Melika Baghi

mbaghi3@gatech.edu | melikabaghi.github.io | Google Scholar |
linkedin.com/in/melika-baghi | github.com/melikabaghi

Research Interests

Reduced-order and surrogate modeling for high-dimensional PDE systems; geometry-aware learning on Riemannian manifolds (Grassmann, Stiefel); uncertainty quantification and calibrated inference; representation learning and multimodal learning; generative and diffusion-based models for scientific systems; statistical learning theory; continual and sequential learning; applications in aerospace engineering, computational fluid dynamics, and computational biomedicine.

Education

Georgia Institute of Technology *Atlanta, GA*
Ph.D. in Industrial & Systems Engineering — System Informatics & Control *Aug 2023 – Present*

- Advisors: Prof. Kamran Paynabar and Prof. Xiao Liu
- Dissertation: *Geometry-Aware Learning Frameworks for Reduced-Order Modeling with Uncertainty Quantification in High-Dimensional PDE Systems*

Georgia Institute of Technology *Atlanta, GA*
M.S. in Statistics *Aug 2023 – May 2026*

- Research focus: *Reduced-Order Modeling with Uncertainty Quantification for High-Dimensional PDE Systems in Digital Twin Applications*
- Advisors: Prof. Kamran Paynabar and Prof. Xiao Liu

Amirkabir University of Technology (Tehran Polytechnic) *Tehran, Iran*
M.S. in Industrial Engineering — Healthcare Systems; GPA: 4.0 / 4.0 (19.35 / 20) *Oct 2021 – Jul 2023*

- Thesis: *Scheduling of Medical Tourists Visiting Medical Centers Using an Agent-Based Approach*
- Advisor: Prof. Hadi Mosadegh
- Direct-entry admitted as an elite student (ranked 3rd of 120 applicants)

Amirkabir University of Technology (Tehran Polytechnic) *Tehran, Iran*
B.S. in Industrial Engineering; GPA: 3.91 / 4.0 (18.65 / 20) *Oct 2018 – Jun 2021*

Amirkabir University of Technology (Tehran Polytechnic) *Tehran, Iran*
B.S. in Aerospace Engineering; GPA: 3.79 / 4.0 (18.14 / 20) *Sep 2016 – May 2020*

- Advisor: Prof. Mehdi Sabzeh Parvar
- Selected as an elite student and approved to pursue a dual degree in Aerospace Engineering (2018)

Publications and Manuscripts

Preprints and Manuscripts Under Review

[1] **M. Baghi**, X. Liu, K. Paynabar. "Constrained Extreme Gradient Boosting for Adapting Reduced-Order Models." *Manuscript under review.* [arXiv:2605.04130] [code]

- Proposes cXGBoost, a gradient-boosting ensemble that learns injective mappings from the Grassmann manifold to a constrained Euclidean space via a quadratically constrained quadratic program (QCQP) at each tree split. Demonstrated on cylinder wake flows, wave propagation, Burgers equation, and Euler–Bernoulli beam dynamics; reduces high-Re reconstruction error by 5× relative to Grassmann interpolation baselines.

[2] **M. Baghi**, H. Mosadegh. "Reducing Waiting Time for Medical Tourists Through Hybrid Agent-Based and Discrete-Event Simulation: A Hospital Case Study." *Under review, Health Care Management Science.* [arXiv:2605.19139] [code] [DOI]

- Develops a hybrid ABS+DES model of an international patient department in AnyLogic. A 256-run two-level fractional factorial design over 16 factors identifies bed capacity, patient-priority rules, channel design, and clinic slot intervals as dominant levers. Against an otherwise identical DES-only counterpart, the hybrid model reduces the average medical-tourist hospital-queue waiting time from 13.7 days to 2.4 days (82%); compatible-section bed sharing reduces the mean inpatient admission queue from 25.9 to 14.8 days (43%, 30 randomized-seed replications, $p < 0.001$).

[3] **M. Baghi**. "Guaranteed Adaptive Modality Acquisition: When the Policy Chooses Its Own Calibration Group." *Manuscript under review.* [arXiv:2608.15520] [code]

- Certifies a prediction when the modalities already observed determine the decision, and otherwise abstains and recommends which modality to acquire next. Provides finite-sample guarantees conditional on the terminal modality pattern created by adaptive acquisition, so coverage holds for every stopping pattern rather than only on average. Evaluated across four multimodal datasets spanning ECG, speech and wearable sensing (PTB-XL, CMU-MOSEI, MHEALTH, IEMOCAP), with reproducible experiments and public code.

[4] **M. Baghi**, X. Liu, K. Paynabar. "Tensor-Based Reduced Order Model." *Manuscript under review.*

- Develops a surrogate modeling framework for parameterized, tensor-valued spatiotemporal simulation outputs. Combines Tucker decomposition, product-Grassmann interpolation of mode-wise factor subspaces, orthogonal Procrustes core alignment, and a Core-of-Cores low-rank surrogate. Achieves 35.6×–564× complexity reduction and 4×–108× wall-clock speedup over flattened Grassmannian baselines across four PDE benchmarks.

[5] **M. Baghi**, X. Liu, K. Paynabar. "Supervised Active-Subspace Input Reduction and Risk-Controlled Field Uncertainty for Tensor-Valued Computer Experiments." *Manuscript under review.*

- Reduces tensor-valued simulation inputs through supervised active subspaces, represents the output via a higher-order decomposition, and places a decomposition-based Gaussian process on the reduced map to deliver field-level predictive uncertainty, with finite-sample subspace-estimation guarantees via matrix perturbation theory (Wedin / Davis–Kahan).

Workshop Papers

[6] **M. Baghi**. "Pooling and Drift in Delayed Bandits." *ML×OR 2026 Workshop (non-archival).* [arXiv:2609.01761] [code]

- An action produces an intermediate state at once while its loss arrives after a fixed delay. State-EXP3 charges one delayed outcome to every action that could have produced that state; the estimation cost is an effective dimension reading the overlap between rows of the action-to-state matrix rather than counting actions or states. Adds the delayed and noisy setting, the coarsening trade-off, and a drift lower bound showing the cost of delayed feedback cannot be removed by pooling. Reduces regret by up to 79% against action-level weighting and by 32–68% against a rate-optimal action-only delayed-bandit baseline in synthetic experiments.

Working Papers and In Preparation

[7] **M. Baghi**, K. Paynabar, X. Liu. "Reliability-Calibrated Stability for Continual Multimodal Representation Learning." *Research proposal; manuscript in preparation.*

- Argues that uniform stability in continual multimodal learning is misspecified when historical cross-modal alignment mixes semantic shared structure with spurious correlations. Proposes a perturbation-based spectral-block reliability estimator on the Grassmannian and a reliability-weighted projection operator for the alignment head, with an excess-risk decomposition identifying the cost of preserving nuisance blocks vs. releasing shared blocks.

[8] "Reference-Measure Conformal Prediction for Subspace-Valued Outputs." *Manuscript in preparation.*

- Distribution-free conformal prediction sets for outputs that are subspaces rather than vectors, built against a reference measure on the Grassmannian.

[9] "Topological Conformal Risk Control: Distribution-Free Uncertainty for Persistent Features." *Manuscript in preparation.*

- Conformal risk control for persistent homology features, giving distribution-free guarantees on topological summaries.

Research Experience

Georgia Institute of Technology *Atlanta, GA*
Ph.D. Research Assistant — Prof. Kamran Paynabar & Prof. Xiao Liu *Aug 2023 – Present*

- **Geometry-Aware Reduced-Order Modeling for Parametric PDE Systems.** Developed novel statistical learning frameworks for constructing fast surrogates of computationally expensive high-fidelity simulations parameterized by PDEs. Proposed stable injective mappings from the Grassmann manifold to a constrained Euclidean space, enabling classical regression and uncertainty quantification over spaces of orthogonal subspaces. Integrated gradient boosting trees (cXGBoost) and projected Gaussian processes as hybrid predictors; demonstrated stability under nonlinear parameter variation on CFD and FEA benchmarks including cylinder wake dynamics, wave propagation, and shock-dominated flows, reducing high-Re reconstruction error by 5× against Grassmann interpolation baselines. Open-source implementation: `github.com/melikabaghi/cXGBoost`.
- **Tensor-Structured Surrogate Modeling on Product Grassmann Manifolds.** Developed a surrogate framework preserving the native multiway structure of spatiotemporal simulation fields. Combines Tucker decomposition of each training snapshot, product-Grassmann interpolation of mode-wise factor subspaces, orthogonal Procrustes core alignment, and a Core-of-Cores low-rank model for aligned-core variation. Achieves 35.6×–564× interpolation complexity reduction and 4×–108× wall-clock speedup vs. flattened Grassmannian baselines while matching accuracy within 0.35 percentage points across four PDE benchmarks.
- **Reliability-Calibrated Continual Multimodal Representation Learning.** Identified that uniform stability in continual multimodal learning can entrench spurious cross-modal shortcuts. Proposed reliability-calibrated stability: stores a compact spectral-block memory of cross-modal covariance, estimates block reliability via perturbation stability on the Grassmannian, refines reliability when a new modality arrives using cross-covariance evidence, and applies reliability-weighted gradient projection to the alignment head. Derived an excess-risk decomposition under a linear-Gaussian model with block subspace separation, perturbation informativeness, and eigengap assumptions.
- **Directed Projects (Georgia Tech, 2023–2025).** In the course of doctoral study, completed three supervised applied projects under Profs. Paynabar and Liu: (i) *Sensitivity-Guided Diffusion Models* — designed diffusion-based generative models constrained to physically meaningful latent subspaces via sensitivity-informed perturbation schedules, implemented in PyTorch; (ii) *Multimodal Phenotype Modeling* — built representation learning pipelines integrating single-cell transcriptomics, longitudinal EHR data, and imaging-derived phenotypes using contrastive learning and transformer-based phenotype extraction from clinical notes; (iii) *Drug Safety and Clinical Risk Modeling* — developed a temporally-aware adverse drug event detection framework from longitudinal EHR data using regularized regression, tree ensembles, and neural embeddings, with calibration and fairness analyses across demographic strata.

Amirkabir University of Technology *Tehran, Iran*
Graduate Research Assistant — Prof. Hadi Mosadegh *Aug 2021 – Jul 2023*

- **Simulation-Based Scheduling Optimization via Hybrid ABS+DES and Machine Learning.** Modeled patient treatment journeys via a hybrid discrete event simulation (DES) and agent-based simulation (ABS) framework in AnyLogic. Represented registration, consultation, admission, bed allocation, and discharge through DES; represented patient, physician, and ward behaviors through agent-based logic including medication adherence, dropout, doctor switching, and emergency escalation. Conducted a 256-run two-level fractional factorial design over 16 operational factors. Against the DES-only baseline, reduced medical-tourist hospital-queue waiting time by 82% (13.7 to 2.4 days) and the mean inpatient admission queue by 43% (25.9 to 14.8 days).

Amirkabir University of Technology *Tehran, Iran*
Undergraduate Research Assistant — Prof. Mehdi Sabzeh Parvar *May 2018 – May 2020*

- **Learning-Based Safety Region Estimation for Nonlinear Dynamical Systems.** Developed a six-degree-of-freedom physics-based simulator in MATLAB/Simulink using system velocity equations and operational constraints, and generated training and test datasets across multiple operating scenarios. Implemented machine learning models in Python, including a three-layer neural network, to predict system behavior and expand the feasible operating range under hybrid inputs; studied nonlinear behavior and distributional robustness under input uncertainty.

Teaching Experience

Georgia Institute of Technology *Atlanta, GA*
Head Teaching Assistant — ISyE 6525: High-Dimensional Data Analytics (Prof. Paynabar)Jan 2024 – Present

- Lead and coordinate a TA team for a graduate course of 150+ students per semester covering nonparametric methods, dimensionality reduction, sparse modeling, manifold learning, regularized regression (OLS, Ridge, LASSO, Adaptive LASSO, Elastic Net), and high-dimensional inference.
- Design advanced problem sets and final projects emphasizing theoretical understanding, reproducible ML workflows, and model validation under distribution shift.
- Mentor student research projects on uncertainty quantification and interpretability; conduct weekly office hours and lead review sessions.
- Recognized as **Outstanding Online TA** (2025, Tech to Teaching Program, Georgia Institute of Technology).

Amirkabir University of Technology *Tehran, Iran*
Teaching Assistant — Principles of Simulation, Industrial Engineering (Prof. Abbas Ahmadi)Sep 2022 – Aug 2023

- Provided hands-on instruction in Python, MATLAB, Simulink, Arena, and AnyLogic. Led review sessions on Monte Carlo methods, discrete event simulation, and stochastic process modeling. Graded assignments and examinations.

Amirkabir University of Technology *Tehran, Iran*
Teaching Assistant — Engineering Dynamics, Aerospace Engineering (Prof. Hamid Reza Ovesy)Feb 2020 – Aug 2023

- Delivered weekly recitations on analytical mechanics, Lagrangian dynamics, and stability analysis. Led office hours and graded assignments and exams.
- **Ranked #1 Teaching Assistant in the department** by student evaluation for three consecutive years.

Professional Experience

Royan Institute *Tehran, Iran*
Course Instructor — Artificial Intelligence Laboratory *Aug 2022 – Aug 2023*

- Designed and delivered a project-based AI curriculum covering rule-based expert systems, hybrid architectures, and applied machine learning in clinical healthcare contexts. Supervised end-to-end student projects from problem formulation through evaluation.

Arvand Pishro (Automotive R&D) *Safadasht, Iran*
Research Intern — R&D Division *Sep 2018 – Nov 2019*

- Designed and validated a closed-loop fuel tank cooling and monitoring system integrating sensor feedback and automated temperature control. Collaborated with multidisciplinary engineering teams to translate requirements into validated prototype systems.

Honors and Awards

- **Most Innovative Approach Award**, NSF Future Manufacturing Data Challenge 2026
- **Conference Fund Award**, Student Government Association, Georgia Institute of Technology — with College of Engineering matching funds 2026
- **Third-year Progress Travel Award**, H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology 2026
- **Travel Award**, Institute for Mathematical and Statistical Innovation (IMSI) — an NSF-funded institute 2025
- **Outstanding Online TA Award**, ISyE 6525, Tech to Teaching Program, Georgia Institute of Technology 2025
- **Conference Fund Award**, Student Government Association, Georgia Institute of Technology 2024
- **Stewart Topper Fellowship**, Georgia Institute of Technology 2023
- **First-year Fellows Travel Award**, H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology 2023
- **Mansour Doctoral Fellowship**, North Carolina State University — granted for exceptional academic promise in Industrial Engineering 2023
- Direct entry to M.S. in Industrial Engineering as an elite student, ranked **3rd of 120** applicants, Amirkabir University of Technology 2021
- Direct entry to M.S. in Aerospace Engineering as an elite student, ranked **2nd of 120** applicants, Amirkabir University of Technology 2021
- Selected as an elite student and approved to pursue a dual degree in Aerospace Engineering, Amirkabir University of Technology 2018
- Ranked in the **top 1%** among 150,000+ participants in the National University Entrance Exam (Physics & Mathematics track) 2016

Technical Skills

- **Programming:** Python, PyTorch, NumPy, SciPy, pandas, scikit-learn, PySpark, R, MATLAB, SQL, Linux, Git

- **Machine Learning:** Deep learning, diffusion generative models, gradient boosting (XGBoost), Gaussian processes, contrastive and representation learning, multimodal learning, LLM-based information extraction, ensemble methods
- **Mathematics & Theory:** Optimization on manifolds, Riemannian geometry (Grassmann, Stiefel, product manifolds), statistical learning theory, high-dimensional inference, linear algebra, tensor decompositions (Tucker, HOSVD), stochastic processes, probabilistic modeling
- **Scientific Computing:** Reduced-order modeling, surrogate modeling, computational fluid dynamics (CFD), finite element analysis (FEA), Lattice Boltzmann methods, Monte Carlo simulation, tensor methods
- **Simulation:** AnyLogic (ABS + DES), Arena, Simulink, Ansys Fluent, GAMS

Selected Academic Projects

- **Predictive Maintenance and Anomaly Detection in Industrial Systems:** Developed an end-to-end machine learning pipeline for sensor-based fault detection, integrating time-series forecasting, feature engineering, and real-time monitoring. Applied ARIMA and LSTM models for trend prediction and ensemble classifiers for anomaly detection in multivariate telemetry; achieved over 15-hour advance fault warnings through hybrid Weibull–state-space degradation modeling with online retraining. Deployed the workflow using Spark for scalable, low-latency inference across industrial sensor networks. *Python, PySpark* (2025)
- **Sensitivity-Guided Diffusion Models for Scientific Dynamical Systems:** Designed diffusion-based generative models constrained to physically meaningful latent subspaces derived from spectral decomposition and sensitivity analysis. Derived sensitivity-informed directional perturbation schedules guiding the forward SDE along high-variance manifold directions to improve sample efficiency and structural fidelity in data-scarce scientific regimes. *Python, PyTorch* (2024–2025)
- **Multimodal Phenotype Modeling for Translational Disease Discovery:** Constructed representation learning pipelines integrating single-cell transcriptomics, longitudinal EHR data, and imaging-derived phenotypes for disease-relevant biological variation modeling. Built structured cellular and patient embeddings using contrastive representation learning and dimensionality reduction; leveraged transformer-based language models for structured phenotype extraction from unstructured clinical notes. *Python, PyTorch* (2024–2025)
- **Machine Learning for Drug Safety and Clinical Risk Modeling:** Developed an adverse drug event detection framework using temporally-aware patient embeddings integrating medication trajectories, laboratory values, and comorbidity profiles from longitudinal EHR data. Applied regularized regression, tree-based ensembles, and neural representation-learning models; conducted calibration, fairness, and robustness analyses across demographic strata. *Python* (2024–2025)
- **Tensor-Structured Learning Pipelines for High-Dimensional Scientific Data:** Designed deep and tensor-structured learning architectures for parameterized scientific fields; investigated representation learning under structural and geometric constraints. *Python, PyTorch* (2024)
- **Internet Access and Female Economic Empowerment: A Cross-National Nonparametric Analysis:** Studied the relation between national internet usage and two indicators of female economic empowerment—employment rate and share of senior and middle management positions—across 92 countries using World Bank Gender Data Portal indicators for 2022. Classified nations by internet access ($\geq$50% vs. $<$50% of population), assessed normality via Shapiro–Wilk and Anderson–Darling tests with histogram and Q–Q diagnostics, and applied two Mann–Whitney U tests with a Bonferroni correction. *Group project*
- **Anomaly Detection via Sliding Nest Window Control Charts:** Developed statistical monitoring frameworks using ARMA models and sliding window control charts for early anomaly detection in industrial sensor streams. *Python* (2023)
- **Financial Loan Default Prediction:** Built and evaluated ensemble classifiers with calibration analysis

for credit risk modeling. *Python* (2023)

- **Food Recommender Expert System:** Designed and implemented a knowledge-based expert system tailored to individual user health characteristics. *Python* (2021–2022)
- **Breast Cancer Prediction Benchmark:** Evaluated and compared machine learning classification methods for disease prediction with emphasis on statistical validity and generalization. *Python* (2021–2022)
- **Airfoil Flow Simulation and CFD Analysis:** Numerical analysis of transonic flow over RAE airfoils; studied nonlinear PDE solution behavior under varying physical parameters. *Ansys Fluent* (2021–2022)
- **Six-Degree-of-Freedom Dynamics Simulator:** Designed a full six-degree-of-freedom physics-based simulator for system identification and control analysis under hybrid inputs. *MATLAB, Simulink* (2018–2020)
- **Nurse Scheduling with Quick-Response Methods:** Formulated and solved mixed-integer programming models for healthcare workforce scheduling. *GAMS* (2020–2021)
- **Healthcare Operations Quality Improvement:** Applied Six Sigma methodology to identify and reduce service quality deficiencies in a healthcare organization. (2021–2022)

Service and Outreach

- Core team member, CANSAT satellite design competition, Amirkabir University of Technology 2019
- Staff organizer, National AerUp Competition, Amirkabir University of Technology 2017, 2018
- Referee, National Aerospace Competitions, Aerospace Engineering Department, Amirkabir University of Technology 2017, 2018

Languages

- **Persian:** Native
- **English:** Excellent (TOEFL iBT 110 / 120: Reading 28, Listening 30, Speaking 25, Writing 27)
- **Korean:** Excellent spoken proficiency