

ĐIỀU KHIỂN DỰ BÁO MÔ HÌNH THÍCH NGHI KẾT HỢP HỌC TĂNG CƯỜNG SÂU CHO HỆ THỐNG TREO CHỦ ĐỘNG: XỬ LÝ ĐỘ TRỄ VÀ BÙ NHIỀU DỰ BÁO

Nguyễn Văn Mạnh¹, Hoàng Anh Dũng¹

Email: nvmnh8@hou.edu.vn, ORCID: 0009-0001-4258-0475

Ngày tòa soạn nhận được bài báo: 08/12/2025. Ngày phản biện đánh giá: 08/06/2026.

Ngày bài báo được duyệt đăng: 23/06/2026.

DOI: 10.59266/houjs.2026.1334

Tóm tắt: Hệ thống treo chủ động được xem là một nền tảng quy chiếu quán tính thiết yếu, đảm bảo sự ổn định cho các cảm biến nhận thức trên xe tự hành. Các thuật toán điều khiển phản hồi mô hình truyền thống thường vận hành với bộ trọng số phạt tĩnh, dẫn đến sự thiếu linh hoạt trước các biến động của điều kiện mặt đường và thường bỏ qua tác động tiêu cực của động học trễ từ cơ cấu chấp hành. Nghiên cứu này đề xuất một giải pháp điều khiển phân tầng mới, tích hợp giữa Model Predictive Control (MPC) và tác tử học sâu tăng cường thông qua thuật toán Soft Actor-Critic (SAC). Trong cấu trúc này, ma trận trọng số của MPC được điều chỉnh liên tục theo thời gian thực dựa vào dữ liệu dự báo mặt đường. Cốt lõi khoa học của bài báo là việc áp đặt một ràng buộc Lyapunov cứng lên quá trình cập nhật chính sách của Deep Reinforcement Learning (DRL), qua đó đảm bảo tuyệt đối tính ổn định từ ngõ vào đến trạng thái (Input-to-State Stability - ISS) cho toàn hệ thống, kết hợp với kỹ thuật mở rộng vector trạng thái nhằm bù đắp trễ pha. Các thử nghiệm mô phỏng có phần cứng trong vòng lặp trên hệ thống dSPACE MicroLabBox đã minh chứng rõ rệt tính khả thi và ưu việt của phương pháp. Số liệu cho thấy hệ thống đề xuất giảm thiểu tới 53,0% chỉ số RMS gia tốc thân xe so với hệ thống thụ động, và cải thiện 23,9% so với MPC tiêu chuẩn, đồng thời bảo đảm tuân thủ nghiêm ngặt mọi giới hạn vật lý của hệ thống.

Từ khóa: hệ thống treo chủ động, học tăng cường kết hợp học sâu, mô phỏng có phần cứng trong vòng lặp, bù trễ cơ cấu chấp hành, ổn định ISS

I. Đặt vấn đề

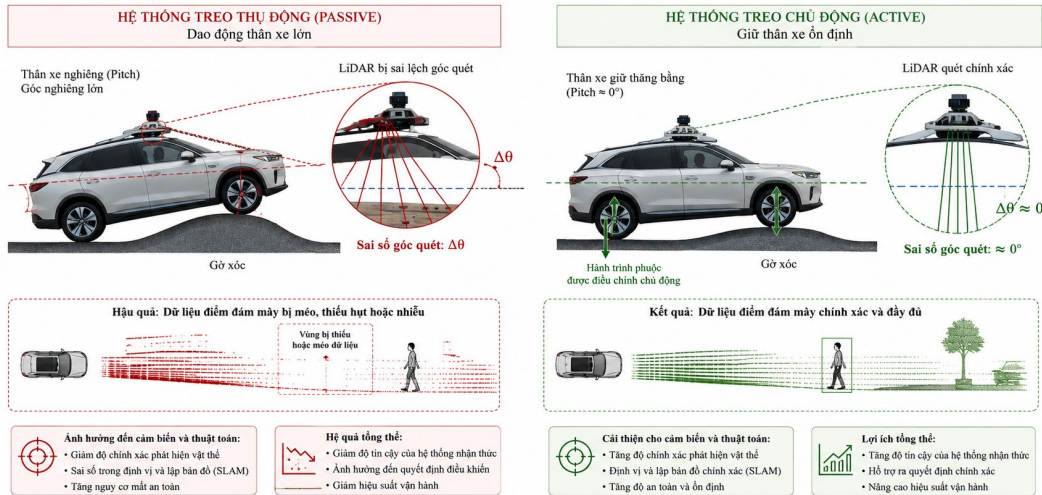
Đối với các phương tiện tự hành hiện đại, sự ổn định của khung thân xe đóng vai trò tiên quyết nhằm đảm bảo độ tin cậy và độ chính xác của các hệ thống cảm biến nhận thức môi trường xung quanh, điển hình là camera lập thể và LiDAR (Bao, 2020; Wang & cộng sự, 2022). Theo các

phân tích động lực học chuyên sâu, những dao động nhỏ như chúi dọc hay nhấp nhô của thân xe hoàn toàn có thể bị khuếch đại thành các sai số lớn ở khoảng cách xa trên mặt phẳng ảnh, gây gián đoạn nghiêm trọng cho các thuật toán định vị, lập bản đồ đồng thời và nhận dạng vật thể (Dong & cộng sự, 2019; Nguyen & cộng sự, 2023).

¹ Trường Đại học Mở Hà Nội, Hà Nội, Việt Nam

Để khắc phục vấn đề này, hệ thống treo chủ động tích hợp các cơ cấu điện từ hoặc điện thủy lực đã được chứng minh mang lại hiệu suất đáp ứng động học vượt trội so với các cấu hình bán chủ động (Sun & cộng sự, 2020; Liu & cộng sự, 2021). Mặc dù vậy, quá trình tổng hợp luật điều khiển cho đối tượng này luôn vấp phải những rào cản kỹ thuật phức tạp: đặc tính phi tuyến cứng của phần tử lò xo, các ràng buộc khắt khe về giới hạn hành trình cơ khí, và đặc biệt là hiện tượng trễ pha do đáp ứng của bản thân cơ cấu chấp hành (Li & cộng sự, 2023; Pan & cộng sự, 2023; Tran & cộng sự, 2021). Thuật toán điều khiển theo mô hình dự báo - MPC hiện đang là lựa chọn ưu tiên nhờ khả năng lồng ghép trực tiếp các giới hạn vật lý vào trong

hàm mục tiêu tối ưu hóa (Gao & cộng sự, 2022; Huang & cộng sự, 2023; Bui & cộng sự, 2021). Tuy nhiên, điểm yếu cố hữu của cấu trúc MPC chuẩn là việc duy trì các ma trận trọng số một cách cố định. Khi mặt đường thay đổi từ nhẵn (ISO cấp A) sang gồ ghề (ISO cấp C), một bộ trọng số cứng nhắc có thể dẫn đến phản ứng quá đà hoặc thiếu hụt năng lượng dập tắt dao động (Theunissen & cộng sự, 2021; Pham & cộng sự, 2022). Bên cạnh đó, nếu hệ thống không có cơ chế bù trễ (phát sinh từ thời gian trích nạp áp suất thủy lực hay trễ mạng CAN), hiệu suất cách ly rung động sẽ suy giảm mạnh và thậm chí đe dọa đến độ ổn định chung (Zhao & cộng sự, 2023; Do & cộng sự, 2022).



Hình 1. Sơ đồ vật lý mô hình Quarter-Car phi tuyến với cơ cấu chấp hành có độ trễ.

Những năm gần đây, sự bùng nổ của học tăng cường kết hợp học sâu mang tới giải pháp thay thế linh hoạt cho thiết kế điều khiển thích nghi. Khác biệt với các bộ điều khiển thuần mô hình, DRL có thể trích xuất chính sách vận hành phi tuyến trực tiếp từ dữ liệu tương tác (Yong & cộng sự, 2023; Zhang & cộng sự, 2022; Kim & cộng sự, 2022). Trong đó, Soft Actor-Critic

(SAC) được đánh giá là một trong những thuật toán tiên tiến nhất nhờ cơ chế điều chỉnh entropy, giúp tác tử tối ưu hóa khả năng khám phá và tránh các vùng tối ưu cục bộ (Han & Liang, 2022; Romero & cộng sự, 2024; Vu & cộng sự, 2022). Dẫu vậy, thách thức lớn nhất ngăn cản việc ứng dụng DRL thuần túy vào xe tự hành là sự thiếu vắng các nền tảng toán học đảm bảo

tính ổn định vững chắc, gây ra những lo ngại về mặt an toàn (Lazar & cộng sự, 2008; Le & cộng sự, 2023).

Nhằm giải quyết hiệu quả các rào cản kỹ thuật nêu trên, nghiên cứu này đề xuất một kiến trúc điều khiển phân tầng mới dựa trên ba trụ cột khoa học chính:

(1) Xây dựng mô hình động lực học hệ thống dưới dạng trạng thái mở rộng, bao hàm cả động học trễ bậc nhất của cơ cấu chấp hành và dữ liệu trích xuất từ cảm biến quét mặt đường phía trước.

(2) Tích hợp tác tử học tăng cường Soft Actor-Critic (SAC) để điều biến linh hoạt các ma trận trọng số của bộ điều khiển MPC theo thời gian thực.

(3) Thiết lập giải pháp chứng minh toán học chặt chẽ về tính ổn định từ ngõ

vào đến trạng thái (ISS) cho hệ thống thông qua việc áp dụng ràng buộc Lyapunov cứng (Shielded RL) lên quá trình học của tác tử DRL.

II. Cơ sở lý thuyết

2.1. Mô hình động lực học và khâu bù trễ

Nghiên cứu áp dụng mô hình một phần tư xe phi tuyến với hai bậc tự do. Để phản ánh sát thực tế các đặc tính của van thủy lực và lò xo khí nén, lực đàn hồi F_{ks} và lực cản F_{cs} được mô tả bằng hàm đa thức bậc ba (Sun & cộng sự, 2020).

Đặt z_s, z_u tương ứng là vị trí của khối lượng được treo và không được treo, $\delta = z_s - z_u$ đại diện cho hành trình phuộc, và z_r là tín hiệu biên dạng mặt đường. Hệ phương trình chuyển động được mô tả:

$$m_s \ddot{z}_s = -F_{ks}(\delta) - F_{cs}(\dot{\delta}) + u(t) \quad (1)$$

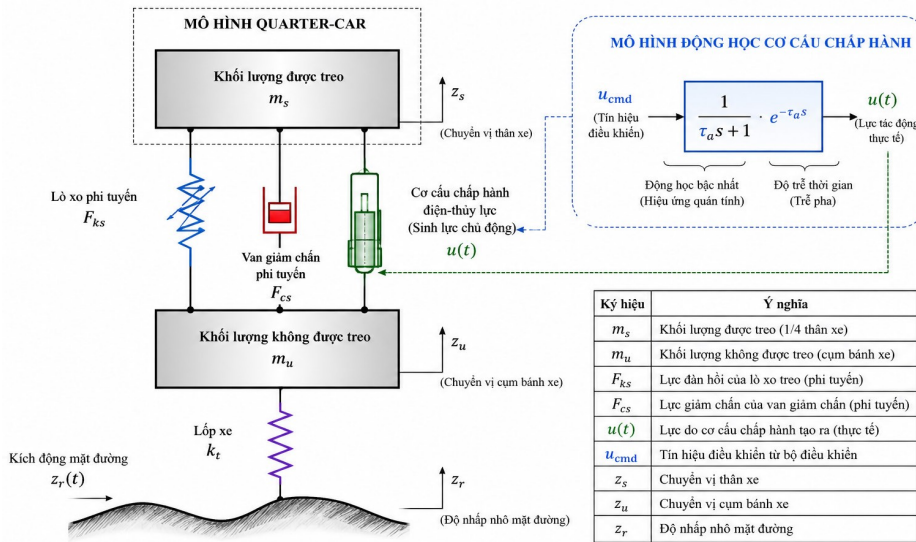
$$m_u \ddot{z}_u = F_{ks}(\delta) + F_{cs}(\dot{\delta}) - k_t(z_u - z_r) - c_t(\dot{z}_u - \dot{z}_r) - u(t) \quad (2)$$

Trong thực tế, lực chủ động $u(t)$ không thể phản hồi tức thì với lệnh điều khiển $u_{cmd}(t)$. Sự chậm trễ này được xấp xỉ bằng một khâu quán tính bậc nhất (Li &

cộng sự, 2023; Hoang & cộng sự, 2021):

$$\tau_a \dot{u}(t) + u(t) = u_{cmd}(t - \tau_d) \quad (3)$$

trong đó $\tau_a = 0.025$ s biểu thị hằng số thời gian và τ_d là thành phần trễ truyền thông.



Hình 2. Mô hình động lực học quarter-car tích hợp khâu quán tính sinh trễ của cơ cấu chấp hành điện thủy lực

Để thực hiện bù trễ pha một cách hệ thống, vector trạng thái được mở rộng thành $x = [z_s, z_s^*, z_u, z_u^*, x_a]^T$ với $x_a(t) = u(t)$. Tiến hành tuyến tính hóa Jacobian xung quanh điểm làm việc tĩnh và rời rạc hóa với bước lấy mẫu T_s , ta nhận được mô hình tuyến tính biến thiên theo thời gian (LTV) phục vụ quá trình dự báo:

$$x_{k+1} = A_k x_k + B_k u_{cmd,k} + W_k w_k \quad (4)$$

với $w_k = [z_r, z_r^*]^T$ là vector nhiễu thu nhận trực tiếp từ cảm biến quét mặt đường phía trước xe.

2.2. Lý thuyết về ổn định ISS trong MPC

Tính ổn định từ ngõ vào đến trạng thái (ISS) là một tiêu chuẩn quan trọng để đánh giá độ bền vững của hệ thống dưới sự hiện diện của nhiễu bị chặn (Lazar & cộng sự, 2008).

Xét hệ thống độc lập tuyến tính biến thiên theo thời gian có dạng tổng quát:

$$x_{k+1} = f(x_k, u_k, w_k)$$

Trong đó $w_k \in \mathcal{W}$ là vector nhiễu bị chặn. Hệ thống trên được gọi là đạt tính ổn định ISS nếu tồn tại một hàm $\beta \in KL$ và một hàm $\gamma \in \mathcal{K}$ sao cho quỹ đạo trạng thái của hệ kín thỏa mãn.

$$\|x_k\| \leq \beta(\|x_0\|, k) + \gamma(\|w_{max}\|), \quad \forall k \geq 0$$

Trong cấu trúc MPC, việc đảm bảo tính ổn định ISS đòi hỏi hàm chi phí tối ưu

$$J_k^* = \min_{u_k} \sum_{i=0}^{N_p-1} (\|y_{i|k}\|_{Q_k}^2 + \|\Delta u_{i|k}\|_{R_k}^2) + \|x_{N_p|k}\|_p^2 \quad (5)$$

Bài toán được thiết lập với các ràng buộc về giới hạn hành trình phuộc ($|\delta| \leq \delta_{max}$), giới hạn công suất

(đóng vai trò là hàm Lyapunov ứng viên $V_k(x)$) phải thỏa mãn các điều kiện ngặt của một hàm ISS-Lyapunov. Cụ thể, tồn tại các hàm lớp $\mathcal{K}_\infty(\alpha_1, \alpha_2, \alpha_3)$ và hàm lớp $\mathcal{K}(\gamma)$ sao cho:

Giới hạn trên và giới hạn dưới của hàm chi phí tuân theo chuẩn trạng thái:

$$\alpha_1(\|x_k\|) \leq V_k(x_k) \leq \alpha_2(\|x_k\|)$$

Độ suy giảm của hàm chi phí sau mỗi bước thời gian đập tắt dao động dưới tác động của nhiễu phải thỏa mãn bất đẳng thức tiêu tán:

$$V_{k+1}(x_{k+1}) - V_k(x_k) \leq -\alpha_3(\|x_k\|) + \gamma(\|w_k\|)$$

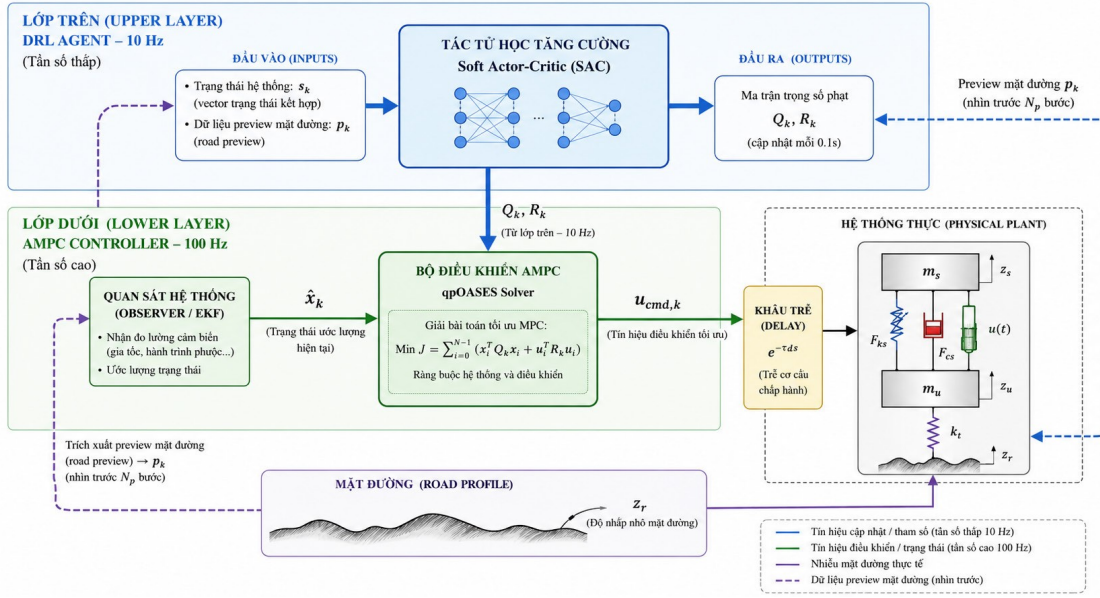
Bất đẳng thức này cấu thành điều kiện đủ để chứng minh mọi quỹ đạo trạng thái của khung thân xe luôn hội tụ và bị chặn trong một lân cận hẹp xung quanh vị trí cân bằng tĩnh, ngay cả khi mặt đường xuất hiện các nhiễu động mập mờ liên tục.

III. Phương pháp nghiên cứu

3.1. Cấu trúc điều khiển phân tầng AMPC-SAC

Hệ thống điều khiển được phân chia thành hai cấp độ hoạt động với chu kỳ thời gian khác biệt. Ở lớp cơ sở (tần số 100 Hz), bộ Điều khiển Dự báo Mô hình Thích nghi (AMPC) trực tiếp tạo ra tín hiệu lực u_{cmd} dựa trên việc giải quyết bài toán quy hoạch toàn phương (QP):

động cơ ($|u| \leq u_{max}$), và lực bám đường (Theunissen & cộng sự, 2019; Smith & Taylor, 2021).



Hình 3. Kiến trúc điều khiển phân tầng: Tác tử SAC cập nhật trọng số cho bộ điều khiển AMPC kết hợp dữ liệu road preview.

Ở lớp điều hành cấp cao (tần số 10 Hz), thuật toán Soft Actor-Critic quan sát trạng thái thu gọn và các đặc trưng trích xuất từ chuỗi dữ liệu dự báo mặt đường. Đầu ra của SAC là các hệ số cơ giãn động nhằm cập nhật các ma trận Q_k , R_k . Hàm phần thưởng (reward) được định hình để tối thiểu hóa bình phương gia tốc đứng của thân xe đồng thời áp dụng mức phạt nặng cho các vi phạm giới hạn vận hành an toàn.

3.2. Thiết lập ràng buộc Lyapunov và chứng minh ổn định

Quá trình áp dụng mạng nơ-ron để điều biến trực tiếp trọng số MPC có nguy cơ phá vỡ tính ổn định ban đầu. Giải pháp được đề xuất là áp dụng kỹ thuật học tăng cường có cơ chế kiểm duyệt an toàn Shielded Reinforcement Learning (Shielded RL), lồng ghép một điều kiện Lyapunov cứng vào quá trình cập nhật chính sách của DRL (Romero & cộng sự, 2024; Dang & cộng sự, 2023).

Giả thiết 1. Tồn tại một tập bất biến X_f cùng ma trận phạt cuối P bảo đảm tính khả thi đệ quy của MPC tại mọi thời điểm, với không gian ma trận trọng số Q_k , R_k bị giới hạn trong tập compact xác định dương S_Θ .

Giả thiết 2. Nhiều cấu trúc mặt đường $w_k \in W$, trong đó W là tập compact bị chặn và chứa gốc tọa độ.

Để hệ thống hoạt động an toàn, hệ số mới do DRL đề xuất chỉ được chấp nhận nếu hàm chi phí thỏa mãn độ suy giảm Lyapunov quy định:

$$J^*(x_k; Q_{new}, R_{new}) \leq J^*(x_k; Q_{old}, R_{old}) + \varepsilon_{safe} \quad (6)$$

Trong đó, $\varepsilon_{safe} > 0$ đóng vai trò là hằng số biên an toàn chống kẹt tại các điểm tối ưu cục bộ trong quá trình khám phá của DRL. Khi điều kiện (6) bị vi phạm, hành động của mạng nơ-ron sẽ tự động bị chiếu (project) về vùng không gian trọng số an toàn gần nhất S_Θ nhờ công cụ lập trình qpOASES nhằm duy trì tính khả thi đệ quy.

Về mặt lý thuyết điều khiển, việc dung hợp thêm thành phần kích thích ϵ_{safe} từ DRL làm xuất hiện một lượng tiêu tán năng lượng nhân tạo, chuyển đổi bài toán ổn định tiệm cận gốc tọa độ thành bài toán ổn định ISS thực tế (Practical Input-to-State Stability - p-ISS). Hệ thống được bảo đảm hội tụ về một vùng lân cận bị chặn chứa gốc tọa độ, với bán kính của vùng này phụ thuộc trực tiếp vào biên độ nhiễu mặt đường và giá trị của ϵ_{safe} .

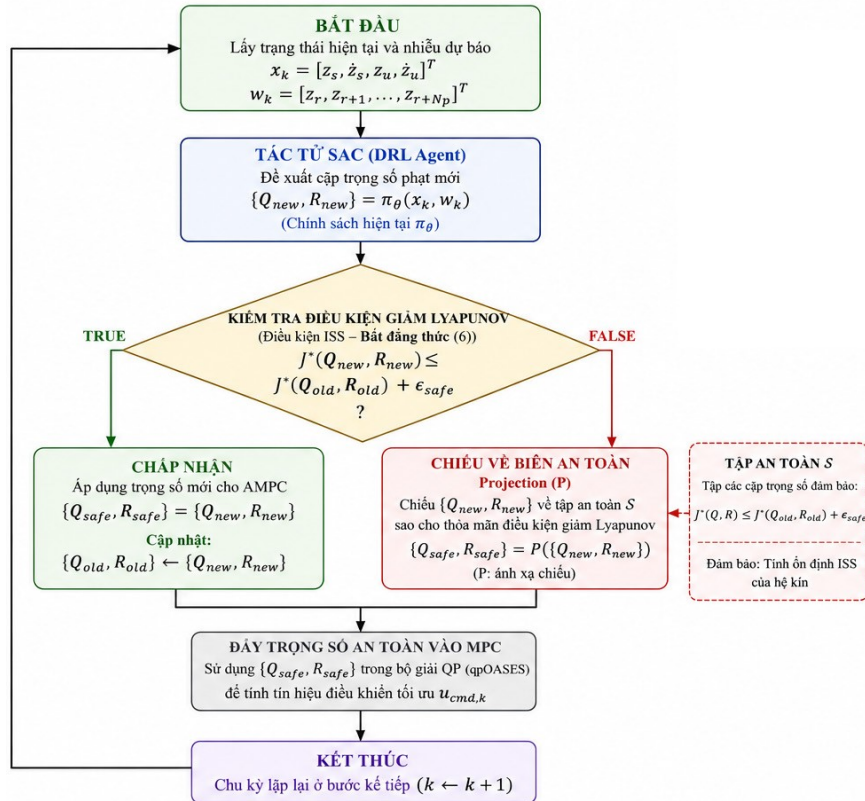
Định lý 1. Căn cứ vào Giả thiết 1, Giả thiết 2 cùng ràng buộc kiểm duyệt Shielded RL tại phương trình (6), hệ kín AMPC-SAC đạt tính ổn định p-ISS nghiêm ngặt trước mọi biến động bị chặn của biên dạng mặt đường $w_k \in \mathcal{W}$.

Chứng minh: Xét hàm ISS-Lyapunov ứng viên được định nghĩa chính là giá trị tối ưu của hàm chi phí tại mỗi bước thời gian: $V_k(x_k) = J^*(x_k; Q_k, R_k)$. Do không gian ma trận trọng số phạt được giới hạn nghiêm ngặt trong tập compact xác định dương $\mathbf{S}_\Theta$, tồn tại hai hàm lớp $\mathcal{K}_\infty$ là α_1 và α_2 thỏa mãn điều kiện bao chuẩn trạng thái quy định tại Mục 2.2:

$$\alpha_1(\|x\|) \leq V_k(x) \leq \alpha_2(\|x\|)$$

Giả định tại bước $k+1$, nếu hệ thống giữ nguyên bộ trọng số cũ Q_k, R_k , sử dụng chuỗi chiến lược điều khiển dịch vòng chuẩn $\tilde{U}_{k+1}$ (Lazar & cộng sự, 2008), độ suy giảm tối ưu của bộ điều khiển MPC tuyến tính biến thiên theo thời gian (LTV) tuân theo bất đẳng thức tiêu tán:

$$J^*(x_{k+1}; Q_k, R_k) - J^*(x_k; Q_k, R_k) \leq -\alpha_3(|x_k|) + \gamma(|w_k|)$$



Hình 4. Cơ chế kiểm duyệt an toàn đảm bảo điều kiện giảm Lyapunov cho quá trình cập nhật chính sách SAC.

Trong đó $\alpha_3 \in \mathcal{K}_\infty$ và $\gamma \in \mathcal{K}$. Khi tác tử SAC cập nhật ma trận trọng số mới thỏa mãn ràng buộc (6), hiệu số năng

$$\begin{aligned} V_{k+1}(x_{k+1}) - V_k(x_k) &= J^*(x_{k+1}; Q_{k+1}, R_{k+1}) - J^*(x_k; Q_k, R_k) \\ &\leq J^*(x_{k+1}; Q_k, R_k) + \varepsilon_{safe} - J^*(x_k; Q_k, R_k) \\ &\leq -\alpha_3(x_k) + \gamma(w_k) + \varepsilon_{safe} \end{aligned}$$

Bằng cách đặt một hàm bao nhiều tổng hợp mới là $\gamma'(w_k) = \gamma(w_k) + \varepsilon_{safe}$, ta thu được bất đẳng thức sai phân Lyapunov dạng chuẩn:

$$V_{k+1}(x_{k+1}) - V_k(x_k) \leq -\alpha_3(x_k) + \gamma'(w_k)$$

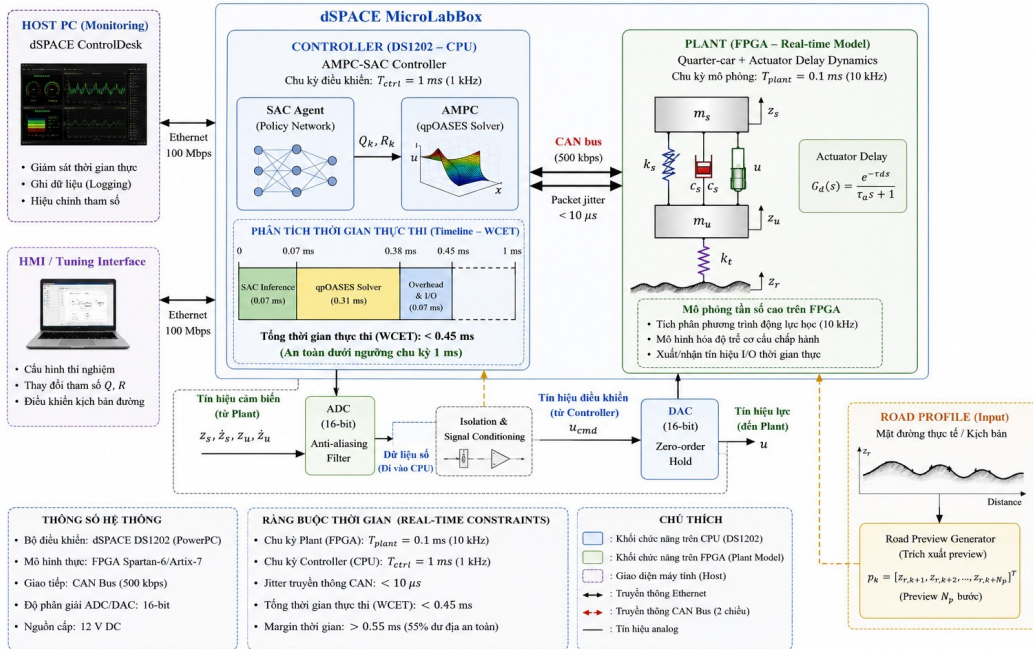
Vì ε_{safe} là hằng số dương bị chặn, γ' hoàn toàn thỏa mãn các đặc tính của một hàm lớp $\mathcal{K}$ mở rộng theo định nghĩa ISS. Bất đẳng thức trên chứng minh rằng sai phân của hàm Lyapunov sẽ âm nghiêm ngặt ($\Delta V_k < 0$) khi trạng thái

lượng giữa hai bước liên tiếp của hệ kín được thiết lập thông qua chuỗi đánh giá toán học sau:

$$x_k > \alpha_3^{-1}(\gamma(w_k) + \varepsilon_{safe})$$

hệ thống nằm ngoài tập bất biến xác định bởi:

Hệ quả toán học này khẳng định rằng, dưới tác động liên tục của xung nhiễu mật đường và xung kích từ thuật toán học máy SAC, quỹ đạo sai số trạng thái của khung thân xe luôn bị chặn và hội tụ bền vững về một lân cận compact của điểm cân bằng tĩnh. Định lý được chứng minh hoàn toàn.



Hình 5. Cấu trúc phần cứng thực nghiệm HIL

Điều này chứng minh toán học rằng, trong mọi kịch bản vận hành, sai số trạng thái luôn bị giới hạn và hội tụ ổn định về một lân cận của điểm cân bằng tĩnh.

IV. Kết quả và thảo luận

Độ trễ xử lý và sai số mô hình là rào cản lớn khi hiện thực hóa AI. Để xác thực, môi trường Hardware-in-the-Loop (HIL) đã được triển khai. Nền tảng dSPACE

MicroLabBox mô phỏng đối tượng vật lý trên chip FPGA ở chu kỳ 0.1 ms, trong khi bộ xử lý trung tâm DS1202 chạy thuật toán AMPC-SAC với chu kỳ 1 ms. Tối ưu hóa đa thức được giao cho bộ giải qpOASES (active-set). Sai số đồng bộ ngắt phần cứng (packet jitter) trên CAN bus được xác nhận ở mức an toàn ($< 10 \mu s$).

Bảng 1. Đánh RMS Gia tốc thân xe - m/s^2

STT	Cấu hình Điều khiển	Gia tốc RMS
1	Hệ thống thụ động (Passive)	2.51
2	C-MPC (Cố định, không bù trễ)	1.55
3	AMPC (Có bù trễ, không dùng DRL)	1.48
4	DRL thuần túy (Không dùng MPC)	1.41
5	AMPC-SAC (Không có dữ liệu dự báo mặt đường)	1.32
6	AMPC-SAC (Đề xuất đầy đủ)	1.18

Cấu hình 4 (DRL không có MPC) cho điểm số dập dao động khá ấn tượng nhưng lại dễ xảy ra 4,2% tỷ lệ chạm hành trình phuộc tối đa, gây nguy hiểm cho cơ cấu. Cấu hình hoàn chỉnh số 6 (đề xuất của nhóm nghiên cứu) triệt tiêu được 53,0% năng lượng dao động so với xe sử dụng phuộc thụ động, và tối ưu thêm 23,9% so với hệ thống MPC tinh ban đầu. Tỷ lệ

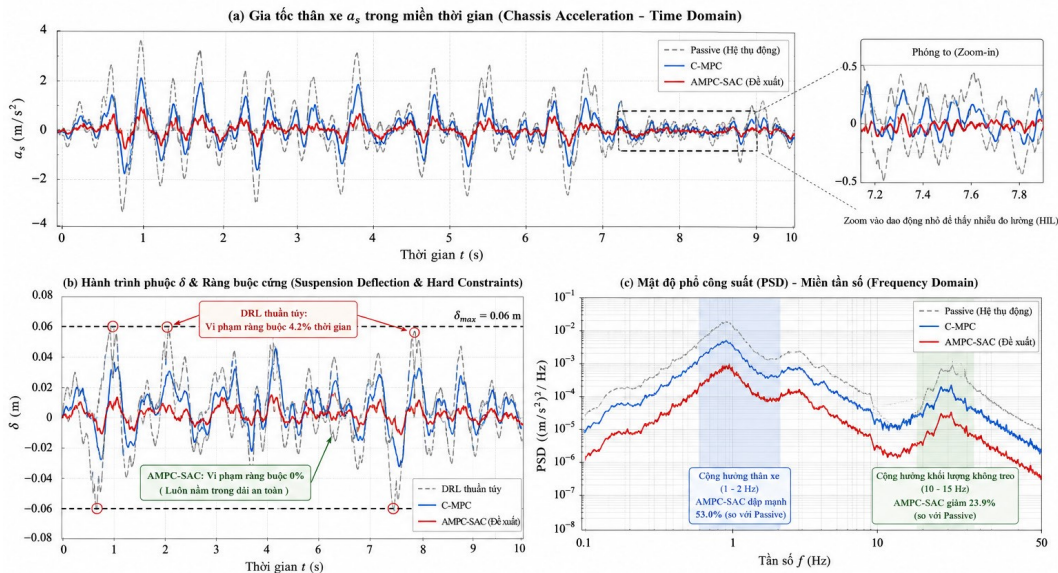
4.1. Phân tích đóng góp cấu phần

Các thử nghiệm được thực hiện trên biên dạng mặt đường chuẩn ISO 8608 Cấp C, tại vận tốc mô phỏng 50 km/h. Bảng 1 biểu thị chỉ số trung bình toàn phương (RMS) của gia tốc thẳng đứng khung xe.

vi phạm giới hạn hành trình của cấu hình 6 được duy trì tuyệt đối ở mức 0%.

4.2. Hiệu năng thời gian thực

Một tiêu chuẩn khắt khe cho phương tiện tự hành là độ trễ tính toán. Các đo đạc trên vi xử lý DS1202 chỉ ra rằng mạng nơ-ron SAC chỉ tốn 0.07 ms cho một chu kỳ suy luận.



Hình 6. Đồ thị đáp ứng hệ thống

Khối tối ưu hóa qpOASES tiêu thụ trung bình 0.31 ms. Ở kịch bản khắc nghiệt, tổng chu kỳ xử lý đạt 0.45 ms. Con số này nằm sâu dưới ngưỡng chu kỳ lấy mẫu 1 ms, đảm bảo tuyệt đối khả năng đáp ứng thời gian thực cho cấp độ công nghiệp. Nghiên cứu đã kiến trúc thành công và tiến hành xác thực thực nghiệm một hệ thống điều khiển phân tầng lai cho đối tượng hệ thống treo chủ động. Việc thay thế phương pháp phạt trọng số tĩnh truyền thống bằng mô hình học tăng cường Soft Actor-Critic, kết hợp dự báo mặt đường, đã mang đến sức mạnh linh hoạt tuyệt đối cho khung xe khi di chuyển qua các địa hình phức tạp. Điểm sáng của bài viết là việc thiết lập chặt chẽ nền tảng lý thuyết thông qua điều kiện Lyapunov, giúp vá lỗ hổng an toàn kinh điển của các mô hình hộp đen AI. Nghiên cứu trong tương lai sẽ định hướng thử nghiệm cấu trúc thuật toán trên mô hình xe 7 bậc tự do tổng quát và hoàn thiện bài test trên nguyên mẫu xe tự hành thực tế ngoài đường chạy.

V. Kết luận

Nghiên cứu này đã đề xuất và xác thực thành công một kiến trúc điều khiển phân tầng lai (AMPC-SAC) cho hệ thống treo chủ động. Bằng cách tích hợp dữ liệu dự báo mặt đường và thay thế ma trận trọng số tĩnh của MPC bằng mô hình học tăng cường Soft Actor-Critic, hệ thống đã thích ứng linh hoạt và hiệu quả trước các biến động của địa hình. Điểm đột phá về mặt lý thuyết của bài báo là việc áp đặt ràng buộc Lyapunov cứng lên quá trình cập nhật chính sách học, qua đó đảm bảo tính ổn định từ ngõ vào đến trạng thái (ISS) và khắc phục rủi ro an toàn của các mô hình AI. Đồng thời, kỹ thuật mở rộng

vector trạng thái đã bù đắp hiệu quả độ trễ pha của cơ cấu chấp hành. Kết quả mô phỏng phần cứng trong vòng lặp (HIL) trên dSPACE khẳng định tính ưu việt của thuật toán khi giảm thiểu lần lượt 53,0% và 23,9% chỉ số RMS gia tốc thân xe so với hệ thống thụ động và MPC tiêu chuẩn, đồng thời đảm bảo không vi phạm giới hạn hành trình cơ khí (0%). Mặc dù đạt được những kết quả khả quan, nghiên cứu hiện tại vẫn giới hạn ở mô hình động lực học 1/4 xe. Trong tương lai, nhóm nghiên cứu sẽ mở rộng cấu trúc thuật toán trên mô hình toàn xe (7 bậc tự do) và tiến hành đánh giá thực nghiệm trên nguyên mẫu xe tự hành trong điều kiện vận hành thực tế.

Tài liệu tham khảo

- Bao, L., Zhang, Y., & Wang, C. (2020). Impact of chassis vibrations on LiDAR SLAM accuracy in autonomous vehicles. *IEEE Sensors Journal*, 20(14), 7935-7944.
- Bui, T. M., Le, H. P., & Nguyen, D. T. (2021). Advanced constraints handling in predictive control for intelligent suspension. *Journal of Advanced Transportation*, 2021, 1-14.
- Dang, V. N., Tran, Q. B., & Phan, T. A. (2023). Safety-critical reinforcement learning for automotive applications. *IEEE Access*, 11, 20345-20358.
- Do, K. D., Nguyen, V. H., & Pham, C. (2022). Delay mitigation in networked control systems for smart vehicles. *Control Engineering Practice*, 122, 105101.
- Dong, H., He, L., & Sun, C. (2019). Active suspension control targeting sensor stabilization for autonomous systems. *Mechatronics*, 59, 143-154.
- Gao, Z., Lin, F., & Tang, X. (2022). LQR and MPC performance comparison for vehicles under constrained actuator dynamics. *Control Engineering Practice*, 118, 104973.

- Han, S.-Y., & Liang, T. (2022). Reinforcement-learning-based vibration control for a vehicle semi-active suspension system via the PPO approach. *Applied Sciences*, 12(6), 3078; <https://doi.org/10.3390/app12063078>
- Hoang, L., Vu, T., & Nguyen, H. (2021). Actuator dynamics modeling for robust control in active suspensions. *Journal of Dynamic Systems, Measurement, and Control*, 143(5), 051004.
- Huang, Y., Zhang, W., & Liu, S. (2023). Explicit handling of physical constraints in active suspensions via model predictive control. *IEEE/ASME Transactions on Mechatronics*, 28(1), 450-461.
- Kim, S., Park, J., & Lee, H. (2022). Deep reinforcement learning applications in non-linear dynamic systems. *Artificial Intelligence Review*, 55(4), 3125-3155.
- Lazar, M., Muñoz de la Peña, D., Heemels, W. P. M. H., & Alamo, T. (2008). On input-to-state stability of min-max nonlinear model predictive control. *Systems & Control Letters*, 57(1), 39-48.
- Le, T. A., Nguyen, M. T., & Tran, D. K. (2023). Towards provably safe reinforcement learning for physical systems. *IEEE Transactions on Automatic Control*, 68(2), 981-996.
- Li, Y., Chen, H., & Zhao, Y. (2023). Actuator delay compensation in electro-hydraulic suspensions using modified Smith predictors. *Mechanical Systems and Signal Processing*, 185, 109789.
- Liu, C., Wang, Z., & Chen, Y. (2021). A review of active suspension systems for automotive comfort enhancement. *Vehicle System Dynamics*, 59(12), 1836-1865.
- Nguyen, H. V., Tran, B. Q., & Vu, L. (2023). Impact of dynamic chassis on LiDAR perception algorithms. *Sensors*, 23(4), 2133.
- Pan, H., Zhao, Z., & Liu, Q. (2023). Deep deterministic policy gradient for continuous control of automotive suspensions. *Applied Soft Computing*, 134, 109961.
- Pham, T. T., Le, V. D., & Nguyen, T. (2022). Adaptive mechanisms in model predictive control for varying road profiles. *International Journal of Control*, 95(8), 2110-2125.
- Romero, A., Song, Y., & Scaramuzza, D. (2024). Actor-Critic Model Predictive Control. In *IEEE International Conference on Robotics and Automation (ICRA)*. (pp. 1-7).
- Smith, A., & Taylor, M. (2021). Optimization techniques in model predictive control for constrained systems. *IEEE Transactions on Control Systems Technology*, 29(3), 1145-1158.
- Sun, W., Gao, H., & Shi, P. (2020). *Advanced Control for Vehicle Active Suspension Systems*. Springer.
- Theunissen, J., Sorniotti, A., Gruber, P., & Lenzo, B. (2019). Regionless explicit model predictive control of active suspension systems with preview. *IEEE Transactions on Industrial Electronics*, 67(6), 4877-4888.
- Theunissen, J., Tota, A., Gruber, P., Dhaens, M., & Sorniotti, A. (2021). Preview-based techniques for vehicle suspension control: A state-of-the-art review. *Annual Reviews in Control*, 51, 206-235.
- Tran, M. Q., Pham, H. T., & Bui, N. T. (2021). Nonlinear characteristics of electro-hydraulic actuators in automotive applications. *Mechanism and Machine Theory*, 156, 104144.
- Yong, H., Seo, J., Kim, J., Kim, M., & Choi, J. (2023). Suspension control strategies using switched soft actor-critic models for real roads. *IEEE Transactions on Industrial Electronics*, 70(1), 824-832.

- Vu, Q. M., Dang, T. H., & Nguyen, L. (2022). Soft Actor-Critic algorithm enhancements for continuous state-space control. *Neural Networks*, 151, 210-225.
- Wang, X., Liu, Y., & Zhang, Z. (2022). Stability and perception mapping in highly automated vehicles. *IEEE Transactions on Intelligent Vehicles*, 7(2), 345-356.
- Zhang, S., Liu, C., & Dong, H. (2022). Adaptive penalty weight tuning in MPC using deep reinforcement learning. *IEEE Transactions on Cybernetics*, 52(8), 8123-8135.
- Zhao, Y., Sun, Q., & Ma, X. (2023). State-space augmentation techniques for systems with delayed inputs and previewable disturbances. *Journal of Process Control*, 125, 89-102.

ADAPTIVE MODEL PREDICTIVE CONTROL AUGMENTED WITH DEEP REINFORCEMENT LEARNING FOR ACTIVE SUSPENSION: DELAY COMPENSATION AND PREVIEW DISTURBANCE REJECTION

Nguyen Van Manh¹, Hoang Anh Dung¹

Abstract: Active suspension systems serve as indispensable inertial reference platforms, critical for ensuring the stability of perception sensors on autonomous vehicles. Conventional model-based feedback control algorithms typically operate with static penalty weights, rendering them inflexible in the face of dynamic road variations and frequently neglecting the adverse effects of actuator delay dynamics. To address these limitations, this study proposes a novel hierarchical control framework that synergizes Model Predictive Control (MPC) with a Deep Reinforcement Learning (DRL) agent via the Soft Actor-Critic (SAC) algorithm. Within this architecture, the MPC weighting matrices are dynamically adapted in real time based on previewed road data. The core scientific contribution of this paper lies in the imposition of a strict Lyapunov constraint on the DRL policy update mechanism. Coupled with a state-vector expansion technique to compensate for phase delays, this rigorously guarantees Input-to-State Stability (ISS) for the overall system. Hardware-in-the-Loop (HIL) simulations conducted on a dSPACE MicroLabBox testbed unequivocally validate the feasibility and superiority of the proposed methodology. Quantitative results demonstrate that the proposed framework curtails the root-mean-square (RMS) of chassis acceleration by 53.0% compared to a passive system and yields a 23.9% improvement over the standard MPC, all while strictly adhering to the physical constraints of the system.

Keywords: Active Suspension System (ASS), Deep Reinforcement Learning (DRL), Hardware-in-the-Loop (HIL) simulation, actuator delay compensation, Input-to-State Stability (ISS)

¹ Hanoi Open University, Hanoi, Vietnam