AI-Assisted Design Optimisation of Permanent Magnet Synchronous Machine for E-Bike

by

Mohammed Abdeldjabar GUESMIA

THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT OF A MASTER'S DEGREE
WITH THESIS IN ELECTRICAL ENGINEERING
M.A.Sc.

MONTREAL, 2ND APRIL, 2026

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

M. Qingsong Wang, Thesis supervisor
Département de Génie électrique, École de Technologie Supérieure

Mrs. Gaixia Zhang, Chair, Board of Examiners
Département de Génie électrique, École de Technologie Supérieure

M. Andy Shih, Member of the Jury
Département de Génie électrique, École de Technologie Supérieure

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND THE PUBLIC

ON 27TH MARCH, 2026

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

FOREWORD

This thesis stems from a practical question frequently encountered in light electric mobility: how to design a compact traction motor that delivers high efficiency and a smooth user experience at the same time. In an e-bike mid-drive system, performance is not only defined by average torque and efficiency, but also by torque quality, since ripple can be felt by the rider and can contribute to noise and vibration. The work presented here was motivated by the need to explore these trade-offs using high-fidelity finite-element models while keeping the overall design process computationally tractable.

The central idea of this thesis is to make multi-objective electromagnetic design exploration more data-efficient and easier to steer. To that end, we develop an automated simulation–optimization workflow and investigate how a retrieval-grounded large language model can be used as a memory-based design assistant to propose candidate geometries and encode qualitative guidance alongside quantitative objectives. Rather than aiming to replace physics-based simulation, the approach is designed to bring human experience into the optimization loop; capturing engineering intuition, past decision rationales, or qualitative preferences in natural language, and using AI to translate that tacit knowledge into consistent, reusable guidance during iterative design.

This document is structured as follows. The introduction sets the stage by presenting the e-bike mid-drive application context, defining the design requirements and constraints, and stating the research objectives and contributions of this work. Then Chapter 1 surveys the relevant literature on PMSM topologies for traction applications, finite-element modeling practices, and multi-objective optimization methods including LLM assisted algorithms, motivating the choices adopted in this thesis. Chapter 2 introduces the application and design constraints, reviews the motor topology and FEM/optimization foundations, details the proposed LLM-guided Bayesian optimization workflow, and presents the resulting trade-offs and comparative analyses. Finally, Chapter 3 summarizes the main contributions and outlines directions for future work.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my research director for their guidance, support, and invaluable advice throughout the course of this thesis. Their expertise, encouragement, and constant availability greatly contributed to the development and completion of this work.

I would also like to thank the members of my jury, Gaixia Zhang and Andy Shih, for accepting to evaluate this thesis and for the time and attention they devoted to reading and assessing my work. Their insights and comments are greatly appreciated.

My thanks also go to GREPCI, the laboratory in which this research was conducted, for providing a stimulating academic environment and the resources necessary to carry out this project under the best possible conditions.

I am also deeply grateful to my family for their unwavering support, patience, and encouragement throughout this journey. Their confidence in me has been a constant source of strength and motivation.

Finally, I would like to extend my gratitude to all those who, in one way or another, contributed to the accomplishment of this thesis.

Optimisation Assistée par IA d'une Machine Synchrone à Aiment Permanent pour Vélo à Assistance Électrique

Mohammed Abdeldjabar GUESMIA

RÉSUMÉ

Cette thèse porte sur le développement d'une démarche d'optimisation multiobjectif, économe en nombre de simulations, pour la conception électromagnétique d'un moteur synchrone à aimants permanents (MSAP) destiné à une transmission mid-drive de vélo à assistance électrique (VAE). Dans ce contexte, le groupe motopropulseur doit concilier compacité, rendement élevé et qualité du couple, puisque l'ondulation de couple influence directement le confort de pédalage, le bruit/vibration et la contrôlabilité. L'évaluation précise de ces compromis repose généralement sur la méthode des éléments finis (MEF), mais l'optimisation (MEF dans la boucle) demeure coûteuse en temps de calcul et limite l'exploration du domaine de conception.

Pour répondre à cette contrainte, un environnement modulaire d'évaluation automatique est mis en place en couplant MATLAB et ANSYS Maxwell afin d'exécuter des campagnes de simulation reproductibles. Le moteur étudié est un MSAP intérieur 48 V, de configuration 48 encoches / 8 pôles, avec une topologie d'aimants enterrés en forme de Δ. La géométrie est paramétrée par des variables géométriques clés des encoches statoriques et par des variables décrivant les dimension des aimants. Le problème est formulé en trois objectifs : maximiser le couple moyen, maximiser le rendement et minimiser l'ondulation de couple.

L'approche proposée combine une optimisation bayésienne (OB) basée sur un processus gaussien et l'enrichit par un grand modèle de langage (GML) ancré par récupération, utilisé comme agent doté de mémoire. En s'appuyant sur une base interne de résultats MEF et des résumés de tendances, le GML propose des candidats et intègre des règles qualitatives en langage naturel pour orienter l'exploration sans réglages manuels répétés des pondérations. Des comparaisons avec une métaheuristique multiobjectif de référence montrent que l'OB guidée par GML atteint une région de Pareto comparable tout en diminuant le nombre d'évaluations MEF coûteuses. Les résultats confirment l'intérêt d'un guidage GML ancré par données pour améliorer l'efficacité et l'interprétabilité de l'optimisation MEF des MSAP dédiés à la mobilité légère.

Mots-clés: moteur synchrone à aimants permanents; méthode des éléments finis; optimisation multiobjectif; optimisation bayésienne; grand modèle de langage; génération augmentée par récupération; vélo à assistance électrique.

AI-Assisted Design Optimisation of Permanent Magnet Synchronous Machine for E-Bike

Mohammed Abdeldjabar GUESMIA

ABSTRACT

This thesis investigates a data-efficient multi-objective optimization workflow for the electromagnetic design of a permanent magnet synchronous motor (PMSM) intended for a mid-drive electric bicycle (e-bike) powertrain. In this application, the motor must meet competing requirements in compactness, efficiency, and torque quality, since torque ripple directly impacts pedaling smoothness, noise/vibration, and controllability. High-fidelity assessment of these trade-offs generally requires finite-element modeling (FEM), yet FEM-driven design exploration is computationally expensive and often limits the number of candidate geometries that can be evaluated.

To address this challenge, a modular simulation–optimization framework is developed to couple MATLAB with ANSYS Maxwell and automate parametric FEM evaluations. The considered motor is a 48 V interior PMSM featuring a 48-slot/8-pole configuration and a Δ-shaped buried-magnet topology. The design is parameterized by key stator-slot dimensions and magnet geometric variables, and the optimization is formulated with three objectives: maximize average torque, maximize efficiency, and minimize torque ripple.

The proposed approach combines Bayesian optimization (BO) with a Gaussian-process surrogate model and augments the BO loop with a retrieval-augmented generation (RAG) large language model (LLM) acting as a memory-based design agent. Using an internal database of prior FEM results and trend summaries, the LLM proposes candidate designs and incorporates qualitative natural-language rules to steer exploration without repeated manual retuning of objective weights. Comparative studies against a reference multi-objective metaheuristic demonstrate that the LLM-guided BO strategy reaches a competitive Pareto region while reducing the number of expensive FEM evaluations. The results support the use of RAG-LLM assistance to improve sample efficiency and interpretability in FEM-based multi-objective PMSM geometry optimization for light electric mobility.

Keywords: permanent magnet synchronous motor; finite-element modeling; multi-objective optimization; Bayesian optimization; large language model; retrieval-augmented generation; e-bike.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AI	Artificial intelligence
BLDC	Brushless DC (brushless direct-current) motor
BO	Bayesian optimization
CAD	Computer-aided design
CPU	Central processing unit
DC	Direct current
DOE	Design of experiments
EMF	Electromotive force
EMI	Electromagnetic interference
FE	Finite element
FEA	Finite element analysis
FEM	Finite element method / finite-element modeling
GP	Gaussian process
IPMSM	Interior permanent magnet synchronous machine
LHS	Latin hypercube sampling
LLM	Large language model
MO-AHA	Multi-objective Artificial Hummingbird Algorithm
MOGWO	Multi-objective Grey Wolf Optimizer
MOPSO	Multi-objective Particle Swarm Optimization

NSGA-II	Non-dominated Sorting Genetic Algorithm II
NVH	Noise, vibration, and harshness
PM	Permanent magnet
PMSM	Permanent magnet synchronous machine
PSO	Particle swarm optimization
RAG	Retrieval-augmented generation
RAM	Random-access memory

LIST OF SYMBOLS AND UNITS OF MEASUREMENTS

Bs_0	Stator-slot geometric parameter (mm)
D_n	Dataset of n observations $\{(x_i, r_i)\}_{i=1}^{n}$
Hs_0	Stator-slot geometric parameter (mm)
Hs_1	Stator-slot geometric parameter (mm)
HV_t	Hypervolume indicator at iteration t
Mt	Magnet thickness parameter (mm)
Mt_V	V-shaped magnet thickness parameter (mm)
Mw	Magnet width parameter (mm)
Mw_V	V-shaped magnet width parameter (mm)
P	Current Pareto set
$P_{\text{core}}(t)$	Core (iron) losses (W)
$P_{\text{cu}}(t)$	Copper (Joule) losses (W)
$P_{\text{mech}}(t)$	Mechanical output power (W)
R_{phase}	Phase resistance (Ω)
$T(t)$	Electromagnetic torque (N·m)
$TR(\mathbf{x})$	Torque ripple objective (%)
$T_{\text{avg}}(\mathbf{x})$	Average torque objective (N·m)
X	Feasible design space defined by bounds/constraints
$i_A(t)$, $i_B(t)$, $i_C(t)$	Phase currents (A)

$p(\cdot)$	Probability (density) function
r_i	Response vector at sample/design x_i
t	Time (s)
x_i	Sampled design point (input)
Θ	Parameter space of θ
$\alpha_1,\ \alpha_2$	Weighting coefficients in the combined predictor
$\bar{\eta}$	Average efficiency over one electrical period (%)
ϵ	Convergence tolerance (stopping threshold)
$\eta(t)$	Instantaneous efficiency (%)
$\hat{r}_k$	Predicted response at candidate design $\tilde{x}_k$
$\mathbf{f}(\mathbf{x})$	Objective vector $\left[T_{\text{avg}}(\mathbf{x}),\ \eta(\mathbf{x}),\ TR(\mathbf{x})\right]^T$
$\mathbf{x}$	Design (decision) vector
$\mathbf{x}_{\min},\ \mathbf{x}_{\max}$	Lower/upper bounds of the design variables
ρ	Sliding-window length in iterations (used for convergence check)
θ	Model parameters (e.g., surrogate hyperparameters)
τ_{e}	Electrical period (s)
$\omega(t)$	Mechanical angular speed (rad/s)
GB	Giga byte
GHz	Giga hertz
$\text{N}\cdot\text{m}$	Newton-metre (torque)

V	Volt
min	Minute
mm^2	Square millimetre
rpm	Revolutions per minute
s	Second

INTRODUCTION

0.1 Background and motivation

The transition toward more energy-efficient and environmentally sustainable transportation is accelerating the development of light electric mobility solutions. Electrical-assisted bicycles (e-bikes) in particular experienced significant growth in many societies around the globe. In this context, the powertrain must satisfy a demanding compromise: compactness, high efficiency, controlled cost, reliability, and torque quality. Mid-drive architectures (motor at the crank) are distinctly attractive because they provide the ability to leverage a built-in mechanical transmission (gearbox) to adapt a high-speed electric machine to typical pedaling cadence seamlessly and compactly, while maintaining good performance across a wide range of loads. Among candidate technologies, the permanent magnet synchronous machine (PMSM) is often preferred for light mobility applications due to its high power density, efficiency. However, optimizing a PMSM for an e-bike is not only about maximizing average torque: torque quality (notably torque ripple) and efficiency directly affect pedaling comfort, noise/vibration, controllability, range, and thermal behavior.

0.2 Problem statement

The electromagnetic performance of a PMSM depends sensitively on geometric design choices (stator slots, tooth opening, magnet geometry, volume and placement, etc.). These parameters are strongly coupled and lead to multi-objective trade-offs: a change that improves average torque may worsen torque ripple, or vice versa, while efficiency is influenced by copper and iron losses. In advanced design practice, accurate assessment of these trade-offs often relies on finite element analysis (FEA/FEM), since analytical models can struggle to capture certain effects (nonlinearities, saturation and geometric details). However, FEM is computationally expensive: each evaluation requires substantial runtime and a well-configured simulation

workflow. This motivates a clear need: a multi-objective optimization workflow capable of reaching Pareto-optimal solutions with a reduced number of FEM evaluations, while remaining flexible enough to incorporate design preferences (e.g., explicitly favoring smoother torque) without repeatedly hand-tuning objective weights or hard constraints.

0.3 Hypothesis and research questions

This thesis investigates the hypothesis that the efficiency of FEM-driven PMSM optimization can be improved by combining both:

1. A Bayesian Optimization (BO) framework guided by a surrogate model (Gaussian process).
2. A large language model (LLM) used as a "memory-enabled" agent that can exploit the history of simulations (through a form of retrieval-augmented generation centered on an internal database of results), and importantly, incorporate qualitative rules expressed in natural language.

The research questions are therefore:

- To what extent can multi-objective Bayesian optimization coupled with an FEM loop efficiently identify a meaningful Pareto front for an e-bike-assist PMSM?
- Does adding an LLM + retrieval agent (based on FEM outputs and trend logs) help reduce the number of FEM evaluations required to reach a Pareto region comparable to a reference multi-objective metaheuristic?

0.4 Thesis objectives

The overall objective is to propose and evaluate an AI-assisted, multi-objective optimization approach for the electromagnetic design of a PMSM intended for a mid-drive e-bike application. More specifically, this thesis aims to:

1. Define and parameterize a PMSM (stator/rotor) architecture suitable for a geared mid-drive powertrain, and formulate the optimization problem with three objectives: maximize average torque, maximize efficiency, and minimize torque ripple.
2. Develop an automated simulation–optimization framework coupling a computation environment (e.g., MATLAB) with an FEM solver (e.g., ANSYS Maxwell) to run reproducible optimization campaigns.
3. Integrate a Bayesian optimization strategy and augment it with an LLM + retrieval module enabling (i) candidate proposal, (ii) trend extraction from the optimization history, and (iii) incorporation of qualitative design rules expressed in natural language.
4. Compare performance of: (a) retrieval-assisted BO without rules, (b) BO assisted by an LLM with qualitative rules, and (c) a reference multi-objective metaheuristic.

0.5 Proposed approach and scope

The work focuses on parametric geometry optimization of a PMSM (key stator/rotor dimensions and topology), evaluated via FEM simulations. The optimization is treated as a gradient-free "black-box" problem: each design point triggers an FEM run followed by post-processing to compute average torque, torque ripple, and efficiency. The LLM is not intended to replace FEM, but to better steer exploration by summarizing and reusing historical trends, proposing candidates consistent with prior outcomes, and injecting qualitative preferences through rules. The scope is deliberately centered on: (i) the quality of multi-objective trade-offs and (ii) efficiency in terms of FEM evaluation count. The validation and evaluation is conducted through FEM and optimizer-to-optimizer comparison within a coherent and reproducible framework.

0.6 Contributions

The main contributions of this thesis can be summarized as follows:

- A modular simulation–optimization coupling framework enabling automated interaction between an optimizer and a parametric FEM model.
- An LLM + retrieval integration exploiting an internal simulation history (trend logs) to improve exploration efficiency, including the injection of qualitative design rules in natural language.
- A comparative evaluation demonstrating the competitiveness of the proposed approach.

0.7 Thesis outline

This thesis is organized as follows. The Introduction presents the research context, motivation, problem statement, objectives, scope, and main contributions. Chapter 1 provides the literature review, covering PMSMs for light electric mobility, torque ripple and key design trade-offs, FEM-based multi-objective optimization, Bayesian optimization, and recent approaches that integrate large language models (LLMs) into Bayesian optimization and related optimization workflows. Chapter 2 corresponds to the peer-reviewed article produced during this research, which reports the methodology, setup, and results that form the core of this work. Finally, Chapter 3 offers a general conclusion, summarizes the main findings, and presents recommendations and directions for future work.

CHAPTER 1

LITERATURE REVIEW

1.1 E-Bike Drives and Motor Systems

Electric-assisted bicycles impose a distinctive set of constraints on the traction motor and its design workflow: tight packaging, stringent efficiency targets (range and thermal limits), and torque quality requirements that directly affect perceived pedaling smoothness and acoustic comfort. Within this context, permanent magnet synchronous machines (PMSMs) and BLDC-type machines are frequently selected due to power density and efficiency, but their suitability depends strongly on how electromagnetic design decisions translate into ripple, noise/vibration, and controllability.

A significant portion of the accessible engineering literature emphasizes hub-motor development and verification workflows. For instance, hub-motor studies commonly follow an iterative loop from preliminary sizing and equivalent-circuit/analytical tools to finite-element validation, highlighting how electrical choices (voltage/current, winding selection, slot fill) quickly become constraints on losses and torque ripple Chawrasia, Das & Chanda (2021). Such works are valuable as they illustrate the practical coupling between winding/slot geometry and performance metrics. However, hub-focused results are not automatically transferable to mid-drive architectures because the drivetrain introduces mechanical filtering, additional NVH paths, and a different torque–speed operating envelope at the motor shaft.

Beyond conventional BLDC/PMSM choices, alternative topologies have been explored for e-bike propulsion. An outer-rotor PM-assisted synchronous reluctance motor (PMaSynRM) was proposed as a substitute for a commercial BLDC unit and was experimentally validated, reporting reduced oscillatory behavior and cogging torque relative to the reference motor while acknowledging operating-point-dependent trade-offs Nasiri-Zarandi, Karami-Shahnani, Sedigh Toulabi & Tessarolo (2023). The key takeaway for mid-drive design is not that one topology universally dominates, but that torque ripple, magnets cost, and manufacturability are

topology-level trade-offs that should be handled explicitly in an optimization framework rather than tuned in an ad hoc manner.

At the system level, pedal-assist quality depends not only on the machine but also on sensing and control. A study on torque measurement and adaptive control under varying external loads (e.g., slope and rider mass) shows that rider-perceived "synergy" is sensitive to how assistance torque is regulated under disturbances Ho *et al.* (2023b). Controller-oriented modeling of BLDC drives likewise provides implementation insights into commutation and closed-loop speed behavior relevant to e-bike motor drives Purnata, Yusuf & Riyanto (2020). These contributions motivate an important design implication: improving electromagnetic torque quality (e.g., reducing ripple and cogging) can simplify control burdens and reduce drivetrain excitation, but it must be pursued while respecting efficiency and torque density constraints.

Overall, existing e-bike motor references provide (i) practical workflows for machine verification and (ii) evidence that torque quality is central to user experience Chawrasia *et al.* (2021); Ho *et al.* (2023b). However, much of the design literature remains hub-centric or control-centric, with limited emphasis on mid-drive PMSM geometry optimization that treats average torque, efficiency, and torque ripple as simultaneous objectives. This motivates a literature-based need for a design approach that is (a) explicitly multi-objective and (b) directly tailored to mid-drive PMSM geometries, while remaining computationally feasible when high-fidelity simulation is required.

1.2 Finite-Element Modeling as a Standard Tool in Electric Machine Design

Finite-element modeling (FEM) is widely adopted in modern electric machine development because it provides a physics-grounded mapping from geometry and material choices to electromagnetic fields, torque waveforms, and loss calculations. In practice, FEM is commonly used as the high-fidelity layer that complements fast analytical estimates, particularly when saturation, slotting permeance harmonics, and spatially resolved losses materially affect performance.

A representative example is the combined FE and Particle Swarm optimization study of a high-speed PM synchronous machine, where FEM is used specifically because analytical approaches, while useful for rapid estimates, can be insufficient for accurately resolving loss components and field distributions in detail Belahcen, Martin, Zaim, Dlala & Kolondzovski (2015). This line of work also illustrates a recurring engineering compromise: designers often rely on 2D FEM to control computational cost, while still requiring time-stepping or otherwise dynamic formulations for torque and loss fidelity Belahcen *et al.* (2015).

FEM also functions as a controlled "virtual experiment" platform when experimental setups are costly or difficult to instrument. For instance, local demagnetization fault modeling in PM-assisted synchronous reluctance machines uses full-machine finite-element analysis to generate interpretable fault signatures from back-EMF and flux indicators, despite the associated computational burden Creux, Haje Obeid, Boileau & Meibody-Tabar (2022). Similarly, synchronous machine parameter identification can be performed via FEM-generated frequency-response data coupled with parameter optimization using algorithms like genetic algorithm, reducing dependence on extensive test campaigns but also revealing practical identifiability and modeling-limit issues Escarela-Perez, Niewierowicz & Campero-Littlewood (2001).

Table 1.1 E-bike traction motor solutions and workflows in the literature

Ref.	**Focus**	**Workflow and validation**	**Engineering takeaway**
Chawrasia *et al.* (2021)	Hub motor design workflow and verification using Motor-CAD	Tool-driven design + validation loop	Illustrates an industrial-style workflow where rapid sizing is combined with higher-fidelity validation, motivating data-efficient optimization when simulations are expensive.
Nasiri-Zarandi *et al.* (2023)	Outer-rotor PM-assisted SynRM for e-bike propulsion	Experimental assessment	Demonstrates a non-BLDC alternative and highlights torque-quality and performance trade-offs that must be handled as coupled objectives.
Dash *et al.* (2023)	System-level e-bike design and analysis	Design/analysis study	Provides broader context (requirements and constraints), motivates why mid-drive operating envelopes and constraints should be made explicit.

Table **??** summarizes representative e-bike traction motor solutions and design workflows, highlighting the topologies and validation approaches reported in the literature. It shows that many studies emphasize hub-drive architectures or general workflows, while the explicit treatment of mid-drive constraints and torque-quality trade-offs is less developed. The literature supports two other points that are essential for this thesis. First, FEM is often the most defensible approach when geometry-driven effects (slotting, saturation, localized losses) govern torque ripple and efficiency Belahcen *et al.* (2015). Second, FEM's key limitation is evaluation cost/time: once FEM becomes the objective-function evaluator inside an optimizer, sample efficiency becomes a primary constraint. Therefore, any credible multi-objective geometry optimization for PMSMs must address not only model fidelity but also how to reduce the number of FEM evaluations needed to reach a useful Pareto front.

1.3 Geometry-Driven Torque, Ripple, and Efficiency Trade-Offs in PM Brushless Machines

Torque production, torque ripple, and efficiency in PM machines are strongly shaped by rotor magnet geometry and stator slot/tooth design. A foundational analytical treatment is provided by the IEEE Transactions series by Zhu and Howe on instantaneous field distributions in brushless PM DC machines, which explicitly shows how stator slotting imposes spatially varying permeance that injects harmonics into the air-gap field and produces ripple-bearing torque components Zhu, Howe, Bolte & Ackermann (1993); Zhu & Howe (1993a,c,b). This work establishes a clear causal chain: *geometry* $\rightarrow$ *air-gap harmonics* $\rightarrow$ *torque ripple/cogging*, and it motivates why simplified lumped corrections can be insufficient when ripple is a primary design concern.

From an optimization and design perspective, multiple studies show that these electromagnetic metrics must be treated as coupled objectives rather than tuned sequentially. A benchmark study on permanent magnet synchronous machine characterization and optimization demonstrates that preferred slot/tooth proportions and copper section depend on operating conditions and loss trade-offs, illustrating that "optimal geometry" is load-dependent rather than universal Sergeant,

Crevecoeur, Dupré & Van den Bossche (2009). A broader review of PMSM optimization methods similarly frames electric machine design as a nonlinear, constrained multi-objective problem in which Pareto fronts naturally arise due to competing objectives and constraints Duan & Ionel (2013). More recent multi-objective studies reinforce that magnet placement and stator geometry influence average torque and ripple simultaneously, and that improvements can be achieved through careful geometry refinement rather than isolated parameter tuning Farshbaf Roomi, Vahedi & Mirnikjoo (2021); Zheng *et al.* (2022); Mao, Niu & Wang (2021).

Table 1.2 FEM integration with optimization in electric machine design

Ref.	**Scope**	**Optimization method**	**Coupling style**	**Optimization intent**
Belahcen *et al.* (2015)	High-speed PMSM	PSO	FEM-in-the-loop	Direct black-box optimization where each candidate is evaluated by FEM; highlights the evaluation-cost bottleneck.
Farshbaf Roomi *et al.* (2021)	V-shape IPMSM	Sensitivity analysis + LHS (DOE) + NSGA-II	DOE-assisted multi-objective FEM workflow	Cost-aware search: variable screening + efficient sampling + multi-objective evolutionary optimization for competing objectives.
Zhang *et al.* (2024)	IPMSM	(IM)MOAHA; compared with MOGWO, MOPSO, NSGA-II	Surrogate-assisted optimization with FEM validation	Multi-objective formulation including torque improvement and ripple reduction; provides comparative evidence across several optimizers.
Duan & Ionel (2013)	General machine design (review)	Survey of surrogate models + deterministic/stochastic search	Method taxonomy	Explains optimizer/surrogate choices under expensive evaluations; motivates BO-style data efficiency for FEM-driven design.

Ripple mitigation strategies (e.g., skewing and back-EMF shaping) further illustrate the practical trade-off structure. For example, cogging-torque reduction via skewing can be effective but is typically bounded because it can introduce other penalties (e.g., reduced torque density or altered ripple spectrum), reinforcing that ripple reduction is rarely "free" in a constrained design Hari Krishnan *et al.* (2024). Complementary BLDC motor reviews emphasize that torque ripple, acoustic noise, and EMI are persistent concerns and that stator slot/tooth parameters and rotor PM structures are recurring design levers Mohanraj *et al.* (2022). In sum, the literature converges

on the idea that ripple, average torque, and efficiency should be jointly optimized, and that FEM-based evaluation is often necessary once slotting and saturation dominate.

Table 1.2 synthesizes how finite-element modeling (FEM) is coupled with optimization across electric machine studies, distinguishing direct FEM-in-the-loop search from DOE/surrogate-assisted workflows. The comparison makes it clear that while FEM is often necessary to capture slotting, saturation, and ripple accurately, its time consuming quickly becomes the limiting factor in multi-objective exploration. This supports the need for data-efficient optimization strategies under a constrained simulation budget. The above objectives trade-offs converge to the same conclusion "each high-fidelity evaluation is expensive." Belahcen *et al.* (2015); Duan & Ionel (2013). As a result, the practical bottleneck in applying multi-objective geometry optimization is not defining objectives, but reaching a competitive Pareto region with a limited FEM budget. This creates a direct need for optimization strategies designed for expensive black-box evaluations.

1.4 Language-Guided Bayesian Optimization for Data-Efficient Electric Motor Optimization

FEM-driven motor design is naturally a constrained black-box optimization problem: objectives such as average torque, torque ripple, and efficiency are computed from simulation outputs, and gradients are often unavailable or unreliable when the workflow includes meshing, nonlinear materials, and time-stepping. Classical metaheuristics and AI-based optimizers have been applied to motor design, demonstrating performance gains but often requiring substantial evaluation counts that can be costly in FEM settings Mopari, Hirekhan & Murthy (2024). Bayesian optimization (BO) is particularly attractive in this regime because it is explicitly designed to achieve good solutions under expensive evaluations by using probabilistic surrogate models to guide sampling.

Recent work has investigated how large language models (LLMs) can improve BO by injecting structured priors, enabling natural-language interaction, and guiding exploration based on contextual “rules” or preferences. LLM-augmented BO has been proposed as a mechanism for

improving BO performance by providing informative suggestions, warm starts, or priors that reflect latent structure in the problem Liu, Astorga, Seedat & van der Schaar (2024). Surveys and empirical studies caution that LLM utility depends on careful grounding and evaluation, but they also support the general premise that language can encode useful domain constraints and preferences in optimization loops Gupta, Hartford & Liu (2025). Several methodological directions are directly relevant to engineering design: (i) integrating language-guided priors into BO Topalis, Schieseck & Gehlhoff (2025), (ii) incorporating interactive natural-language feedback during optimization Kobalczyk *et al.* (2025), (iii) enforcing trustworthiness and reliability in LLM-in-the-loop BO Chang, Azvar, Okwudire & Al Kontar (2025), and (iv) leveraging long-context reasoning or agentic assistance to generate candidates and structure optimization workflows Yang *et al.* (2025b); Sayeed, Soneji, Baird & Sparks (2025); Zhang, Choong, Madhawa & Ozawa (2025).

Table 1.3 categorizes these recent LLM-enhanced Bayesian optimization approaches by where language enters the optimization loop. This organization clarifies which LLM roles are most relevant when evaluations are expensive, as in FEM-based motor design. It also provides a concrete basis for selecting an LLM mechanism that improves sample efficiency without sacrificing interpretability or reliability.

For electric motor geometry optimization, these developments suggest a concrete opportunity: LLMs can be used not as replacements for FEM, but as interfaces for encoding qualitative expert knowledge (e.g., “wider slot opening tends to reduce ripple”) and as memory-based agents that summarize the optimization history and propose promising regions to sample next. If grounded in prior simulation results, language guidance can help reduce wasted evaluations in dominated regions while preserving interpretability of design trends.

Despite rapid progress in LLM-augmented BO, the motor-design literature still lacks well-scoped demonstrations that simultaneously satisfy all of the following requirements:

- **Mid-drive e-bike specificity:** a PMSM topology and variable set aligned with a geared mid-drive operating envelope.

Table 1.3 LLM-enhanced Bayesian optimization: where language enters the loop and why it matters for expensive design evaluations

Ref.	**LLM role in BO**	**Signal type**	**Why it matters for FEM-expensive engineering design**
Liu *et al.* (2024)	Language-mediated BO guidance / improved proposing	Optimization history (textualized)	Targets cold-start efficiency: better early proposals can dominate outcomes when only a small number of FEM evaluations are feasible.
Chang *et al.* (2025)	Trustworthy hybrid: LLM guidance with calibrated BO core	Context + surrogate uncertainty	Keeps uncertainty-aware BO behavior while still injecting structured guidance, aligned with engineering needs for reliability.
Topalis *et al.* (2025)	Language-informed priors / constraints for BO	Text priors → BO priors	Provides a route to encode qualitative design rules without manually crafting kernels/constraints for every new geometry problem.
Kobalczyk *et al.* (2025)	Interactive feedback: language → utility shaping	Human feedback → scalar utility	Matches real design workflows where preferences are iterative (e.g., "too noisy", "prefer smoother torque") and can guide search efficiently.
Zhang *et al.* (2025)	LLM-assisted BO pipeline logic in scientific optimization	Domain text + outcomes	Demonstrates feasibility of integrating LLM guidance within iterative optimization loops similar in structure to repeated simulation/validation cycles.

- **True multi-objective formulation with torque ripple:** explicit joint optimization of average torque and efficiency while minimizing torque ripple, evaluated via FEM rather than low-order surrogates alone.
- **Data efficiency under FEM cost:** evidence that the optimization reaches a competitive Pareto region using fewer FEM evaluations than a strong baseline optimizer.
- **Interpretable incorporation of qualitative preferences:** a mechanism to inject human design rules in natural language without repeatedly re-tuning objective weights or hard constraints.

The accompanying article (*AI-Assisted Bayesian Optimization of a Permanent Magnet Synchronous Motor for E-Bike Applications*) is positioned to address these gaps by combining (i) an FEM-based evaluation loop, (ii) multi-objective Bayesian optimization, and (iii) an LLM-assisted, retrieval-grounded mechanism that can incorporate qualitative rules and reuse optimization history to improve sample efficiency Guesmia *et al.* (2026). This literature review therefore motivates

the core methodological choice of the thesis: a data-efficient, interpretable multi-objective optimization framework for FEM-based mid-drive PMSM geometry design.

CHAPTER 2

AI-ASSISTED BAYESIAN OPTIMIZATION OF A PERMANENT MAGNET SYNCHRONOUS MOTOR FOR E-BIKE APPLICATIONS

Mohammed Abdeldjabar Guesmia[1] , Chuan Pham[1] , Ya-Jun Pan[2] , Kim Khoa Nguyen[1] , Kamal Al-Haddad[1] , Qingsong Wang[1]

[1] Département de génie électrique, École de technologie supérieure (ÉTS), Montréal, QC, H3C 1K3, Canada
[2] Department of Mechanical Engineering, Dalhousie University, Halifax, NS, B3H 4R2, Canada

Article published in the journal "Machines" in February 2026.

Abstract

This paper presents an artificial intelligence (AI)-assisted multi-objective topology optimization of a 48 V interior permanent magnet synchronous motor (PMSM) intended for mid-drive e-bike applications. The machine features a 48-slot, 8-pole stator–rotor combination with Δ-shaped three buried magnets per pole, and is coupled to a multi-stage gearbox that adapts its high-speed, low-torque output to a human-scale crank speed. The design problem simultaneously maximizes average torque and efficiency while minimizing torque ripple by varying key stator slot dimensions and magnet geometries. A modular MATLAB–ANSYS Maxwell framework is developed in which finite element simulations are driven by a Bayesian optimization (BO) loop augmented by a large language model (LLM) with retrieval-augmented generation (RAG). The LLM acts as a memory-based agent that proposes candidates, shapes Gaussian Process priors, and incorporates natural language rules expressing qualitative design knowledge. Two AI-assisted trials are compared against a multi-objective Artificial Hummingbird Algorithm benchmark, RAG + BO with and without natural language input. All three methods converge to a similar Pareto region with average torque around 5.4–5.7 Nm, torque ripple of approximately 12.8–14.2%, and efficiency near 93.3–93.6%, suitable for geared e-bike drives. The LLM-guided trial achieves this performance with a 20.1% reduction in simulation expenses relative to the BO baseline and by about 48% compared to the Artificial Hummingbird Algorithm. The results demonstrate that integrating LLM guidance into Bayesian optimization improves sample

efficiency while providing interpretable design trends for PMSM topologies tailored for light electric vehicles.

Keywords: bayesian optimization; finite element modeling; large language model; multi-objective optimization; permanent magnet synchronous machine; retrieval-augmented generation.

2.1 Introduction

The increasing demand for energy-efficient and environmentally friendly transportation has accelerated research into advanced electric propulsion systems, with permanent magnet synchronous motors emerging as a key technology in light electric vehicles such as e-bikes and scooters Chawrasia, Das, Chanda & Banerjee (2020); Dash *et al.* (2023); Chawrasia *et al.* (2021). PMSM motors offer superior efficiency, compactness, high torque-to-weight ratio, and reduced maintenance compared to traditional brushed machines and synchronous reluctance machines Duan & Ionel (2013); Nasiri-Zarandi *et al.* (2023); Wu, Huang, Zou, Ren & Gerada (2024); Ho *et al.* (2023a); Ho *et al.* (2023b); Mohanraj *et al.* (2022), making them ideal for sustainable mobility applications. Among the inherited drawbacks of a PMSM is the torque ripple elevated by the winding type, slot/pole combination, and design choices Hari Krishnan *et al.* (2024); Duan & Ionel (2013); Chen, Meng, Zhang, Li & Qu (2025); Nasiri-Zarandi *et al.* (2023); Wu *et al.* (2024). Accurate prediction of motor performance parameters such as back-EMF, torque constant, and inductance often relies on finite element method (FEM) analysis, which has become a standard tool for electromagnetic field evaluation in permanent magnet machines Chen *et al.* (2025); Zhao, Liang, Li & Qu (2025).

Given the wide range of strategies in which an electrical propulsion system designing process is approached, optimization frameworks and algorithms have been wildly used in electric machine design, parameter identification Knebl, Bárta, Bramerdorfer & Vítek (2021); Parouha & Verma (2021); Yang, Liu, Niu, Lyu & Chau (2025a); Escarela-Perez *et al.* (2001); Belahcen *et al.* (2015), topology Duan & Ionel (2013); Zheng *et al.* (2022), and control optimization. Multi-objective

formulations of FEM-based topology optimizations have proven essential in capturing trade-offs Zhao *et al.* (2022); Knebl *et al.* (2021); Zheng *et al.* (2022), maximizing average torque and efficiency while minimizing torque ripple. FEM evaluation provides accurate insights into flux distributions, torque ripple, and losses, yet each evaluation is time consuming compared to analytical models, which are less accurate in return. This significantly reduces the number of FEM evaluations needed for the optimization algorithms to converge Duan & Ionel (2013); Zheng *et al.* (2022); Mopari *et al.* (2024).

In recent years, artificial intelligence (AI) has emerged as a transformative enabler in electric motor optimization. AI-assisted methods combine surrogate models or machine learning with evolutionary or Bayesian optimization to accelerate FEM-based design optimization Mopari *et al.* (2024); Liu *et al.* (2024). Unlike purely heuristic algorithms, AI-driven frameworks can learn correlations from prior simulations to predict performance trends and guide exploration toward promising design regions Mopari *et al.* (2024); Liu *et al.* (2024). This reduces computational expenses and enables the discovery of novel topologies that may not arise through conventional search processes.

More recently, large language models (LLMs) have been increasingly investigated as decision-making components for experimental design and Bayesian optimization, providing a natural language interface for encoding expert knowledge, guiding exploration policies, and improving sample efficiency. Recent work has evaluated the practical effectiveness and limitations of LLM-driven experimental design in scientific domains Gupta *et al.* (2025). In parallel, LLM-guided Bayesian optimization frameworks have been proposed for complex multi-objective optimization Zhang *et al.* (2025), and reasoning-enhanced Bayesian optimization strategies have been introduced to better exploit long-context inference and iterative hypothesis refinement Yang *et al.* (2025b). Additional efforts have focused on improving the reliability and safety of LLM-in-the-loop optimization through trustworthy BO formulations Chang *et al.* (2025), incorporating interactive natural language feedback from users within the optimization loop Kobalczyk *et al.* (2025), and integrating language-guided priors in Bayesian optimization to formalize domain preferences and constraints Topalis *et al.* (2025). Moreover, agentic

retrieval-augmented assistants have recently been introduced to translate natural language optimization goals into executable Bayesian optimization pipelines and reusable code templates Sayeed *et al.* (2025).

This work presents a topology optimization study of a 48 V PMSM motor tailored for e-bike applications. Both rotor and stator geometries are optimized using an AI-assisted Bayesian optimization framework coupled with FEM analysis. The AI module identifies performance trends across the design space and generates refined candidate solutions for Bayesian optimization using retrieval-augmented generation. It also benefits from user input in natural language format using large language models to change user input into optimization policies, which enables efficient searching process. The optimization goal is to simultaneously maximize average torque and efficiency while minimizing torque ripple, yielding a Pareto set of optimal operating points.

To implement this optimization, a framework is created to couple ANSYS Maxwell (2022R2) to MATLAB (2023b) Guzman *et al.* (2023); this allows the interchangeability of optimization algorithms applied to the FEM model and still support any parametric FEM model provided. To assess the effectiveness of the proposed AI-assisted Bayesian optimization, its performance is evaluated against a multi-objective version of the Artificial Hummingbird Algorithm (MOAHA), selected as a benchmark due to its demonstrated suitability for electric machine topology optimization Zhang *et al.* (2024); Zhao *et al.* (2022). In the recent literature, MOAHA (and its multi-objective variants) has been successfully applied to the optimization of a PMSM topology closely related to the one investigated in this work Zhang *et al.* (2024), and it was systematically compared to established metaheuristics such as PSO and NSGA-II, showing strong convergence and competitive Pareto performance specifically for this class of machines Zhang *et al.* (2024). Therefore, adopting MOAHA as the reference optimizer provides a meaningful and literature-supported baseline for benchmarking the efficiency and solution quality of the proposed approach.

This investigates whether large language model (LLM) guidance can improve the sample efficiency of multi-objective topology optimization for PMSMs. Specifically, how effectively an

AI-assisted Bayesian optimization framework coupled with FEM can optimize rotor and stator geometric parameters, and whether retrieval-augmented LLM guidance reduces the number of FEM evaluations required to reach Pareto-optimal solutions compared to a benchmark optimizer previously applied to a closely related problem.

The integration of AI with multi-objective optimization contributes a scalable and intelligent framework for electric motor design, supporting the development of high-performance, energy-efficient electric propulsion systems.

2.2 Methodology

2.2.1 Motor Configuration and Design Variables

The e-bike assistance powertrain is a mid-drive system in which the electric motor transmits its mechanical output to the rider's pedal interface through a built-in gearbox, which serves as a critical mechanism for torque adaptation and rotational speed conditioning (Figure 2.1).

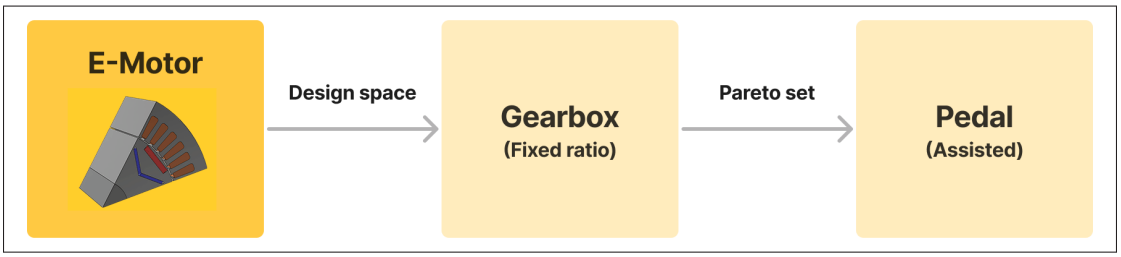
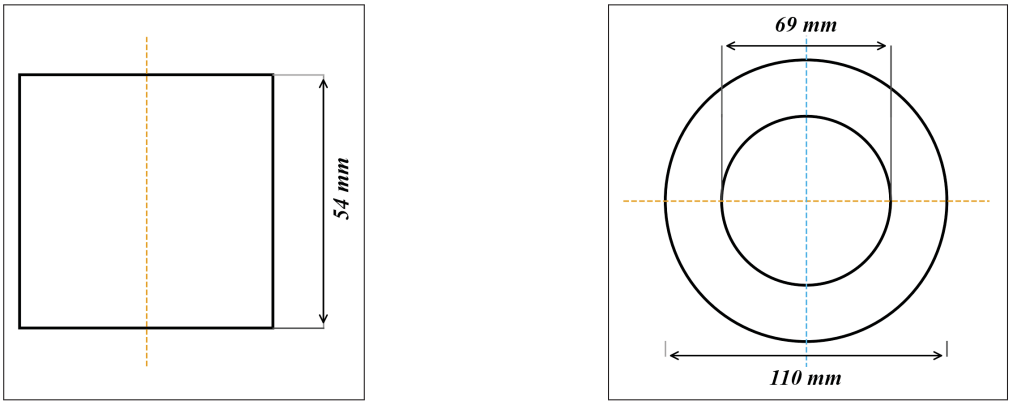


Figure 2.1 E-bike assistance powertrain

The electric motor is a permanent magnet synchronous machine that operates at high rotational velocities and low torque relative to the user pedaling cadence and the torque required for assistance. Therefore, motor characteristics that are optimal for electromagnetic efficiency are incompatible with the mechanical demands of human-scale crank motion. To reconcile this mismatch, the motor shaft output is routed into the multi-stage gearbox, which performs fixed torque amplification while reducing rotational speed to approximately 100 rpm, suitable for

human pedaling frequency. This arrangement enables the drivetrain to maintain efficient motor operation across variable load conditions, while also ensuring smooth torque delivery to the crankset and pedal spindle.

The machine main dimensions are illustrated in Figure 2.2. The motor has 48 slots, 8 poles, and a Δ-type interior magnet topology, where each pole contains three buried permanent magnet segments arranged in a Δ-shaped pattern (Figure 2.3). This approach is widely used to enhance flux concentration and provide mechanical robustness. The 48-slot stator accommodates a distributed three-phase winding to smooth electromagnetic torque and reduce cogging effects.



a) Axial length b) Diameters

Figure 2.2 Electric motor main dimensions

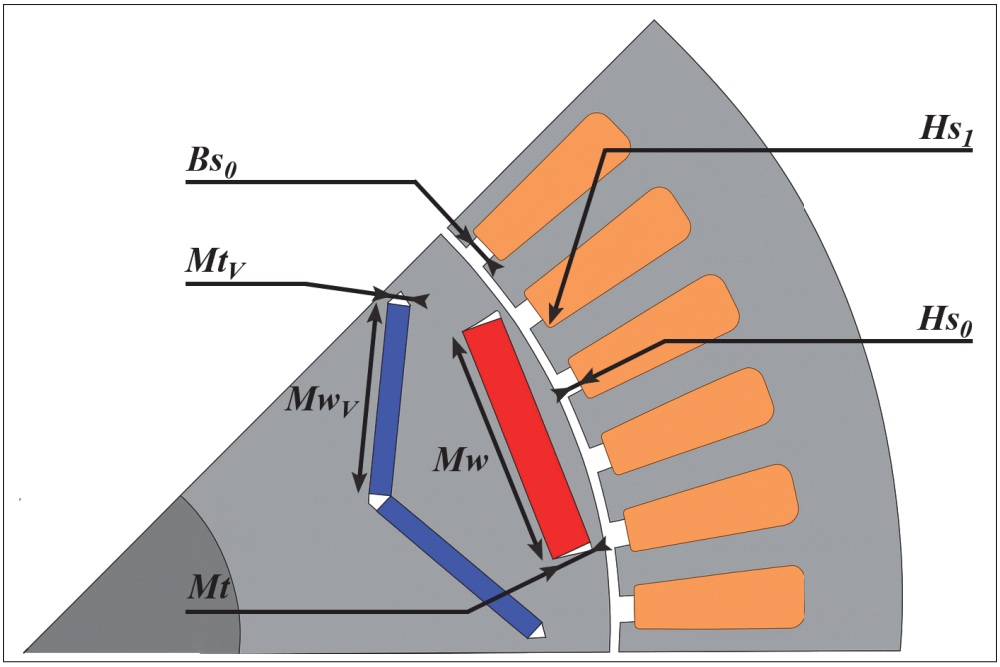
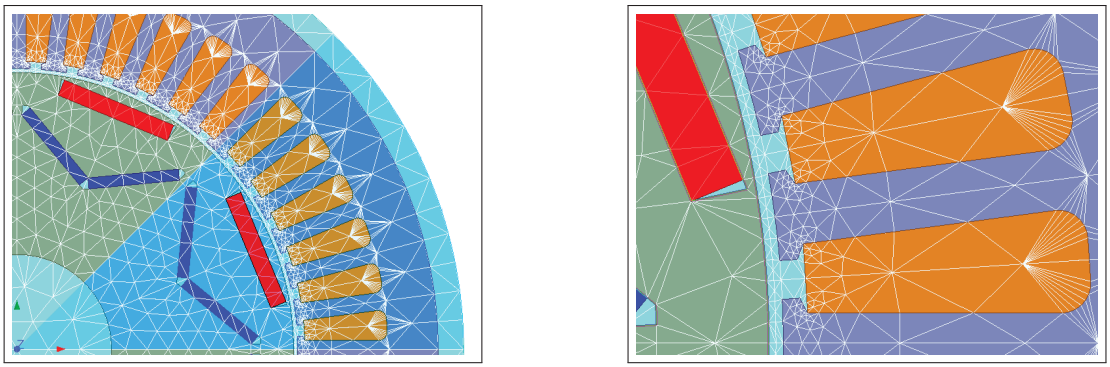


Figure 2.3 FEM model and parameters

2.2.2 Finite Element Model

Finite element model: The electromagnetic performance of the machine was evaluated using a 2D finite element model implemented in *Ansys Maxwell* 2022 R2. The model represents a periodic sector of 45° corresponding to one pole (Figure 2.4). A transient formulation was adopted to compute the motor torque and losses under current excitation, and symmetry was exploited to reduce the simulation time.



a) Global view b) Zoomed-in view

Figure 2.4 Mesh images

Material models and losses: Nonlinear magnetic behavior was considered using B–H curves for M19_24G electrical steel (from the built-in *Ansys Maxwell* materials library), while magnets were modeled with NdFe35.

The machine efficiency $\bar{\eta}$ is defined as the time-averaged efficiency measured over the last electrical period of an FEM evaluation period. It is computed as the ratio of the mechanical output power P_{mech} to the total converted power, including copper losses P_{cu} computed using Equation (2.3) where R_{phase} is the phase resistance, and core losses P_{core} obtained from the *Ansys Maxwell* built-in core loss model (hysteresis, eddy-current, and excess losses). Here, $i_A(t)$, $i_B(t)$, and $i_C(t)$ are the instantaneous phase currents.

$$\eta(t) = \frac{P_{\text{mech}}(t)}{P_{\text{mech}}(t) + P_{\text{cu}}(t) + P_{\text{core}}(t)} \tag{2.1}$$

$$\bar{\eta} = \frac{1}{\tau_{\text{e}}} \int_{t}^{t+\tau_{\text{e}}} \eta(t)\, dt \tag{2.2}$$

$$P_{\text{cu}}(t) = R_{\text{phase}} \left(i_A^2(t) + i_B^2(t) + i_C^2(t) \right) \tag{2.3}$$

The mechanical power is calculated using Equation (2.4), where $T(t)$ is the instantaneous torque and $\omega(t)$ is the rotor speed.

$$P_{\text{mech}}(t) = T(t)\,\omega(t) \tag{2.4}$$

Mesh strategy and discretization: The domain was discretized using triangular elements with adaptive refinement concentrated in regions with high field gradients, particularly the air gap, magnet edges, and slot tooth tips. The final mesh contains 23,938 elements, with a minimum element area of 1.78×10^{-9} mm^2. Two mesh snapshots are provided in Figure 2.4a (global view) and Figure 2.4b (zoom-in of the air gap/slots) to illustrate the refinement strategy; Table 2.1 summarizes the mesh information.

Table 2.1 Mesh statistics by object

Object	Num Elements	Min Edge Length (mm)	Max Edge Length (mm)	Min Elem Area (mm^2)	Max Elem Area (mm^2)	Mean Elem Area (mm^2)
Band	1390	0.000500	0.009427	1.38×10^{-7}	3.11×10^{-5}	1.33×10^{-6}
Coil (×1)	98	0.000271	0.003554	5.46×10^{-8}	4.17×10^{-6}	1.17×10^{-6}
V_Magnet_1 (×1)	81	0.000650	0.002724	3.29×10^{-7}	2.11×10^{-6}	6.85×10^{-7}
V_Magnet_2 (×1)	85	0.000650	0.002724	3.29×10^{-7}	2.11×10^{-6}	6.52×10^{-6}
Magnet (×1)	86	0.000772	0.002872	4.86×10^{-7}	2.77×10^{-6}	1.25×10^{-6}
Region	3744	0.000022	0.003000	2.95×10^{-9}	2.58×10^{-6}	1.14×10^{-6}
Rotor	4184	0.000714	0.009459	2.64×10^{-7}	1.30×10^{-5}	2.36×10^{-6}
Stator	7872	0.000022	0.007310	1.78×10^{-9}	1.83×10^{-5}	1.44×10^{-6}
Total	23938	0.000022	0.009427	1.78×10^{-9}	3.11×10^{-5}	1.25×10^{-6}

2.2.3 MATLAB–ANSYS Coupling Framework

The proposed optimization environment is structured around a MATLAB–ANSYS coupling that automates the design evaluation loop. MATLAB acts as the supervisory layer, coordinating design-space exploration, invoking FE-based simulations, and interpreting the results at each iteration. The framework is intentionally modular, allowing easy replacement of optimization algorithms and FEM models; it also enables integration with Python (version 3.13.5) AI scripts.

As illustrated in Figure 2.5, the optimization algorithm proposes a new set of design variables and calls MATLAB for an FE evaluation. MATLAB then modifies and launches dedicated Visual Basic scripts, which update the parametric model in *Ansys Maxwell* and run the electromagnetic simulation for the current design.

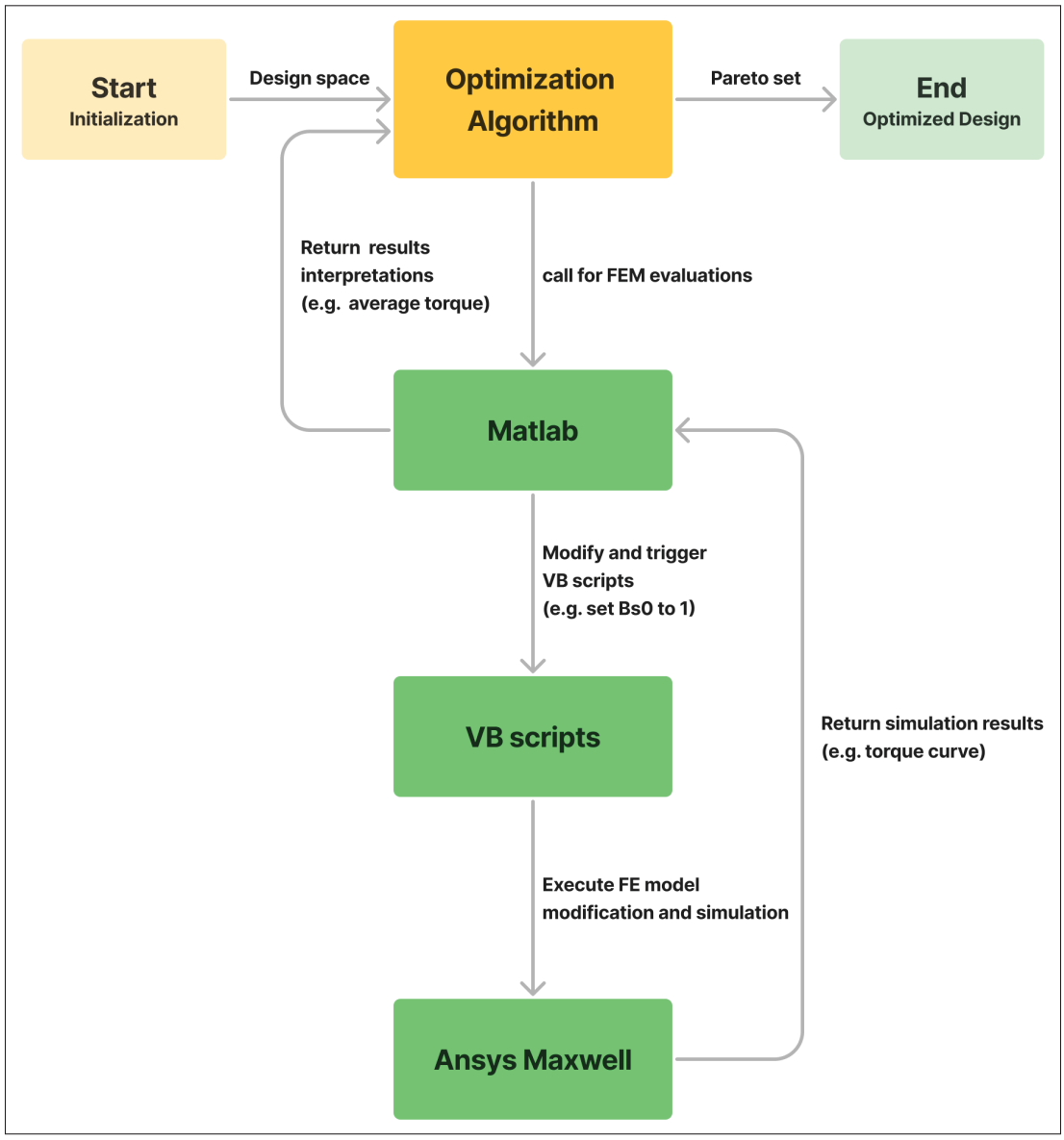


Figure 2.5 Finite element model optimization framework workflow

Once the simulation is finished, *Ansys Maxwell* returns the electromagnetic results (e.g., torque curves) to MATLAB. These results are post-processed into key performance indicators such as average torque, torque ripple, and efficiency, and the metrics are then sent back to the optimization algorithm. The optimizer uses this feedback to update its search strategy and iterates the loop until a set of optimized motor geometries is obtained.

2.2.4 Optimization Problem Formulation

The PMSM topology optimization problem is formulated as a constrained multi-objective design task, in which a set of geometric variables is tuned to improve the electromagnetic performance of the e-bike motor. The decision vector is defined as

$$\mathbf{x} = [Bs_0,\ Hs_0,\ Hs_1,\ Mt,\ Mw,\ Mt_V,\ Mw_V]^T \tag{2.5}$$

The first set of parameters $\{Bs_0,\ Hs_0,\ Hs_1\}$ represents the stator teeth/slot-opening geometry, which contributes to the torque-ripple characteristics. The remaining parameters—Mt, Mw, Mt_V, and Mw_V—represent the magnet dimensions, which affect the torque generated by the machine.

2.2.4.1 Design Constraints and Feasible Space

The optimization is performed under box constraints defined by practical geometric limits. These bounds establish the feasible design space:

$$\mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max} \tag{2.6}$$

where the lower and upper values for each parameter are given in Table 2.2, which directly defines the constraints for the optimization study.

Table 2.2 Motor design parameters and their corresponding lower and upper bounds used in the optimization study

Parameter	Lower Bound (mm)	Upper Bound (mm)
Bs_0	0.6	1.8
Hs_0	0.3	0.84
Hs_1	0.0	0.6
Mt	1.74	3.0
Mw	4.8	8.4
Mt_V	1.5	1.8
Mw_V	8.4	12.8

These bounds were selected to satisfy mechanical manufacturability and electrical constraints (e.g., minimum slot-opening dimensions, wire gauge considerations, and non-interference constraints during parametric model updates).

2.2.4.2 Objective Functions

Each candidate design $\mathbf{x}$ is evaluated using the MATLAB–ANSYS Maxwell loop, where FEM simulations return electromagnetic outputs (torque waveform and losses). These results are post-processed into key performance indicators used as objectives: average torque, torque ripple, and efficiency. The goal is to identify a Pareto-optimal set of motor geometries that simultaneously:

- maximizes average torque, $T_{\text{avg}}(\mathbf{x})$;
- maximizes efficiency, $\eta(\mathbf{x})$;
- minimizes torque ripple, $TR(\mathbf{x})$.

This multi-objective formulation can be expressed as Equation (2.7) and the objective vector as Equation (2.8), where X is the feasible region defined by the parameter bounds in Table 2.2.

The optimization objective is therefore a three-objective trade-off problem, where improving torque capability conflicts with ripple minimization, and efficiency interacts with both.

$$\max_{\mathbf{x}\in X} \Big(T_{\text{avg}}(\mathbf{x}),\ \eta(\mathbf{x})\Big) \qquad \text{and} \qquad \min_{\mathbf{x}\in X} TR(\mathbf{x}), \tag{2.7}$$

$$\mathbf{f}(\mathbf{x}) = \big[T_{\text{avg}}(\mathbf{x}),\ \eta(\mathbf{x}),\ TR(\mathbf{x})\big]^T \tag{2.8}$$

Because the performance mapping is obtained through the MATLAB–ANSYS simulation environment, the function is treated as a black box, providing only objective evaluations without gradient information; therefore, gradient-based optimization is not directly applicable.

2.2.5 AI-Assisted Bayesian Optimization

The motivation behind this integration is to provide the designer with a prompt where they can request design specifications, provide experience-based insights, or impose constraints qualitatively through natural language. The prompted AI agent can then execute actions, assess the results, and refine its policy for subsequent actions (Figure 2.6).

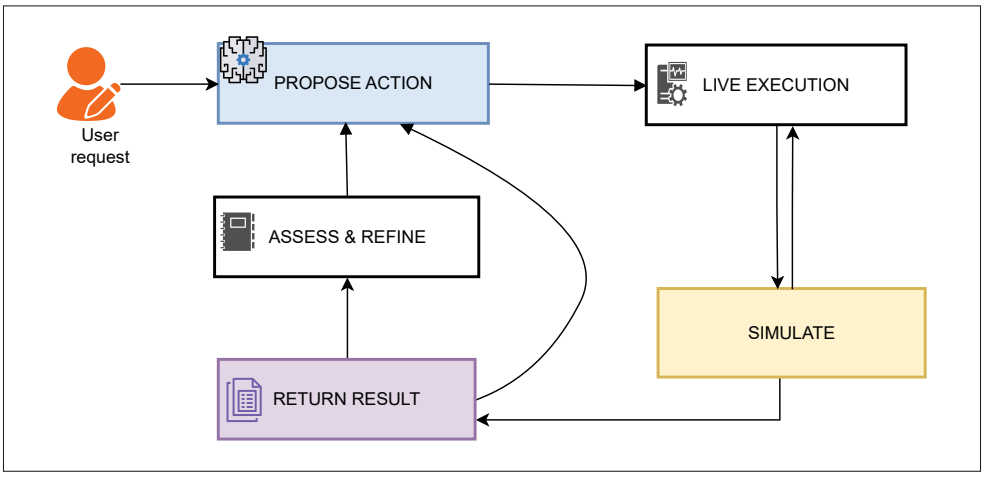


Figure 2.6 AI agent loop

Bayesian optimization (BO) provides an approximation model to find candidates $x \in \mathcal{X}$. Specifically, an approximation model can be written as

$$p(r \mid x; D_n) = \int_{\Theta} p(r \mid x, \theta; D_n)\, p(\theta \mid x; D_n)\, d\theta \tag{2.9}$$

where $D_n = \{(x_i, r_i)\}_{i=1}^{n}$ is the observed dataset (knowledge base), and θ is a latent variable capturing the relationship between x and the response r. In standard BO, the prior $p(\theta)$ is often chosen to be simple (e.g., a broad Gaussian). With an LLM augmented by retrieval, human expertise or historical trends can be used to define a more informative prior. For example, the LLM may infer that increasing Mw tends to increase torque. In this work, the incorporation of prior knowledge is expressed as

$$p(\theta \mid D_n, x) \propto p(D_n \mid \theta, x)\, p(\theta) \tag{2.10}$$

where $p(\theta)$ can be enhanced using insights extracted from the optimization history.

A Gaussian process (GP) surrogate is used as the approximation method and is fitted to D_n, where x_i are parameter sets (e.g., $\{Bs_0 : 1,\ Hs_0 : 0.5,\ \ldots\}$) and r_i are the objective values obtained from FEM runs. The LLM is used to propose priors $p(\theta)$ and/or refine predictions. For example, for a new candidate $x' = \{Bs_0 : 1.2,\ Hs_0 : 0.6,\ \ldots\}$, the GP may predict $\hat{r} = 0.85 \pm 0.1$, while the LLM adjusts this estimate based on observed trends (e.g., "wider Bs_0 reduces ripple") to improve early predictions.

RAG + BO: A high-level overview of the architecture is as follows:

- **Initialization:** The LLM suggests initial parameter sets based on the problem description and logged human knowledge from previous runs. This replaces random initialization and can improve convergence speed.
- **Candidate sampling:** After a few iterations, the LLM proposes new candidates to Pareto-optimize the cost vector

$$\mathbf{C}(x) = \left[T_{\text{avg}}(x),\ \eta(x),\ TR(x)\right]^{T} \tag{2.11}$$

targeting non-dominated solutions based on the current knowledge base D_n and the allowable parameter ranges. In-context learning with historical data is used to propose diverse candidate sets guided by observed trends (e.g., increasing Mw boosts torque).

- **Approximation modeling:** The LLM is combined with a GP surrogate to support candidate selection $x \in \mathcal{X}$ using the knowledge base D_n and latent variables θ.
- **Evaluation:** Candidates are evaluated through the black-box MATLAB–ANSYS simulation loop, which updates the objective values after execution. For the multi-objective case, the LLM can extract Pareto-relevant trade-offs from the updated results.
- **Stopping:** The algorithm stops when the Pareto front becomes stable over multiple iterations.

Figure 2.7 illustrates the workflow of the LLM-assisted optimization loop. The agent component mediates between the simulation environment, the knowledge base, and the LLM. At the start of each iteration, the agent prompts the LLM with the current dataset and past simulation history to suggest parameter sets and sampling settings. These candidates are passed to a hybrid GP–LLM module, which uses a GP surrogate and LLM-based reasoning to select the next design to evaluate via an acquisition function. The agent then triggers simulations, collects the resulting performance metrics, and updates the knowledge base. The LLM analyzes the new results relative to previous runs, allowing the system to progressively refine its understanding of the design space and improve subsequent optimization steps.

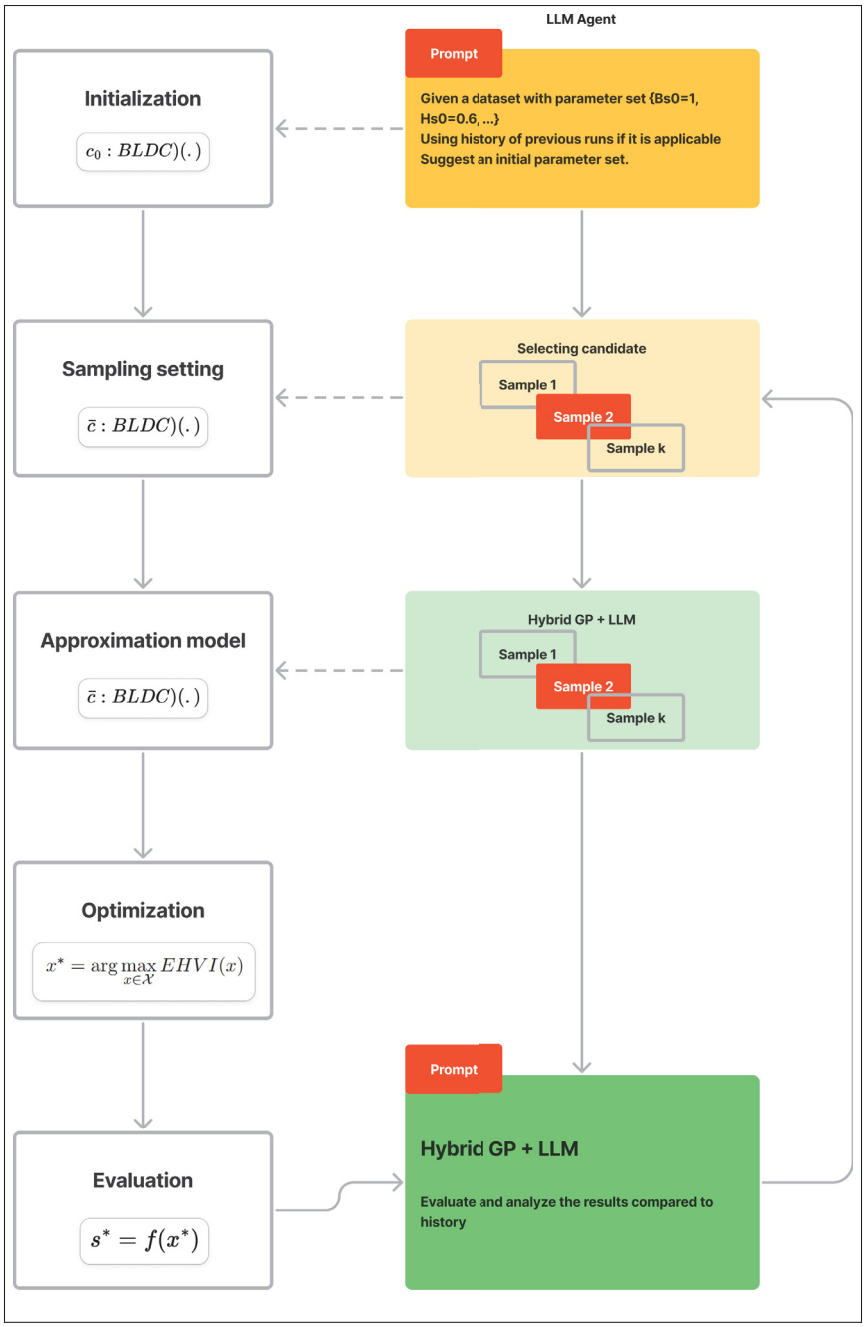


Figure 2.7 LLM-assisted FEM optimization framework

In this work, the retrieval mechanism is applied in a closed loop where retrieval is performed exclusively from the internal knowledge base D_n (historical FEM simulation results) and from LLM-generated memory logs containing trend summaries and qualitative design rules. External

unstructured documents (e.g., textbooks or research papers) are not included in the retrieval corpus. While this differs from classic RAG applications that augment LLMs with large external knowledge bases, the use of vector-based semantic retrieval to ground LLM generations in the optimization history aligns with the core principles of retrieval-augmented generation and enables the LLM to serve as a memory-based design agent.

2.2.6 RAG–BO–LLM-Based Algorithm for Multi-Objective Electric Motor Optimization

Leveraging the advantages of large language models (LLMs) (e.g., a fine-tuned GPT-4 class model or similar), we build an adaptive agent that injects qualitative design knowledge into the Bayesian optimization (BO) loop, improving sample efficiency beyond a standard Gaussian process (GP) surrogate. To address limitations of pre-trained LLM knowledge and reduce hallucinations, the agent is restricted to an internal optimization corpus consisting of human interactive guidance, design constraints, and prior simulation results (e.g., log files and evaluations). Unlike traditional RAG systems that retrieve from external documents, this approach relies on a dynamic, internal knowledge base, summarized as follows.

- **Retrieved information and corpus:** The primary corpus is a dynamic knowledge base

$$D_n = \{(x_i, r_i)\}_{i=1}^{n} \tag{2.12}$$

  where x_i is a vector of design parameters (e.g., $[Bs_0,\ Hs_0,\ \ldots,\ Mw_V]$) and

$$r_i = \left[T_{\text{avg}}(x_i),\ \eta(x_i),\ TR(x_i)\right]^T \tag{2.13}$$

  is the corresponding multi-objective vector from FEM evaluations. This corpus is stored as a vector database for efficient similarity search and updated after each iteration. Retrieval also includes memory logs, such as logs from previous optimization runs, serialized summaries of trends, Pareto analyses, and user-provided natural language rules (Table 2.3). Queries to the corpus use semantic similarity (e.g., cosine distance in an embedding space) to fetch

the top-k most relevant entries, such as past designs with similar objective trade-offs or parameter ranges. The output of this step is a shortlist of candidates and context that the LLM uses to propose actions within the optimization loop.

- **Integration into the optimization loop:** Retrieval is invoked at key steps (see Figure 2.7 and Algorithm 2.1):
  - **Initialization (Lines 7–8 in Algorithm 2.1):** The LLM is prompted with the problem description (maximize average torque and efficiency, minimize torque ripple), augmented by retrieved memory logs from prior runs (if available). This produces initial parameter sets, replacing random sampling for faster convergence.
  - **Candidate sampling (Lines 16–18):** After computing the current non-dominated Pareto set

$$r' = \text{NonDominated}\left(\{r_i\}_{i=1}^{|D_n|}\right) \tag{2.14}$$

    the LLM is prompted with retrieved subsets of D_n (selected by similarity to r') and the natural language rules (if provided). For example, the prompt can request $K = 10$ diverse candidates near the Pareto front. The LLM outputs candidate sets $\{\tilde{x}_k\}_{k=1}^{K}$, which are then filtered to satisfy feasibility within bounds X.
  - **Approximation modeling (Lines 19–22):** A GP is fitted to D_n to predict $\hat{r}_k$ for each $\tilde{x}_k$. Retrieval augments this step by extracting trends from D_n (e.g., parameter–objective correlations) and prompting the LLM to refine the GP prediction for each candidate. The combined prediction is modeled as

$$\hat{r}_k = \alpha_1 \times \text{GP}(\tilde{x}_k) + \alpha_2 \times \text{LLM}(\tilde{x}_k) \tag{2.15}$$

    where α_1 and α_2 are tunable weights.
  - **Evaluation and analysis (Line 28):** When trade-offs are detected (e.g., changes in hypervolume), retrieval fetches similar past Pareto analyses from memory logs to prompt the LLM to analyze the new result r^* relative to D_n and suggest policy refinements. During optimization, a human operator may monitor trends and optionally inject run-time knowledge to accelerate convergence.

- **Stopping:** The algorithm stops when the Pareto front stabilizes (or a specified threshold criterion is met).

• **Effects on the RAG + BO-based model:**

- **Priors $p(\theta)$:** Retrieved trends from D_n and natural language rules inform the GP prior (e.g., biasing the mean function toward regions where wider slots reduce ripple), consistent with

$$p(\theta \mid D_n, x) \propto p(D_n \mid \theta, x)\, p(\theta) \tag{2.16}$$

where $p(\theta)$ is shaped by LLM-summarized knowledge and can be updated dynamically to initialize optimization parameters more effectively than random initialization.

- **Constraints:** Natural language rules act as soft constraints by biasing candidate generation, without hard enforcement in the definition of the feasible set X.
- **Acquisition strategies:** The acquisition step is indirectly influenced because LLM-generated candidates $\tilde{x}_k$ pre-filter the search space, focusing evaluations on promising, knowledge-aligned points without modifying the underlying BO algorithm.

Table 2.3 Natural language input (rules)

Rule ID	**Rule expression**
1	Hs_1 has the lowest effect on torque ripple.
2	Mt_V has the lowest effect on the average torque.
3	The average torque increases when Mw and Mw_V increase.
4	The torque ripple decreases when Bs_0 increases.
5	The torque ripple decreases when Hs_0 increases.

This integration makes the framework interpretable (via natural language prompts and memory logs) and scalable, as the corpus grows with additional runs, enabling transfer learning across similar motor designs.

Furthermore, the design architecture mitigates LLM hallucinations or incorrect physical guidance. First, the LLM does not directly control the final decision: it proposes candidates or adjusts priors, while the final selection (e.g., "select the best candidate using expected improvement") remains driven by the GP-based acquisition function. Thus, even if the LLM suggests a poor setting within the allowable range, the GP assigns it low expected improvement if past data indicates poor performance. Second, the RAG mechanism conditions LLM responses on observed simulation results (D_n) and operator-provided rules (e.g., Table 2.3), grounding the LLM in empirical data. Third, if an LLM-suggested design performs poorly, it is recorded in D_n and becomes less favorable in subsequent GP predictions. Finally, the weights α_1 and α_2 limit the influence of LLM-based adjustments when approaching optimal solutions; incorrect guidance is logged so that the agent can avoid repeating similar mistakes.

2.3 Results and Discussion

2.3.1 AI-Assisted Optimization Without Natural Language Input

The first trial uses the RAG + BO to search for optimal solutions without any natural language input. The results presented in Figure 2.8 demonstrate a clear convergence pattern toward high-efficiency operating conditions, with the Pareto-optimal solutions forming a compact frontier in the multi-objective design space. In the three-dimensional view (Figure 2.8a), the Pareto set occupies a narrow region characterized by simultaneously low torque ripple, moderate average torque, and high efficiency.

Algorithm 2.1 RAG + BO for multi-objective electric motor optimization

Input: Parameter bounds $\mathcal{X}$ (box constraints from Table 2.2); maximum iterations N; candidates per iteration K; knowledge base $D \leftarrow \emptyset$
Output: Pareto-optimal set $\mathcal{X}_P$ (non-dominated designs from final D)

`/* Initialization */`
Prompt the LLM with the problem description, bounds $\mathcal{X}$, and retrieved memory logs from prior runs;
Generate initial parameter sets $\{x_i^0\}_{i=1}^5$ via the RAG-augmented LLM;
for $i \leftarrow 1$ **to** 5 **do**
  Evaluate x_i^0 using MATLAB–ANSYS FEM simulation;
  $r_i^0 \leftarrow \left[T_{\text{avg}}(x_i^0),\ \eta(x_i^0),\ -TR(x_i^0)\right]^\top$;
  $D \leftarrow D \cup \{(x_i^0, r_i^0)\}$;
end for
`/* Main loop */`
for $t \leftarrow 1$ **to** N **do**
  `/* Candidate sampling */`
  $r' \leftarrow \text{NonDominated}\left(\{r_i\}_{i=1}^{|D|}\right)$;
  Retrieve top-k similar entries from D (semantic similarity on embeddings);
  Prompt LLM with D, r', retrieved trends, and rules to generate K diverse candidates near the Pareto front;
  Obtain feasible candidates $\{\tilde{x}_k\}_{k=1}^K \subset \mathcal{X}$;
  `/* Approximation modeling */`
  Fit independent GP surrogates to each objective using D;
  Retrieve trends from D; prompt LLM to refine GP means for each $\tilde{x}_k$;
  for $k \leftarrow 1$ **to** K **do**
    $\hat{r}_k \leftarrow \alpha_1 \times \text{GP}(\tilde{x}_k) + \alpha_2 \times \text{LLM}(\tilde{x}_k)$;
  end for
  `/* Acquisition and evaluation */`
  $x^* \leftarrow \arg\max_k \ \text{EHVI}(\hat{r}_k \mid r')$;
  Run MATLAB–ANSYS FEM for x^* to obtain $r^* \leftarrow \left[T_{\text{avg}}(x^*),\ \eta(x^*),\ -TR(x^*)\right]^\top$;
  $D \leftarrow D \cup \{(x^*, r^*)\}$;
  `/* Analysis and memory update */`
  Retrieve similar past analyses; prompt LLM to analyze trade-offs of r^* vs. D;
  Update memory logs with LLM summary;
  `/* Stopping criterion */`
  Compute dominated hypervolume $HV(P) = \text{volume}(\bigcup_{r \in P}[r, r_{\text{ref}}])$, where P is the current Pareto set;
  if $\left|HV_t - HV_{t-\rho}\right| / HV_{t-\rho} < \epsilon$ *over the last* ρ *iterations* **then**
    break;
  end if
end for
Extract final Pareto front via non-dominated sorting on D;
return $\mathcal{X}_P \leftarrow \{x_i \mid r_i$ is non-dominated in $D\}$;

The compact frontier of the Pareto-optimal solutions in the objective space indicates convergence toward a consistent trade-off surface. To verify that this clustering is not an artifact of narrow

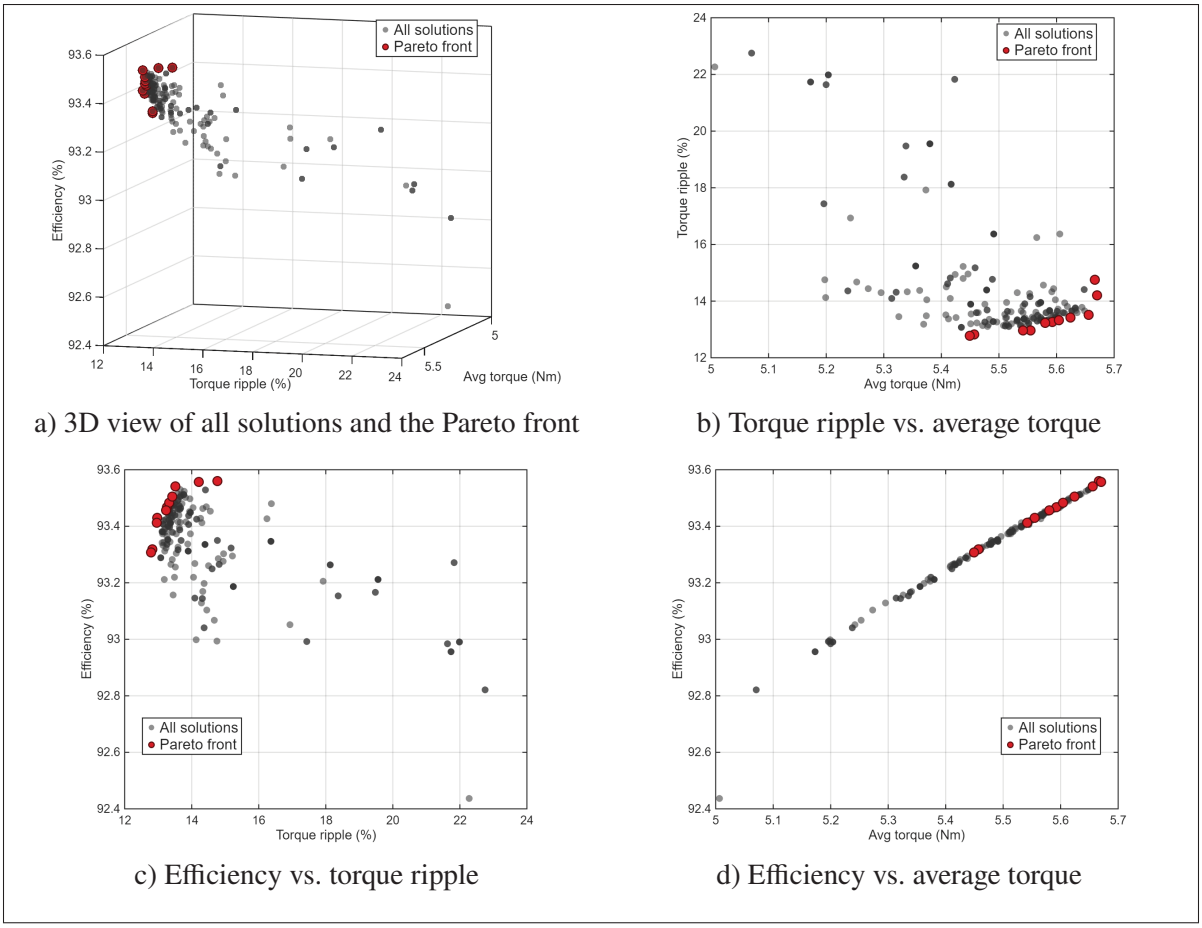


Figure 2.8 Optimization design points without natural language input

design-variable bounds, Table 2.4 reports the explored ranges of each design parameter across all evaluated samples, showing that the optimizer sampled broadly within the feasible domain before concentrating near the non-dominated region.

Table 2.4 Exploration coverage by variable

Variable	Min Explored	Min Bound	Max Explored	Max Bound	Range Explored	Bound Range
Bs_0 (mm)	0.90	0.60	1.80	1.80	0.90	1.20
Hs_0 (mm)	0.44	0.30	0.84	0.84	0.40	0.54
Hs_1 (mm)	0.18	0.00	0.50	0.60	0.32	0.60
Mt (mm)	1.80	1.74	3.00	3.00	1.20	1.26
Mw (mm)	6.02	4.80	8.34	8.40	2.32	3.60
Mt_V (mm)	1.56	1.50	1.80	1.80	0.24	0.30
Mw_V (mm)	10.20	8.40	12.80	12.80	2.60	4.40

The two-dimensional projections (Figure 2.8b, 2.8c) offer deeper insight into the relationships between the objectives. The torque ripple versus average torque plot (Figure 2.8b) shows that the Pareto-optimal designs are concentrated toward higher torque-ripple values within the explored region, indicating a trade-off where increased average torque is obtained at the expense of increased torque ripple. Meanwhile, Figure 2.8c highlights a reversed trend between torque ripple and efficiency, where the Pareto points cluster in the upper-left region, showing that within the explored designs, lower torque-ripple values tend to coincide with higher efficiency.

Overall, these results demonstrate that the combination of RAG and BO is able to identify feasible design candidates and to reveal the underlying structure of the performance landscape. The overall design-point distribution and the Pareto front knee shape indicate convergence behavior and suggest that the chosen objective formulations are well posed for capturing electric motor performance trade-offs.

2.3.2 AI-Assisted Optimization with Natural Language Input

In the second optimization trial, the AI-assisted Bayesian optimization framework was augmented with natural language input, where qualitative, experience-based rules (Table 2.3) allow the LLM agent to translate design preferences into explicit search policies for the permanent magnet

machine. Within the MATLAB–ANSYS coupling, the LLM operates on top of the GP surrogate. In the present trial, this interaction was steered by language instructions that emphasize the effect of specific variables on the objectives (e.g., which variable has the lowest effect on torque, or which has the highest effect on torque ripple).

Such a configuration aligns with recent work on LLM-enhanced Bayesian optimization, where natural language priors are used to bias the acquisition process toward user-preferred regions of the design space.

The resulting distribution of design points in the second trial shows a structured exploration of the three-dimensional objective space, similar to the first trial but with fewer explored configurations. The cloud of gray solutions spans a broad efficiency range, down to approximately 92.3%, indicating a sub-optimal starting region for the optimization process relative to the previous trial. However, the Pareto-optimal set (red markers) collapses into a compact cluster at simultaneously low torque ripple (≈13–14%) and high efficiency (≈93.4–93.6%), with only modest variations in average torque after 182 iterations (Table 2.5). Compared with the first trial, which required 242 iterations (Table 2.5), the LLM input resulted in a 20.1% reduction in simulation expense. This suggests that LLM-guided candidate generation is more effective at sampling along the relevant trade-off surface, rather than spending evaluations in regions that are sub-optimal in all three objectives.

Table 2.5 Number of iterations of the first and second trials

Trial ID	**Optimization technique**	**Number of iterations**	**Number of candidates**
1	RAG + BO without LLM inputs	242	10
2	RAG + BO with LLM inputs	182	10

A closer comparison of the two trials indicates that natural language input reorients the Pareto set toward designs with slightly lower average torque but systematically reduced torque ripple and marginally improved efficiency. In the first trial, the frontier extended toward designs

with higher torque at the cost of increased ripple, consistent with an optimization policy that weighs torque and efficiency more symmetrically against ripple. In contrast, the second trial concentrates Pareto points near the low-ripple edge of the feasible region, indicating that the LLM successfully internalized the designer's qualitative preference and encoded it into the sampling strategy.

From a methodological perspective, the second trial illustrates the principal advantage of embedding an LLM within the optimization loop: the ability to inject high-level engineering knowledge and soft constraints without manually redefining objective weights or hard bounds. The memory-augmented, retrieval-based prompting allows the agent to reason over prior FEM evaluations and express this knowledge in new candidate proposals. This knowledge-assisted behavior is consistent with recent studies showing that LLMs can act as adaptive priors or meta-optimizers, improving sample efficiency compared with purely data-driven BO in complex design tasks.

2.3.3 Comparative Performance Evaluation of the Proposed Method and the MO-AHA

To assess the benefit of the proposed AI-assisted Bayesian optimization, its results are compared with those obtained using the multi-objective version of the Artificial Hummingbird Algorithm (MO-AHA). Both optimizers operate on the same MATLAB–ANSYS framework, FEM model, design variables, and bounds defined in Sections 2.2 and 2.3, and use the same multi-objective formulation that simultaneously maximizes average torque and efficiency while minimizing torque ripple.

From the 3D scatter plots of the second trial (AI-assisted Bayesian optimization with LLM input; Figure 2.9) and Trial 3 (MO-AHA; Figure 2.10), both methods ultimately identify Pareto-optimal designs located in a similar region in the objective space. This indicates that, in terms of best-found designs, AI-assisted Bayesian optimization is competitive with the MO-AHA benchmark, achieving comparable trade-offs between smoothness, efficiency, and torque capability.

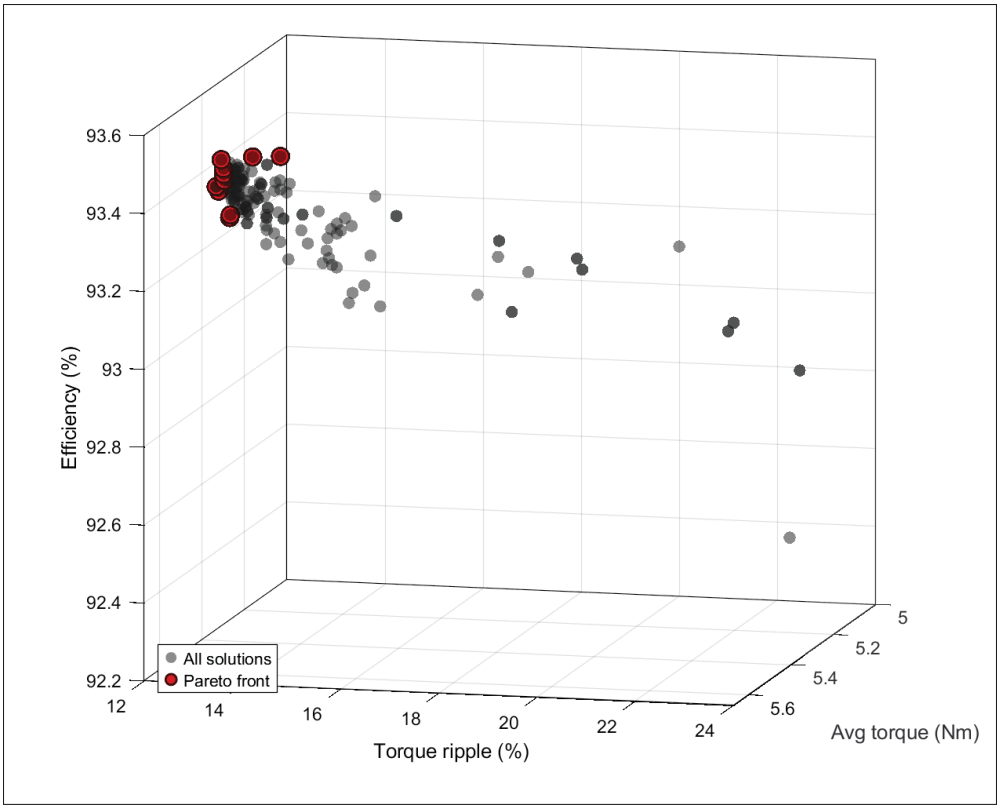


Figure 2.9 RAG + BO with natural language input results

However, the search dynamics of the two methods differ. In Trial 3, the MO-AHA generates a wide cloud of solutions that spans a wide range of torque ripple (roughly 13–26%) and efficiency (about 92–93.6%), as well as a broad interval of average torque values. The resulting scatter shows a clear global trend from high-ripple and low-efficiency designs toward the low-ripple and high-efficiency region where the Pareto front resides, reflecting broad exploration from the meta-heuristic. In contrast, LLM-assisted Bayesian optimization in Trial 2 produces a denser concentration of design points close to the final Pareto cluster, with fewer evaluations spent in clearly dominated regions. This behavior is consistent with the language-encoded rules in Table 2.3, which bias the search toward the targeted objectives.

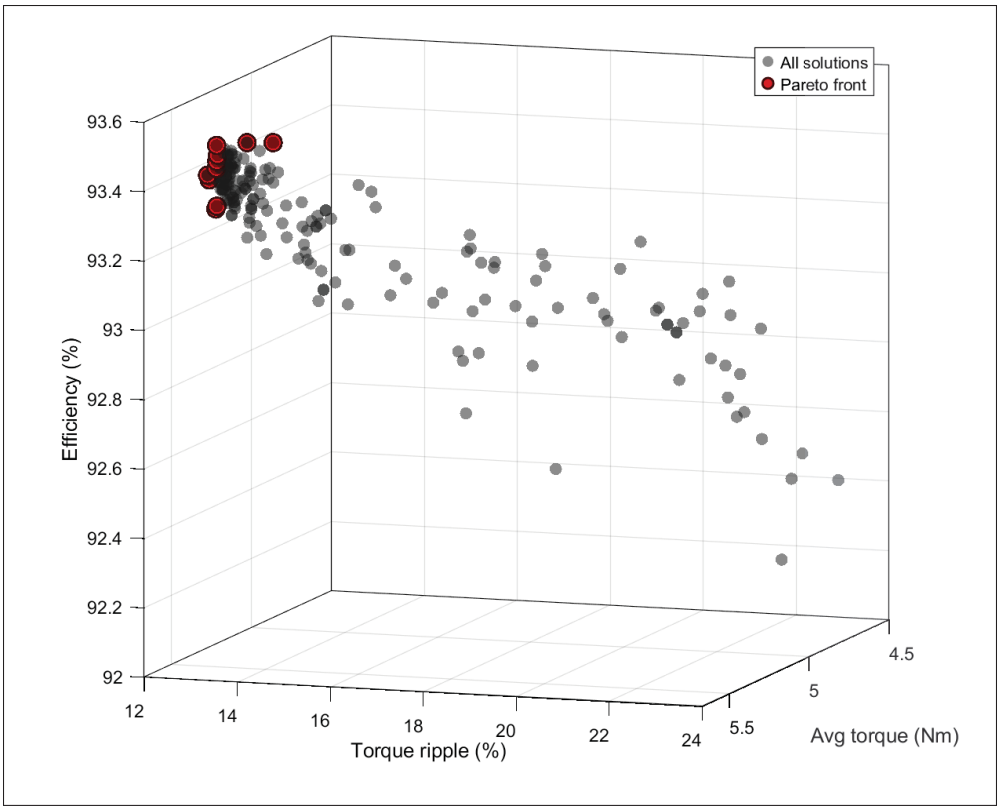


Figure 2.10 MO-AHA results

The number of FEM evaluations required quantifies the difference in sample efficiency for each trial. LLM-guided Bayesian optimization converged to its compact Pareto set after 182 iterations (Table 2.6), whereas the MO-AHA required 351 iterations to reach a similar quality of solutions (Table 2.6). Given the high computational cost of each FEM call, this corresponds to an approximate 48% reduction in simulation effort when using the proposed AI-assisted approach. The improved efficiency arises from exploiting both the surrogate model and the prior knowledge injected through natural language rules and retrieval-augmented memory, which allows the optimizer to focus quickly on the most promising region of the design space.

Table 2.6 Number of iterations (Trials 2 and 3)

Trial ID	Optimization technique	Number of iterations
2	RAG + BO with LLM inputs	182
3	MO-AHA	351

All simulations and optimization runs were executed on a workstation equipped with an Intel Xeon CPU E5-2670 @ 2.60 GHz (eight physical cores / 16 threads), 32 GB RAM, running Windows 10. The FEM analysis represents the dominant computational cost: the average wall-clock time per FEM evaluation was 1 min 0.62 s (min–max: 59.87 s–1 min 1.87 s) for the mesh settings described in Section 2.2.

Overall, the comparison shows that the MO-AHA provides good global coverage of the search space and serves as a robust benchmark. Nevertheless, AI-assisted Bayesian optimization achieves comparable Pareto performance with substantially fewer simulations; the time reduction is mainly due to the reduced number of iterations required for convergence. For applications such as electric motor design, where FEM evaluations dominate the computational cost of the optimization process, this combination of high-quality solutions and fewer evaluations highlights the practical advantage of integrating LLM guidance into the optimization loop.

2.3.4 Optimal Designs

The optimal design set summarized in Table 2.7 shows that the RAG + BO framework consistently identifies Pareto-optimal configurations that frequently lie near the upper bounds of several design variables. Most optimal candidates adopt large slot-opening and slot-height values ($Bs_0 \approx$ 1.6–1.8 mm, $Hs_0 \approx$ 0.6–0.84 mm) together with large magnet dimensions ($Mt \approx$ 3.0 mm, $Mw \approx$ 8.3–8.34 mm, $Mt_V \approx$ 1.7–1.8 mm, Mw_V close to its upper limit). Across the solutions, the average torque remains tightly clustered around 5.5–5.7 Nm while torque ripple and efficiency span approximately 12.8–14.2% and 93.3–93.6%, respectively,

illustrating a well-defined Pareto trade-off. Designs such as Design 4 achieve high torque and efficiency at the cost of larger ripple, whereas designs such as Design 6 reduce ripple but slightly sacrifice torque and efficiency. This behavior is consistent with the multi-objective trends observed in Sections 2.3–2.3 (Subsections 2.3), where the Pareto-front knee reflects a preference for smoother torque at moderate torque levels suitable for the geared e-bike drivetrain.

Table 2.7 Optimal motor designs

Design	Bs_0	Hs_0	Hs_1	Mt	Mw	Mt_V	Mw_V	**Avg. torque (Nm)**	**Ripple (%)**	**Eff. (%)**
1	1.80	0.84	0.42	3.00	8.34	1.80	12.80	5.66	13.51	93.54
2	1.80	0.84	0.42	3.00	8.34	1.80	11.40	5.46	12.83	93.32
3	1.80	0.84	0.42	3.00	8.34	1.80	12.08	5.56	12.97	93.43
4	1.63	0.62	0.38	2.99	8.33	1.78	12.76	5.67	14.21	93.56
5	1.69	0.84	0.37	2.99	8.33	1.75	12.41	5.59	13.26	93.47
6	1.79	0.83	0.41	2.95	8.34	1.75	11.48	5.45	12.78	93.31
7	1.79	0.83	0.38	3.00	8.33	1.72	12.02	5.54	12.96	93.41
8	1.72	0.77	0.39	3.00	8.33	1.73	12.39	5.60	13.37	93.48
9	1.73	0.76	0.40	2.99	8.34	1.80	12.49	5.62	13.42	93.50

Given that the Pareto-optimal candidates exhibit a narrow spread in average torque, average torque becomes a practical discriminator when manually selecting one optimal design. Efficiency improvements are highly desirable given the focus on range and power consumption in e-bikes, provided they do not come at the expense of torque ripple. In this context, Design 9 is an appropriate choice, since it preserves near-maximum efficiency (93.50%) while maintaining a comparatively low torque ripple (13.42%) at a representative torque level (5.62 Nm). This selection deliberately avoids the highest-efficiency option (Design 4, 93.56%), which has noticeably larger ripple (14.21%) for a marginal efficiency gain. Overall, Design 9 offers a balanced compromise within the reported Pareto set.

2.4 Conclusions

This work proposed and evaluated an AI-assisted optimization framework for the topology design of a Δ-type interior PMSM dedicated to an e-bike mid-drive powertrain. By coupling a MATLAB–ANSYS Maxwell simulation environment with a Bayesian optimization scheme enhanced by an LLM-based, retrieval-augmented agent, the study addressed a three-objective problem that seeks to maximize average torque and efficiency while minimizing torque ripple. Across all trials, the optimizer identified feasible geometries that deliver moderate average torque (≈5.3–5.7 Nm), low torque ripple (≈12.8–14.2%), and high efficiency (≈93.3–93.6%), which are compatible with the requirements of a geared e-bike drivetrain.

The comparison between the purely data-driven RAG + BO baseline and the LLM-guided RAG + BO configuration highlights the benefit of incorporating natural language design knowledge directly into the search process. When qualitative rules on the influence of slot and magnet parameters were provided, the LLM reoriented the exploration toward smoother torque and high-efficiency designs and concentrated samples near the relevant trade-off surface. This led to convergence in 182 iterations, representing a 20.1% reduction in simulations relative to the 242 iterations required by the baseline trial without language guidance.

Against the multi-objective Artificial Hummingbird Algorithm benchmark, AI-assisted Bayesian optimization achieved a comparable Pareto front while requiring substantially fewer simulations (182 versus 351 iterations, a reduction of about 48%). MO-AHA offered broad global exploration, but the proposed approach exploited both surrogate model information and LLM-encoded priors to avoid clearly dominated regions of the design space. This improved sample efficiency is particularly valuable in electric machine design, where each high-fidelity FEM evaluation is computationally expensive.

The resulting optimal designs also provide physical insight. Most Pareto-optimal solutions push the stator slot-opening, slot height, and magnet dimensions close to their upper bounds, confirming that increased magnet volume and wider slots promote torque production while maintaining acceptable ripple and efficiency. The fact that the best candidates cluster near these

limits, without clear signs of magnetic saturation, suggests that relaxing some geometric bounds in future studies could unlock additional torque and efficiency gains.

Author Contributions: Conceptualization, M.A.G.; software, M.A.G. and C.P.; validation, Q.W., Y.-J.P., K.A.H., and K.K.N.; resources, M.A.G., C.P., Q.W., and K.K.N.; writing—original draft preparation, M.A.G.; writing—review and editing, Q.W., Y.-J.P., and K.A.H.; visualization, M.A.G. and C.P.; supervision, Q.W. and K.K.N.; project administration, Q.W. and K.K.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by The Natural Sciences and Engineering Research Council of Canada (NSERC).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

CONCLUSION AND RECOMMENDATIONS

3.1 Conclusion

The electromagnetic design of a mid-drive e-bike traction motor is ultimately validated at the rider interface: torque must be delivered smoothly, efficiently, and within strict constraints on package volume. In this thesis, the optimization of an interior permanent magnet synchronous machine (IPMSM) was therefore treated as a genuinely multi-objective design problem, where torque generation capability (average torque), torque quality (torque ripple), and efficiency must be optimized simultaneously rather than sequentially. The selected Δ-type IPMSM architecture and the chosen design variables (slot features and magnet dimensions) provided an appropriate design space to expose the trade-offs most relevant to e-bike propulsion.

A decisive practical challenge in this setting is that high-fidelity finite-element modeling (FEM) is not merely a final verification step; it is the objective-function evaluator that resolves slotting effects, saturation, and harmonic content with a level of confidence not generally achievable with analytical models. Consequently, the limiting factor becomes the number of FEM evaluations that can be afforded before the design study becomes impractically time consuming. The thesis adopted sample efficiency as a primary success criterion for the proposed optimization framework (achieving a competitive Pareto region with fewer FEM calls).

To address this bottleneck, an automated optimization environment was developed by coupling a MATLAB–ANSYS Maxwell workflow with a multi-objective Bayesian optimization (BO) strategy based on Gaussian-process (GP) surrogate modeling. Critically, the framework was augmented with a retrieval-augmented large language model (LLM) agent. The role of the LLM was not to replace physics-based simulation, but to steer the optimization more intelligently by leveraging the accumulated optimization history as a retrievable corpus and by incorporating qualitative engineering knowledge through natural-language rules. In this architecture, the

LLM proposes candidate regions and biases priors in a controlled manner, while the GP-based acquisition function and FEM evaluation remain the final arbiters of performance. In doing so, the framework integrates human-readable intent and machine-learned memory into a physically grounded optimization loop.

Across optimization trials, the proposed methodology identified feasible motor geometries delivering moderate average torque, low torque ripple, and high efficiency. Representative Pareto-optimal solutions achieved average torque in the range of $\approx$ 5.3–5.7 N·m, torque ripple of approximately $\approx$ 12.8–14.2%, and efficiency of roughly $\approx$ 93.4–93.6%. The clustering of solutions in a compact region of the objective space indicates that the selected design variables and bounds were sufficient to locate a stable trade-off region suitable for the intended drivetrain, while still revealing clear trends: optimal candidates frequently adopted relatively large slot-opening and slot-height values together with magnet dimensions near the upper bounds of the explored range.

Most importantly, the integration of the LLM into the BO loop reduced the number of FEM evaluations required to reach a compact Pareto set of comparable quality. In direct comparison, the LLM-guided configuration converged after 182 iterations versus 242 iterations for the corresponding BO trial without natural-language input, representing a 20.1% reduction in simulation expense. Relative to a metaheuristic benchmark (MO-AHA), the LLM-guided BO approach reached similar Pareto quality in 182 iterations versus 351 iterations, yielding a substantially improved simulation budget. In the context of FEM-in-the-loop motor optimization, these reductions translate directly into enhanced practicality and scalability of the design process.

Beyond the specific e-bike motor case study, the thesis demonstrates a broader methodological contribution: LLMs can be incorporated into engineering optimization frameworks in a way that remains auditable and physically grounded. When conditioned by retrieval over empirical simulation outcomes and constrained to influence candidate generation and prior shaping, the

LLM becomes a useful mechanism for encoding design heuristics and allocating expensive simulations more effectively. As such, the proposed FEM based AI-assisted framework provides a viable direction for accelerating high-fidelity machine design under tight computational budgets and integrating AI into motor design.

3.2 Recommendations and Future Work

1. **Integrate coupled-physics constraints and reliability margins.** The present electromagnetic objectives should be complemented with thermal constraints and mechanical stress. NVH proxies would also improve the connection between torque-ripple metrics and perceived ride quality.
2. **Adopt a staged fidelity strategy: 2D exploration with selective 3D validation.** While 2D FEM is appropriate for iterative optimization, selected Pareto candidates should be re-evaluated using higher-fidelity models (more accurate 3D models) to quantify end effects such axial leakage and more accurate copper losses. A two-stage workflow, 2D BO for exploration and sparse 3D validation for finalists is recommended.
3. **Experimental validation.** Fabrication and bench testing of one or more Pareto-optimal designs would provide the strongest validation of the predicted improvements in torque ripple and efficiency.
4. **Natural-language motor design interface.** Develop an end-to-end CAD–FEM–optimization interface in which the user specifies requirements in natural language (e.g., voltage, torque, envelope, cost and ripple limits), and an LLM automatically generates a valid parametric motor geometry, constraints, and initial priors. The system should compile the prompt into a reproducible machine-readable design specification, launch FEM evaluations, and iteratively refine the geometry through Bayesian optimization with retrieval over prior designs.

BIBLIOGRAPHY

Hanselman, D. C. (1994). *Brushless Permanent-Magnet Motor Design*. New York: McGraw-Hill.

Hanselman, D. C. (2003). *Brushless Permanent Magnet Motor Design* (ed. 2). Cranston, RI: The Writers' Collective.

LIST OF REFERENCES

Belahcen, A., Martin, F., Zaim, M. E.-H., Dlala, E. & Kolondzovski, Z. (2015). Combined FE and Particle Swarm algorithm for optimization of high speed PM synchronous machine. *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 34(2), 475–484. doi: 10.1108/COMPEL-07-2014-0168.

Chang, C.-Y., Azvar, M., Okwudire, C. & Al Kontar, R. [arXiv preprint arXiv:2505.14756]. (2025). LLINBO: Trustworthy LLM-in-the-Loop Bayesian Optimization.

Chawrasia, S. K., Das, A., Chanda, C. K. & Banerjee, S. (2020). Design, analysis and comparative study of Hub motor for an electric bike. *Michael Faraday IET International Summit (MFIIS-2020)*.

Chawrasia, S. K., Das, A. & Chanda, C. K. (2021). Design and Analysis of Electric bike Hub-Motor using Motor-CAD. *2020 3rd International Conference on Energy, Power and Environment: Towards Clean Energy Technologies (ICEPE)*. doi: 10.1109/ICEPE50861.2021.9404475.

Chen, H., Meng, Y., Zhang, Q., Li, D. & Qu, R. (2025). Design and Analysis of a New Asymmetric Consequent-Pole Flux Reversal Dual-PM Excited Machine with Trapezoidal PMs. *IEEE Transactions on Transportation Electrification*, 11(4), 8702–8713. doi: 10.1109/TTE.2025.3544433.

Creux, J., Haje Obeid, N., Boileau, T. & Meibody-Tabar, F. (2022, October). Local Demagnetization Fault Detection in PMASynRM based on Finite Element Modeling and Characterisation. *IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society*. doi: 10.1109/IECON49645.2022.9968829.

Dash, A., Yadav, D., Sahu, A. K., Fuad, M., Sharma, S. & Chatterjee, S. (2023). Design and Analysis of Electric Bike(E-Bike). *2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*. doi: 10.1109/ICSTSN57873.2023.10151644.

Duan, Y. & Ionel, D. M. (2013). A Review of Recent Developments in Electrical Machine Design Optimization Methods With a Permanent-Magnet Synchronous Motor Benchmark Study. *IEEE Transactions on Industry Applications*, 49(3), 1268-1275. doi: 10.1109/TIA.2013.2252597.

Escarela-Perez, R., Niewierowicz, T. & Campero-Littlewood, E. (2001). Synchronous Machine Parameters from Frequency-Response Finite-Element Simulations and Genetic Algorithms. *IEEE Transactions on Energy Conversion*, 16(2), 198–203. doi: 10.1109/60.921473.

Farshbaf Roomi, F., Vahedi, A. & Mirnikjoo, S. (2021). Multi-Objective Optimization of Permanent Magnet Synchronous Motor Based on Sensitivity Analysis and Latin Hypercube Sampling assisted NSGAII. *2021 12th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC)*. doi: 10.1109/PEDSTC52094.2021.9405918.

Guesmia, M. A., Pham, C., Pan, Y.-J., Nguyen, K. K., Al-Haddad, K. & Wang, Q. (2026). AI-Assisted Bayesian Optimization of a Permanent Magnet Synchronous Motor for E-Bike Applications. *Machines*, 14(2), 160. doi: 10.3390/machines14020160.

Gupta, R., Hartford, J. & Liu, B. (2025). LLMs for Bayesian Optimization in Scientific Domains: Are We There Yet? *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 15482-15510.

Guzman, S., Deilami, S., Chen, E., Bruce, T., Schröder, T. & Wang, Y. (2023). Development of a MATLAB–GAMS Framework to Improve the Operation of Power-to-Gas Considering the Real-Time Fluctuating Electricity Price. *Resources*, 12(3), 35. doi: 10.3390/resources12030035.

Hari Krishnan, G., Chandrasekar, P., Prabhu, S., Sudharshan Reddy, L., Chennakesava, M. & Pramod Kalyan, P. (2024). Cogging Torque Reduction Strategies in Six-Phase BLDC Motors: Skewing and Back EMF Tuning for EV Applications. *2024 10th International Conference on Electrical Energy Systems (ICEES)*. doi: 10.1109/ICEES61253.2024.10776880.

Ho, P.-J., Yi, C.-P., Lin, Y.-J., Chung, W.-D., Chou, P.-H., Sie, B.-H. & Yang, S.-C. (2023a). Motor Torque Control of Electric Assist Bike Considering External Resistances. *IECON 2023 – 49th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1–7. doi: 10.1109/IECON51785.2023.10312189.

Ho, P.-J., Yi, C.-P., Lin, Y.-J., Chung, W.-D., Chou, P.-H. & Yang, S.-C. (2023b). Torque Measurement and Control for Electric-Assisted Bike Considering Different External Load Conditions. *Sensors*, 23(10), 4657. doi: 10.3390/s23104657.

Knebl, L., Bárta, J., Bramerdorfer, G. & Vítek, O. (2021). Multi-Objective Optimization of a Line-Start Synchronous Machine Using a Self-Organizing Algorithm. *IEEE Transactions on Magnetics*, 57(6), 1–4. doi: 10.1109/TMAG.2021.3056403. Art. no. 9344813.

Kobalczyk, K., Lin, Z. J., Letham, B., Zhao, Z., Balandat, M. & Bakshy, E. [arXiv preprint arXiv:2510.17671]. (2025). LILO: Bayesian Optimization with Interactive Natural Language Feedback.

Liu, T., Astorga, N., Seedat, N. & van der Schaar, M. (2024). Large Language Models to Enhance Bayesian Optimization. *International Conference on Learning Representations (ICLR)*.

Mao, Y., Niu, S. & Wang, Q. (2021). Design and Optimization of a Slot-PM-assisted Doubly-salient Machine Based on Saturation Assuaging. *Chinese Journal of Electrical Engineering*, 7(3), 65-72. doi: 10.23919/CJEE.2021.000026.

Mohanraj, D., Aruldavid, R., Verma, R., Sathiyasekar, K., Barnawi, A. B., Chokkalingam, B. & Mihet-Popa, L. (2022). A Review of BLDC Motor: State of Art, Advanced Control Techniques, and Applications. *IEEE Access*, 10, 54833-54869. doi: 10.1109/ACCESS.2022.3175011.

Mopari, S. S., Hirekhan, S. R. & Murthy, P. K. (2024). Optimization of Electric Motor Design Using AI-Based Algorithm. *Nanotechnology Perceptions*, 20(7), 2139-2158.

Nasiri-Zarandi, R., Karami-Shahnani, A., Sedigh Toulabi, M. & Tessarolo, A. (2023). Design and Experimental Performance Assessment of an Outer Rotor PM-Assisted SynRM for the Electric Bike Propulsion. *IEEE Transactions on Transportation Electrification*, 9(1), 727–738.

Parouha, R. P. & Verma, P. (2021). State-of-the-Art Reviews of Meta-Heuristic Algorithms with Their Novel Proposal for Unconstrained Optimization and Applications. *Archives of Computational Methods in Engineering*, 28, 4049–4115. doi: 10.1007/s11831-021-09532-7.

Purnata, H., Yusuf, M. & Riyanto, S. D. (2020). Closed Loop System BLDC Motor For Electric-Bike. *2020 International Conference on Applied Science and Technology (iCAST)*. doi: 10.1109/iCAST51016.2020.9557642.

Sayeed, H. M., Soneji, A., Baird, S. & Sparks, T. [ChemRxiv preprint]. (2025). Honegumi RAG Assistant: An Agentic System for Accelerating Bayesian Optimization Adoption in Experimental Sciences.

Sergeant, P., Crevecoeur, G., Dupré, L. & Van den Bossche, A. (2009). Characterization and optimization of a permanent magnet synchronous machine. *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 28(2), 272-285. doi: 10.1108/03321640910929218.

Topalis, P., Schieseck, M. & Gehlhoff, F. (2025). LangBO: A Framework for Language-Guided Prior Integration in Bayesian Optimization. *2025 IEEE 30th International Conference on Emerging Technologies and Factory Automation (ETFA)*. doi: 10.1109/ETFA65518.2025.11205749.

Wu, Y., Huang, H., Zou, T., Ren, X. & Gerada, C. (2024). Design of a 1200 Nm High Torque Density In-Wheel Magnetic Geared Motor for Electrical Vehicles. *2024 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pp. 1–6. doi: 10.1109/VPPC63154.2024.10755333.

Yang, C., Liu, W., Niu, S., Lyu, J. & Chau, K. T. (2025a). Parameter-Tuning-Free Two-Step Identification of Mechanical Parameters for PMSM Drives. *IEEE Transactions on Industrial Electronics*, 72(12), 12378–12392. doi: 10.1109/TIE.2025.3577312.

Yang, Z., Wang, D., Ge, L., Wang, B., Fu, T. & Li, Y. [arXiv preprint arXiv:2505.12833]. (2025b). ReasoningBO: Enhancing Bayesian Optimization with the Long-Context Reasoning Power of LLMs.

Zhang, S., Yan, H., Yang, L., Zhao, H., Du, X. & Zhang, J. (2024). Optimization Design of Permanent Magnet Synchronous Motor Based on Multi-Objective Artificial Hummingbird Algorithm. *Actuators*, 13(7), 243. doi: 10.3390/act13070243.

Zhang, Y., Choong, J. J., Madhawa, K. & Ozawa, K. (2025). AutoLead: An LLM-Guided Bayesian Optimization Framework for Multi-Objective Lead Optimization. *bioRxiv*. doi: 10.1101/2025.08.19.671029.

Zhao, W., Zhang, Z., Mirjalili, S., Wang, L., Khodadadi, N. & Mirjalili, S. M. (2022). An Effective Multi-Objective Artificial Hummingbird Algorithm with Dynamic Elimination-Based Crowding Distance for Solving Engineering Design Problems. *Computer Methods in Applied Mechanics and Engineering*, 398, 115223. doi: 10.1016/j.cma.2022.115223.

Zhao, Y., Liang, Z., Li, D. & Qu, R. (2025). Design Considerations of Spoke-Array PM Vernier Machine for Different Power Levels. *IEEE Transactions on Industry Applications*, 61(4), 6149–6159. doi: 10.1109/TIA.2025.3545017.

Zheng, S., Zhu, X., Xu, L., Xiang, Z., Quan, L. & Yu, B. (2022). Multi-Objective Optimization Design of a Multi-Permanent-Magnet Motor Considering Magnet Characteristic Variation Effects. *IEEE Transactions on Industrial Electronics*, 69(4), 3428-3438. doi: 10.1109/TIE.2021.3073311.

Zhu, Z. Q. & Howe, D. (1993a). Instantaneous Magnetic Field Distribution in Brushless Permanent Magnet DC Motors, Part II: Armature-Reaction Field. *IEEE Transactions on Magnetics*, 29(1), 136-142.

Zhu, Z. Q. & Howe, D. (1993b). Instantaneous Magnetic Field Distribution in Permanent Magnet Brushless DC Motors, Part IV: Magnetic Field on Load. *IEEE Transactions on Magnetics*, 29(1), 152-158.

Zhu, Z. Q. & Howe, D. (1993c). Instantaneous Magnetic Field Distribution in Brushless Permanent Magnet DC Motors, Part III: Effect of Stator Slotting. *IEEE Transactions on Magnetics*, 29(1), 143-151.

Zhu, Z. Q., Howe, D., Bolte, E. & Ackermann, B. (1993). Instantaneous Magnetic Field Distribution in Brushless Permanent Magnet DC Motors, Part I: Open-circuit Field. *IEEE Transactions on Magnetics*, 29(1), 124-135.