ROZSZERZONE STRESZCZENIE ROZPRAWY DOKTORSKIEJ

Aktywna kontrola transmisji hałasu z wykorzystaniem uczenia głębokiego przez struktury zamknięte (enkapsulowane)

Kandydat: Alkahf Aboutiman
Promotorzy: Prof. Dr. Hamid Reza Karimi (Politecnico di Milano), Prof. Francesco Ripamonti (Politecnico di Milano), Prof. dr hab. inż. Marek Pawełczyk (Politechnika Śląska), Dr inż. Stanisław Wrona (Politechnika Śląska)
Cykl doktorski 38 (2023–2026) — doktorat podwójny (cotutelle) między Politecnico di Milano a Politechniką Śląską

---

1. Zakres rozprawy

Niniejsza rozprawa dotyczy aktywnej kontroli hałasu (Active Noise Control, ANC) w złożonych środowiskach wibroakustycznych, ze szczególnym uwzględnieniem struktur zamkniętych (enkapsulowanych) — obudów takich jak kabiny samochodowe, kadłuby samolotów, przedziały kolejowe czy obudowy przemysłowe, w których wewnętrzna przestrzeń musi zostać akustycznie odizolowana od środowiska zewnętrznego. Struktury zamknięte charakteryzują się silnym zachowaniem modalnym, wrażliwością na otwory i warunki brzegowe, a często wykazują dynamikę fazy nieminimalnej (non-minimum phase, NMP), która zasadniczo ogranicza skuteczność klasycznych adaptacyjnych algorytmów ANC. W praktyce takie obudowy często zawierają funkcjonalne otwory — służące wentylacji, przeprowadzeniu okablowania, regulacji termicznej lub przepływowi powietrza — które dodatkowo modyfikują akustyczne ścieżki transmisji i wprowadzają dodatkowe efekty nieliniowe oraz dyfrakcyjne, niemożliwe do uchwycenia przez klasyczne wąskopasmowe modele sterowania.

Zakres pracy obejmuje trzy uzupełniające się kierunki badawcze:

1. Fizyczne i matematyczne modelowanie zamkniętych systemów wibroakustycznych, uzyskane za pomocą analizy metodą elementów skończonych oraz eksperymentalnej identyfikacji systemu, oraz projektowanie strategii sterowania opartych na modelu — w szczególności filtracji Kalmana — zdolnych radzić sobie z dynamiką fazy nieminimalnej, której klasyczne filtry adaptacyjne nie są w stanie skompensować.
2. Hybrydowe architektury uczenia głębokiego łączące lekkie konwolucyjne sieci neuronowe z klasyczną filtracją adaptacyjną w celu radzenia sobie z nieliniowościami wprowadzanymi przez przetworniki i otwory w obudowie, oceniane

zarówno za pomocą metryk obiektywnych (tłumienie w dB, widma redukcji hałasu), jak i subiektywnych metryk percepcyjnych skoncentrowanych na człowieku (uciążliwość, jakość afektywna, głośność).

3. Kontrolery oparte na głębokim transferowym uczeniu ze wzmocnieniem (deep transfer reinforcement learning, DTRL), zbudowane na proponowanej architekturze FxMDNoiseNet, które adaptują się online do silnie niestacjonarnych, zmiennych w czasie warunków akustycznych panujących wewnątrz poruszającego się pojazdu, w tym do niedopasowania ścieżki wtórnej spowodowanego zmienną liczbą pasażerów i konfiguracją kabiny.

Poza samymi algorytmami sterowania, rozprawa kładzie wyraźny nacisk na wykonalność obliczeniową w kontekście wdrożeń wbudowanych czasu rzeczywistego — wymóg powracający we wszystkich trzech kierunkach badawczych, uzasadniający systematyczne stosowanie lekkich architektur sieciowych (rzędu kilku tysięcy parametrów trenowalnych), operacji zgodnych z arytmetyką stałoprzecinkową oraz niskich częstotliwości próbkowania odpowiednich dla procesorów sygnałowych stosowanych w motoryzacji.

Badania zostały przeprowadzone w ramach doktoratu podwójnego między Politecnico di Milano (Wydział Inżynierii Mechanicznej) a Politechniką Śląską, w ramach projektu IN-NOVA finansowanego z programu Horyzont Europa w ramach działań Marie Skłodowska-Curie (umowa grantowa nr 101073037). Staże naukowe zrealizowano na University College London (badania subiektywnej percepcji ANC, z Prof. Jian Kangiem i Dr. Francesco Alettą), w instytucji partnerskiej w Polsce, oraz w firmie Müller-BBM Active Sound Technology GmbH w Niemczech, gdzie przemysłowe praktyki kontroli hałasu motoryzacyjnego wpłynęły na projekt eksperymentalny ostatniego kierunku badawczego.

2. Cele rozprawy i teza

Nadrzędnym celem rozprawy jest poprawa odporności, adaptacyjności oraz skuteczności percepcyjnej systemów aktywnej kontroli hałasu działających w złożonych, nieliniowych i zmiennych w czasie środowiskach wibroakustycznych — przy czym głównym scenariuszem aplikacyjnym jest kontrola hałasu w kabinie pojazdu, ze względu na połączenie silnego sprzężenia modalnego, nieliniowości przetworników oraz nieustannie zmieniających się warunków eksploatacji (obciążenie silnika, nawierzchnia drogi, prędkość pojazdu, liczba pasażerów).

Główna teza broniona w niniejszej pracy głosi, że klasyczne, czysto adaptacyjne algorytmy ANC (np. FxNLMS) są strukturalnie ograniczone w konfrontacji z dynamiką fazy nieminimalnej, nieliniowościami przetworników oraz niestacjonarnymi warunkami akustycznymi typowymi dla struktur zamkniętych, oraz że ograniczenia te można

systematycznie przezwyciężyć poprzez połączenie modelowania fizycznego ze sterowaniem opartym na uczeniu maszynowym — a konkretnie estymacji stanu za pomocą filtru Kalmana, hybrydowej selekcji/adaptacji filtrów opartej na uczeniu głębokim (rodziny SFANC/GFANC) oraz głębokiego transferowego uczenia ze wzmocnieniem (DDPG/TD3) — przy jednoczesnym zachowaniu niskiej złożoności obliczeniowej kontrolerów, umożliwiającej ich wdrożenie w czasie rzeczywistym na urządzeniach wbudowanych.

Twierdzeniem drugorzędnym, lecz istotnym metodologicznie, jest to, że obiektywne metryki tłumienia (wyrażone w dB) są konieczne, lecz niewystarczające do oceny praktycznej wartości systemu ANC, oraz że reakcja percepcyjna człowieka — uciążliwość, jakość afektywna i postrzegana głośność — musi być oceniana bezpośrednio poprzez kontrolowane eksperymenty słuchowe, aby zweryfikować, czy poprawa algorytmiczna przekłada się na rzeczywisty wzrost komfortu pasażera.

Szczegółowe cele wynikające z tej tezy są następujące:

- Scharakteryzowanie, za pomocą modelowania metodą elementów skończonych oraz identyfikacji systemu, zachowania fazy nieminimalnej zamkniętej struktury wibroakustycznej oraz określenie ilościowe jego wpływu na skuteczność klasycznych adaptacyjnych algorytmów ANC.
- Zaprojektowanie i walidacja strategii ANC opartej na filtrze Kalmana, poprawiającej odporność i tłumienie w warunkach dynamiki NMP w porównaniu z FxNLMS, przy jednoczesnym określeniu ilościowym związanego z tym narzutu obliczeniowego.
- Opracowanie hybrydowych architektur ANC opartych na uczeniu głębokim (SFANC-NLMS, GFANC-THFxNLMS), które pozostają efektywne obliczeniowo, jednocześnie poprawiając odporność na nieliniowości głośników oraz otwory w obudowie.
- Ocena, poprzez kontrolowane eksperymenty słuchowe z udziałem ludzi, w jaki sposób skuteczność algorytmiczna ANC przekłada się na postrzeganą uciążliwość, jakość afektywną i postrzeganą głośność, z wykorzystaniem zwalidowanych protokołów psychoakustycznych (ISO 15666) oraz odpowiedniej wielowymiarowej analizy statystycznej.
- Zaprojektowanie lekkiego neuronowego kontrolera ANC (FxMDNoiseNet) oraz jego integracja z głębokim uczeniem ze wzmocnieniem (DDPG, TD3) i uczeniem transferowym w celu uzyskania kontrolera opartego na głębokim transferowym uczeniu ze wzmocnieniem (FxMDNoiseNet-DTRL), zwalidowanego w realistycznych, eksperymentalnie zmierzonych warunkach motoryzacyjnych, w tym przy niedopasowaniu ścieżki wtórnej oraz dynamice liniowej zmiennej w czasie (LTV).

- Ocena, we wszystkich trzech kierunkach badawczych, kompromisu między skutecznością sterowania a kosztem obliczeniowym, wyrażonym w liczbie parametrów trenowalnych, liczbie operacji mnożenia-akumulacji (MAC) oraz kompatybilności z wbudowanymi procesorami sygnałowymi stosowanymi w motoryzacji.

3. Opis przedmiotu badań

3.1 Aktywna kontrola hałasu w zamkniętych systemach o fazie nieminimalnej

Pierwszy kierunek badawczy dotyczy ANC w zamkniętej strukturze wibroakustycznej — laboratoryjnym „noise boxie" reprezentatywnym dla kabiny samochodowej — modelowanej metodą elementów skończonych (MES) w celu uchwycenia sprzężonej dynamiki strukturalno-akustycznej. Domena strukturalna została zdyskretyzowana za pomocą elementów powłokowych opartych na teorii płyt Reissnera–Mindlina z formułowaniem MITC (Mixed Interpolation of Tensorial Components), podczas gdy wnęka akustyczna została zamodelowana w ramach zunifikowanego frameworku elementów skończonych wibroakustyki. Tłumienie strukturalne wprowadzono za pomocą izotropowego współczynnika strat $\eta = 0{,}003$, zgodnego z typowymi wartościami raportowanymi dla rozważanego materiału płyty. Uzyskane funkcje odpowiedzi częstotliwościowej (FRF) między źródłem pierwotnym, źródłem wtórnym (sterującym) a mikrofonem błędu zostały obliczone, a następnie zwalidowane względem pomiarów eksperymentalnych, uzyskanych przy użyciu skalibrowanej aparatury, w tym mikrofonu o czułości 37,37 mV/Pa, skalibrowanego kalibratorem akustycznym Brüel & Kjær 4231 przy 1 kHz i 94 dB.

Analiza biegunów i zer zidentyfikowanych ścieżek pierwotnej, wtórnej i referencyjnej ujawnia zera zlokalizowane poza okręgiem jednostkowym w dziedzinie dyskretnej — czyli zachowanie fazy nieminimalnej. Ta cecha NMP wprowadza opóźnienia fazowe i ograniczenia przyczynowości, które — jak wykazano zarówno analitycznie, jak i eksperymentalnie — zasadniczo ograniczają osiągalne tłumienie klasycznego algorytmu Filtered-x Normalised Least Mean Squares (FxNLMS), którego optymalny filtr sterujący formalnie wymaga niestabilnej inwersji ścieżki wtórnej, gdy ta ostatnia posiada charakter fazy nieminimalnej.

Aby przezwyciężyć to ograniczenie, przeprowadzono dyskretną identyfikację ścieżek akustycznych w przestrzeni stanów oraz zaproponowano strategię ANC opartą na filtrze Kalmana, w której poprawiona estymacja stanu kompensuje niepewność wynikającą z dynamiki NMP bez konieczności jawnej inwersji ścieżki wtórnej. Wszystkie symulacje w ramach tego badania przeprowadzono przy częstotliwości próbkowania 8 kHz, zgodnej z

pasmem zainteresowania dla niskoczęstotliwościowego hałasu motoryzacyjnego. Złożoność obliczeniowa obu podejść została przeanalizowana pod względem liczby operacji zmiennoprzecinkowych mnożenia i dodawania (FLOP) na próbkę, potwierdzając, że podejście oparte na filtrze Kalmana, choć bardziej wymagające niż FxNLMS, pozostaje kompatybilne z implementacją czasu rzeczywistego na wbudowanym sprzęcie motoryzacyjnym, kosztem umiarkowanego, lecz uzasadnionego wzrostu obciążenia obliczeniowego.

3.2 Hybrydowa kontrola ANC oparta na uczeniu głębokim: ocena obiektywna i subiektywna

Drugi kierunek badawczy wprowadza hybrydowe architektury łączące lekką jednowymiarową konwolucyjną sieć neuronową (1D CNN) z klasyczną filtracją adaptacyjną. W ramach SFANC-NLMS sieć CNN wybiera, klatka po klatce, najbardziej odpowiedni wstępnie wytrenowany filtr sterujący z banku filtrów stałych, inspirowanego metodą Selective Fixed-filter Active Noise Control (SFANC), podczas gdy etap NLMS udoskonala online współczynniki wybranego filtru. Ta hybrydowa konstrukcja łączy niskie opóźnienie selekcji filtru — pozwalające kontrolerowi na natychmiastową reakcję na zmiany w widmie hałasu — ze stabilnością i zdolnością precyzyjnego dostrajania właściwą adaptacyjnemu udoskonalaniu.

Klasyfikator CNN wytrenowano na syntetycznym zbiorze danych obejmującym 84 000 próbek szerokopasmowego szumu losowego, podzielonym na 80 000 próbek treningowych, 2000 próbek walidacyjnych oraz 2000 próbek testowych. Trening przeprowadzono z wykorzystaniem optymalizatora Adam ze współczynnikiem uczenia 10^{-4}, rozmiarem paczki (batch size) 250, przez łącznie 30 epok, z regularyzacją L2 zastosowaną do wszystkich parametrów trenowalnych w celu ograniczenia przeuczenia i zachowania lekkiej architektury odpowiedniej do implementacji wbudowanej. Uzyskany model osiąga wysoką dokładność klasyfikacji przy znacznej redukcji liczby parametrów trenowalnych w porównaniu z większymi architekturami referencyjnymi, co potwierdza jego przydatność do zastosowań ANC czasu rzeczywistego na sprzęcie wbudowanym. Podczas działania systemu przychodzący hałas jest próbkowany z częstotliwością 16 kHz, przy czym CNN analizuje jednosekundowe ramki i przekazuje indeks wybranego filtru do kontrolera czasu rzeczywistego, który następnie w sposób ciągły dostraja parametry filtru z częstotliwością próbkowania; walidacja z wykorzystaniem nagrań ze zbioru danych miejskich dźwięków SONYC potwierdziła zdolność generalizacji wykraczającą poza syntetyczny rozkład treningowy.

Poza tłumieniem obiektywnym, ten kierunek badawczy analizuje, w jaki sposób skuteczność ANC jest postrzegana przez słuchaczy — wymiar często pomijany w czysto

algorytmicznych badaniach nad ANC. Wykorzystując makietę wnętrza pojazdu w formie noise boxu, wyposażoną w symulator głowy i tułowia (Head and Torso Simulator, HATS) oraz skalibrowane odtwarzanie przez słuchawki, zrekrutowano łącznie 35 uczestników (16 mężczyzn i 19 kobiet, w wieku od 23 do 62 lat, średni wiek około 31 lat), poddanych ekspozycji na kombinacje rodzaju hałasu (motocykl, ulica, pociąg), poziomu ciśnienia akustycznego (55, 65, 72 dB(A)) oraz warunku sterowania (brak kontroli, NLMS, SFANC-NLMS). Dwóch uczestników zostało wykluczonych z ostatecznej analizy z powodu niewiarygodnych wzorców odpowiedzi (współczynnik korelacji Spearmana na poziomie 0,8 w połączeniu z wartością p przekraczającą 20% w kontrolach spójności powtórzeń bodźców). Przed głównym eksperymentem uczestnicy wypełnili kwestionariusze demograficzne oraz przeszli ocenę wstępną obejmującą skalę indywidualnej wrażliwości na hałas (Individual Noise Sensitivity, INS) Weinsteina, w celu uwzględnienia zmienności międzyosobniczej oraz potencjalnych czynników zakłócających, takich jak stres i nastrój.

Uczestnicy oceniali postrzeganą uciążliwość (PAY) za pomocą standaryzowanej skali jednopozycyjnej ISO 15666, jakość afektywną (PAQ) za pomocą dwuwymiarowej siatki afektywnej (pobudzenie/walencja) oraz postrzeganą głośność (PLN) metodą estymacji wielkości. Analizę statystyczną przeprowadzono z wykorzystaniem trójczynnikowej ANOVA z powtarzanymi pomiarami (3WR-ANOVA) dla PAY i PLN, z rodzajem hałasu, poziomem hałasu i warunkiem ANC jako czynnikami wewnątrzobiektowymi; dla PAQ zastosowano zamiast tego trójczynnikową permutacyjną wielowymiarową analizę wariancji z powtarzanymi pomiarami (3WR-PERMANOVA), ze względu na naruszenia wielowymiarowej normalności potwierdzone testem Mardii. Tam, gdzie wykryto istotne efekty główne lub interakcje, zastosowano porównania parami post-hoc. Wykazano, że SFANC-NLMS konsekwentnie przewyższa NLMS we wszystkich trzech wymiarach percepcyjnych, niezależnie od rodzaju i poziomu hałasu, a zależność między obiektywnym tłumieniem a postrzeganym komfortem okazała się nietrywialna — w szczególności zawartość niskoczęstotliwościowa (20–500 Hz) okazała się dominującym czynnikiem postrzeganego dyskomfortu, nawet gdy ogólne tłumienie szerokopasmowe było zbliżone między warunkami — co uzasadnia potrzebę oceny skoncentrowanej na człowieku jako niezbędnego uzupełnienia metryk opartych na decybelach w projektowaniu systemów ANC.

Bazując na tych wynikach, opracowano metodę Generative Fixed-filter ANC połączoną z etapem FxNLMS uwzględniającym przetwornik/harmoniczne (GFANC-THFxNLMS), mającą na celu przeciwdziałanie nieliniowym zniekształceniom wprowadzanym przez głośnik wtórny, co jest szczególnie istotne dla struktur zamkniętych z otworami. Konfiguracja eksperymentalna zastosowana w tym badaniu obejmowała obudowę z kołowym otworem o promieniu 0,05 m, trzy mikrofony umieszczone wewnątrz obudowy oraz głośnik wtórny

umieszczony centralnie za otworem. Identyfikacja systemu potwierdziła, że ścieżka pierwotna wykazuje silne zachowanie rezonansowe poniżej 350 Hz, wynikające ze sprzężonej dynamiki między wibrującą płytą a wyciekiem akustycznym przez otwór, podczas gdy ścieżka wtórna, sterowana przez głośnik wewnętrzny, wykazuje gładszą, bardziej tłumioną odpowiedź, zgodną z bardziej bezpośrednią propagacją akustyczną. Sieć CNN dynamicznie wybiera i liniowo łączy podfiltry rozłożone z szerokopasmowego filtru sterującego, podczas gdy etap THFxNLMS kompensuje nieliniowości głośnika w czasie rzeczywistym, wykorzystując mechanizm skalowania gradientu, który dostosowuje efektywny rozmiar kroku w zależności od chwilowego poziomu hałasu, poprawiając stabilność i zbieżność w warunkach niestacjonarnych. Ostateczny wytrenowany model CNN dla tego badania zawiera 6,74 tys. parametrów i wymaga około 19,33 miliona operacji mnożenia-akumulacji (MAC) na jednosekundową ramkę przy częstotliwości próbkowania 3 kHz — koszt obliczeniowy wystarczająco niski, aby model mógł być w całości przechowywany w pamięci on-chip na wbudowanych procesorach sygnałowych, takich jak Analog Devices ADSP-SC598, minimalizując opóźnienie wnioskowania.

3.3 Głębokie transferowe uczenie ze wzmocnieniem dla ANC we wnętrzach pojazdów

Trzeci i najbardziej zaawansowany kierunek badawczy dotyczy silnie niestacjonarnych, nieliniowych warunków akustycznych występujących w poruszającym się pojeździe, gdzie obciążenie silnika, wymuszenie drogowe, prędkość pojazdu oraz liczba pasażerów nieustannie zmieniają ścieżki transmisji akustycznej. Zaproponowano lekki kontroler w postaci sieci neuronowej, FxMDNoiseNet, działający w architekturze filtered-x. Jego rdzeń składa się z ułożonych w stos modułów Atrous Scaled Convolution (ASC) o stopniowo zwiększających się współczynnikach dylatacji, połączonych z przyczynowymi i punktowymi jednowymiarowymi splotami oraz połączeniami rezydualnymi/skip, umożliwiającymi efektywną wieloskalową ekstrakcję cech w dziedzinie czasu z surowego sygnału audio, przy jednoczesnym utrzymaniu bardzo niskiej całkowitej liczby parametrów — decyzja projektowa bezpośrednio podyktowana ograniczeniami czasu rzeczywistego i sprzętu wbudowanego, zidentyfikowanymi w całej rozprawie. Dane treningowe dla tego etapu pochodziły z odpowiedzi impulsowych zarejestrowanych w rzeczywistym pojeździe wyposażonym w głośniki niskotonowe zamontowane w drzwiach oraz fabrycznie zainstalowany subwoofer umieszczony w bagażniku, przy czym sygnały pierwotnie zarejestrowano z częstotliwością 48 kHz, a następnie przepróbkowano do 3 kHz w celu redukcji złożoności obliczeniowej, zgodnie z niskoczęstotliwościową zawartością istotną dla kontroli hałasu w kabinie. Optymalizację przeprowadzono z wykorzystaniem optymalizatora Adam ze współczynnikiem uczenia 2×10^{-4} oraz zanikiem wag (weight decay) 10^{-4}, w połączeniu ze schedulerem współczynnika uczenia typu step (współczynnik zaniku 0,5 co 5 epok) oraz inicjalizacją wag Xaviera w celu zapewnienia stabilnej

zbieżności, z zastosowaniem tej samej metodologii podziału 80 000/2000/2000 na zbiór treningowy/walidacyjny/testowy, co w poprzednim kierunku badawczym.

Aby poprawić adaptacyjność w warunkach zmiennych w czasie, na bazie FxMDNoiseNet zintegrowano warstwę głębokiego uczenia ze wzmocnieniem (RL). Problem ANC sformułowano jako sekwencyjne zadanie podejmowania decyzji (proces decyzyjny Markowa), w którym w każdym dyskretnym kroku czasowym n agent obserwuje stan s_n, wyprowadzony z estymowanego zakłócenia, i generuje ciągłą akcję sterującą $a_n = \pi\phi(s_n)$, stosowaną poprzez ścieżkę wtórną w celu wytworzenia sygnału antyhałasu; wynikowy błąd resztkowy e(n) jest wykorzystywany do obliczenia nagrody r(n) = −|e(n)|, tak że maksymalizacja skumulowanej nagrody odpowiada minimalizacji energii hałasu resztkowego w czasie. Zbadano dwa algorytmy RL dla sterowania ciągłego typu off-policy actor–critic: Deep Deterministic Policy Gradient (DDPG) oraz jego bardziej stabilny następca, Twin Delayed DDPG (TD3). TD3 rozwiązuje dwa dobrze znane ograniczenia DDPG — obciążenie przeszacowaniem (overestimation bias) funkcji wartości oraz akumulację błędu wynikającą z bootstrapowanych celów krytyka — za pomocą trzech uzupełniających się mechanizmów: (i) Clipped Double Q-learning, w którym trenowane są dwa niezależne krytyki Qθ1 i Qθ2, a do wyznaczenia celu Bellmana wykorzystywana jest minimalna z ich estymacji, łagodząc przeszacowanie wartości; (ii) wygładzanie polityki docelowej (target policy smoothing), w którym do akcji docelowej dodawany jest przycięty szum gaussowski, regularyzujący estymację wartości względem ostrych, podatnych na błędy szczytów funkcji Q; oraz (iii) opóźnione aktualizacje polityki, w których krytycy są aktualizowani w każdym kroku środowiska, podczas gdy aktor (i sieci docelowe) są aktualizowane jedynie co d iteracji, co pozwala estymacji wartości ustabilizować się, zanim zostanie wykorzystana do poprawy polityki.

W obu konfiguracjach, DDPG i TD3, sieć polityki RL jest inicjalizowana wstępnie wytrenowanymi wagami FxMDNoiseNet, zamiast inicjalizacji losowej — strategia uczenia transferowego, która znacząco przyspiesza zbieżność, poprawia stabilność treningu i zmniejsza ilość danych interakcji wymaganych z (kosztownym w symulacji) rzeczywistym środowiskiem akustycznym, prowadząc do zaproponowanego frameworku FxMDNoiseNet-DTRL. Eksploracja podczas treningu jest kierowana przez adaptacyjny proces szumu gaussowskiego skalowanego obwiednią, którego odchylenie standardowe jest modulowane przez chwilową obwiednię sygnału zakłócenia, umożliwiając bardziej agresywną eksplorację podczas stanów przejściowych o wysokiej energii oraz bardziej konserwatywne zachowanie w okresach cichych. Sieci krytyków są trenowane w celu minimalizacji błędu Bellmana z regularyzacją $\ell 2$, pobierając mini-paczki (mini-batche) z bufora powtórek (replay buffer) zapełnianego podczas interakcji z (symulowanym, a następnie eksperymentalnie zmierzonym) środowiskiem akustycznym.

Framework został oceniony z wykorzystaniem eksperymentalnie zmierzonych odpowiedzi impulsowych zarejestrowanych w rzeczywistym pojeździe, zarówno w warunkach liniowych stacjonarnych w czasie (LTI), jak i liniowych zmiennych w czasie (LTV), w tym z celowym niedopasowaniem ścieżki wtórnej — trenując kontroler przy użyciu odpowiedzi impulsowej pustego pojazdu, a oceniając go przy użyciu odpowiedzi impulsowej zmierzonej z dwoma pasażerami — w celu jawnego przetestowania odporności na niepewność modelowania. Warunki te ściśle odzwierciedlają rzeczywistą eksploatację pojazdu, w której rzeczywista ścieżka wtórna rzadko jest znana dokładnie i zmienia się nieustannie wraz z obciążeniem pojazdu i konfiguracją kabiny.

4. Główne wyniki i wnioski

We wszystkich trzech kierunkach badawczych wyniki konsekwentnie pokazują, że połączenie modelowania fizycznego ze sterowaniem opartym na uczeniu maszynowym prowadzi do istotnej poprawy względem klasycznego adaptacyjnego ANC, przy jednoczesnym zachowaniu zwartości kontrolerów wystarczającej do zastosowań czasu rzeczywistego na wbudowanym sprzęcie motoryzacyjnym.

Podsumowanie głównych wyników ilościowych:

Badanie	Porównywane metody	Kluczowy wynik ilościowy
Artykuł A — struktura zamknięta NMP	ANC oparte na filtrze Kalmana vs. FxNLMS	16,92 dB vs. 2,72 dB tłumienia (pasmo 0–450 Hz)
Artykuł B — percepcja subiektywna	SFANC-NLMS vs. NLMS (N = 33 ważnych uczestników)	Istotna poprawa PAY, PAQ i PLN we wszystkich testowanych warunkach (3WR-ANOVA / 3WR-PERMANOVA, $p < 0,05$)
Artykuł C — GFANC-THFxNLMS	GFANC-THFxNLMS vs. GFANC vs. THFxNLMS	9,49 dB vs. 8,01 dB vs. 6,74 dB tłumienia
Artykuły D–E — FxMDNoiseNet-DTRL	FxMDNoiseNet-DTRL (DDPG/TD3) vs. FxMDNoiseNet vs. FxNLMS	Poprawiony NMSE i tłumienie PSD w warunkach LTI i LTV przy niedopasowaniu ścieżki wtórnej

Dla zamkniętej struktury o fazie nieminimalnej (Artykuł A) strategia ANC oparta na filtrze Kalmana osiągnęła tłumienie do 16,92 dB w paśmie 0–450 Hz, w porównaniu z zaledwie 2,72 dB dla FxNLMS. Porównania sygnału błędu resztkowego w dziedzinie czasu oraz spektrogramów potwierdzają, że FxNLMS pozostawia istotną energię resztkową skoncentrowaną wokół częstotliwości rezonansowych i antyrezonansowych sprzężonego

systemu strukturalno-akustycznego, podczas gdy podejście oparte na filtrze Kalmana istotnie tłumi tę energię resztkową w całym analizowanym paśmie, kosztem umiarkowanego wzrostu obciążenia obliczeniowego, który pozostaje kompatybilny z implementacją czasu rzeczywistego. Wyniki te potwierdzają, że dokładna estymacja stanu jest niezbędna, gdy ścieżki akustyczne wykazują dynamikę fazy nieminimalnej, której klasyczne filtry adaptacyjne, ze swojej konstrukcji, nie są w stanie skompensować.

Badanie subiektywne (Artykuł B) wykazało, że SFANC-NLMS jest konsekwentnie preferowany względem NLMS pod względem postrzeganej uciążliwości i postrzeganej głośności, przy czym istotność statystyczna została potwierdzona trójczynnikową ANOVA z powtarzanymi pomiarami (dla PAY i PLN) oraz permutacyjną wielowymiarową analizą wariancji (dla PAQ). Choć wyższe poziomy ciśnienia akustycznego zgodnie z oczekiwaniami ogólnie zwiększają postrzeganą uciążliwość, nie odwracają one względnej hierarchii algorytmów ANC — SFANC-NLMS pozostaje warunkiem preferowanym we wszystkich testowanych rodzajach i poziomach hałasu. Zawartość niskoczęstotliwościowa, a nie samo tłumienie szerokopasmowe, okazała się dominującym predyktorem postrzeganego dyskomfortu, co wzmacnia wartość podejścia hybrydowego, opartego na uczeniu maszynowym, wykraczającą poza czysto obiektywne metryki skuteczności oparte na decybelach, i potwierdza metodologiczny wybór połączenia oceny algorytmicznej i psychoakustycznej w całej rozprawie.

Hybrydowy framework GFANC-THFxNLMS (Artykuł C) osiągnął redukcję hałasu na poziomie 9,49 dB, przewyższając zarówno sam GFANC (8,01 dB), jak i sam THFxNLMS (6,74 dB), w zamkniętej strukturze z otworem akustycznym — konfiguracji, w której nieliniowości głośnika i efekty dyfrakcyjne są szczególnie wyraźne. Wyniki te dowodzą, że połączenie selekcji filtrów opartej na CNN z filtracją adaptacyjną uwzględniającą nieliniowości poprawia odporność na nieliniowości głośników bez utraty efektywności obliczeniowej, potwierdzając hipotezę, że architektury hybrydowe mogą przewyższać zarówno czystą selekcję filtrów stałych, jak i czystą filtrację adaptacyjną rozpatrywane osobno.

Wreszcie, framework FxMDNoiseNet-DTRL (Artykuły D i E) przewyższył zarówno samodzielny kontroler oparty na uczeniu głębokim (FxMDNoiseNet), jak i klasyczny FxNLMS w realistycznych warunkach motoryzacyjnych, w tym przy niedopasowaniu ścieżki wtórnej i dynamice LTV — najbardziej wymagającym scenariuszu testowym rozważanym w rozprawie, łączącym składowe hałasu silnikowego i drogowego. W dziedzinie częstotliwości FxMDNoiseNet-DTRL konsekwentnie osiągał najsilniejsze tłumienie w paśmie niskich częstotliwości (50–250 Hz), gdzie dominuje energia hałasu silnikowego, co potwierdza jego odporność w celowaniu w najbardziej energetyczne składowe akustyczne pomimo błędów modelowania celowo wprowadzonych poprzez warunki LTV i niedopasowanie ścieżki

wtórnej. Strategia inicjalizacji poprzez uczenie transferowe — zasilenie polityki RL wstępnie wytrenowanymi wagami FxMDNoiseNet zamiast wag losowych — okazała się kluczowym czynnikiem umożliwiającym osiągnięcie stabilnej zbieżności w ramach praktycznej liczby interakcji treningowych, co ma bezpośrednie implikacje dla wykonalności wdrażania kontrolerów ANC opartych na RL w rzeczywistych cyklach rozwoju motoryzacyjnego, w których interakcja z rzeczywistym środowiskiem akustycznym jest kosztowna.

Podsumowując, rozprawa prezentuje spójną progresję — od modelowania fizycznego dynamiki fazy nieminimalnej, przez hybrydową filtrację opartą na uczeniu głębokim, zwalidowaną eksperymentami percepcji człowieka, aż po głębokie transferowe uczenie ze wzmocnieniem — która systematycznie rozszerza zakres działania systemów ANC w kierunku realistycznych, nieliniowych, zmiennych w czasie środowisk motoryzacyjnych, przy jednoczesnym poszanowaniu na każdym etapie ograniczeń obliczeniowych wdrożeń wbudowanych czasu rzeczywistego. We wszystkich trzech kierunkach liczba parametrów trenowalnych została utrzymana w zakresie kilku tysięcy, a koszt obliczeniowy był konsekwentnie wyrażany w kategoriach kompatybilnych z istniejącymi wbudowanymi procesorami sygnałowymi stosowanymi w motoryzacji, co bezpośrednio wspiera praktyczne znaczenie proponowanych metod wykraczające poza badania symulacyjne.

Rozprawa wskazuje trzy główne kierunki dalszych badań. Po pierwsze, przejście od formułowania sterowania próbka po próbce do formułowania blokowego pozwoliłoby kontrolerowi lepiej wykorzystać dłuższe zależności czasowe w sygnale zakłócenia oraz zredukować obciążenie obliczeniowe poprzez przetwarzanie wsadowe, kosztem wprowadzenia dodatkowego opóźnienia algorytmicznego, które wymagałoby dokładnej charakteryzacji. Po drugie, zbadanie bardziej zaawansowanych algorytmów uczenia ze wzmocnieniem dla sterowania ciągłego — takich jak Distributed Distributional Deep Deterministic Policy Gradient (D4PG) lub inne nowsze metody off-policy uwzględniające dystrybucyjną estymację wartości — mogłoby dodatkowo poprawić szybkość zbieżności oraz odporność na niepewność modelowania wykraczającą poza to, co osiągnięto z TD3. Po trzecie, implementacja proponowanego frameworku FxMDNoiseNet-DTRL na wbudowanym sprzęcie czasu rzeczywistego, zamiast walidacji offline z wykorzystaniem wcześniej zarejestrowanych odpowiedzi impulsowych, pozwoliłaby na ocenę podejścia w warunkach rzeczywistych opóźnień, szumu czujników oraz nieliniowości sprzętowych, zamykając lukę między obecną walidacją opartą na symulacji a pełnym wdrożeniem w pojeździe.

5. Odniesienia do głównych publikacji autora

Rozprawa jest przedstawiona jako zbiór pięciu załączonych, tematycznie powiązanych artykułów (Artykuły A–E), wymienionych poniżej. Z badań nie wynikły żadne patenty;

wszystkie wyniki zostały upowszechnione poprzez publikacje recenzowane oraz materiały konferencyjne.

Artykuły załączone, stanowiące podstawę rozprawy:

Artykuł A. A. Aboutiman, R. Shams, H.R. Karimi, F. Ripamonti, M. Pawełczyk, „Active noise control in encapsulated structures with non-minimum phase characteristics using a Kalman filter approach", *Journal of Sound and Vibration*, 615 (2025) 119187.

Artykuł B. A. Aboutiman, Z. Rachman, T. Oberman, F. Aletta, J. Kang, H.R. Karimi, F. Ripamonti, „Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study", *Applied Acoustics*, 239 (2025) 110823.

Artykuł C. A. Aboutiman, K.S.A. Maamoun, H.R. Karimi, F. Ripamonti, „Hybrid deep learning–based active noise control for encapsulated structures with openings", *Expert Systems With Applications*, 309 (2026) 131247.

Artykuł D. A. Aboutiman, J. Kirchhof, F. Ripamonti, „Transfer Learning-Enhanced DDPG for Adaptive Active Noise Control in Vehicle Interiors", 39th Conference on Neural Information Processing Systems (NeurIPS 2026), w recenzji.

Artykuł E. A. Aboutiman, J. Kirchhof, H.R. Karimi, F. Ripamonti, „Deep transfer reinforcement learning for active noise control in vehicle interiors", *Nature Machine Intelligence*, złożony do publikacji, 2026.

Pozostałe publikacje kandydata:

- K.S.A. Maamoun, A. Aboutiman, A.M. Ali, „Synergistic use of inertial actuators for advanced noise control in transparent patched vibrating plates", *Measurement*, 2025.
- C. Liang, A. Aboutiman, J. Kirchhof, F. Ripamonti, H.R. Karimi, „Experimental Evaluation of Simultaneous Perturbation Stochastic Approximation Algorithms for Active Engine Noise Control in Automotive Applications", *Applied Acoustics*.
- A. Aboutiman, H.R. Karimi, F. Ripamonti, „A hybrid learning-based active noise control algorithm for encapsulated structures with linear time-varying characteristics", 31st International Congress on Sound and Vibration (ICSV31), 2025 — **nagroda za najlepszy referat (1. miejsce)**.
- A. Aboutiman, H.R. Karimi, F. Ripamonti, „Hybrid Learning-based Active Noise Control in Encapsulated Structures", International Conference on Machine Learning and Artificial Intelligence for Audio, 2025.

- A. Aboutiman, H.R. Karimi, F. Ripamonti, „Frequency-Based Modeling and Active Noise Control Design for Encapsulated Structures", ICSV30, 2024.

6. Podsumowanie własnego wkładu kandydata

Kandydat, Alkahf Aboutiman, jest pierwszym i korespondencyjnym autorem wszystkich pięciu załączonych artykułów (A–E) oraz — we współpracy z promotorami — opracował ogólny program badawczy łączący modelowanie fizyczne, hybrydowe uczenie głębokie oraz głębokie transferowe uczenie ze wzmocnieniem w kontekście aktywnej kontroli hałasu.

Konkretny wkład własny kandydata obejmuje:

- Opracowanie modeli elementów skończonych oraz modeli w przestrzeni stanów zamkniętej struktury wibroakustycznej, identyfikację jej charakterystyk fazy nieminimalnej poprzez analizę biegunów i zer eksperymentalnie zmierzonych funkcji odpowiedzi częstotliwościowej, oraz zaprojektowanie, implementację i eksperymentalną walidację strategii ANC opartej na filtrze Kalmana, wraz z analizą złożoności obliczeniowej potwierdzającą jej wykonalność w czasie rzeczywistym (Artykuł A).
- Zaprojektowanie i przeprowadzenie subiektywnych eksperymentów słuchowych — obejmujących rekrutację uczestników, projekt protokołu eksperymentalnego zgodnego z ISO 15666 oraz zwalidowaną metodologię siatki afektywnej, przygotowanie skalibrowanych bodźców z wykorzystaniem symulatora głowy i tułowia oraz zastosowanie trójczynnikowej ANOVA i PERMANOVA z powtarzanymi pomiarami — przeprowadzonych we współpracy z University College London podczas stażu naukowego, oraz analizę wynikowych danych percepcyjnych dotyczących uciążliwości, jakości afektywnej i głośności (Artykuł B).
- Zaprojektowanie, implementację i eksperymentalną walidację hybrydowych algorytmów SFANC-NLMS oraz GFANC-THFxNLMS, w tym modułu selekcji filtrów i łączenia podfiltrów opartego na CNN, etapu adaptacyjnego THFxNLMS uwzględniającego nieliniowości z adaptacyjnym skalowaniem gradientu, oraz odpowiedniego pipeline'u treningowego i analizy wykonalności wdrożenia wbudowanego (Artykuł C).
- Zaprojektowanie lekkiej architektury FxMDNoiseNet, opartej na przyczynowych, punktowych i rozszerzonych (atrous) splotach, oraz jej integrację z głębokim uczeniem ze wzmocnieniem opartym na DDPG i TD3 oraz z uczeniem transferowym, a także zaprojektowanie i przeprowadzenie kampanii eksperymentalnej w rzeczywistym pojeździe — obejmującej pomiar odpowiedzi impulsowej w

warunkach LTI i LTV przy niedopasowaniu ścieżki wtórnej — wykorzystanej do walidacji uzyskanego frameworku FxMDNoiseNet-DTRL (Artykuły D i E).

- Współpromotorstwo dwóch studentów studiów magisterskich (Maxime Zaman i Reza Shams), którzy przyczynili się do powiązanych aspektów badań, oraz współpracę z Dr. Chao Liangiem, Dr. Zulfi Rachmanem i Dr. Khaledem Saidem Ahmedem Maamounem przy powiązanych publikacjach rozszerzających podstawowe metody rozprawy na pokrewne problemy sterowania wibroakustycznego.

W przypadku publikacji wieloautorskich, oświadczenie o indywidualnym procentowym i merytorycznym udziale wszystkich współautorów jest dołączone osobno do wniosku o nadanie stopnia doktora, zgodnie z wymogiem §2(4) Regulaminu nadawania stopnia doktora Politechniki Śląskiej. Zgodnie z tym samym przepisem, w przypadkach gdy uzyskanie indywidualnych oświadczeń od wszystkich współautorów nie było możliwe, oświadczenie zostało zamiast tego przedłożone przez autora korespondencyjnego (kandydata) lub odpowiedniego kierownika projektu/zespołu, stosownie do sytuacji.