پایدارساز در قانون به‌روزرسانی شبکه نقاد استفاده‌شده است. درحالی‌که بایستی توجه کرد در

[1] Safe RL
[2] Control Barrier Function
[3] Fault-Tolerant Optimal Attitude Tracking Control
[4] Quadrotor
[5] Identifier
[6] Critic
[7] Experience Replay
[8] Lyapunov

قانون به‌روزرسانی پیشنهادی، از تخمین دینامیک سیستم استفاده می‌شود که توسط شناساگر به‌دست‌آمده است.
در ساختار پیشنهادی، آشکارسازی وقوع عیب بدون نیاز به هیچ‌گونه بانک فیلتر و یا مدل ثانویه و تنها با مقایسه مقدار خطای باقی‌مانده معادله HJB نسبت یک آستانه از پیش تعیین‌شده انجام می‌شود. این مقدار آستانه سطح مجاز خطای معادله HJB را تعیین می‌کند و مبتنی بر شرط توقف فرآیند یادگیری انتخاب می‌شود.
با استفاده قضیه لیاپانوف اثبات می‌شود که کلیه سیگنال‌های سیستم حلقه بسته و همچنین خطای وزن‌های هر دو شبکه عصبی شناساگر و نقاد، پایدار فراگیر یکنواخت[1] (UUB) هستند. در مقایسه با کارهای مرتبط پیشین، مهم‌ترین دستاوردها و مزایای روش پیشنهادی را می‌توان به‌صورت زیر خلاصه نمود:
۱) برخلاف [۲۴] که در آن از هرگونه قید سیستمی صرف‌نظر شده است، قیود ورودی و حالت مبتنی بر روش SRL با معرفی یک مسئله پایدارسازی نامقید برای سیستم افزوده در طراحی کنترل‌کننده در نظر گفته‌شده است.
۲) در مقایسه با شناساگرهای عصبی تطبیقی [۲۰ و ۲۱]، یک قانون به‌روزرسانی جدید مبتنی بر ER برای شبکه شناساگر ارائه‌شده است که ضریب فراموشی در آن متغیر می‌باشد و باعث بهبود قابل‌توجه در مشخصه‌های حالت گذرا و خطای حالت ماندگار شبکه می‌شود.
۳) یک قانون به‌روزرسانی جدید با نرخ یادگیری متغیر برای شبکه نقاد ارائه‌شده است که در مقایسه با حالت ثابت باعث افزایش سرعت همگرایی وزن‌های شبکه می‌شود [۲۴].
۴) در مقایسه با روش‌های FTC نظارتی [۲۷] و [۲۸]، مزیت اصلی این روش را می‌توان در عدم نیاز به کنترل‌کننده از پیش طراحی‌شده یا بانک فیلتر جهت آشکارسازی، تشخیص و جبران‌سازی عیوب دانست.
۵) وقوع عیب جدید بدون نیاز به هیچ‌گونه اطلاعات اضافی از دینامیک سیستم و صرفاً با مقایسه باقی‌مانده معادله HJB با یک آستانه از پیش تعیین‌شده آشکارسازی می‌شود.
در ادامه مقاله، ابتدا مدل ریاضیاتی دینامیک پرنده در قسمت ۲ بیان می‌شود. سپس در قسمت ۳ به تعریف مسئله، فرضیات حاکم بر آن و هدف از طراحی FTOATC پرداخته می‌شود. سپس در قسمت ۴ روش پیشنهادی ارائه می‌گردد. در انتها نیز نتایج حاصل از شبیه‌سازی روش مذکور و نتیجه‌گیری به ترتیب در قسمت‌های ۵ و ۶ آورده شده است.

۲- مدل دینامیکی کوادروتور

پرنده‌های کوادروتور دارای دینامیک غیرخطی شش درجه آزادی نسبتاً پیچیده‌ای هستند به‌طوری‌که پارامترهای فروانی، رفتار حرکتی آن‌ها را تحت تأثیر قرار می‌دهد. به‌منظور ساده‌سازی تحلیل رفتاری و طراحی سیستم کنترل، مدل‌سازی دینامیک آن‌ها تحت فرضیات ساده‌کننده زیر انجام می‌پذیرد که فاصله چشم‌گیری با واقعیت ندارد و با آن سازگار است:

۱) بدنه کوادروتور و پروانه موتورها صلب در نظر گرفته می‌شود.
۲) ساختار هندسی و جرمی متقارن است.
۳) ماتریس ممان اینرسی متقارن است.
۴) مرکز جرم و مرکز مختصات محلی بر هم منطبق و در وسط بدنه پرنده قرار دارد.

به‌منظور محاسبه معادلات دینامیکی کوادروتور با استفاده از روش لاگرانژ[2] لازم است نیروها و گشتاورهای اعمالی به پرنده مشخص شوند که در ادامه به آن‌ها پرداخته‌شده است.

۲-۱- نیروها و گشتاورهای ناشی از هر روتور

شکل **۱** ساختار کلی پرنده کوادروتور را به همراه نیروهای وارد بر آن در دستگاه مختصات نشان می‌دهد.
هر یک از روتورهای پرنده شامل یک نیروی بالابرنده عمود بر صفحه چرخش روتورها و یک ممان القایی در خلاف جهت چرخش روتورها است. نیروی بالابرنده هر یک از روتورها طبق رابطه زیر قابل‌محاسبه است:

$$T_i = b\Omega_i^2, i = 1, 2, 3, 4 \qquad (۱)$$

که در آن b ضریب ثابت نیروی بالابرنده و Ω_i سرعت زاویه‌ای ملخ هر یک از روتورها است.

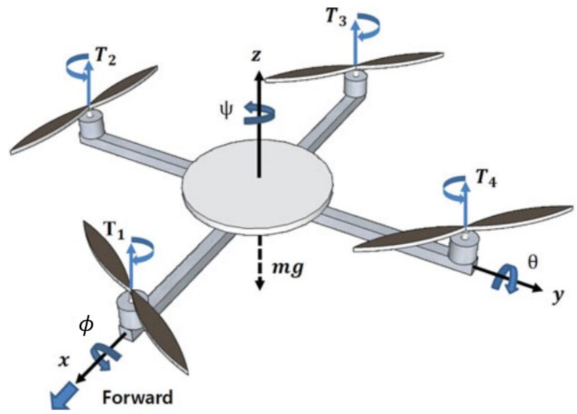
[1] Uniformly Ultimately Bounded

[2] Lagrange

ممان آیرودینامیکی ناشی از روتورها به‌صورت زیر محاسبه می‌شود:

$$\tau_i = d\Omega_i^2, \qquad i = 1,2,3,4 \tag{۲}$$

که در آن d ضریب ثابت درگ ملخ است.



شکل (۱): ساختار پرنده کوادروتور در دستگاه مختصات [۲۹].

۲-۲- دینامیک دورانی

معادلات دینامیک دورانی کوادرتور به‌صورت زیر به دست می‌آید [۲۹]:

$$\ddot{\phi} = \frac{I_y - I_z}{I_x}\dot{\theta}(t)\dot{\psi}(t) + \frac{l}{I_x}u_1 - \frac{J_r}{I_x}\Omega\dot{\theta} \tag{۳}$$

$$\ddot{\theta} = \frac{I_z - I_x}{I_y}\dot{\psi}(t)\dot{\phi}(t) + \frac{l}{I_y}u_2 - \frac{J_r}{I_y}\Omega\dot{\phi} \tag{۴}$$

$$\ddot{\psi} = \frac{I_x - I_y}{I_z}\dot{\phi}(t)\dot{\theta}(t) - \frac{l}{I_z}u_3 \tag{۵}$$

که در آن ϕ، θ و ψ به ترتیب زوایای غلت، پیچ و چرخ می‌باشند که در شکل **۱** نشان داده‌شده‌اند. l فاصله مرکز دوران هر روتور از مرکز ثقل کوادروتور است. $I_i(i = x, y, z)$ مرکز اینرسی کوادروتور در راستای محورهای مختصات و J_r اینرسی هر ملخ است. u_i ها سیگنال‌های کنترلی ورودی مجازی می‌باشند و به ترتیب مقدار گشتاور غلت، پیچ و چرخ را نشان می‌دهند. مقدار u_i ها و Ω به‌صورت زیر به دست می‌آیند:

$$\Omega = \Omega_1 + \Omega_3 - \Omega_2 - \Omega_4 \tag{۶}$$

$$\begin{aligned} u_1 &= b(\Omega_4^2 - \Omega_2^2) \\ u_2 &= b(\Omega_3^2 - \Omega_1^2) \\ u_3 &= d(\Omega_2^2 + \Omega_4^2 - \Omega_1^2 - \Omega_3^2) \end{aligned} \tag{۷}$$

که در آن u_1، u_2 و u_3 به ترتیب گشتاورهای زوایای غلت، پیچ و چرخ می‌باشند.

۳- تعریف مسئله

در این بخش مدل دینامیکی کوادروتور در معرض عیب عملگر و اجزا، به‌صورت یک سیستم غیرخطی با فرم افاین[۱] نسبت به ورودی بیان می‌شود و همچنین فرضیات لازم در دینامیک سیستم و عیوب جهت حل مسئله کنترل‌کننده ردیاب بهینه تحمل‌پذیر عیب ارائه می‌گردد. سپس مسئله FTOATC برای سیستم حاصل معرفی می‌شود.

۳-۱- بیان مدل کوادروتور به‌صورت سیستم غیرخطی افاین

بردار حالت x_1 و x_2 را به‌صورت $x_1 = [\phi \quad \theta \quad \psi]^T$ و $x_2 = [\dot{\phi} \quad \dot{\theta} \quad \dot{\psi}]^T$ در نظر بگیرید. با صرف‌نظر از Ω با توجه به کوچک بودن مقدار آن و تعریف $x(t) = [x_1(t), x_2(t)]^T$، مدل دینامیکی (۳) تا (۵) را می‌توان به فرم غیرخطی افاین نسبت به ورودی به‌صورت زیر بازنویسی کرد:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= f\big(x(t)\big) + g\big(x(t)\big)u(t) \end{aligned} \tag{۸}$$

که در آن، $g = diag\left[l/I_x, l/I_y, l/I_z\right]$ ، $u = [u_1, u_2, u_3]$ و $f(x)$ به صورت زیر است:

$$f(x) = \Big[\big((I_y - I_z)/I_x\big)x_{22}x_{23}, \big((I_z - I_x)/I_y\big)x_{23}x_{21}, \big((I_x - I_y)/I_z\big)x_{21}x_{22}\Big]^T$$

در اینجا فرض شده است که بردار حالت $x(t)$ قابل‌اندازه‌گیری است و $f(x)$ و $g(x)$ ناشناخته بوده به ترتیب دینامیک‌های داخلی[۲] و ورودی[۳] هستند.

فرض ۱: $f\big(x(t)\big) + g\big(x(t)\big)u(t)$ در طول هر مجموعه بسته دلخواه $\aleph \in R^n$ پیوسته و لیپشیتز، $f(0) = 0$ و سیستم پایدارپذیر است. همچنین $f(x)$ نیز در طول $\aleph$ پیوسته و لیپشیتز به‌صورت $\|f(x)\| \le b_f\|x\|$ و $g(x)$ محدود و به‌صورت $\|g(x)\| \le b_g$ است که b_f و b_g اعداد مثبت مشخصی هستند.

[1] Affine
[2] Drift
[3] Input

۳-۲- عیب اجزا و عملگر

در این مقاله فرض شده است که پرنده در معرض عیوب کاهش کارایی نسبی [۱] (PLOE) عملگر و اجزاء قرار دارد. زمانی که عیب PLOE در عملگرها اتفاق می‌افتد، خروجی عملگر $u^F(t)$ از ورودی کنترلی $u(t)$ فاصله می‌گیرد که می‌توان آن را به‌صورت زیر مدل‌سازی نمود:

$$u^F(t) = \big(I - \varrho(t)\big)u(t) \qquad (۹)$$

که در آن $u^F(t) = [u_1^F(t), \quad \cdots \quad, u_m^F(t)]^T$ است و $u_i^F(t)$ به سیگنال خروجی عملگر معیوب iام اشاره دارد. علاوه بر این، $\varrho(t) = diag\{\varrho_1(t), \ \cdots, \varrho_m(t)\}$ است و $\varrho_i(t)$ نرخ کاهش کارایی متغیر با زمان در عملگر iام است.

عیوب اجزاء باعث تغییر در ساختار و مشخصه‌های سیستم می‌شوند. ازاین‌رو دینامیک سیستم در معرض عیوب PLOE عملگر و اجزاء را می‌توان به‌صورت زیر نوشت:

$$\dot{x}(t) = f_F(x(t), t) + g_{F(t)}(x(t), t)\Big(\big(I - \varrho(t)\big)u(t)\Big) \qquad (۱۰)$$

که در آن $f_F(x(t), t)$ و $g_F(x(t), t)$ به ترتیب دینامیک‌های ناشناخته انحراف و ورودی پرنده در حالت معیوب هستند.

فرض ۲: عیب کاهش کارایی عملگر محدود است به این معنی که برای $i = 1, 2, 3$، $0 < \varrho_i(t) \le 1$.

فرض ۳: با توجه مشخصه‌های زمانی، عیوب را می‌توان به سه دسته ناگهانی [۲]، تدریجی [۳] و تکرارشونده [۴] تقسیم نمود [۳۰]، که در سیستم (۱۰) تنها دو حالت ناگهانی و تکرارشونده در نظر گرفته‌شده است. ازاین‌رو عیب عملگر $\varrho_i(t)$ به‌صورت تکه‌ای ثابت و عیب اجزا $f_F(x(t), t)$ و $g_F(x(t), t)$ به‌صورت تکه‌ای پیوسته از سمت راست هستند که نقاط ناپیوستگی آن‌ها در لحظات وقوع عیب ایجاد می‌شوند.

t_{F_k} را kامین لحظه تغییر میزان عیب در سیستم در نظر بگیرید. بر طبق فرض ۳، سیستم (۱۰) را در طول بازه زمانی $\big[t_{F_k}, t_{F_{k+1}}\big)$ می‌توان به‌صورت مستقل از زمان زیر نوشت:

$$\dot{x}(t) = f_F\big(x(t)\big) + g_F\big(x(t)\big)\big((I - \varrho)u(t)\big) \qquad (۱۱)$$

بدون از دست دادن کلیت روابط، مسئله طراحی FTOATC برای سیستم (۱۱) در طول باز زمانی $\big[t_{F_k}, t_{F_{k+1}}\big)$ انجام می‌شود و سپس به‌تمامی $t \ge 0$ توسعه داده می‌شود.

۳-۳- مسئله FTOATC مقید

فرض ۴: بردار حالت $x_d(t)$ مسیر مرجع محدود برای سیستم (۱۱) است که به ازای شرایط اولیه محدود توسط تابع پیوسته لیپشیتز $h_d(.) \in R^n$ با $h_d(0) = 0$ به‌صورت زیر تولید می‌شود:

$$\dot{x}_d(t) = h_d\big(x_d(t)\big) \qquad (۱۲)$$

ازاین‌رو خطای ردیابی به‌صورت زیر قابل‌تعریف است:

$$e_d(t) \triangleq x(t) - x_d(t) \qquad (۱۳)$$

لذا بر اساس (۱۱)، (۱۲) و (۱۳)، دینامیک خطای ردیابی به‌صورت زیر قابل‌محاسبه است:

$$\dot{e}_d = \dot{x} - \dot{x}_d = f_F(e_d + x_d) + g_F(x_d + e_d)\big((I - \varrho)u\big) - H(x_d) \qquad (۱۴)$$

هدف این مقاله پیدا کردن قانون کنترل $u(t)$ به نحوی است که سیستم (۱۱) بتواند ردیابی مسیر مرجع $x_d(t)$ را ضمن کمینه کردن تابع هزینه موردنظر طراح تضمین دهد.

از سوی دیگر $\mathcal{S} \in R^n$ را مجموعه حالات ایمن سیستم در نظر بگیرید. در اینجا فرض می‌شود که مجموعه $\mathcal{S}$ را بتوان به‌صورت یک مجموعه چندبر [۵] با رابطه زیر بیان نمود:

$$\mathcal{S} = \{x | c_i(x) \ge 0, i = 1, \dots, N_c\} \qquad (۱۵)$$

که در آن $c_i(x)$ یک تابع مشتق‌پذیر نسبت به x است و $c_i(x) \ge 0$ به iامین مرز از مجموعه $\mathcal{S}$ یا در واقع iامین قید ایمنی اشاره دارد.

همچنین درصورتی‌که حد اشباع عملگرها برابر u_m باشد، می‌توان مجموعه ورودی‌های کنترلی مقید را به‌صورت زیر تعریف نمود:

$$\mathcal{U}_c = \Big\{u \big| \sup_t |u_i| < u_m, \ i = 1, \dots, m\Big\} \qquad (۱۶)$$

---

[1] Partial Loss of Effectiveness
[2] Abrupt
[3] Incipient
[4] Intermittent
[5] Polytope

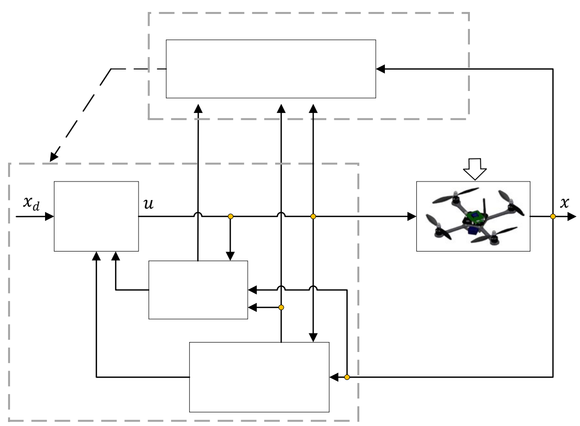
درنتیجه مسئله طراحی FTOATC برای سیستم (۱۴) را در حضور قید ورودی $\mathcal{U}_c$ و قید حالت $\mathcal{S}$ می‌توان به‌صورت عمومی زیر بیان نمود:

$$\min_{u\in(\mathcal{U}_a\cap\mathcal{U}_c)} J(e_d,u) = \int_0^{\infty} e^{-\gamma(\tau-t)} r\big(e_d(\tau),u(\tau)\big)d\tau \qquad (۱۷)$$
$$s.t.\ (۱۲), (۱۴), x(0)=x_0, x\in\mathcal{S}$$

در ادامه به حل مسئله (۱۷) بدون نیاز به شناخت دینامیک سیستم پرداخته می‌شود.

۴- طراحی FTOATC

شکل **۲** ساختار روش پیشنهادی را نشان می‌دهد. در این ساختار، شبکه شناساگر وظیفه تخمین دینامیک ناشناخته سیستم و شبکه نقاد وظیفه تخمین مقدار تابع ارزش را بر عهده دارد. سپس قانون FTOATC مبتنی بر تخمین به‌دست‌آمده از دینامیک سیستم و مقدار تابع ارزش محاسبه می‌شود.



شکل (۲): ساختار روش FTOATC پیشنهادی.

در این روش، هر دو شبکه عصبی در یک زمان به‌روزرسانی می‌شوند که در مقایسه با روشی ترتیبی، یک الگوریتم آموزش هم‌زمان را نتیجه می‌دهند. شروع فرآیند آموزش در این روش با شناسایی وقوع عیب و یا تغییر مقدار آن صورت می‌پذیرد که توسط واحد آشکارسازی عیب تشخیص داده می‌شود.

۴-۱- شناساگر عصبی تطبیقی

بر اساس فرض ۱، سیستم ناشناخته (۱۱) را می‌توان به‌صورت زیر تخمین زد:

$$\dot{x} = W_1^T\phi_1(x,u) + \varepsilon_1 \qquad (۱۸)$$

که در آن $W_1=[w_{11}^T, w_{12}^T]\in R^{(k_{w_{11}}+k_{w_{12}})\times n}$ و $\phi_1(x,u)=[\phi_{11}^T(x), u^T\phi_{12}^T(x)]^T\in R^{k_{w_{11}}+k_{w_{12}}}$ و $\phi_{11}\in R^{k_{w_{11}}}$ و $\phi_{12}\in R^{k_{w_{12}}\times m}$ توابع پایه شبکه عصبی و ε_1 خطای تخمین هستند. با توجه به اینکه بردار وزن‌های ایده‌آل W_1 ناشناخته است، رابطه (۱۱) می‌تواند به صورت زیر تخمین زده شود:

$$\hat{x} = \widehat{W}_1^T\phi_1(x,u) \qquad (۱۹)$$

به‌منظور ارائه قانون به‌روزرسانی برای تخمین $\widehat{W}_1$، متغیرهای فیلتر شده x_f و ϕ_{1f} به‌صورت زیر تعریف‌شده‌اند:

$$\begin{cases} k\dot{x}_f + x_f = x \\ k\dot{\phi}_{1f} + \phi_{1f} = \phi_1 \end{cases} \qquad (۲۰)$$

که $k>0$ یک عدد مثبت ثابت است. با استفاده از (۱۹) می‌توان دینامیک بردار حالت فیلتر شده x_f را به‌صورت زیر بازنویسی کرد:

$$\dot{x}_f = \frac{x-x_f}{k} = {W_1}^T\phi_{1f}(x,u) + \varepsilon_{1f} \qquad (۲۱)$$

که در آن $\varepsilon_{1f}\in R^n$ فیلتر شده خطای تخمین ترکیب‌شده ε_1 طبق رابطه $k\dot{\varepsilon}_{1f}+\varepsilon_{1f}=\varepsilon_1$ است.

با نمادگذاری مشابه، $\hat{x}\in R^n$ و $\hat{x}_f\in R^n$ را به ترتیب به‌عنوان بردار تخمین حالت x و فیلتر شده آن در نظر بگیرید، که با توجه به (۲۰) داریم: $k\dot{\hat{x}}_f+\hat{x}_f=\hat{x}$. برای هر $l,k_1,k_2>0$، ماتریس‌های پایه فیلتر شده $P\in R^{(k_{w_{11}}+k_{w_{12}})\times(k_{w_{11}}+k_{w_{12}})}$ و $D\in R^{(k_{w_{11}}+k_{w_{12}})\times n}$ به‌صورت زیر تعریف می‌شوند:

$$\begin{cases} \dot{P} = -\left(\frac{l+\|e_1\|^{k_1}}{1+\|e_2\|^{k_2}}\right)P + \phi_{1f}\phi_{1f}^T \\ \dot{D} = -\left(\frac{l+\|e_1\|^{k_1}}{1+\|e_2\|^{k_2}}\right)D + \phi_{1f}\left[\frac{x-x_f}{k}\right]^T \end{cases} \qquad (۲۲)$$

که در آن $P(0)=0$ و $D(0)=0$ است. علاوه بر این، e_1 بردار خطای تقریب حالت فیلتر شده به‌صورت $e_1=x_f-\hat{x}_f$ و e_2 بردار خطای حالت فیلتر شده به‌صورت $e_2=x-x_f$ است. همچنین ماتریس کمکی $M\in R^{(k_{w_{11}}+k_{w_{12}})\times n}$ که تابعی از P و D است را به‌صورت زیر نظر بگیرید:

$$M = P\widehat{W}_1 - D \qquad (۲۳)$$

بردار خطای حالت فیلتر شده $\hat{x}_f$ به‌صورت $k\dot{\hat{x}}_f+\hat{x}_f=\hat{x}$ تعریف می‌شود که در آن $\hat{x}$ بردار تخمین حالت x است. سپس قانون به‌روزرسانی برای $\widehat{W}_1$ را می‌توان به‌صورت زیر طراحی کرد:

$$\dot{\widehat{W}}_1 = -\Gamma\left(M + \sum_{j=1}^{N_s} M_j\right) = -\Gamma\left(P\widehat{W}_1 - D + \sum_{j=1}^{N_s} P_j\widehat{W}_1 - \sum_{j=1}^{N_s} D_j\right) \quad (۲۴)$$

که در آن $\Gamma \in R^{(k_{w_{11}}+k_{w_{12}})\times(k_{w_{11}}+k_{w_{12}})}$ یک ماتریس بهره یادگیری مثبت معین است و سرعت همگرایی وزن‌ها را به مقادیر واقعی‌شان را مشخص می‌کند. علاوه بر این، $M_j \triangleq M(t_j)$، $P_j \triangleq P(t_j)$ و $D_j \triangleq D(t_j)$ به داده‌های دریافت و ذخیره‌شده در حافظه در زمان $t_j < t$ که توسط طراح تعیین‌شده، اشاره دارد. قابل‌توجه است که N_s تعداد زمان‌های ذخیره‌سازی داده است که توسط طراح تعیین می‌شود.

توجه ۱: ایده ER مبتنی بر استفاده از داده‌های گذشته ذخیره‌شده به‌منظور بررسی آسان‌تر برقراری شرط برانگیختگی کامل (PE) است [۳۱]، که در ادامه نشان داده می‌شود که یک شرط لازم برای همگرایی وزن‌ها شبکه به مقادیر واقعی‌شان محسوب می‌شود. علاوه بر این، اثبات می‌گردد که استفاده از رویکرد ER، باعث افزایش سرعت همگرایی و کاهش باند خطای وزن‌های شبکه شناساگر می‌شود.

توجه ۲: مقدار k در رابطه (۲۰)، پهنای باند فیلتر را مشخص می‌کند که به‌منظور حفظ مقاومت نسبت به نویز اندازه‌گیری بایستی کمینه مقدار در نظر گرفته شود. همچنین $(l + \|e_1\|^{k_1})/(1 + \|e_2\|^{k_2})$ در رابطه (۲۲)، مقدار ضریب فراموشی و بهره حالت ماندگار را مشخص می‌کند. بنابراین انتخاب مقادیر مناسب برای l، k_1 و k_2 به یک مصالحه بین خطای تخمین، سرعت همگرایی و مقاومت احتیاج دارد.

تعریف ۱: تابع $\phi_1(x(t), u(t))$ را برانگیخته کامل گوییم درصورتی‌که اعداد مثبت ثابت α_1، α_2 و T_1 برای تمامی $t \geq 0$ به نحوی وجود داشته باشد که:

$$\alpha_1 I \leq \int_t^{t+T_1} \phi_1(x(\tau), u(\tau))\phi_1(x(\tau), u(\tau))d\tau \leq \alpha_2 I \quad (۲۵)$$

می‌توان نشان داد که مثبت معین بودن ماتریس P در (۲۲) یک شرط ضروری برای همگرایی خطای تخمین $\widetilde{W}_1$ است که صورت $\widetilde{W}_1 = W_1 - \widehat{W}_1$ تعریف می‌شود. $\lambda_{max}(.)$ و $\lambda_{min}(.)$ را به ترتیب به‌عنوان بزرگ‌ترین و کوچک‌ترین مقدار ویژه ماتریس‌های متناظر با آن در نظر بگیرید. لم ۱ رابطه بین PE بودن بردار ϕ_1 و مثبت معین بودن P را نشان می‌دهند.

لم ۱ [۳۲]: درصورتی‌که بردار $\phi_1(x,u)$ در (۱۹)، PE باشد، ماتریس P در (۲۲) مثبت معین خواهد بود؛ به این یعنی که عدد مثبت σ به نحوی وجود دارد که $\lambda_{min}(P) > \sigma > 0$.

قضیه ۱: سیستم (۱۹) با قانون به‌روزرسانی (۲۴) که رابطه (۱۱) را تخمین می‌زند در نظر بگیرید. اگر بردار ϕ_1 دارای ویژگی PE باشد، سپس خطای تخمین $\widetilde{W}_1$ زمانی که خطای تخمین ε_1 صفر باشد پایدار مجانبی نمایی خواهد بود و زمانی که خطای تخمین ε_1 صفر نباشد UUB خواهد بود.

اثبات: با جایگذاری (۲۱) در (۲۲) و سپس انتگرال‌گیری از آن، می‌توان دریافت که:

$$D = PW_1 - v \quad (۲۶)$$

که در آن $v = -\int_0^t e^{-\left(\frac{l+\|e_1\|^{k_1}}{1+\|e_2\|^{k_2}}\right)(t-\tau)} \phi_{1f}(\tau)\varepsilon_{1f}^T(\tau)d\tau$. از آنجایی که ϕ_{1f}، ε_{1f}، e_1 و e_2 همگی محدود هستند می‌توان نتیجه گرفت که خطای v نیز محدود است؛ یا عبارت دیگر عدد مثبت $\bar{B}_v$ وجود دارد که $\|v\| \leq \bar{B}_v$.

سپس با جایگذاری (۲۶) در (۲۳) خواهیم داشت:

$$M = P\widehat{W}_1 - PW_1 - v = -P\widetilde{W}_1 + v \quad (۲۷)$$

رابطه (۲۷) نشان می‌دهد که ماتریس M حاوی اطلاعات خطای تخمین وزن‌های شبکه است. با جایگذاری (۲۷) در (۲۴) و در نظر گرفتن $\dot{\widetilde{W}}_1 = -\dot{\widehat{W}}_1$، دینامیک خطای وزن‌های شبکه به‌صورت زیر به دست می‌آید:

$$\dot{\widetilde{W}}_1 = \Gamma\left(-P\widetilde{W}_1 + v - \sum_{j=1}^{N_s} P_j\widetilde{W}_1 + \sum_{j=1}^{N_s} v_j\right) \quad (۲۸)$$

حال تابع لیاپانوف کاندید را برای دینامیک خطای وزن‌های شبکه به‌صورت زیر نظر بگیرید:

$$V_1 = \frac{1}{2}tr\left(\widetilde{W}_1^T\Gamma^{-1}\widetilde{W}_1\right) \quad (۲۹)$$

بنابراین مشتق زمانی آن به‌صورت زیر قابل‌محاسبه خواهد بود:

$$\dot{V}_1 = tr\left(\widetilde{W}_1^T\Gamma^{-1}\dot{\widetilde{W}}_1\right) = -tr\left(\widetilde{W}_1^T P_T\widetilde{W}_1\right) + tr\left(\widetilde{W}_1^T v_T\right) \quad (۳۰)$$

که در آن $P_T \triangleq \left(P + \sum_{j=1}^{N_s} P_j\right)$ و $v_T \triangleq \left(v + \sum_{j=1}^{N_s} v_j\right)$ است. قابل استنتاج است که v_T نیز مانند v محدود است، به‌طوری‌که عدد مثبت $\bar{B}_{v_T}$ وجود دارد که $\|v_T\| \leq \bar{B}_{v_T}$.

الف) زمانی $\varepsilon_1 = 0$ باشد؛ در این حالت $v_T = 0$ است و درنتیجه (۲۹) به‌صورت زیر قابل بازنویسی است:

$$\dot{V}_1 = -tr\left(\widetilde{W}_1^T P_T \widetilde{W}_1\right) < -\lambda_{min}(P_T)\left\|\widetilde{W}_1\right\|^2 \leq -\beta_1 V_1 \quad (۳۱)$$

که با توجه مثبت معین بودن P، $\beta_1 = 2\lambda_{min}(P_T)/\lambda_{max}(\Gamma^{-1})$ یک عدد مثبت است. لذا طبق قضیه لیاپانوف خطای تخمین به‌صورت نمایی به صفر همگرا می‌شود.

ب) زمانی $\varepsilon_1 \neq 0$ باشد، رابطه (۲۹) را می‌توان به‌صورت زیر بازنویسی کرد:

$$\dot{V}_1 = -tr\left(\widetilde{W}_1^T P_T \widetilde{W}_1\right) + tr\left(\widetilde{W}_1^T v_T\right) \leq -\left\|\widetilde{W}_1\right\|\left(\lambda_{min}(P_T)\left\|\widetilde{W}_1\right\| - \bar{B}_{v_T}\right) \quad (۳۲)$$

با توجه به (۳۲)، برای اینکه $\dot{V}_1$ منفی باشد باید نامساوی زیر برآورده شود:

$$\left\|\widetilde{W}_1\right\| > \frac{\bar{B}_{v_T}}{\lambda_{min}(P_T)} \quad (۳۳)$$

سپس با توجه به قضیه لیاپانوف توسعه داده‌شده، خطای تخمین در بیرون دایره به شعاع $\mathcal{R} \triangleq \frac{\bar{B}_{v_T}}{\lambda_{min}(P_T)}$، پایدار UUB خواهد بود.

■

توجه ۳: در مقایسه با [۲۱]-[۲۳]، ضریب فراموشی و بهره حالت ماندگار در قانون به‌روزرسانی (۲۴) متغیر و تابعی از مقدار لحظه‌ای خطای تخمین حالت فیلتر شده e_1 و خطای حالت فیلتر شده e_2 است. خطای e_1 نسبت به نویزهای فرکانس بالا غیر حساس و مستقل است و مقدار آن در انتهای شناسایی به صفر همگرا می‌شود. از سوی مقابل، خطای e_2 کاملاً نسبت به نویزهای اندازه‌گیری حساس می‌باشد و مقدار آن در هنگام وجود نویز افزایش می‌یابد. در مقایسه با قانون به‌روزرسانی پایه با ضریب فراموشی ثابت l، قانون به‌روزرسانی مبتنی بر ER بیان‌شده با ضریب فراموشی متغیر $(l + \|e_1\|^{k_1})/(1 + \|e_2\|^{k_2})$ ضمن حفظ مقاومت نسبت به نویز اندازه‌گیری باعث بهبود سرعت همگرایی و کاهش محدوده خطا وزن‌های شبکه می‌شود؛ زیرا:

۱) بر اساس رابطه (۲۲)، به‌منظور کاهش تأثیر داده‌های نویزی در فرآیند یادگیری، لازم است که ضریب فراموشی در هنگام وجود نویز به‌اندازه کافی کوچک باشد.

۲) بر اساس رابطه (۳۲)، سرعت همگرایی وزن‌های شبکه شناساگر با افزایش کوچک‌ترین مقدار ویژه ماتریس P_T بهبود میابد. بنابراین استفاده از روش ER با افزایش مقادیر ویژه ماتریس P_T نسبت ماتریس P، باعث افزایش سرعت همگرایی شبکه شناساگر می‌شود.

۳) بر اساس رابطه (۳۱)، باند خطای حالت ماندگار وزن‌های شبکه شناساگر با خطای تخمین v_T رابط مستقیم و با کوچک‌ترین مقدار ویژه ماتریس P_T رابطه معکوس دارد. بنابراین لازم است ضریب فراموشی در طول دوره آموزش و انتهای آن به ترتیب مقداری بزرگ و کوچک داشته باشد.

۴-۲- FTOATC مقید

ازآنجایی‌که دینامیک سیستم (۱۱) ناشناخته در نظر گرفته‌شده است، از تخمین شناسایی‌شده در رابطه (۱۹) در طراحی FTOATC استفاده می‌شود. با تعریف بردار حالت افزوده $z = [\hat{e}_d^T \quad x_d^T]^T$، دینامیک سیستم افزوده به‌صورت زیر خواهد بود:

$$\dot{z} = \hat{F}(z) + \hat{G}(z)u \quad (۳۴)$$

که در آن $\hat{F} \in R^{2n}$ و $\hat{G} \in R^{2n\times m}$ به ترتیب دینامیک‌های انحراف و ورودی سیستم افزوده با رابطه زیر هستند:

$$\hat{F}(z) = \begin{pmatrix} \widehat{w}_1\psi_1(x_d + e_d) - H(x_d) \\ H(x_d) \end{pmatrix}$$

$$\hat{G}(z) = \begin{pmatrix} \widehat{w}_2\psi_2(x_d + e_d) \\ 0 \end{pmatrix}$$

همان‌طور که در [۳۳] نشان داده‌شده است به‌راحتی قابل‌اثبات است که مسئله طراحی کنترل‌کننده پایدارساز برای سیستم افزوده (۳۴)، معادل مسئله کنترل‌کننده ردیاب برای سیستم (۱۱) خواهد بود.

فرض ۵: پیرو فرض‌های ۱ و ۴، می‌توان نتیجه گرفت که سیستم افزوده (۳۴) پیوسته لیپشیتز بوده و بر روی هر مجموعه بسته $\mathcal{X} \in R^{2n}$ شامل مبدأ پایدار پذیر می‌باشد و $F(0) = 0$ است. علاوه بر این، $\hat{F}(z) \leq \bar{B}_{\hat{F}}\|z\|$ و $\hat{G}(z) \leq \bar{B}_{\hat{G}}$ که در آن $\bar{B}_{\hat{F}}$ و $\bar{B}_{\hat{G}}$ اعداد مثبت ثابت هستند.

حل مسئله بهینه‌سازی مقید (۱۷) در حالت کلی بسیار دشوار است. ازاین‌رو به‌منظور تسهیل آن، رویکردهای مختلفی جهت تبدیل مسئله مقید به مسئله نامقید ارائه‌شده است که به ترتیب با انتخاب توابع هزینه و اشباع بر روی حالات و ورودی‌ها سعی در تضمین این قیود دارند.

در این زمینه، به‌منظور تضمین قیود اشباع ورودی در مسئله (۱۷)، تابع عمومی هموار زیر در تابع هزینه مورداستفاده قرارگرفته است [۳۳-۳۵]:

$$U(u) = 2\int_0^u \left(u_m \Phi^{-1}(v/u_m)\right)^T R\,dv = 2u_m \sum_{i=1}^m \int_0^{u_i}\left(\Phi^{-1}(v_i/u_m)\right)^T R_i\,dv_i \qquad (۳۵)$$

که در آن $R \in R^{m\times m}$ یک ماتریس قطری مثبت معین و $\varphi \in R^m$ یک تابع فرد اکیداً صعودی محدود است به‌طوری‌که $\Phi(.) < 1$. در این مقاله $\Phi(.) = \tanh(.)$ در نظر گرفته‌شده است.

همچنین به‌منظور تضمین قیود ایمنی حالات، بر اساس روش SRL، یک جمله CBF به تابع هزینه افزوده می‌شود. ازاین‌رو با در نظر گرفتن هزینه خطای ردیابی به فرم مربعی $Q(z) = e_d^T Q_1 e_d$، مسئله FTOATC مقید (۱۷) برای سیستم (۱۱) تبدیل به مسئله پایداری سازی نامقید زیر برای سیستم افزوده (۳۴) می‌شود:

$$\min_{u\in\mathcal{U}_a} J(z,u) = \int_0^\infty e^{-\gamma(\tau)}\Big(z^T(\tau)Q_T z(\tau) + U\big(u(\tau)\big) + B\big(z(\tau)\big)\Big)d\tau \qquad (۳۶)$$
$$s.t.\ (۳۴),\ x(0) = x_0$$

که در آن $B(z)$ یک تابع CBF است که در فرض زیر معرفی‌شده است.

فرض ۵: $B(z)$ یک تابع مثبت معین با ویژگی‌های زیر است:

۱) $B(z)\to\infty$ زمانی که $z \to \partial\mathcal{S}$، که در آن $\partial\mathcal{S}$ مرز ناحیه $\mathcal{S}$ است.

۲) $B(z)$ به ازای $\forall z \in \mathcal{S}$، اکیداً کاهشی است.

با بهره‌گیری از روش ارائه‌شده در [۱۹]، تابع $B(z)$ به‌صورت زیر در نظر گرفته می‌شود:

$$B(z) = \sum_{i=1}^{N_c} -log\left(\frac{\eta_i c_i(x_d+e_d)}{\eta_i c_i(x_d+e_d)+1}\right) \qquad (۳۷)$$

که در آن η_i ها به ازای $i = 1,\dots,N_c$، اعداد مثبت ثابتی هستند که محدوده اثرگذاری حداکثری $B(z)$ را در نزدیکی iامین مرز ایمنی مشخص می‌کند.

بر اساس (۳۶)، برای کنترل‌کننده قابل‌قبول $u(z)$، تابع ارزش به‌صورت زیر تعریف می‌شود:

$$\mathrm{V}\big(z(t)\big) = \int_t^\infty e^{-\gamma(\tau - t)}\Big(z^T(\tau)Q_T z(\tau) + U\big(u(\tau)\big) + B(z)\Big)d\tau \qquad (۳۸)$$

با مشتق‌گیری از رابطه (۳۸) در طول مسیر سیستم افزوده، همیلتونین متناظر با آن به‌صورت زیر قابل‌تعریف است:

$$H\big(z,u,\nabla V(z)\big) = \nabla V^T(z)\big(\hat F(z) + \hat G(z)u\big) - \gamma V(z) + z^T Q_T z + U(u) + B(z) \qquad (۳۹)$$

که در آن $\nabla V(z) = \partial V(z)/\partial Z$ است. با استفاده از شرط ایستایی $\frac{\partial H}{\partial u} = 0$، قانون کنترل بهینه $u^*(z)$ به فرم بسته ریاضیاتی زیر بدست می‌آید:

$$u^*(z) = -u_m\, tanh\left(\frac{1}{2u_m}R^{-1}\hat G^T(z)\nabla V^*(z)\right) \qquad (۴۰)$$

که در آن $V^*(z)$ تابع ارزش بهینه است. بنابراین با استفاده از (۳۵) خواهیم داشت:

$$U\big(u(t)\big) = 2u_m\int_0^{-u_m\tanh A(z)} \tanh^{-T}\left(\frac{v}{u_m}\right) R\, dv = 2u_m^2 A^T(z) R \tanh\big(A(z)\big) + u_m^2\sum_{i=1}^m R_i \ln\big[1-\tanh^2\big(A_i(z)\big)\big] \qquad (۴۱)$$

که در آن $A(z) = [A_1(z),\dots,A_m(z)]^T = \left(\frac{1}{2u_m}\right)R^{-1}\hat G(z)^T\,\nabla V^*(z)$ و به ازای $i=1,\dots,m$، $A_i(z)\in R$ است. سپس با جایگذاری (۴۰) و (۴۱) در (۳۹) و ساده‌سازی آن، معادله HJB متناظر به صورت زیر بدست می‌آید:

$$\nabla V^{*T}(z)\hat F(z) - \gamma V^*(z) + z^T Q_T z + u_m^2\sum_{i=1}^m R_i\ln[1-\tanh^2 A_i(z)] + B(z) = 0 \qquad (۴۲)$$

با توجه به هموار بودن تابع ارزش بهینه $V^*(z)$، بر اساس قضیه تخمین مرتبه بالا وایرشتراس [۳۶]، می‌توان آن را با استفاده از شبکه عصبی تک لایه نقاد به‌صورت زیر تخمین زد:

$$V^*(z) = W_2^T\phi_2(z) + \varepsilon_2(z) \qquad (۴۳)$$

که در آن $W_2\in R^{N_2}$ بردار وزن‌های ایده‌آل شبکه نقاد بهینه، $\phi_2\in R^{N_2}$ بردار توابع پایه مستقل خطی، N_2 تعداد توابع پایه و $\varepsilon_2\in R$ خطای تخمین است. همچنین مشتق آن به‌صورت $\frac{\partial V^*(z)}{\partial z} = \nabla\phi_2^T(z)W_2 + \nabla\varepsilon_2(z)$ قابل محاسبه است.

فرض ۶: بردار وزن‌های ایده‌آل W_2 و بردار توابع پایه شبکه نقاد $\phi_2(z)$ و گرادیان آن $\nabla\phi_2(z)$ محدود هستند به‌طوری‌که اعداد مثبت $\bar B_{W_2}$، $\bar B_{\phi_2}$ و $\bar B_{\nabla\phi_2}$ وجود دارد که $\|W_2\| < \bar B_{W_2}$، $\|\phi_2\| < \bar B_{\phi_2}$ و $\|\nabla\phi_2\| < \bar B_{\nabla\phi_2}$. همچنین خطای تخمین $\varepsilon_2(z)$ گرادیان آن $\nabla\varepsilon_2(z)$ محدود هستند به‌طوری‌که اعداد مثبت $\bar B_{\varepsilon_2}$ و $\bar B_{\nabla\varepsilon_2}$ وجود دارد که $\|\varepsilon_2\| < \bar B_{\varepsilon_2}$ و $\|\nabla\varepsilon_2\| < \bar B_{\nabla\varepsilon_2}$.

با استفاده از (۴۳)، قانون کنترل بهینه (۴۰) به‌صورت زیر قابل بازنویسی است:

$$u^* = -u_m \tanh\left(\tau(z) + \varepsilon_{2u}(z)\right) \qquad (۴۴)$$

که در آن $\tau(z) = [\tau_1(z), \dots, \tau_m(z)]^T = \frac{1}{2u_m} R^{-1}\hat{G}^T(z)\nabla\phi_2^T(z)W_2$ و به ازای $i = 1, \dots, m$ ، $\tau_i(z) \in R$ و $\varepsilon_{2u}(z) = [\varepsilon_{2u_1}(z) \quad \dots \quad \varepsilon_{2u_m}(z)]^T = \frac{1}{2u_m}R^{-1}\hat{G}^T(z)\nabla\varepsilon_2(z)$ و $\varepsilon_{2u_i}(z) \in R$ است. با استفاده قضیه مقدار میانگین [۳۷]، می‌توان نوشت:

$$u^* = u_m \tanh\left(\tau(z)\right) + \varepsilon_{u^*} \qquad (۴۵)$$

که در آن $\varepsilon_{u^*} = -u_m\left(I_m - (tanh^2(q_1))\right)\varepsilon_{2u}$ و $q_1 = [q_{11}, \dots, q_{1m}]^T$ و به ازای $i = 1, \dots, m$، $q_{1i} \in R$ و بین $A_i(z)$ و $\tau_i(z)$ انتخاب می‌شود.

با جایگذاری (۴۳) و (۴۵) در (۴۲) می‌توان نتیجه گرفت:

$$W_2^T\nabla\phi_2(z)\hat{F}(z) - \gamma W_2^T\phi_2(z) + +z^TQ_Tz + u_m^2\sum_{i=1}^m R_i\ln[1 - \tanh^2\tau_i(z)] + B(z) + \varepsilon_{HJB} = 0 \qquad (۴۶)$$

که در آن ε_{HJB} خطای تخمین HJB ناشی از خطای تخمین شبکه نقاد است و رابطه‌ای به‌صورت زیر دارد:

$$\varepsilon_{HJB} = \nabla\varepsilon_2^T(z)\hat{F}(z) + \sum_{i=1}^m \frac{2u_m^2}{q_{2i}}\tanh(q_{3i})\left(\tanh^2(q_{3i}) - 1\right)\varepsilon_{2u_i} - \gamma\varepsilon_2(z) \qquad (۴۷)$$

که در آن $q_2 = [q_{21}, \dots, q_{2m}]^T$ و $q_3 = [q_{31}, \dots, q_{3m}]^T$ به‌طوری‌که به ازای $i = 1, \dots, m$، $q_{2i} \in R$ و بین $1 - \text{tANH}^2\left(A_i(z)\right)$ و $1 - \tanh^2\left(\tau_i(z)\right)$ و همچنین $q_{3i} \in R$ و بین $A_i(z)$ و $\tau_i(z)$ انتخاب می‌شوند.

با توجه به اینکه W_2 ناشناخته است، $V^*(z)$ به‌صورت زیر تخمین زده می‌شود:

$$\hat{V}(z) = \hat{W}_2^T\phi_2(z) \qquad (۴۸)$$

که $\hat{W}_2$ تقریب وزن‌های ایده‌آل W_2 است. با استفاده از (۴۸)، تخمین قانون کنترل بهینه $\hat{u}$ به‌صورت زیر به دست می‌آید:

$$\hat{u} = -u_m\tanh\left(\hat{\tau}(z)\right) \qquad (۴۹)$$

که در آن $\hat{\tau}(z) = [\hat{\tau}_1(z), \dots, \hat{\tau}_m(z)] = \left(\frac{1}{2u_m}\right)R^{-1}\hat{G}^T(z)\nabla\phi_2^T(z)\hat{W}_2$ و ازای $i = 1, \dots, m$، $\hat{\tau}_i(z) \in R$ است. سپس با جایگذاری (۴۸) و (۴۹) در (۴۶)، تخمین همیلتونین به‌صورت زیر به دست می‌آید:

$$\hat{H}(z, \hat{W}_2) = \hat{W}_2^T\nabla\phi_2(z)\hat{F}(z) - \gamma\hat{W}_2^T\phi_2(z) + z^TQ_Tz + u_m^2\sum_{i=1}^m R_i\ln[1 - \tanh^2\hat{\tau}_i(z)] + B(z) \triangleq \hat{e}\left(z(t)\right) \qquad (۵۰)$$

که در آن $\hat{e}\left(z(t)\right)$ خطای باقی‌مانده بلمن است. درصورتی‌که خطای تقریب وزن‌های شبکه را به‌صورت $\widetilde{W}_2 = W_2 - \hat{W}_2$ در نظر بگیریم، با استفاده از (۴۶) و (۵۰) خطای باقی‌مانده بلمن $\hat{e}(z)$ به‌صورت زیر به دست می‌آید:

$$\hat{e}\left(z(t)\right) = -\widetilde{W}_2^T\nabla\phi_2(z)\hat{F}(z) - \gamma\widetilde{W}_2^T\phi_2(z) + u_m^2\left[\sum_{i=1}^m R_i\ln[1 - \tanh^2\hat{\tau}_i(z)] - \sum_{i=1}^m R_i\ln[1 - \tanh^2\tau_i(z)]\right] - \varepsilon_{HJB} \qquad (۵۱)$$

قابل‌توجه است که برای هر $\xi \in R$، رابطه زیر برقرار است:

$$\ln[1 - \tanh^2\xi] = ln\,4 - 2\xi\text{sgn}(\xi) - 2\ln\left(1 + \exp\left(-2\xi\text{sgn}(\xi)\right)\right) \qquad (۵۲)$$

که در آن sgn (.) تابع علامت است. بنابراین با استفاده از (۵۲)، می‌توان (۵۱) را به‌صورت زیر بازنویسی نمود:

$$\begin{aligned}\hat{e}\left(z(t)\right) &= -\widetilde{W}_2^T\nabla\phi_2(z)\hat{F}(z) - \gamma\widetilde{W}_2^T\phi_2(z) + 2u_m^2R\left[\tau^T(z)\text{sgn}\left(\tau(z)\right) - \hat{\tau}^T(z)\text{sgn}\left(\hat{\tau}(z)\right)\right] + u_m^2\Delta_\tau - \varepsilon_{HJB} \\ &= -\widetilde{W}_2^T\nabla\phi_2(z)\hat{F}(z) - \gamma\widetilde{W}_2^T\phi_2(z) + u_m\left[W_2^T\nabla\phi_2(z)\hat{G}(z)\text{sgn}\left(\tau(z)\right) - \hat{W}_2^T\nabla\phi_2(z)\hat{G}(z)\text{sgn}\left(\hat{\tau}(z)\right)\right] + u_m^2\Delta_\tau - \varepsilon_{HJB} \\ &= -\widetilde{W}_2^T\left(\nabla\phi_2(z)\hat{F}(z) + \gamma\phi_2(z) - \nabla\phi_2(z)\hat{G}(z)\text{sgn}\left(\hat{\tau}(z)\right)\right) + u_mW_2^T\nabla\phi_2(z)\hat{G}(z)\left[\text{sgn}\left(\tau(z)\right) - \text{sgn}\left(\hat{\tau}(z)\right)\right] + u_m^2\Delta_\tau - \varepsilon_{HJB} \\ &= -\widetilde{W}_2^T\left(\nabla\phi_2(z)\hat{F}(z) + \gamma\phi_2(z) - \nabla\phi_2(z)\hat{G}(z)\text{sgn}\left(\hat{\tau}(z)\right)\right) + \rho(z)\end{aligned} \qquad (۵۳)$$

که در آن

$$\Delta_\tau = 2\sum_{i=1}^m R_i\ln\frac{1 + \exp\left(-2\tau_i(z)\text{sgn}\left(\tau_i(z)\right)\right)}{1 + \exp\left(-2\hat{\tau}_i(z)\text{sgn}\left(\hat{\tau}_i(z)\right)\right)}$$

$$\rho(z) = u_mW_2^T\nabla\phi_2(z)\hat{G}(z)\left[\text{sgn}\left(\tau(z)\right) - \text{sgn}\left(\hat{\tau}(z)\right)\right] + u_m^2\Delta_\tau - \varepsilon_{HJB}$$

قبل از پرداختن به چگونگی محاسبه ورودی کنترلی بهینه، فرض زیر موردنیاز خواهد بود.

فرض ۷: فرض کنید $J_1(z)$ یک تابع لیاپانوف کاندید شعاعی نامحدود برای سیستم (۳۴) باشد به‌طوری‌که $\nabla J_1^T(z)\left(\hat{F}(z) + \hat{G}(z)u^*\right) \le 0$ و $\nabla J_1 = \frac{\partial J_1}{\partial z}$ باشد. در این صورت ماتریس مثبت معین $\vartheta(z) \in R^{2n\times 2n}$ وجود دارد به‌طوری‌که

$$\nabla J_1^T(z)\left(\hat{F}(z) + \hat{G}(z)u^*\right) = -\nabla J_1^T(z)\vartheta(z)\nabla J_1(z) \qquad (۵۴)$$

که $\|\vartheta(z)\| = 0$ اگر و تنها اگر $z = 0$ و $\vartheta(z) \to \infty$ زمانی که $z \to \infty$. علاوه بر این برای هر مجموعه بسته $\mathcal{X} \in R^{2n}$، اعداد مثبت $\underline{B}_\vartheta$ و $\overline{B}_\vartheta$ وجود دارد به‌طوری‌که $\underline{B}_\vartheta \le \vartheta(z) \le \overline{B}_\vartheta$.

توجه ۴: اغلب فرض می‌شود که خروجی سیستم حلقه بسته $\hat{F}(z)+\hat{G}(z)u^*$ بر روی مجموعه بسته $\mathcal{X}\in R^{2n}$ محدود است [۲۶]، به این معنی که عدد مثبت $\wp$ وجود دارد به‌طوری‌که $\|\hat{F}(z)+\hat{G}(z)u^*\|\le \wp$. درحالی‌که در اینجا فرض می‌شود که خروجی سیستم حلقه بسته از بالا محدود به تابعی از حالت z باشد به‌طوری‌که عدد مثبت q وجود دارد که $\|\nabla J(\hat{F}(z)+\hat{G}(z)u^*)\|\le q\|\nabla J_1\|^2$. قابل‌توجه است که $J_1^T\big(\hat{F}(z)+\hat{G}(z)u^*\big)\le 0$ به معنی امکان‌پذیر بودن (۵۴) است.

بر اساس (۵۳)، می‌توان نتیجه گرفت با کمینه شدن $E=\frac{1}{2}\hat{e}^2$، بردار وزن‌های $\widehat{W}_2$ به W_2 همگرا می‌شود. از سوی دیگر قابل‌توجه است کمینه‌سازی E به‌تنهایی پایداری سیستم را در طول آموزش تضمین نمی‌دهد. به‌منظور کاهش این محدودیت، از یک جمله پایدارساز در قانون به‌روزرسانی شبکه نقاد استفاده‌شده است. ازاین‌رو قانون به‌روزرسانی $\widehat{W}_2$ مبتنی بر الگوریتم گرادیان نزولی نرمال شده به‌صورت زیر با نرخ یادگیری متغیر در نظر گرفته‌شده است:

$$\begin{aligned}\dot{\widehat{W}}_2 &= -\alpha_1\big(|\hat{e}(z(t))|+l_2\big)\left(\frac{\bar{\sigma}}{m_s}\right)\hat{e} + \alpha_1\Xi(z,\hat{u})\left(\frac{1}{2}\nabla\phi_2(z)\hat{G}(z)R^{-T}\big[I_m - \mathfrak{B}(\hat{\tau}(z))\big]\hat{G}^T(z)\nabla J_1\right) + \alpha_1\big(|\hat{e}(z(t))| + l_2\big)\bigg((K_1\bar{\sigma}^T - K_2)\widehat{W}_2 \\ &\quad + u_m\nabla\phi_2(z)\hat{G}(z)R^{-T}\big[\tanh(\hat{\tau}(z)) - \operatorname{sgn}(\hat{\tau}(z))\big]\left(\frac{\bar{\sigma}^T}{m_s}\right)\widehat{W}_2\bigg)\end{aligned} \qquad (۵۵)$$

که در آن $\bar{\sigma}=\frac{\hat{\sigma}^T}{(\hat{\sigma}^T\hat{\sigma}+1)}$، $\hat{\sigma}=\nabla\phi_2(z)\big(\hat{F}(z)-\hat{G}(z)\hat{u}\big)$ – $\gamma\phi_2(z)$، $m_s=\hat{\sigma}^T\hat{\sigma}+1$، $\mathfrak{B}(\hat{\tau}(z))=\operatorname{diag}\{\tanh^2\hat{\tau}_i(z)\}$، $i=1,\dots,m$ و J_1 در فرض ۷ تشریح شده است. همچنین $\alpha_1>0$، $0\le l_2<1$، $K_1\in R^{N_2}$ و $K_2\in R^{N_2\times N_2}$ پارامترهای ثابت قابل تنظیم هستند و $\Xi(z,\hat{u})$ تابع تکه‌ای ثابت و نشان‌دهنده وضعیت پایداری سیستم حلقه بسته است که رابطه‌ای به‌صورت زیر دارد:

$$\Xi(z,\hat{u}) = \begin{cases} 0 & if\ \nabla J_1^T\dot{z} = \nabla J_1^T(\hat{F}(z) - u_m\hat{G}(z)\tanh(\hat{\tau}(z))) < 0 \\ 1 & else \end{cases} \qquad (۵۶)$$

بنابراین با استفاده از (۵۵) و (۵۶) دینامیک خطای تقریب وزن‌های شبکه $\widetilde{W}_2$ به‌صورت زیر به دست می‌آید:

$$\begin{aligned}\dot{\widetilde{W}}_2 &= \alpha_1\big(|\hat{e}(z(t))| + l_2\big)\frac{\bar{\sigma}}{m_s}\Big(-\widetilde{W}_2^T\hat{\sigma} + u_m\widetilde{W}_2^T\nabla\phi_2(z)\hat{G}(z)R^{-T}\big[\operatorname{sgn}(\hat{\tau}(z)) - \tanh(\hat{\tau}(z))\big] + \rho(z)\Big) \\ &\quad - \alpha_1\Xi(z,\hat{u})\left(\frac{1}{2}\nabla\phi_2(z)\hat{G}(z)R^{-T}\big[I_m - \mathfrak{B}(\hat{\tau}(z))\big]\hat{G}^T(z)\nabla J_1\right) + \alpha_1\big(|\hat{e}(z(t))| + l_2\big)\bigg((K_2 - K_1\bar{\sigma}^T)\widehat{W}_2 \\ &\quad + u_m\nabla\phi_2(z)\hat{G}(z)R^{-T}\big[\operatorname{sgn}(\hat{\tau}(z)) - \tanh(\hat{\tau}(z))\big]\left(\frac{\bar{\sigma}}{m_s}\right)\widehat{W}_2\bigg)\end{aligned} \qquad (۵۷)$$

قضیه ۲: مسئله بهینه‌سازی (۳۶) را به همراه معادله HJB (۴۲) برای سیستم افزوده (۳۴) در نظر بگیرید. با فرض برقراری فرض ۴ و ۶ و PE بودن بردار توابع پایه $\phi_1(x,u)$، درصورتی‌که قانون‌های به‌روزرسانی (۲۴) و (۵۵) به ترتیب جهت آموزش هم‌زمان وزن‌های شبکه‌های شناساگر و نقاد مورداستفاده قرار گیرند، آنگاه تخمین قانون FTOATC در (۴۹) ضمن تضمین قیود ورودی و حالت به جواب مسئله بهینه‌سازی همگرا می‌شود و بردار حالت z، خطای تقریب وزن‌های شبکه شناساگر $\widetilde{W}_1$ و شبکه نقاد $\widetilde{W}_2$، UUB خواهند بود.

اثبات: تابع لیاپانوف کاندید را به‌صورت زیر نظر بگیرید:

$$J = J_1(z) + \frac{1}{2}tr\big(\widetilde{W}_1^T\Gamma^{-1}\widetilde{W}_1\big) + \frac{1}{2\alpha_1}\widetilde{W}_2^T\widetilde{W}_2 \qquad (۵۸)$$

که $J_1(z)$ در فرض ۷ تشریح شده، Γ ماتریس بهره یادگیری شبکه شناساگر در رابطه (۲۴) و α_1 نرخ یادگیری پایه در شبکه نقاد طبق رابطه (۵۵) است. لذا مشتق تابع لیاپانوف در طول مسیر سیستم حلقه بسته به‌صورت زیر خواهد بود:

$$\begin{aligned}\dot{J} &= \nabla J_1^T\big(\hat{F}(z) - u_m\hat{G}(z)\tanh(\hat{\tau}(z))\big) + \\ &\quad tr\left(\widetilde{W}_1^T\Gamma^{-1}\dot{\widetilde{W}}_1\right) + \dot{\widetilde{W}}_2^T\alpha_1^{-1}\widetilde{W}_2\end{aligned} \qquad (۵۹)$$

با بهره‌گیری از (۵۷)، جمله آخر رابطه (۵۹) به‌صورت زیر قابل بازنویسی است:

$$\begin{aligned}\dot{\widetilde{W}}_2^T\alpha_1^{-1}\widetilde{W}_2 &= \big(|\hat{e}(z(t))| + l_2\big)\Big(-\widetilde{W}_2^T\hat{\sigma} + u_m\widetilde{W}_2^T\nabla\phi_2(z)\hat{G}(z)R^{-T}\mathfrak{R}(z) + \rho(z)\Big)\frac{\bar{\sigma}^T}{m_s}\widetilde{W}_2 - \\ &\quad \Xi(z,\hat{u})\left(\frac{1}{2}\nabla J_1^T\hat{G}(z)\big[I_m - \mathfrak{B}(\hat{\tau}(z))\big]R\hat{G}^T(z)\nabla\phi_2^T(z)\right)\widetilde{W}_2 + \big(|\hat{e}(z(t))| +\end{aligned} \qquad (۶۰)$$

$$l_2)\left(u_m\widetilde{W}_2^T\nabla\phi_2(z)\hat{G}(z)R^{-T}\Re(z)\left(\frac{\bar{\sigma}^T}{m_s}\right)\widehat{W}_2 + \widetilde{W}_2^T\left(K_2\widehat{W}_2 - K_1\bar{\sigma}^T\widehat{W}_2\right)\right) = \left(\left|\hat{e}(z(t))\right| + l_2\right)\left(-\widetilde{W}_2^T\bar{\sigma}\bar{\sigma}^T\widetilde{W}_2 + \alpha(z)\bar{\sigma}^T\widetilde{W}_2 + \widetilde{W}_2^T\beta(z)\right) - \Xi(z,\hat{u})\left(\frac{1}{2}\nabla J_1^T\hat{G}(z)[I_m - \mathfrak{B}(\hat{\tau}(z))]R\hat{G}^T(z)\nabla\phi_2^T(z)\right)\widetilde{W}_2 + \left(\left|\hat{e}(z(t))\right| + l_2\right)\widetilde{W}_2^T\left(K_2\widehat{W}_2 - K_1\bar{\sigma}^T\widehat{W}_2\right)$$

که در آن $\alpha(z) = \frac{\rho(z)}{m_s}$، $\beta(z) = u_m\nabla\phi_2(z)\hat{G}(z)R^{-T}\Re(z)\left(\frac{\bar{\sigma}^T}{m_s}\right)\widetilde{W}_2$ و $\Re(z) = sgn(\hat{\tau}(z)) - tanh(\hat{\tau}(z))$ است.

بر اساس تعریف $\widetilde{W}_2$، جمله آخر رابطه (۶۰) را می‌توان به‌صورت زیر بازنویسی کرد:

$$\widetilde{W}_2^T\left(K_2\widehat{W}_2 - K_1\bar{\sigma}^T\widehat{W}_2\right) = \widetilde{W}_2^T K_2 W_2 - \widetilde{W}_2^T K_2\widetilde{W}_2 - \widetilde{W}_2^T K_1\bar{\sigma}^T W_2 + \widetilde{W}_2^T K_1\bar{\sigma}^T\widetilde{W}_2 \qquad (۶۱)$$

با تعریف $\mathcal{Y} = [\widetilde{W}_2^T\bar{\sigma} \quad \widetilde{W}_2^T]^T$ می‌توان رابطه (۶۰) را به‌صورت زیر بازنویسی کرد:

$$\dot{\widetilde{W}}_2^T\alpha_1^{-1}\widetilde{W}_2 = \left(\left|\hat{e}(z(t))\right| + l_2\right)(-\mathcal{Y}^T\mathcal{M}\mathcal{Y} + \mathcal{Y}^T\mathcal{N}) - \Xi(z,\hat{u})\left(\frac{1}{2}\nabla J_1^T\hat{G}(z)[I_m - \mathfrak{B}(\hat{\tau}(z))]R\hat{G}^T(z)\nabla\phi_2^T(z)\right)\widetilde{W}_2 \qquad (۶۲)$$

که در آن

$$\mathcal{M} = \begin{bmatrix} I & -\frac{1}{2}K_1^T \\ -\frac{1}{2}K_1 & K_2 \end{bmatrix}$$

$$\mathcal{N} = \begin{bmatrix} \alpha(z) \\ \beta(z) + K_2W_2 - K_1\bar{\sigma}^T W_2 \end{bmatrix}$$

به‌راحتی می‌توان نشان داد که ماتریس $\mathcal{N}$ محدود است به‌طوری‌که عدد مثبت $\bar{B}_{\mathcal{N}}$ وجود دارد که $\|\mathcal{N}\| \le \bar{B}_{\mathcal{N}}$. رابطه (۵۵) نشان می‌دهد که خطای باقی‌مانده $\hat{e}$ تابعی از $\widetilde{W}_2$ است، ازاین‌رو می‌توان نوشت:

$$\left|\hat{e}(z(t))\right| + l_2 = \left|-\widetilde{W}_2^T\left(\nabla\phi_2(z)\hat{F}(z) + \gamma\phi_2(z) - \nabla\phi_2(z)\hat{G}(z)\mathrm{sgn}(\hat{\tau}(z))\right) + \rho(z)\right| + l_2 \le \left|\widetilde{W}_2^T\left(\nabla\phi_2(z)\hat{F}(z) + \gamma\phi_2(z) - \nabla\phi_2(z)\hat{G}(z)\mathrm{sgn}(\hat{\tau}(z))\right)\right| + |\rho(z)| + l_2 \le h_1\left\|\widetilde{W}_2^T\right\| + h_2 \qquad (۶۳)$$

که در آن

$$h_1 = \left\|\nabla\phi_2(z)\hat{F}(z)\right\| + \gamma\|\phi_2(z)\| + \left\|\nabla\phi_2(z)\hat{G}(z)\mathrm{sgn}(\hat{\tau}(z))\right\|$$
$$h_2 = |\rho(z)| + l_2$$

بر اساس تعریف $\mathcal{N}$ با انتخاب مناسب K_1 و K_2 و همچنین یک مقدار بسیار کوچک برای l_2 می‌توان تضمین داد که :

$$h_2 = |\rho(z)| + l_2 \le \bar{B}_{\mathcal{N}} \qquad (۶۴)$$

علاوه بر این بر اساس تعریف ماتریس $\mathcal{Y}$ می‌توان نوشت:

$$\left|\hat{e}(\widetilde{W}_2)\right| + l_2 \le h_1\|\mathcal{Y}\| + h_2 \qquad (۶۵)$$

با جایگذاری (۶۲) در (۵۹) و با استفاده از (۶۵) و انتخاب K_1 و K_2 به‌نحوی‌که ماتریس $\mathcal{M}$ مثبت معین باشد، داریم:

$$\dot{J} \le \nabla J_1^T\left(\hat{F}(z) + \hat{G}(z)\hat{u}\right) - \Xi(z,\hat{u})\left(\frac{1}{2}\nabla J_1^T\hat{G}(z)[I_m - \mathfrak{B}(\hat{\tau}(z))]R\hat{G}^T(z)\nabla\phi_2^T(z)\right)\widetilde{W}_2 + (h_1\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}})(-\lambda_{min}(\mathcal{M})\|\mathcal{Y}\|^2 + \bar{B}_{\mathcal{N}}\|\mathcal{Y}\|) + tr\left(\widetilde{W}_1^T\Gamma^{-1}\dot{\widetilde{W}}_1\right) \qquad (۶۶)$$

بر اساس (۵۶)، رابطه (۶۶) بایستی در دو حالت $\Xi(z,\hat{u}) = 0$ و $\Xi(z,\hat{u}) = 1$ موردبررسی قرار بگیرد.

الف) $\Xi(z,\hat{u}) = 0$: با توجه به اینکه $\|z\| > 0$ طبق شرط PE تضمین داده می‌شود، می‌توان نتیجه گرفت که عدد مثبت $\dot{z}_{min}$ وجود دارد به‌طوری‌که $\|\dot{z}\| > \dot{z}_{min} > 0$. ازاین‌رو (۶۶) رابطه زیر را نتیجه می‌دهد:

$$\dot{J} \le -\|\nabla J_1\|\dot{z}_{min} + (h_1\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}})(-\lambda_{min}(\mathcal{M})\|\mathcal{Y}\|^2 + \bar{B}_{\mathcal{N}}\|\mathcal{Y}\|) - \|\widetilde{W}_1\|\left(\lambda_{min}(P)\|\widetilde{W}_1\| - \bar{B}_{v_T}\right) = -\|\nabla J_1\|\dot{z}_{min} + \|\mathcal{Y}\|(-\lambda_{min}(\mathcal{M})h_1\|\mathcal{Y}\|^2 + d_1\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}}^2) - \lambda_{min}(P_T)\left(\|\widetilde{W}_1\| - \frac{\bar{B}_{v_T}}{2\lambda_{min}(P_T)}\right)^2 + \Pi \qquad (۶۷)$$

که در آن $d_1 = \bar{B}_{\mathcal{N}}h_1 - \lambda_{min}(\mathcal{M})\bar{B}_{\mathcal{N}}$ و $\Pi = \frac{\bar{B}_{v_T}^2}{4\lambda_{min}(P_T)}$ است. بنابراین برای تضمین $\dot{J} \le 0$ بایستی حداقل یکی از دو نامساوی (۶۸) یا (۶۹) به همراه نامساوی (۷۰) برقرار باشند:

$$\|\nabla J_1\| \ge \frac{\Pi}{\dot{z}_{min}} \triangleq B_{\nabla J_1} \qquad (۶۸)$$

$$\|\widetilde{W}_1\| \ge \sqrt{\frac{\Pi}{\lambda_{min}(P_T)}} + \frac{\bar{B}_{v_T}}{2\lambda_{min}(P_T)} \triangleq B_{\widetilde{W}_1} \qquad (۶۹)$$

$$\|\mathcal{Y}\| > \frac{d_1}{2\lambda_{min}(\mathcal{M})h_1} + \sqrt{\frac{d_1^2}{4\lambda_{min}^2(\mathcal{M})h_1^2} + \frac{\bar{B}_{\mathcal{N}}^2}{\lambda_{min}(\mathcal{M})h_1}} \qquad (۷۰)$$

بر اساس قضیه بسط استاندارد لیاپانوف، می‌توان نتیجه گرفت که ∇J_1، $\widetilde{W}_1$ و $\mathcal{Y}$ به ترتیب با محدوده‌های متناظر تعریف‌شده در (۶۸)، (۶۹) و (۷۰)، UUB هستند.

ب) $\Xi(z,\hat{u}) = 1$: در این حالت کنترل‌کننده جاری ممکن است پایداری سیستم حلقه بسته را تضمین ننماید.

با استفاده از بسط تیلور و در نظر گرفتن تنها اولین جمله آن می‌توان نوشت:

$$tanh\left(\hat{\tau}(z)\right) + \frac{1}{2u_m}\left[I_m - \mathfrak{B}\left(\hat{\tau}(z)\right)\right] R\hat{G}^T(z)\nabla\phi_2^T(z)\widetilde{W}_2 = tanh\left(\tau(z)\right) - O\left(\left(\tau(z) - \hat{\tau}(z)\right)^2\right) \quad (۷۱)$$

که در آن $O\left(\left(\tau(z) - \hat{\tau}(z)\right)^2\right)$ به جملات با مرتبه بالاتر از ۱ اشاره دارد که می‌توان نشان داد اعداد مثبت b_1 و b_2 وجود دارند به‌طوری‌که $\left\|O\left(\left(\tau(z) - \hat{\tau}(z)\right)^2\right)\right\| \le b_1 + b_2\|\widetilde{W}_2\|$.

درنتیجه با جایگذاری (۷۱) در (۶۶) و با استفاده از (۴۵) خواهیم داشت:

$$\dot{J} \le \nabla J_1^T\left(\hat{F}(z) + \hat{G}(z)u^*\right) - \nabla J_1^T\hat{G}(z)\,\varepsilon_{u^*} - u_m\nabla J_1^T\hat{G}(z)O\left(\left(\tau(z) - \hat{\tau}(z)\right)^2\right) + (h_1\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}})(-\lambda_{min}(\mathcal{M})\|\mathcal{Y}\|^2 + \bar{B}_{\mathcal{N}}\|\mathcal{Y}\|) + tr\left(\widetilde{W}_1^T\Gamma^{-1}\dot{\widetilde{W}}_1\right) \quad (۷۲)$$

با استفاده از فرض ۷ و ساده‌سازی ریاضیاتی، رابطه (۷۲) را می‌توان به‌صورت زیر نوشت:

$$\dot{J} \le -\lambda_{min}\left(\vartheta(z)\right)\|\nabla J_1\|^2 + \bar{B}_{\hat{G}}\left(u_m b_1 + \bar{B}_{\varepsilon_{u^*}}\right)\|\nabla J_1\| + \bar{B}_{\hat{G}}u_m b_2\|\nabla J_1\|\|\widetilde{W}_2\| + (h_1\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}})(-\lambda_{min}(\mathcal{M})\|\mathcal{Y}\|^2 + \bar{B}_{\mathcal{N}}\|\mathcal{Y}\|) - \lambda_{min}(P_T)\left(\|\widetilde{W}_1\| - \frac{\bar{B}_{v_T}}{2\lambda_{min}(P_T)}\right)^2 + \frac{\bar{B}_{v_T}^2}{4\lambda_{min}(P_T)} \quad (۷۳)$$

که در آن $\bar{B}_{\varepsilon_{u^*}}$ حد بالای ε_{u^*} است به‌طوری‌که $\|\varepsilon_{u^*}\| \le \bar{B}_{\varepsilon_{u^*}}$.

اعداد مثبت μ_i را به ازای $i = 1, 2$ در نظر بگیرید به‌طوری‌که $\mu_1 + \mu_2 = 1$ و $0 < \mu_i < 1$ باشد. با توجه به اینکه $\|\widetilde{W}_2\|^2 < \|\mathcal{Y}\|^2$ است می‌توان (۷۳) را به‌صورت زیر بازنویسی کرد:

$$\dot{J} \le -\mu_1\lambda_{min}\left(\vartheta(z)\right)\|\nabla J_1\|^2 + \bar{B}_{\hat{G}}\left(u_m b_1 + \bar{B}_{\varepsilon_{u^*}}\right)\|\nabla J_1\| + \frac{\left(\bar{B}_{\hat{G}}u_m b_2\right)^2}{4\mu_2\lambda_{min}\left(\vartheta(z)\right)}\|\mathcal{Y}\|^2 + (h_1\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}})(-\lambda_{min}(\mathcal{M})\|\mathcal{Y}\|^2 + \bar{B}_{\mathcal{N}}\|\mathcal{Y}\|) - \lambda_{min}(P)\left(\|\widetilde{W}_1\| - \frac{\bar{B}_{v_T}}{2\lambda_{min}(P)}\right)^2 + \frac{\bar{B}_{v_T}^2}{4\lambda_{min}(P)} = -\mu_1\lambda_{min}\left(\vartheta(z)\right)\left(\|\nabla J_1\| - \frac{\bar{B}_{\hat{G}}\left(u_m b_1 + \bar{B}_{\varepsilon_{u^*}}\right)}{2\mu_1\lambda_{min}\left(\vartheta(z)\right)}\right)^2 + \|\mathcal{Y}\|(-\lambda_{min}(\mathcal{M})h_1\|\mathcal{Y}\|^2 + d_2\|\mathcal{Y}\| + \bar{B}_{\mathcal{N}}^2) - \lambda_{min}(P_T)\left(\|\widetilde{W}_1\| - \frac{\bar{B}_{v_T}}{2\lambda_{min}(P_T)}\right)^2 + \Upsilon \quad (۷۴)$$

که در آن

$$d_2 = (-\lambda_{min}(\mathcal{M})\bar{B}_{\mathcal{N}} + h_1\bar{B}_{\mathcal{N}}) + \frac{(\bar{B}_{\hat{G}}u_m b_2)^2}{4\mu_2\lambda_{min}\left(\vartheta(z)\right)},$$

$$\Upsilon = \frac{\bar{B}_{v_T}^2}{4\lambda_{min}(P_T)} + \frac{\bar{B}_{\hat{G}}^2\left(u_m b_1 + \bar{B}_{\varepsilon_{u^*}}\right)^2}{4\mu_1\lambda_{min}\left(\vartheta(z)\right)}$$

بنابراین برای اینکه $\dot{J} \le 0$ باشد بایستی حداقل یکی از دو نامساوی (۷۵) یا (۷۶) به همراه نامساوی (۷۷) برقرار باشد:

$$\|\nabla J_1\| > \frac{\bar{B}_{\hat{G}}\left(u_m b_1 + \bar{B}_{\varepsilon_{u^*}}\right)}{2\mu_1\lambda_{min}\left(\vartheta(z)\right)} + \sqrt{\frac{\Upsilon}{\mu_1\lambda_{min}\left(\vartheta(z)\right)}} \triangleq B'_{\nabla J_1} \quad (۷۵)$$

$$\|\widetilde{W}_1\| > \frac{\bar{B}_{v_T}}{2\lambda_{min}(P_T)} + \sqrt{\frac{\Upsilon}{\lambda_{min}(P_T)}} \triangleq B'_{\widetilde{W}_1} \quad (۷۶)$$

$$\|\mathcal{Y}\| > \frac{d_2}{2\lambda_{min}(\mathcal{M})h_1} + \sqrt{\frac{d_2^2}{4\lambda_{min}^2(\mathcal{M})h_1^2} + \frac{\bar{B}_{\mathcal{N}}^2}{\lambda_{min}(\mathcal{M})h_1}} \triangleq B'_{\mathcal{Y}} \quad (۷۷)$$

لذا می‌توان نتیجه گرفت که ∇J_1، $\widetilde{W}_1$ و $\mathcal{Y}$، UUB هستند.
بنابراین برای هر دو حالت الف و ب، $\|\nabla J_1\|$، $\|\widetilde{W}_1\|$ و $\|\mathcal{Y}\|$ با محدوده‌های نهایی $\bar{B}_{\nabla J_1} \triangleq \max\left(B_{\nabla J_1}, B'_{\nabla J_1}\right)$، $\bar{B}_{\widetilde{W}_1} \triangleq \max\left(B_{\widetilde{W}_1}, B'_{\widetilde{W}_1}\right)$ و $\bar{B}_{\mathcal{Y}} \triangleq \max\left(B_{\mathcal{Y}}, B'_{\mathcal{Y}}\right)$ ، UUB هستند.
بر اساس فرض ۷، تابع لیاپانوف کاندید J_1 یک تابع شعاعی نامحدود و پیوسته مشتق‌پذیر نسبت به z است. ازاین‌رو می‌توان از محدود بودن ∇J_1، محدود بودن $\|z\|$ را نتیجه گرفت. لذا $\|z\|$ محدود به عدد مثبت $\bar{B}_z$ است که بر اساس $\bar{B}_{\nabla J_1}$ قابل‌تعیین است. همچنین از محدود بودن $\|\mathcal{Y}\|$ و $\|\widetilde{W}\|^2 < \|\mathcal{Y}\|^2$، می‌توان نتیجه گرفت که خطای تقریب وزن‌های شبکه نقاد $\|\widetilde{W}_2\|$ محدود به عدد مثبت $\bar{B}_{\widetilde{W}_2}$ است.
بر اساس (۴۰) و (۴۹) و همچنین محدود بودن $\|\widetilde{W}_2\|$ و $\|\nabla\phi_2\|$ می‌توان نتیجه گرفت که:

$$\|\hat{u} - u^*\| \le \left\|-\frac{1}{2}R^{-1}\hat{G}^T(z)\nabla\phi_2^T(z)\widetilde{W}_2\right\| \le \frac{1}{2}\lambda_{max}(R^{-1})\bar{B}_{\hat{G}}\bar{B}_{\nabla\phi_2}\bar{B}_{\mathcal{Y}} \triangleq \epsilon_u \quad (۷۸)$$

که $\bar{B}_{\nabla\phi_2}$ حد بالای $\nabla\phi_2(z)$ است به‌طوری‌که $\|\nabla\phi_2\| \le \bar{B}_{\nabla\phi_2}$. لذا درستی قضیه ۲ بدین‌صورت اثبات گردید.
■

۴-۳- آشکارسازی مبتنی بر HJB

حالتی را در نظر بگیرید که قانون کنترل با سیستم در انطباق

است که به‌عنوان وضعیت انطباق شناخته می‌شود. با توجه به اینکه FTOATC طراحی‌شده در دو حالت مقید و نامقید تابع هزینه عملکرد متناظر را کمینه می‌کند، بر اساس روش ارائه‌شده می‌توان دریافت که خطای باقی‌مانده HJB یا درواقع خطای بلمن $\hat{e}(z)$، در شرایط انطباق مقداری نزدیک به صفر دارد. لذا فرض کنید که نامساوی زیر در وضعیت انطباق برای خطای $\hat{e}(z)$ برقرار باشد:

$$|\hat{e}(z)| < \mathcal{L} \tag{۷۹}$$

که $\mathcal{L}$ مقدار از پیش تعیین‌شده بر اساس شرط توقف فرآیند یادگیری FTOATC است. لذا (۷۹) را می‌توان به‌عنوان یک آزمون معقولیت در نظر گرفت [۳۸]، که اولین زمانی که نامساوی نقض شود زمان آشکارسازی عیب خواهد بود.

توجه ۵: حساسیت آشکارسازی عیب رابطه مستقیمی با اثرگذاری عیب عملگر و اجزا بر مقدار معادله HJB دارد. در این رویکرد، تنها عیوبی قابل آشکارسازی هستند که باعث افزایش مقدار خطای معادله HJB یا به‌عبارت‌دیگر کاهش عملکرد سیستم شوند. از سوی دیگر بایستی توجه کرد که $\hat{e}(z(t)) \geq \mathcal{L}$ لزوماً به معنی ناپایداری سیستم نخواهد بود.

۵- شبیه‌سازی

در این قسمت، به‌منظور بررسی عملکرد و کارایی روش پیشنهادی در کنترل وضعیت پرنده کوادروتور، نتایج شبیه‌سازی حاصل از به‌کارگیری آن بر روی مدل غیرخطی پرنده با استفاده از نرم‌افزار متلب آورده شده است. مدل دینامیکی پرنده کوادروتور مطابق رابطه (۸) و پارامترهای آن مطابق با جدول **۱** در نظر گرفته‌شده است.

جدول (۱): پارامترهای مدل دینامیکی پرنده [۲۹].

پارامتر	مقدار	پارامتر	مقدار
$I_x\ (kg.m^2)$	۰/۰۰۷۷۸	l (m)	۰/۳۳
$I_y\ (kg.m^2)$	۰/۰۰۷۷۸	$J_r\ (kg.m^2)$	۰/۰۰۶
$I_z\ (kg.m^2)$	۰/۰۱۴۴		

دینامیک پرنده (۸) در معرض هر دو عیب عملگر و اجزا به‌صورت زیر در نظر گرفته‌شده است:

$$\dot{x}_1(t) = x_2(t) \tag{۸۲}$$

$$\dot{x}_2(t) = \upsilon(t)f\big(x(t)\big) + g\big(x(t)\big)\Big(\big(I - \varrho(t)\big)u(t)\Big)$$

که $\upsilon(t) = diag\{\upsilon_1(t), \cdots, \upsilon_3(t)\}$ و $\varrho(t) = diag\{\varrho_1(t), \cdots, \varrho_3(t)\}$ رفتار زمانی به صورت زیر دارند:

$$\upsilon_1(t) = \begin{cases} 1 & 0 < t < 120 \\ 0.3 & t \geq 120 \end{cases} \tag{۸۳}$$

$$\varrho_1(t) = \begin{cases} 1 & 0 < t < 195 \\ 0.4 & t \geq 195 \end{cases} \tag{۸۴}$$

و برای $i = 2, 3$ در تمامی زمان‌ها $\upsilon_i(t) = 1$ و $\varrho_i(t) = 1$ است. پارامترهای تنظیم تابع هزینه به‌صورت $R = 1$، $Q = 20$ و $\gamma = 5$ در نظر گرفته‌شده است. مسیر مرجع زاویه غلتش $\phi_d(t)$ به‌صورت تکه‌ای ثابت با رفتار زمانی زیر تعریف‌شده است:

$$\phi_d(t) = \begin{cases} 0.5 & t < 120 \\ -0.25 & 120 < t < 195 \\ 0.25 & t > 195 \end{cases}$$

در اینجا مطلوب است سیستم کنترل وضعیت پرنده بتواند ضمن ردیابی بهینه مسیر مرجع زاویه غلتش $\phi_d(t)$ در حضور عیوب عملگر و اجزا، قیود ورودی $|u(t)| < 0.2$ و نرخ غلتش $|\dot{\phi}(t)| < 0.5$ را نیز تضمین نماید.

پارامترهای تنظیم شبکه شناساگر به‌صورت

$$l_1 = 0.5$$
$$l_2 = 0.8$$
$$k_1 = 1$$
$$k_2 = 0.1$$
$$k = 0.005$$
$$\Gamma = 40$$

بردار توابع پایه آن به‌صورت $\phi_1(x, u) = \big[\dot{\phi}\ \dot{\theta}\ \dot{\psi}\ \dot{\phi}\dot{\theta}\ \dot{\phi}\dot{\psi}\ \dot{\theta}\dot{\psi}\ u^T\big]^T$ و $W_1(0)$ به‌صورت یک بردار تصادفی با ابعاد مناسب در نظر گرفته‌شده است.

بردار حالت سیستم افزوده در طراحی کنترل‌کننده به‌صورت زیر در نظر گرفته‌شده است:

$$z = [z_1 \ldots z_7]^T$$
$$= [\phi - \varphi_d \quad \theta \quad \psi \quad \dot{\phi} \quad \dot{\theta} \quad \dot{\psi} \quad \phi_d]^T$$

همچنین پارامترهای تنظیم شبکه نقاد به‌صورت $\alpha_1 = 10$ و $\eta_1 = \eta_2 = 200$ و بردار توابع پایه آن شامل تمامی تک‌جمله‌ای‌های از درجه ۲ تا ۴ به‌صورت زیر است:

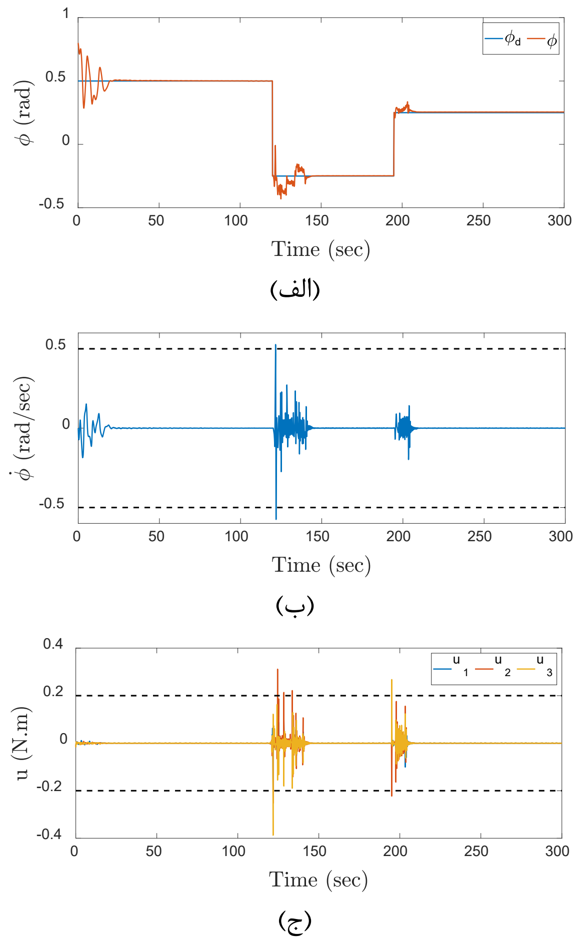
$$\varphi_2(z) = \Big[z_1^2 \quad z_1 z_2 \quad \dots \quad z_2^2 \quad z_2 z_3 \quad z_7^2 \quad z_1^4 \quad \dots \quad z_7^4 \quad -\log\left(\frac{\eta_1(10-z_4))}{\eta_1(10-z_4)+1}\right) \quad -\log\left(\frac{\eta_2(z_4+10))}{\eta_2(z_4+10)+1}\right)\Big]$$

به‌منظور ارزیابی کارایی روش ارائه‌شده در ردیابی مسیر مرجع غلتش $\phi_d(t)$، روش مقید مذکور با روش مستقل از مدل نامقید ارائه‌شده در [۲۴] مورد مقایسه قرارگرفته است. در هر دو روش، شرط توقف آموزش، برقراری هر سه نامساوی $\hat{e}(z(t)) < 0.05$، $\|\hat{W}_1\| < 0.1$ و $\|\hat{W}_2\| < 0.01$ در نظر گرفته‌شده است. به‌منظور تسهیل برآورده شدن شرط PE بودن در همگرایی شبکه شناساگر، نویز جستجو $n(t)$ با رابطه زیر در هر دو روش به سیگنال کنترل $\hat{u}$ افزوده‌شده است:

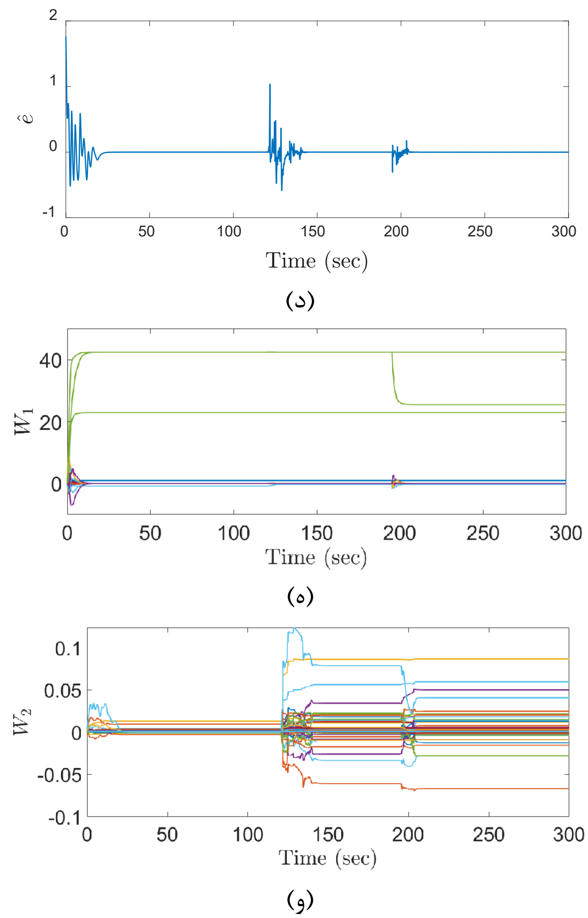
$$n(t) = 0.01\sin(5t) + 0.005\cos(10t) + 0.01\sin^2(t)$$

عملکرد روش نامقید [۲۴] و عملکرد روش مقید پیشنهادی در آموزش کنترل‌کننده وضعیت بهینه و همچنین جبران‌سازی اثر عیوب به ترتیب در شکل‌های ۳ و ۴ نشان داده‌شده است. همان‌طور که در شکل ۳ نشان شده است اگرچه پرنده تحت روش [۲۴] قادر است مسیر مرجع را در هر دو حالت سالم و معیوب ردیابی نماید ولی ورودی کنترلی قید اشباع عملگر پرنده را رعایت ننموده است و همچنین قید نرخ تغییرات زاویه غلتش نیز نقض شده است. درحالی‌که همان‌طور که در شکل ۴ نشان داده‌شده است، در روش پیشنهادی علاوه بر ردیابی مسیر مرجع، قیود مذکور نیز رعایت شده است. با مقایسه رفتار وزن‌های شبکه‌های شناساگر در قسمت (ه) این دو شکل می‌توان دریافت که استفاده از رویکرد پاسخ تجربه در قانون به‌روزرسانی شبکه شناساگر، باعث بهبود در سرعت همگرایی آن شده است. قابل‌توجه است که نرم خطای وزن‌ها یعنی $\|\tilde{W}_1\|$ در انتهای دوره آموزش در شکل ۳ کمتر از 0.05 و در شکل ۴ کمتر از 0.02 بوده است. از سوی دیگر با توجه به قسمت (د) در هر یک از این دو شکل می‌توان دریافت که تغییر میزان عیب باعث افزایش خطای معادله HJB گردیده است که طبق رابطه (۷۹) با در نظر گرفتن $\mathcal{L} = 0.05$ آشکارسازی رخداد عیوب اجزا و عملگر در شکل ۳ به ترتیب در زمان‌های $120.05\ s$ و $195.07\ s$ و در شکل ۴ به ترتیب در زمان‌های $120.06\ s$ و $195.04\ s$ صورت گرفته است.

با توجه به شکل ۴ می‌توان نتیجه گرفت که روش ارائه‌شده در این مقاله توانسته است ضمن تضمین قیود عملکردی حالات و ورودی، مسیر مرجع مطلوب را بدون نیاز به شناخت قبلی از دینامیک پرنده در حضور عیوب عملگر و اجزا ردیابی نماید. همچنین از مقایسه شکل ۳ و ۴ می‌توان دریافت که استفاده از رویکرد ER در قانون به‌روزرسانی شبکه شناساگر و نرخ یادگیری متغیر در قانون به‌روزرسانی شبکه نقاد باعث افزایش سرعت فرآیند آموزش و به‌تبع جبران‌سازی اثر عیب نسبت به روش ارائه‌شده در [۲۴] گردیده است.



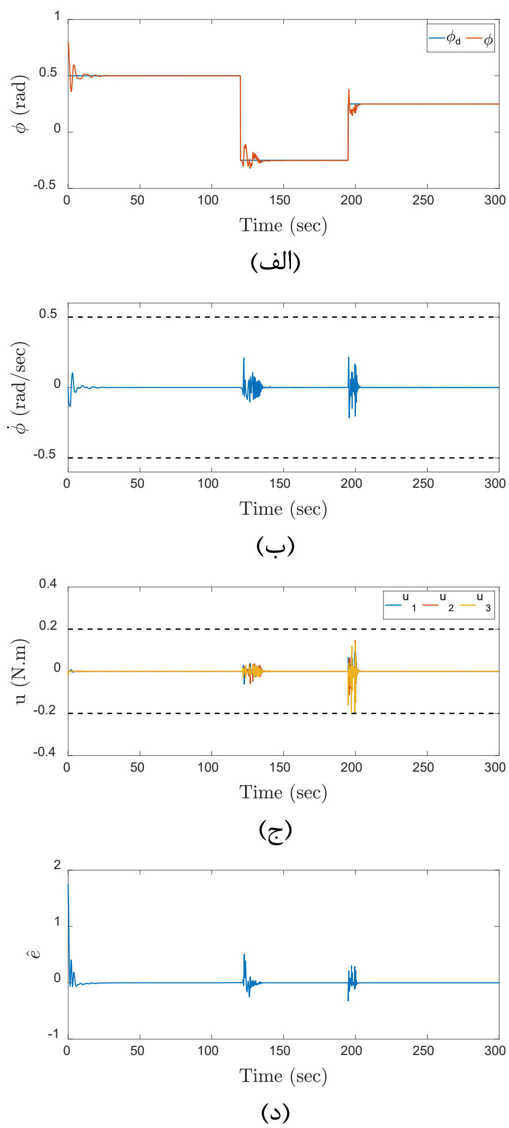
شکل (۳): رفتار زمانی سیستم حلقه بسته با به‌کارگیری روش نامقید [۲۴]: الف) سیگنال مرجع مطلوب زاویه غلتش و خروجی سیستم؛ ب) نرخ تغییرات زاویه غلتش؛ ج) سیگنال کنترل ورودی، د) خطای باقی‌مانده معادله HJB؛ ه) وزن‌های شبکه نقاد؛ و) وزن‌های شبکه نقاد.



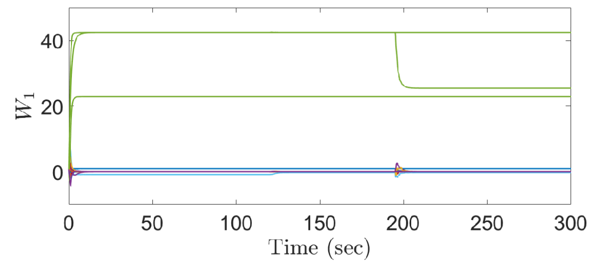
ادامه شکل (۳).

۶- جمع‌بندی و نتیجه‌گیری

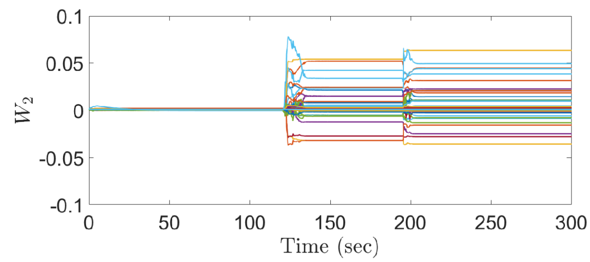
در این مقاله یک روش جدید مبتنی SRL برای طراحی FTOATC مقید در پرنده کوادروتور ارائه‌شده است که به شناخت قبلی از دینامیک پرنده احتیاج ندارد. روش پیشنهادی مبتنی بر یک ساختار شبکه عصبی دوگانه متشکل از شبکه‌های شناساگر و نقاد توسعه داده‌شده است و در مقایسه با ساختارهای تخمین شبکه عصبی سه‌گانه، با توجه به عدم نیاز به شبکه عملگر، باعث کاهش بار محاسباتی و افزایش سرعت همگرایی شده است. در مقایسه با شناساگرهای عصبی تطبیقی پیشین، متغیر در نظر گرفتن ضریب فراموشی و استفاده از روش ER در قانون به‌روزرسانی شبکه شناساگر باعث افزایش سرعت همگرایی و کاهش خطای حالت ماندگار گردیده است. به‌منظور حذف نیازمندی به شناخت کنترل‌کننده پایدارساز اولیه، در قانون به‌روزرسانی شبکه نقاد از یک جمله پایدارساز استفاده‌شده است که امکان جبران‌سازی اثر عیب را بدون نیاز به شناخت قبلی از مقدار عیب و دینامیک سیستم فراهم می‌کند. با استفاده از قضیه لیاپانوف نشان داده شد که خطای ردیابی و خطای وزن‌های شبکه شناساگر و نقاد تحت روش پیشنهادی پایدار UUB خواهند بود و نتایج حاصل از شبیه‌سازی، کارایی آن را در حلقه کنترل وضعیت سیستم خودخلبان نشان می‌دهد.



شکل (۴): رفتار زمانی سیستم حلقه بسته با به‌کارگیری روش مقید ارائه‌شده: الف) سیگنال مرجع مطلوب زاویه غلتش و خروجی سیستم؛ ب) نرخ تغییرات زاویه غلتش؛ ج) سیگنال کنترل ورودی؛ د) خطای باقی‌مانده معادله HJB؛ ه): وزن‌های شبکه نقاد؛ و) وزن‌های شبکه نقاد.



(ه)



(و)

ادامه شکل (۴).

۷- مراجع

[1] Amirani MZ, Bigdeli N, Haeri M. Time varying formation control of unmanned aerial vehicle multi-agent systems with unknown leader input. Journal of Aerospace Mechanics.2021;17(2):53-69. **DOR** https://dorl.net/dor/20.1001.1.26455323.1400.17.2.4.7.

[2] Mahdavi F, Shamaghdari S. Optimal formation control for unmanned aerial vehicle teams with collision avoidance constraint and unknown dynamics. Journal of Aerospace Mechanics. 2023;19(1): 61-79. **DOR** https://dorl.net/dor/20.1001.1.26455323.1402.19.1.5.0.

[3] Zhao W, Liu H, Lewis FL. Data-driven fault-tolerant control for attitude synchronization of nonlinear quadrotors. IEEE Transactions on Automatic Control. 2021;66(11):5584-91. **DOI** https://doi.org/10.1109/TAC.2021.3053194.

[4] Amin AA, Hasan KM. A review of fault tolerant control systems: advancements and applications. Measurement. 2019;143:58-68. DOI https://doi.org/10.1016/j.measurement.2019.04.083.

[5] Roshanravan S, Sobhani Gendeshmin B, Shamaghdari S. Design of an actuator fault-tolerant controller for an air vehicle with nonlinear dynamics. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering. 2019;233(10):3534-46. **DOI** https://doi.org/10.1177/0954410018801254.

[6] Jiang J, Yu X. Fault-tolerant control systems: A comparative study between active and passive approaches. Annual Reviews in control. 2012;36(1):60-72. **DOI** https://doi.org/10.1016/j.arcontrol.2012.03.005.

[7] Rudin K, Ducard GJ, Siegwart RY. Active fault-tolerant control with imperfect fault detection information: Applications to UAVs. IEEE Transactions on Aerospace and Electronic Systems. 2019;56(4):2792-805.

[8] Lan J, Patton RJ. A new strategy for integration of fault estimation within fault-tolerant control. Automatica. 2016;69:48-59.

[9] Roshanravan S, Shamaghdari S. Simultaneous fault detection and isolation and fault-tolerant control using supervisory control technique: asynchronous switching approach. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering. 2020;234(8):900-11. **DOI** https://doi.org/10.1177/0959651819893891.

[10] Ruan Z, Yang Q, Ge SS, Sun Y. Performance-guaranteed fault-tolerant control for uncertain nonlinear systems via learning-based switching scheme. IEEE Transactions on Neural Networks and Learning Systems. 2020;32(9):4138-50. **DOI** https://doi.org/10.1109/TNNLS.2020.3016954.

[11] Li L, Luo H, Ding SX, Yang Y, Peng K. Performance-based fault detection and fault-tolerant control for automatic control systems. Automatica. 2019;99:308-16. **DOI** https://doi.org/10.1016/j.automatica.2018.10.047.

[12] Cheng W, Zhang K, Jiang B. Hierarchical Structure-Based Fixed-Time Optimal Fault-Tolerant Time-Varying Output Formation Control for Heterogeneous Multiagent Systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2023;53(8):4856-66.. **DOI** https://doi.org/10.1109/TSMC.2023.3257426.

[13] Bardi M, Dolcetta IC. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations: Springer; 1997. **DOI** https://doi.org/10.1007/978-0-8176-4755-1.

[14] Lewis FL, Vrabie D. Reinforcement learning and adaptive dynamic programming for feedback control. IEEE circuits and systems magazine. 2009;9(3):32-50. **DOI** https://doi.org/10.1109/MCAS.2009.933854.

continuous-time systems with actuator constraints. International Journal of Control. 2022;95(8):2005-23. **DOI** https://doi.org/10.1080/00207179.2021.1890824.

[24] Roshanravan S, Shamaghdari S. Adaptive fault-tolerant tracking control for affine nonlinear systems with unknown dynamics via reinforcement learning. IEEE Transactions on Automation Science and Engineering. 2022;21(1):569-80. **DOI** https://doi.org/10.1109/TASE.2022.3223702.

[25] Dierks T, Jagannathan S, editors. Optimal control of affine nonlinear continuous-time systems. Proceedings of the 2010 American control conference; 2010: IEEE. **DOI** https://doi.org/10.1109/ACC.2010.5531586.

[26] Liu D, Yang X, Wang D, Wei Q. Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints. IEEE transactions on cybernetics. 2015;45(7):1372-85. **DOI** https://doi.org/10.1109/TCYB.2015.2417170.

[27] Yang H, Jiang B, Staroswiecki M. Supervisory fault tolerant control for a class of uncertain nonlinear systems. Automatica. 2009;45(10):2319-24. **DOI** https://doi.org/10.1016/j.automatica.2009.06.019.

[28] Ma H-J, Xu L-X, Yang G-H. Multiple environment integral reinforcement learning-based fault-tolerant control for affine nonlinear systems. IEEE Transactions on Cybernetics. 2019;51(4):1913-28. **DOI** https://doi.org/10.1109/TCYB.2018.2889679.

[29] Choi YC, Ahn HS. Nonlinear control of quadrotor for point tracking: Actual implementation and experimental tests. IEEE/ASME Transactions on Mechatronics. 2014;20(3):1179-92. **DOI** https://doi.org/10.1109/TMECH.2014.2329945.

[30] Edwards C, Lombaerts T, Smaili H. Fault tolerant flight control. Lecture notes in control and information sciences. 2010;399:1-560. **DOI** https://doi.org/10.1007/978-3-642-11690-2.

[31] Modares H, Lewis FL, Naghibi-Sistani M-B. Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks. IEEE Transactions on neural networks and learning systems. 2013;24(10):1513-25. **DOI** https://doi.org/10.1109/TNNLS.2013.2276571.

[15] Huang J, Zeng W, Xiong H, Noack BR, Hu G, Liu S, Xu Y, Cao H. Symmetry-Informed Reinforcement Learning and its Application to Low-Level Attitude Control of Quadrotors. IEEE Transactions on Artificial Intelligence. 2023;5(3):1147-61. **DOI** https://doi.org/10.1109/TAI.2023.3249683.

[16] Bernini N, Bessa M, Delmas R, Gold A, Goubault E. Reinforcement learning with formal performance metrics for quadcopter attitude control under non-nominal contexts. Engineering Applications of Artificial Intelligence. 2024; 127: 107090. **DOI** https://doi.org/10.1016/j.engappai.2023.107090.

[17] Zhu Y, Lian S, Zhong W, Meng, W. Reinforcement learning method for quadrotor attitude control based on expert information. 8th International Conference on Automation, Control and Robotics Engineering (CACRE); 2023: IEEE. **DOI** https://doi.org/10.1109/CACRE58689.2023.10208497.

[18] Yang Y, Vamvoudakis KG, Modares H, Yin Y, Wunsch DC. Safe intermittent reinforcement learning with static and dynamic event generators. IEEE Transactions on Neural Networks and Learning Systems. 2020;31(12):5441-55. **DOI** https://doi.org/10.1109/TNNLS.2020.2967871.

[19] Marvi Z, Kiumarsi B. Safe reinforcement learning: A control barrier function optimization approach. International Journal of Robust and Nonlinear Control. 2021;31(6):1923-40. **DOI** https://doi.org/10.1002/rnc.5132.

[20] Al-Tamimi A, Lewis FL, Abu-Khalaf M. Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof. IEEE Transactions on Systems, Man, and Cybernetics, Part B. 2008;38(4):943-9. **DOI** https://doi.org/10.1109/TSMCB.2008.926614.

[21] Lv Y, Na J, Yang Q, Wu X, Guo Y. Online adaptive optimal control for continuous-time nonlinear systems with completely unknown dynamics. International Journal of Control. 2016;89(1):99-112. **DOI** https://doi.org/10.1080/00207179.2015.1060362.

[22] Lv Y, Na J, Zhao X, Huang Y, Ren X. Multi-H∞ controls for unknown input-interference nonlinear system with reinforcement learning. IEEE Transactions on Neural Networks and Learning Systems. 2021. **DOI** https://doi.org/10.1109/TNNLS.2021.3130092.

[23] Mishra A, Ghosh S. Simultaneous identification and optimal tracking control of unknown

[32] Na J, Mahyuddin MN, Herrmann G, Ren X, Barber P. Robust adaptive finite-time parameter estimation and control for robotic systems. International Journal of Robust and Nonlinear Control. 2015;25(16):3045-71. **DOI** https://doi.org/10.1002/rnc.3247.

[33] Modares H, Lewis FL. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. Automatica. 2014;50(7):1780-92. **DOI** https://doi.org/10.1016/j.automatica.2014.05.011.

[34] Abu-Khalaf M, Lewis FL. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. Automatica. 2005;41(5):779-91. **DOI** https://doi.org/10.1016/j.automatica.2004.11.034.

[35] Modares H, Lewis FL, Naghibi-Sistani M-B. Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems. Automatica. 2014;50(1):193-202. **DOI** https://doi.org/10.1016/j.automatica.2013.09.043.

[36] Stone M. The generalized Weierstrass approximation theorem. Mathematics Magazine. 1948;21(5): 237-254.

[37] Rudin W. Principles of mathematical analysis1953.

[38] Ding SX. Model-based fault diagnosis techniques: design schemes, algorithms, and tools: Springer Science & Business Media; 2008.