

Sujosh Nag

Durham, NH | nagsujosh2004@gmail.com | sujoshnag.com | linkedin.com/in/sujoshnag | github.com/nagsujosh

Research Interests

Safe and reliable control of long-horizon manipulation by learned visuomotor policies, built on latent dynamics that carry stability guarantees.

Education

University of New Hampshire

B.S. in Computer Science

Durham, NH

Expected May 2027

- **Honors:** Dean's Scholar 2023–2026; International Engagement & Service Scholar.
- **Relevant Coursework:** Reinforcement Learning, Machine Learning, Artificial Intelligence, Foundations of Neural Networks, Mobile Robotics, Linear Algebra, Applied Regression Analysis, Numerical Methods, Statistics for Engineers.

Publications

- Riad Ahmed, **Sujosh Nag**, Moniruzzaman Akash, Mostafa Hussein, and Momotaz Begum. “**Trajectory-Consistent Flow Matching for Robust Visuomotor Policy Learning.**” IEEE/RSJ International Conference on Intelligent Robots and Systems (**IROS 2026**). arXiv:2605.08511.
Flow-matching policies optimize a pointwise velocity objective but integrate that field at inference, and the mismatch compounds along the trajectory. The paper closes it with auxiliary rectified-flow velocity regression, multi-step consistency training on the integrated displacement of trajectory segments, velocity-field regularization for temporal smoothness, and Runge-Kutta integration in place of Euler; the ablation shows none of the four suffices alone.

Manuscripts Under Review

- **Sujosh Nag**, Kaito Namatame, Riad Ahmed, Moniruzzaman Akash, Meisam Khorshidi, and Momotaz Begum. “**MoK: A Mixture of Koopman Operators for Learning Visuomotor Policies.**” Under review.
MoK trains a bank of Koopman operators per task and learns to compose them with no segmentation labels: a sparse gate weights the bank at each step, and the selected operators apply in sequence, so composition order itself separates one dynamical regime of a skill from the next, where prior policies fit a single operator to the whole task. Every composition the gate can form is contractive by construction.
- **Sujosh Nag** and Momotaz Begum. “**Where Should a Skill End? Discovering Skills from How the Robot's Action Moves Objects.**” Under review.
A planner composes skills, so a skill boundary should be set by what a planner reasons about, the effect of an action on objects; boundaries read off appearance or arm motion instead fall wherever the arm moves, separating a reach from a retreat that both leave the object at rest. The paper takes a skill to be an effect that recurs across demonstrations, and merges two of them whenever one model of their effect predicts held-out demonstrations about as well as two.
- Moniruzzaman Akash, **Sujosh Nag**, Sajay Arthanat, and Momotaz Begum. “**Hand2Robot: An End-user Friendly Demonstration Collection Framework to Facilitate Home Deployment of Imitation Learning.**” Under review.
Converting egocentric human video into robot training data is by now established; H2R supplies what bare-hand recordings usually leave out. Six grasp-type-specific contact frames place the parallel jaws for precision and power grasps alike, a recognizer predicts the manipulation phases that time grasp and release, and calibrated surface electromyography yields a grip-force target, so the exported policy commands how firmly to hold an object as well as where to move it.
- Riad Ahmed, **Sujosh Nag**, and Momotaz Begum. “**MIND: Mechanistic Interpretability for Neural Diagnosis of Visuomotor Policies.**” Under review.
MIND trains a sparse autoencoder on the activations of a vision-language-action policy and names each basis feature by a manipulation concept that holds whenever it fires, retaining only names that survive held-out episodes and a closed-loop clamp. A linear monitor over those features flags a failing execution and attributes the alarm to the concepts responsible.

Research Experience

Cognitive Assistive Robotics Lab (CARL), University of New Hampshire

Undergraduate Researcher / Advisor: Prof. Momotaz Begum

Durham, NH

August 2025 – Present

- Train and evaluate **visuomotor manipulation policies** (ACT, Diffusion Policy, DP3, flow matching, $\pi_0/\pi_{0.5}$, Open-VLA, NVIDIA Isaac GR00T) in simulation and on hardware, and build the surrounding infrastructure for demonstration collection, training, closed-loop rollout, and failure analysis.
- Run the lab's real-robot studies on **Franka FR3**, **Boston Dynamics Spot**, **Hello Robot Stretch 3/4**, and **ABB YuMi**, spanning single-arm, mobile-manipulation, and bimanual settings.
- Built and maintain the **Franka + Deoxys** stack for teleoperation, demonstration collection, and policy evaluation, with SpaceMouse and Meta Quest interfaces, multi-camera rigs, and **Inspire dexterous-hand** control driven by Quest hand tracking.
- Supported an **NIH R01 assistive-robotics deployment** in a participant's home.

Independent Research Projects

Structured Belief-Guided Diffusion Policy

Conditioned a diffusion policy on a belief over task phase and progress predicted from image observations, taking as given the skill boundaries that define those phases. The belief helped in Meta-World and on real Franka tasks, yet under temporal perturbation the policy recovered less reliably than an unconditioned diffusion policy. That negative result moved my attention upstream, to where skill boundaries should come from at all.

Belief-PDDL: Neuro-Symbolic Planning Under Partial Observability

Planning framework that maintains factorized Bayesian beliefs over grounded predicates, projects them through CP-SAT constraints to drop inconsistent states, and plans over Top-K feasible worlds with active sensing; robust to heavy observation noise in Blocksworld and ALFWorld.

Course Research

Semantic Observation Repair for Decentralized Multi-Agent Path Finding

GitHub | AI course

Inference-time repair of corrupted observations for a frozen decentralized PPO policy in POGEMA, across eight corruption types (noise, dropout, bit flips, FGSM/PGD). Over 4,000 episodes, semantic projection to valid observation grids raised mean success from **68.8% to 88.2%** and recovered FGSM success from **0% to 97.5%** without retraining.

AAPR: Adversary-Adaptive Privacy-Preserving Representations

GitHub | ML course

Suppressed biometric attributes (speaker identity, gender) in learned representations while keeping task utility, via a variational information bottleneck filter trained through a gradient reversal layer against a periodically re-initialized adversary. Gender leakage fell to chance but identity stayed fully recoverable.

Teaching Experience

Teaching Assistant – CS 733: Mobile Robotics

University of New Hampshire

Fall 2026

Teach students **ROS 2** in simulation and on the **Hello Robot Stretch 3**, covering robot programming, mapping, localization, navigation, and AprilTag-based perception.

Teaching Assistant – CS 518: Introduction to Software Engineering

Spring 2025

Supported team software projects and assignments, answered student questions, and graded coursework.

Professional Experience

Sustainability Roundtable Inc.

Boston, MA

Software Engineer Intern

June 2025 – August 2025

Built production features for the SBER 3.0 platform in React and TypeScript, plus internal automation that cut manual workflow time by roughly 75%.

University of New Hampshire InterOperability Laboratory

Durham, NH

Senior Developer

October 2023 – March 2025

Built test automation for the lab's **NVMe conformance test suite**, which qualifies datacenter SSDs and NVMe-oF targets for the NVM Express Integrator's List: contributed 70+ conformance tests across NVMe and its PCIe transport, resolved 120+ issues, led a test-result comparison tool, and mentored junior developers.

Technical Skills

Robot Learning: Diffusion Policy, DP3, flow matching, $\pi_0/\pi_{0.5}$, OpenVLA, NVIDIA Isaac GR00T, Koopman models, sparse autoencoders, PointNet, R3M, DINOv2, SAM2.

Robotics & Simulation: ROS 2, Deoxys, MuJoCo, robosuite, robomimic, Meta-World, SpaceMouse/Meta Quest teleoperation.

Perception & Data: RealSense RGB-D, point clouds, multi-camera calibration, AprilTags, iPhone ARKit/LiDAR, IMU, HDF5/Zarr datasets.

Software: Python, PyTorch, C++, Swift, Docker, Linux, Git.

Leadership: Founder & Chair, Desi Student Association, UNH; Treasurer, Bangladeshi Student Association, UNH.