

Dynamics-Compliant Trajectory Diffusion for Super-Nominal Payload Manipulation [SUPPLEMENTARY MATERIAL]

Anonymous Author(s)

Affiliation

Address

email

1 Other Design Considerations

Input Normalization. The trajectories in our training set are normalized to the range $[-1, 1]$ to improve training stability. Given that payload masses in our dataset, $m \in \mathcal{D}$, are bounded within well-defined minimum and maximum values, we apply an analogous normalization transformation to map these payload values to $[-1, 1]$. This normalization ensures trajectory and payload features exist in comparable numerical ranges, which empirically improves training stability and convergence by preventing scale mismatches in the network’s learning dynamics.

Conditioning Mechanism. We consider two global conditioning mechanisms to incorporate payload encoding and diffusion timestep information into ϵ_θ [1]. The first approach uses additive conditioning, where the global feature vector formed by a concatenation of the payload encoding and diffusion timestep embedding is projected through a learned embedding layer and added to the output of each convolutional block. The second approach implements an affine transform through Feature-wise Linear Modulation (FiLM) [2], where the global feature vector determines scale and bias parameters that modulate outputs from the convolutional layers. This method provides more expressive control over the generated trajectories due to its ability to selectively amplify or attenuate features based on payload requirements, thereby giving higher planning success rates relative to the additive approach.

2 Evaluation

2.1 Evaluation Hardware

All model training and evaluations are performed on a system with an Intel Xeon W5-2445 processor (3.10 GHz), 64 GB of RAM, and a single NVIDIA GeForce RTX 4090 GPU.

2.2 Problem Domain and Experimental Setup

We evaluate our approach on tabletop pick-and-place tasks using dumbbells as payloads to enable controlled experimentation. Our trained models and experiments assume that the payload’s center of mass corresponds to its geometric center, and that the payload is rigidly attached to the robot’s end-effector. The evaluation comprises 500 planning problems in a tabletop manipulation scenario using the 7DoF Franka Emika Panda robot arm, which has a manufacturer-specified nominal payload capacity of 3kg. Each test scenario consists of randomly sampled start and goal joint configurations, with the end-effector positioned above a table at a fixed height ($z = 0.2m$) for each terminal state

Payloads used for real-world evaluation

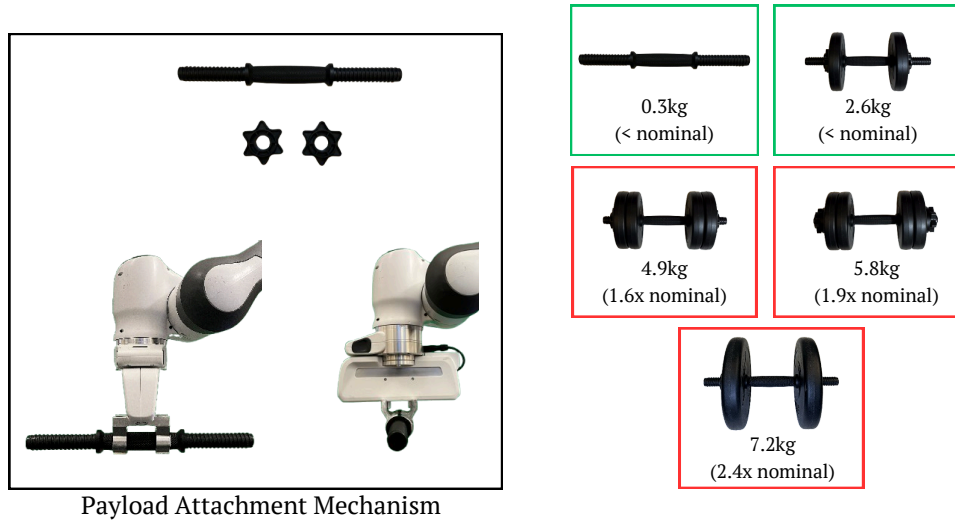


Figure 1: **Red** indicates super-nominal payloads. **Green** indicates payloads within the nominal capacity. See attached video for demos.

Objects used for real-world experimentation

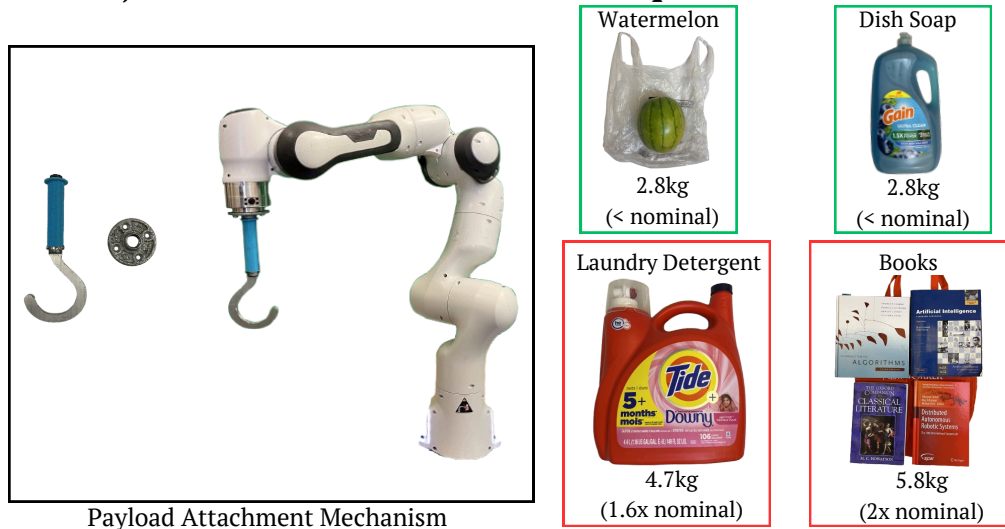


Figure 2: **Red** indicates super-nominal payloads. **Green** indicates payloads within the nominal capacity. See attached video for demos.

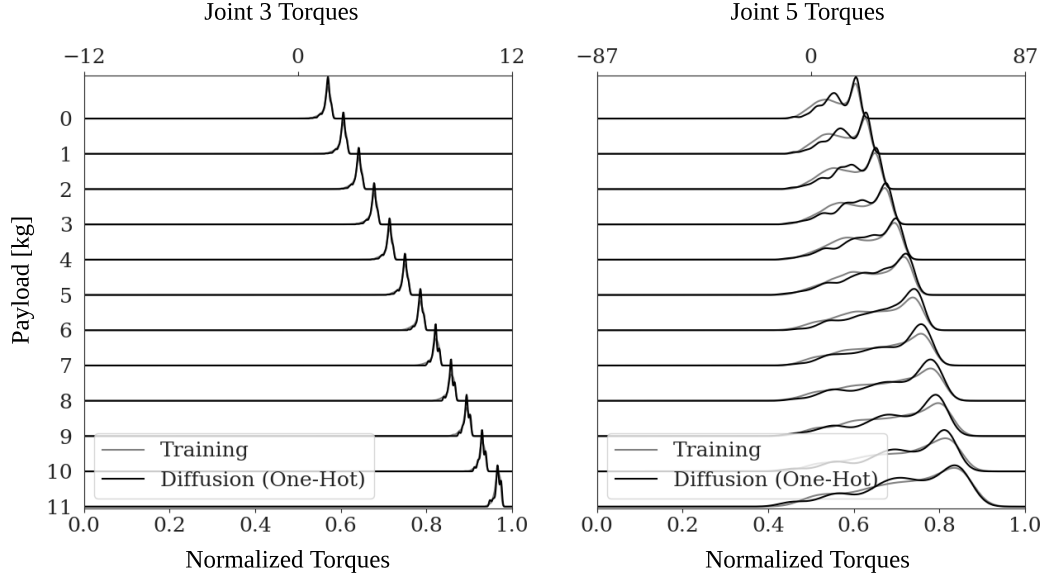


Figure 3: Distribution of joint torques generated by one-hot diffusion at varying payload masses, shown for the elbow (joint 3) and wrist (joint 5). The wrist joint exhibits broader torque distributions compared to the elbow joint, reflecting its position in the kinematic chain.

and oriented downward to facilitate object grasping from the table surface. We initialize all start and goal states with zero joint velocities and accelerations.

2.3 Super-Nominal Payload Manipulation as a Novel Capability

Torque Distribution. The ridgeline plots in Figure 3 demonstrate systematic changes in torque distributions for the elbow (joint 3) and wrist (joint 5) with increasing payload mass. Both joints exhibit distributions that shift progressively toward more positive torques, consistent with their role in counteracting gravitational forces on the payload. The wrist joint displays notably broader distributions compared to the elbow joint across all payload conditions, likely due to its position further along the kinematic chain where it experiences the compound effects of proximal joint motions and must operate with smaller torque limits.

The diffusion model generates torque distributions that closely align with the training data distribution across all joints, including joint 3 and 5, and this consistency holds across different payload masses. This alignment suggests two key insights. First, the model has successfully learned to replicate the statistical properties of valid trajectories, capturing the underlying dynamics constraints and super-nominal operational behavior. Second, diffusion maintains stable performance under more constrained conditions as payload mass increases, especially with limited availability of valid training trajectories that satisfy torque constraints for heavier payloads.

Reachable Workspace and Joint Failures. The heatmaps in Figure 4 and Figure 6 visualize the robot’s accessible workspace in the x-y plane at a fixed height of $z = 0.2m$, with each pixel representing a $1.73 \times 1.73cm$ bin (area = $2.98 cm^2$). The heatmap intensity indicates the success rate of generating dynamically-feasible trajectories for end-effector poses starting or ending at each spatial location. This fine-grained spatial resolution reveals how the robot’s operational boundaries progressively contract under increasing payload conditions, quantifying the relationship between load capacity and accessible workspace. The reachable workspace area at nominal capacity (3kg) spans $1.578 m^2$, reducing to $1.455 m^2$ at $2\times$ nominal load (6kg)—a 92.2% retention. At $3\times$ nominal capacity (9kg), the robot maintains a reachable area of $1.066 m^2$ or 67.6% of its nominal workspace. This demonstrates significant workspace preservation even under substantial payload increases.

The 3×3 joint failure maps (Figure 4), corresponding to subsections of the workspace reachability map, indicate which joints operate near their torque limits in these sections of the workspace. Joints

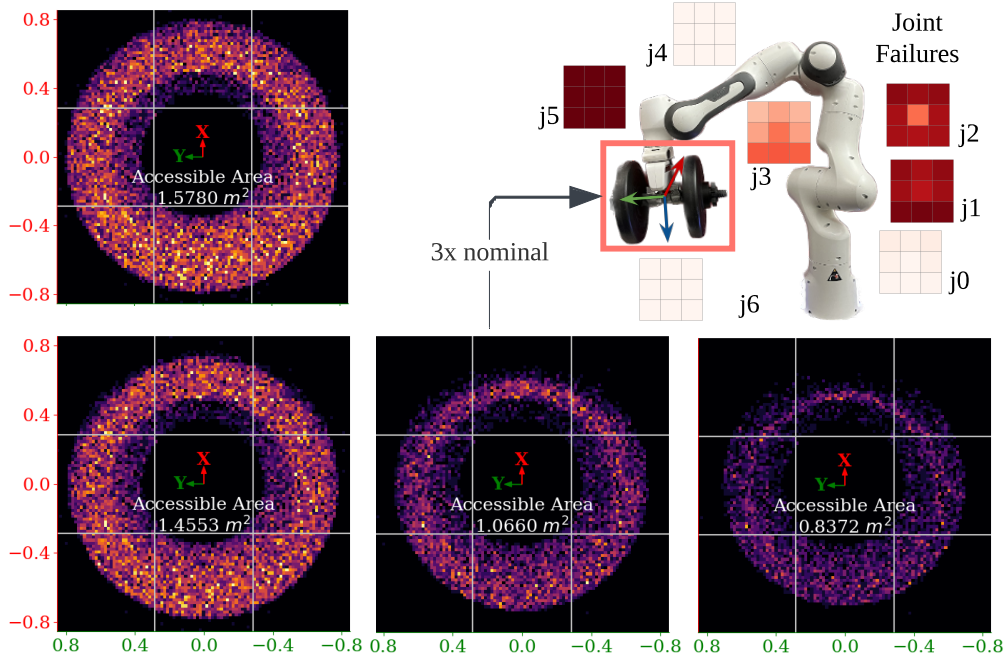


Figure 4: Analysis of diffusion success, workspace reachability, and joint constraints across a range of payloads. Heatmaps: success rate distribution across the robot’s reachable workspace, where color intensity indicates the probability of successful trajectory completion for trajectories starting or ending at each location in the x-y plane ($z = 0.2\text{m}$). Each pixel in the heatmap represents a $1.73 \times 1.73\text{cm}$ spatial bin, providing high-resolution insight into the robot’s operational boundaries and joint-specific limitations. Top right: A jointwise 3×3 grid of spatial patches corresponding to the 9kg heatmap indicating which joints reach their torque limits in different regions of the workspace.

0, 4, and 6 operate well within their torque limits. In contrast, joints 1, 2, 3, and 5 frequently operate close to their torque limits, as their axes lie predominantly perpendicular to the gravity vector, requiring them to generate substantial counteracting torques for payload support. Joint 3 displays a characteristic failure pattern, suggesting limited maneuverability in the negative-x region of the task space. This pattern reveals an inherent asymmetry in the robot’s mechanical design across the x-axis. Joint 5, despite its relatively limited torque capacity compared to the base-mounted joint 1, bears significant responsibility for maintaining payload position and orientation throughout trajectories. This is particularly noteworthy given that joint 1 is primarily engineered to support the weight of the subsequent kinematic chain links and payload.

2.4 Real-World Trajectory Execution

We validate our approach on a Franka Emika Panda robot arm using trajectories generated by our diffusion model ($\Delta t = 0.15\text{s}$, equivalent to 6.67Hz). The generated trajectories were executed through the robot’s velocity control interface, with linear interpolation to match the required 1kHz control frequency.¹ Notably, the diffusion-generated trajectories required no additional post-processing or filtering beyond this basic interpolation. Our real-world testing encompassed payload masses ranging from 0 to 7.2kg —Figure 5 presents the joint torque measurements recorded across 15 trajectory executions at maximum payload for different start and goal joint configurations. While our simulation work demonstrates successful planning capability for payloads up to 11kg with an approximate

¹The accompanying video material demonstrates the robot manipulating super-nominal payload masses and various real-world objects.

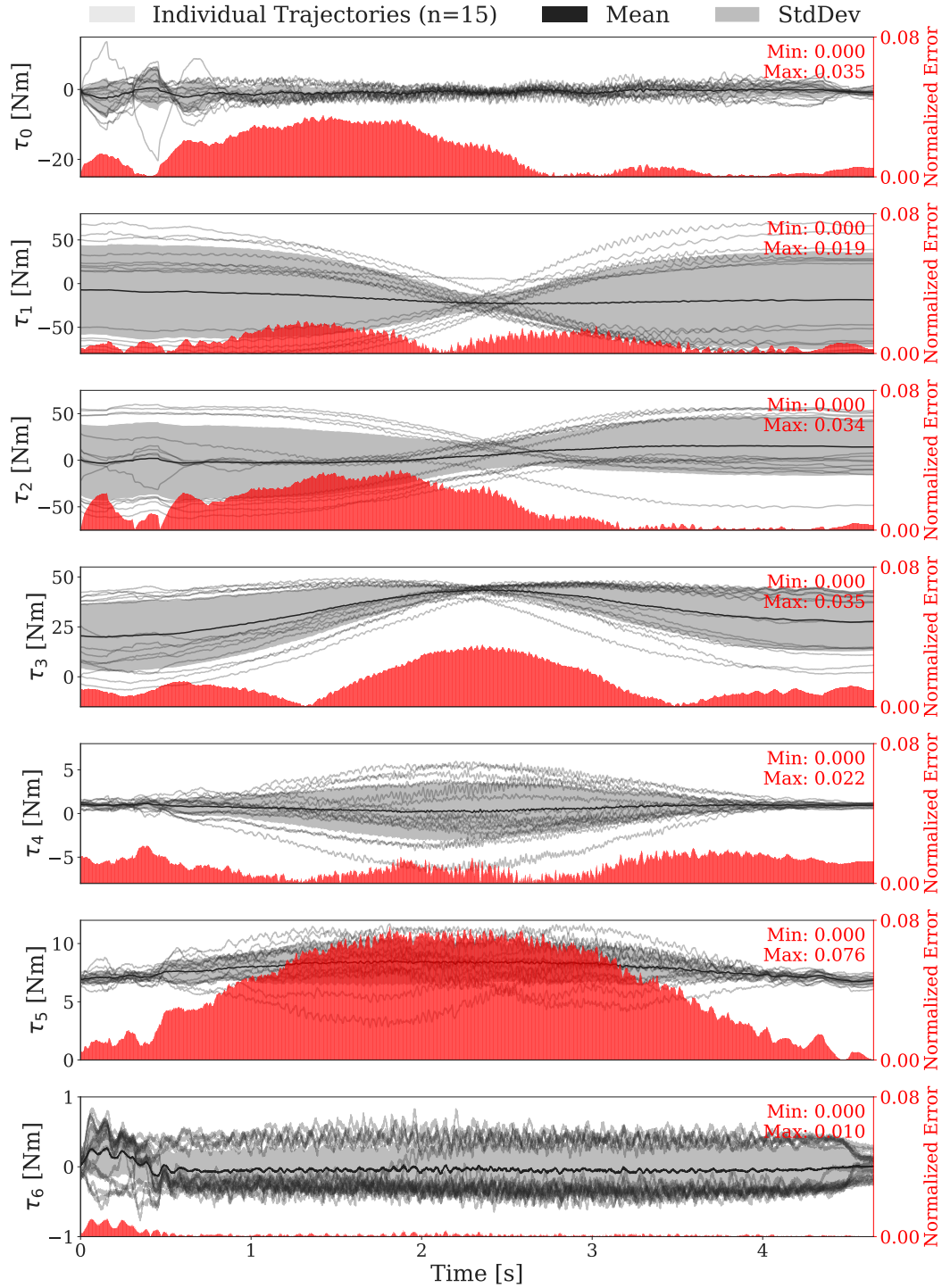


Figure 5: Joint torques and tracking errors during trajectory execution with a 7.2kg payload on the Franka Emika Panda robot. Grey shaded regions show mean \pm one standard deviation of joint torques across 15 different trajectories, while red bars indicate the normalized tracking error for each joint. Note the consistent $\approx 8\text{Nm}$ loading on joint 5 and increased tracking errors in joints 1, 3, and 5 due to gravitational loading effects.

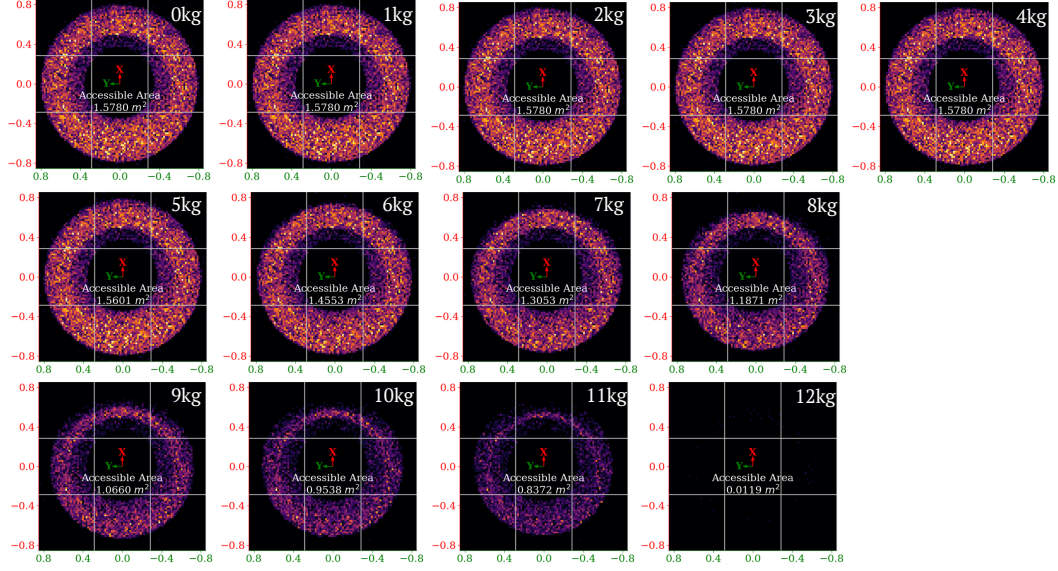


Figure 6: These heatmaps visualize the robot’s reachable workspace resulting from diffusion trajectories generated under different payload attachments for pick-and-place to/from $z = 0.2\text{m}$. Even at $3\times$ nominal payload capacity, we see a 67.6% retention of accessible workspace.

77 success rate of 20% (Figure 4), we constrained our physical experiments to 7.2kg to maintain safe
 78 operating conditions.

79 The measured joint torques exhibit distinct patterns reflecting the robot’s kinematic chain and load-
 80 bearing constraints. Joints 1-3 demonstrate substantial torque variations during dynamic motion
 81 phases, consistent with their primary role in payload manipulation. Joint 5 exhibits notably con-
 82 sistent torque readings (approximately 8Nm) across trials, primarily managing gravitational loads
 83 to maintain end-effector orientation. Joints 0 and 6, which mainly control vertical axis rotation,
 84 show minimal torque requirements throughout the trajectories. A characteristic feature emerges at
 85 the trajectory apex (2 – 3 seconds), where torque standard deviation significantly decreases across
 86 all joints. This reduction corresponds to a mechanically advantageous configuration where the robot
 87 momentarily pauses vertical motion, resulting in more consistent torque distributions across multiple
 88 executions.

89 The normalized tracking error (shown in red in Figure 5) exhibits the highest magnitudes in joints 1,
 90 3, and 5—those perpendicular to the payload’s z-axis. This increased tracking error can be attributed
 91 to: 1) heightened sensitivity to mass and center-of-mass estimation errors due to significant gravita-
 92 tional loading at these joints; 2) higher joint velocities leading to increased joint friction modeling
 93 uncertainties [3]; and, 3) joint compliance effects, where high torques cause elastic deformation in
 94 the transmission elements leading to position deviations beyond encoder measurements. Our phys-
 95 ical hardware experiments demonstrate successful execution of diffusion-planned trajectories with
 96 payloads beyond the derived operational threshold, though safety considerations ultimately bound
 97 rigorous real-world testing.

98 **References**

- 99 [1] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song. Diffusion policy: Vi-
100 suomotor policy learning via action diffusion. In *Proceedings of Robotics: Science and Systems*
101 (*RSS*), 2023.
- 102 [2] E. Perez, F. Strub, H. De Vries, V. Dumoulin, and A. Courville. Film: Visual reasoning with
103 a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*,
104 volume 32, 2018.
- 105 [3] C. Gaz, M. Cagnetti, A. Oliva, P. R. Giordano, and A. De Luca. Dynamic identification of the
106 franka emika panda robot with retrieval of feasible parameters using penalty-based optimization.
107 *IEEE Robotics and Automation Letters*, 4(4):4147–4154, 2019.