MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Constrained Sampling MPC for Safe Contact-Rich Control: From Exploration to Precision via Hybrid Refinement

Wang, Chenghao; Romeres, Diego; Schperberg, Alexander; Li, Na; Wang, Yebin

TR2026-119 September 02, 2026

Abstract

Safe robotic control in contact-rich environments requires navigating non-convex optimization landscapes while enforcing safety constraints and achieving task-level precision. Annealing-based sampling MPC methods enable fast exploration of complex solution spaces but lack principled constraint handling and struggle to reach the tight tolerances required by precise manipulation tasks. We propose Constrained Annealing-based Sampling MPC (CAS-MPC), which integrates inequality constraints via a primal–dual weight reshaping scheme that preserves sample diversity, and a hybrid refinement strategy that switches to Sequential Linear–Quadratic MPC for millimeter-level precision once near the goal. On a Unitree Go2 quadruped and a Fetch mobile manipulator, CAS-MPC maintains a 1.0 alive-trajectory ratio during obstacle avoidance, while the hybrid controller reduces average convergence time by 47% over SLQ-MPC and average position error by 71% over CAS-MPC alone across five SE(3) reaching tasks.

IFAC World Congress 2026

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Constrained Sampling MPC for Safe Contact-Rich Control: From Exploration to Precision via Hybrid Refinement ⋆

Chenghao Wang * **Diego Romeres** ** **Alexander Schperberg** **
Na Li *** **Yebin Wang** **

* *Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA (e-mail: wang.chengh@northeastern.edu)*
** *Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA (e-mails: {romeres, schperberg, yebinwang}@merl.com)*
*** *School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA (e-mail: nali@seas.harvard.edu)*

Abstract: Safe robotic control in contact-rich environments requires navigating non-convex optimization landscapes while enforcing safety constraints and achieving task-level precision. Annealing-based sampling MPC methods enable fast exploration of complex solution spaces but lack principled constraint handling and struggle to reach the tight tolerances required by precise manipulation tasks. We propose Constrained Annealing-based Sampling MPC (CAS-MPC), which integrates inequality constraints via a primal–dual weight reshaping scheme that preserves sample diversity, and a hybrid refinement strategy that switches to Sequential Linear–Quadratic MPC for millimeter-level precision once near the goal. On a Unitree Go2 quadruped and a Fetch mobile manipulator, CAS-MPC maintains a 1.0 alive-trajectory ratio during obstacle avoidance, while the hybrid controller reduces average convergence time by 47% over SLQ-MPC and average position error by 71% over CAS-MPC alone across five SE(3) reaching tasks.

Keywords: Model Predictive Control, Sampling-based Control, Constraint Satisfaction, Mobile Manipulation, Contact-Rich Control, Hybrid Control

1. INTRODUCTION

Robotic systems operating in unstructured environments must balance global exploration of non-convex cost landscapes with satisfaction of safety constraints and task-specific accuracy in contact-rich scenarios. Traditional gradient-based MPC methods like Sequential Linear-Quadratic (SLQ) MPC often get trapped in local minima when initialized far from feasible solutions, while sampling-based variants such as annealing-based sampling MPC mitigate this through stochastic parallel rollouts with dual-loop annealing that balances exploration and convergence (Xue et al., 2024). However, annealing-based sampling MPC does not explicitly handle inequality constraints, exploring the action space without bias toward feasible regions, and its inherent stochasticity makes it difficult to converge within the tight tolerances required for precision-oriented tasks such as manipulation or assembly.

To address these gaps, we propose Constrained Annealing-based Sampling MPC (CAS-MPC) with a hybrid refinement mechanism. CAS-MPC extends unconstrained annealing methods with a primal-dual weight reshaping scheme that augments rollout costs with adaptive soft penalties based on maximum constraint violations, progressively biasing the sampling distribution toward safe regions without the sample inefficiency of hard rejection. To overcome the precision limits inherent to stochastic sampling, we further extend CAS-MPC into a hybrid architecture that automatically switches to SLQ-MPC for gradient-based refinement when a weighted task residual indicates proximity to the goal, with warm-start trajectory transfer ensuring smooth handover. The result is a unified controller that transitions from fast, robust global exploration to precise local feedback control, adapting seamlessly to tasks ranging from coarse locomotion to fine SE(3) manipulator reaching.

1.1 Related Work

Sampling-based MPC optimizes a finite-horizon control sequence by sampling perturbed action trajectories, rolling out dynamics, scoring their costs, and updating the nominal sequence via cost-weighted averaging. In contrast, optimization-based approaches such as offline NLP formulations (Dai et al., 2025) and online gradient-based methods like SLQ-MPC (Pankert and Hutter, 2020) embed explicit constraints but typically rely on reduced-order models or local linearization, and can get trapped in local minima in cluttered environments. Sampling-based approaches avoid these limitations by leveraging high-fidelity simulators directly and exploring non-convex landscapes stochastically.

⋆ This work was done while C. Wang was a research intern and N. Li was a visiting scholar at MERL, Cambridge, MA 02139, USA.

Within this family, MPPI (Williams et al., 2016) is widely used due to its parallelizability and reliance on rollouts only. Subsequent work improves sampling efficiency through covariance shaping (Yin et al., 2022; Yi et al., 2024), robust formulations (Williams et al., 2018), biased samplers (Trevisan and Alonso-Mora, 2024; Askari et al., 2025), or learned traversability (Dong et al., 2025). Predictive Sampling in MJPC (Howell et al., 2022) demonstrates a competitive real-time shooting baseline.

A key challenge in applying sampling-based MPC to high-dimensional problems like full-order robotic systems is the exploration–exploitation trade-off: large sampling variance provides better global coverage but yields high-variance solutions, while small variance enables precise local convergence but risks entrapment in poor local minima. Annealing-based methods such as DIAL-MPC (Xue et al., 2024) address this through a dual-loop annealing process that progressively reduces sampling noise. However, DIAL-MPC does not explicitly handle inequality constraints, potentially producing unsafe trajectories in contact-rich settings.

1.2 Contributions

The primary contributions of this work are:

- proposing Constrained Annealing-based Sampling MPC (CAS-MPC) that integrates primal-dual weight reshaping for efficient constraint satisfaction in contact-rich robotic control.
- introducing an extension of CAS-MPC with Sequential Linear-Quadratic (SLQ) MPC refinement for tasks demanding precision, featuring an automatic switching mechanism and warm-start transfer that enables both global exploration and precise local feedback control.
- demonstrating the efficacy of our approach through simulation and hardware experiment.

2. PRELIMINARY AND PROBLEM STATEMENT

2.1 Problem Formulation

We consider a robotic system with dynamics:

$$\dot{x} = f(x, u), \tag{1}$$

where $x \in \mathcal{X} \subset \mathbb{R}^{n_x}$ is the state, $u \in \mathcal{U} \subset \mathbb{R}^{n_u}$ is the control input, and $f : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^{n_x}$ represents the system dynamics.

To achieve desired behaviors while respecting physical limitations and safety requirements, we seek control inputs that optimize a given task objective over a finite time horizon. This naturally leads to a constrained optimal control formulation. Specifically, at each time step t, we formulate the following finite-horizon optimal control problem:

$$\begin{aligned} \min_{u_{t:t+H}} \quad & \sum_{h=0}^{H} \ell(x_{t+h}, u_{t+h}) + \ell_f(x_{t+H+1}) \\ \text{s.t.} \quad & x_{t+h+1} = f(x_{t+h}, u_{t+h}), \quad h = 0, \dots, H, \\ & g(x_{t+h}, u_{t+h}) \geq 0, \quad h = 0, \dots, H, \end{aligned} \tag{2}$$

where $\ell : \mathcal{X} \times \mathcal{U} \to \mathbb{R}_+$ is the running cost that encodes the task objective, $\ell_f : \mathcal{X} \to \mathbb{R}_+$ is the terminal cost, and $g : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^{n_g}$ encodes inequality constraints such as state bounds, control limits, and collision avoidance.

2.2 Sampling-Based MPC

In unconstrained MPPI, perturbations are added to the nominal sequence and the update is a softmax-weighted average:

$$U^+ = U \; + \; \frac{\displaystyle\sum_{i=1}^{N_W} \exp\!\left(-\frac{J(U + [W]_i)}{\tau}\right) [W]_i}{\displaystyle\sum_{j=1}^{N_W} \exp\!\left(-\frac{J(U + [W]_j)}{\tau}\right)}, \tag{3}$$

where $U = u_{t:t+H}$ is the nominal control sequence defined over time instants $\{t, t+1, \cdots, t+H\}$, $\{[W]_i\}_{i=1}^{N_W} \sim \mathcal{N}(0, \Sigma_{t:t+H})$ are i.i.d. Gaussian perturbations, $J(\cdot)$ is the cumulative rollout cost for a control sequence (e.g., $J(U) = \sum_{h=0}^{H} \ell(x_{t+h}, u_{t+h}) + \ell_f(x_{t+H+1})$), N_W is the number of samples, and $\tau > 0$ is the temperature controlling the sharpness of the softmax weights. However, this unconstrained formulation cannot guarantee constraint satisfaction. A common remedy is trajectory filtering, where samples violating constraints are discarded before computing the weighted update; yet this approach can drastically reduce the effective sample count near constraint boundaries, leading to high-variance updates or even solver failure when most or all trajectories are rejected. Furthermore, the fixed covariance in standard MPPI limits exploration-exploitation balance: large noise improves exploration while small noise favors exploitation. Our approach addresses the first issue through primal-dual weight reshaping, which preserves all samples while adaptively steering the distribution toward feasible regions, and leverages dual-loop annealing to balance exploration and exploitation.

2.3 Dual-Loop Annealing for Sampling-Based MPC

Building on (Xue et al., 2024), we adopt a dual-loop annealing schedule to address the exploration–exploitation trade-off in non-convex landscapes: (i) a *trajectory-level* outer loop of N refinement stages that progressively shrinks the overall sampling covariance, and (ii) an *action-level* inner loop that shapes the per-step variance along the horizon, assigning smaller variance to near-term actions and larger variance to later ones. A practical instantiation combines the two schedules into per-step isotropic covariances:

$$\Sigma_{t+h}^{(i)} = \sigma_0^2 \, \exp\!\Big(-\frac{N-i}{\beta_1 N} - \frac{H-h}{\beta_2 H}\Big) I_{d_u}, \quad h = 0, \dots, H, \tag{4}$$

where $\sigma_0 > 0$ is a base scale, I_{d_u} is the $d_u \times d_u$ identity, $i \in \{1, \dots, N\}$ is the annealing-stage index, and $\beta_1, \beta_2 > 0$ control the outer-stage shrinkage and horizon-level allocation, respectively. Larger β_2 assigns relatively larger variance to later actions in the sequence (which will be revised more times) and smaller variance to near-term actions. The block-diagonal sequence covariance is $\Sigma_{t:t+H}^{(i)} = \mathrm{blkdiag}\big(\Sigma_{t+0}^{(i)}, \dots, \Sigma_{t+H}^{(i)}\big)$. The isotropic choice

simplifies implementation and reduces tuning complexity at the cost of uniform exploration across all control dimensions.

2.4 Sequential Linear-Quadratic MPC

Sequential Linear-Quadratic (SLQ) MPC is a gradient-based method for trajectory optimization that achieves precise local control through iterative refinement (Pankert and Hutter, 2020; Neunert et al., 2016; Grandia et al., 2019). Given a nominal rollout $\{x_k^{\text{nom}}, u_k^{\text{ff}}\}_{k=0}^{H-1}$, SLQ-MPC linearizes the dynamics to form a time-varying LQ subproblem with $A_k = I + \Delta t\, \partial_x f$ and $B_k = \Delta t\, \partial_u f$, then solves it via Riccati recursion to obtain a time-varying affine policy:

$$u_{t+h}(x) = u_{t+h}^{\text{ff}} + K_{t+h}\big(x_{t+h} - x_{t+h}^{\text{nom}}\big). \tag{5}$$

For task costs, we define a residual $r(x) \in \mathbb{R}^{n_r}$ encoding task-specific error (e.g., end-effector pose error), with cost $\ell(x,u) = \frac{1}{2} r(x)^\top W r(x) + \frac{1}{2} u^\top R u$.

Inequality constraints $h_i(x,u) \geq 0$ are handled within the SLQ framework via relaxed log-barrier functions (Grandia et al., 2019), augmenting the stage cost as $\hat{\ell}(x,u) = \ell(x,u) + \mu \sum_{i=1}^{N_{\text{in}}} B(h_i(x,u))$, where $\mu > 0$ is the barrier weight and $B(\cdot)$ is a relaxed log-barrier with quadratic extension near the constraint boundary to maintain differentiability.

3. METHOD

This section addresses two challenges: ensuring constraint satisfaction during exploration via CAS-MPC with primal-dual weight reshaping (Section 3.1), and achieving precision via a hybrid extension that transitions to gradient-based refinement (Section 3.2).

3.1 Constrained Annealing-Based Sampling MPC

While the dual-loop annealing schedule (Section 2) improves convergence to high-quality solutions compared to fixed-covariance MPPI by balancing exploration and exploitation, it does not effectively ensure constraint satisfaction: the annealed sampling explores the action space without bias toward feasible regions, potentially producing unsafe trajectories that violate collision-avoidance constraints or state bounds. We extend the framework with a primal-dual approach that incorporates constraints through soft penalties while maintaining computational efficiency and sample diversity.

For each sampled trajectory i with rollout $\{x_{t+h}^i\}_{h=0}^{H}$ generated by control sequence $U + [W]_i$, we compute the worst-case constraint violation:

$$\phi_i \;=\; \max_{h=0,\dots,H} \big[-\, g(x_{t+h}^i) \big]_+, \tag{6}$$

where $[z]_+ = \max\{z, 0\}$ is the hinge function. This metric captures the maximum violation across the entire horizon, ensuring that trajectories with any constraint violation are penalized. We augment the original cost with a weighted penalty term:

$$\tilde{J}_i \;=\; J\big(U + [W]_i\big) \;+\; \lambda\, \phi_i, \tag{7}$$

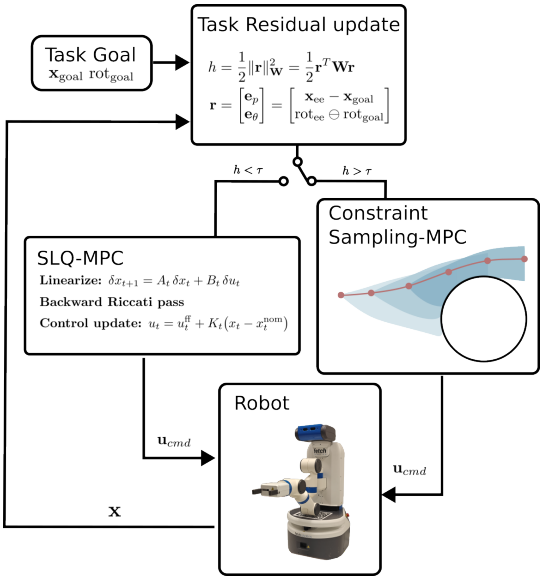

Fig. 1. Overview of the hybrid architecture: Given a task goal, the controller computes a task residual measuring the distance to the goal. When this residual is large (far from goal), CAS-MPC performs global exploration. When the residual is small (near optimal), the controller switches to SLQ-MPC for precise local refinement using Riccati-based feedback control.

where $\lambda \geq 0$ is the dual variable that adapts to enforce constraint satisfaction. The importance weights become:

$$\omega_i \;=\; \frac{\exp\big(-\tilde{J}_i/\tau\big)}{\sum_{j=1}^{N_W} \exp\big(-\tilde{J}_j/\tau\big)}\,. \tag{8}$$

The control update aggregates the weighted perturbations:

$$U^+ \;=\; U \;+\; \sum_{i=1}^{N_W} \omega_i\, [W]_i, \tag{9}$$

The dual variable is updated to drive constraint satisfaction:

$$\lambda \;\leftarrow\; \Big[\lambda \;+\; \eta \sum_{i=1}^{N_W} \omega_i\, \phi_i\Big]_+, \qquad \eta > 0, \tag{10}$$

where η is the dual learning rate. This update increases λ when the weighted average violation is positive, thereby increasing the penalty on infeasible trajectories in subsequent iterations.

The primal-dual formulation offers key advantages over alternative constraint handling strategies. Trajectory filtering, which discards infeasible samples, corresponds to setting $\phi_i = \infty$ in our formulation, driving importance weights $\omega_i \to 0$ in (8). While this achieves hard constraint enforcement, it suffers from two limitations: (1) near constraint boundaries, most samples may be rejected, leading to sample starvation and high-variance updates; and (2) infeasible samples still carry useful information about the cost landscape that is lost when discarded. Pure penalty methods with fixed weights avoid sample starvation but require careful tuning of penalty values and cannot adapt to varying constraint difficulties. Our primal-dual approach addresses both issues by using finite, adaptive penalties that down-weight but do not eliminate infeasible samples,

Algorithm 1 Constrained Annealing-based Sampling MPC with Primal-Dual Updates

Require: Horizon H; simulation length T; annealing stages $i = N, \dots, 1$; nominal sequence $U^{(0)}_{t:t+H} = \mathbf{0}$; temperature τ; sample count N_W; schedules β_1, β_2; dual learning rate η

Initialise dual variable $\lambda \leftarrow 0$
for $t \leftarrow 0$ to T **do**
  Shift plan: $U^{(0)}_{t:t+H} \leftarrow [u^{(1)}_{t+1}, \dots, u^{(1)}_{t+H|t-1}, \mathbf{0}]$
  for $i \leftarrow N$ down to 1 **do**
    for $h \leftarrow 0$ to H **do**
      $\Sigma^{(i)}_{t+h} \leftarrow \sigma_0^2 \exp\left(-\frac{N-i}{\beta_1 N} - \frac{H-h}{\beta_2 H}\right) I_{d_u}$
    $\Sigma^{(i)}_{t:t+H} \leftarrow \mathrm{diag}\left(\Sigma^{(i)}_t, \dots, \Sigma^{(i)}_{t+H}\right)$
    Sample $[W]_k \sim \mathcal{N}\left(\mathbf{0}, \Sigma^{(i)}_{t:t+H}\right), k = 1 : N_W$
    Roll out dynamics, obtain J_k and g_k
    Compute violation $\phi_k \leftarrow \max_h\{0, -g(x^k_{t+h})\}$
    Augmented cost $\tilde{J}_k \leftarrow J_k + \lambda\phi_k$
    Importance weights $\omega_k = \frac{e^{-\tilde{J}_k/\tau}}{\sum_j e^{-\tilde{J}_j/\tau}}$
    Primal update: $U^{(i)} \leftarrow U^{(i-1)} + \sum_k \omega_k [W]_k$
    Dual ascent: $\lambda \leftarrow \max\{0, \lambda + \eta \sum_k \omega_k \phi_k\}$
  Apply first control $u_t \leftarrow U^{(N)}_t$ to system; observe next state x_{t+1}; $t \leftarrow t+1$

extracting information from all rollouts while progressively steering the distribution toward feasibility through the iterative dual update mechanism.

The complete procedure is summarized in Algorithm 1.

3.2 Hybrid Architecture for Precision Tasks

Hybrid Switching Logic Fig. 1 illustrates the hybrid architecture. Given a task goal specifying the desired position x_{goal} and orientation rot_{goal}, the controller computes a task residual $r(x)$ combining position error $e_p = x_{\text{ee}} - x_{\text{goal}}$ and orientation error $e_\theta = \log\big(\text{rot}^{-1}_{\text{goal}}\,\text{rot}_{\text{ee}}\big)^\vee \in \mathbb{R}^3$, denoted compactly as $\text{rot}_{\text{ee}} \ominus \text{rot}_{\text{goal}}$, where $\log(\cdot)^\vee$ is the SO(3) logarithm map. The transition between CAS-MPC and SLQ-MPC is governed by a weighted proximity metric:

$$h(x) \triangleq \|r(x)\|_W = \sqrt{r(x)^\top W\, r(x)}, \tag{11}$$

where $r(x)$ is the task residual used in SLQ-MPC (as defined in Section 2). Given a threshold $\tau_{\text{sw}} > 0$, the hybrid controller operates in CAS-MPC (exploration mode) when $h(x_t) > \tau_{\text{sw}}$, leveraging sampling-based methods to navigate the non-convex landscape. When $h(x_t) \le \tau_{\text{sw}}$, the controller switches to SLQ-MPC (refinement mode) for precise local convergence via Riccati-based feedback.

Warm-Start When switching from CAS-MPC to SLQ-MPC, state continuity is preserved through warm-starting: the final trajectory from CAS-MPC $\{x^{\text{nom}}_k, u^{\text{ff}}_k\}_{k=0}^{H-1}$ initializes the SLQ solver, with feedback gains set to zero ($K_k = 0$) for stable initial rollout. This warm-start mechanism is crucial for ensuring smooth transition from CAS-MPC. We note that this switching scheme does not formally guarantee recursive feasibility of SLQ-MPC at the handover, particularly if the threshold is reached near constraint boundaries; in practice, choosing τ_{sw} such that the residual is dominated by goal-tracking error rather than constraint slack avoids this regime, and a principled treatment via terminal sets or invariance-based switching is left to future work. At runtime, the controller evaluates $h(x_t)$ at every step, runs Algorithm 1 while $h(x_t) > \tau_{\text{sw}}$, and on the first step where $h(x_t) \le \tau_{\text{sw}}$ performs the warm-start transfer above and switches to SLQ iterations (linearization, Riccati backward pass, forward rollout with line search) for all subsequent steps, applying $u_t = u^{\text{ff}}_t + K_t(x_t - x^{\text{nom}}_t)$.

4. EXPERIMENT

4.1 Simulation Setup

We validate our approach in simulation using two robotic platforms that demonstrate the framework's adaptability to different precision requirements. For tasks where precision is not critical, we use a Unitree Go2 quadruped performing dynamic locomotion that focuses primarily on constraint satisfaction, using CAS-MPC-only control. For precision-critical tasks, we employ a Fetch mobile manipulator performing SE(3) goal pose reaching with the hybrid CAS-SLQ-MPC extension. All simulations are conducted using MJX physics engine, which provides the parallel sampling capabilities required by sampling-based methods.

4.2 Hardware Setup

We validate the hybrid framework on a Fetch mobile manipulator performing SE(3) end-effector reaching tasks. To ensure safety during initial hardware validation, we conduct fixed-base experiments where the mobile base remains stationary while the 8-DOF arm (including 1-DOF torso) reaches target poses approximately 1.6m from the robot's base, requiring near-full arm extension and a specific gripper orientation that tests both positional and rotational alignment. We compare CAS-MPC-only, SLQ-MPC-only, and the proposed hybrid approach, with three trials per method. The control system runs at 50 Hz; for the hybrid controller, mode switching uses $\tau_{\text{sw}} = 5$ with a 0.2 s minimum dwell time between transitions.

4.3 Simulation Results

CAS-MPC for Quadruped Locomotion We compare primal-dual constraint handling against a trajectory-filtering baseline that discards rollouts violating constraints before the MPPI update; both share the same annealing schedule and cost. The task is forward trotting at 1.0 m/s past a circular obstacle of radius 0.5 m placed 1.5 m ahead, with a 0.3 m safety margin around the obstacle. Figure 4 shows that while both methods avoid the obstacle, the alive ratio for trajectory selection rapidly decreases and approaches zero, whereas the primal-dual method maintains an alive ratio of 1.0 throughout.

Hybrid Refinement for Precise Mobile Manipulation We further evaluate SE(3) goal-pose reaching with the Fetch manipulator, reaching a target end-effector pose while avoiding a cylindrical obstacle. Figures 2 and 3 show the

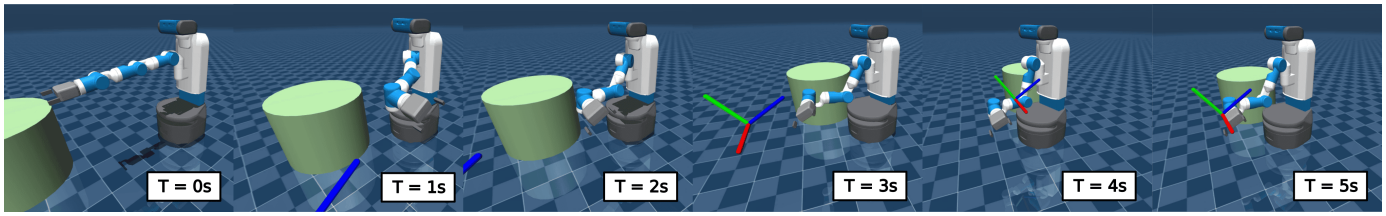

Fig. 2. Simulation snapshot of robot trajectory with PD update from T=0s to T=5s: The robot navigates around the obstacle using primal-dual update (T=1s-2s) and reaches its goal configuration by T=5s.

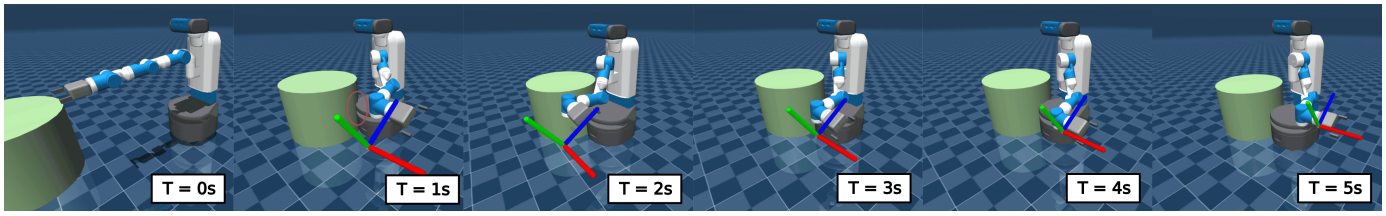

Fig. 3. Simulation snapshot of unconstrained robot trajectory from T=0s to T=5s: The robot collides with the obstacle at T=1s and struggles to reach the goal configuration by T=5s.

Table 1. Performance comparison across five SE(3) goal-reaching tasks. Mean ± std over 5 trials. Time-to-1.0 (s) measures convergence speed; Pos. Error (m) and Rot. Error (rad) measure precision when $h < 1.0$.

Goal	Time-to-1.0 (s) ↓			Pos. Error (m) ↓			Rot. Error (rad) ↓		
	SLQ	CAS-MPC	Hybrid	SLQ	CAS-MPC	Hybrid	SLQ	CAS-MPC	Hybrid
1	6.47 ± 2.73	$\mathbf{3.11 \pm 0.23}$	3.42 ± 0.37	0.063 ± 0.023	0.193 ± 0.023	$\mathbf{0.043 \pm 0.003}$	0.131 ± 0.011	0.128 ± 0.015	$\mathbf{0.119 \pm 0.006}$
2	5.93 ± 1.93	$\mathbf{3.17 \pm 0.60}$	3.77 ± 0.70	0.100 ± 0.072	0.157 ± 0.020	$\mathbf{0.040 \pm 0.011}$	0.148 ± 0.033	0.117 ± 0.023	$\mathbf{0.122 \pm 0.008}$
3	4.68 ± 0.92	5.61 ± 3.40	$\mathbf{3.20 \pm 0.56}$	0.040 ± 0.003	0.210 ± 0.012	$\mathbf{0.077 \pm 0.068}$	0.117 ± 0.003	0.127 ± 0.035	$\mathbf{0.135 \pm 0.034}$
4	7.73 ± 3.16	$\mathbf{3.15 \pm 0.22}$	3.21 ± 0.43	0.105 ± 0.096	0.199 ± 0.036	$\mathbf{0.066 \pm 0.068}$	0.138 ± 0.025	0.122 ± 0.016	$\mathbf{0.130 \pm 0.043}$
5	7.34 ± 3.52	6.39 ± 3.27	$\mathbf{3.39 \pm 0.30}$	0.086 ± 0.072	0.175 ± 0.033	$\mathbf{0.048 \pm 0.029}$	0.131 ± 0.012	0.119 ± 0.014	$\mathbf{0.124 \pm 0.019}$
Avg	6.36 ± 2.70	4.05 ± 2.38	$\mathbf{3.40 \pm 0.50}$	0.075 ± 0.060	0.187 ± 0.030	$\mathbf{0.055 \pm 0.042}$	0.133 ± 0.021	0.123 ± 0.021	$\mathbf{0.126 \pm 0.024}$

qualitative difference: with primal-dual handling the robot navigates around the obstacle and reaches the goal by T=5s, whereas without it the robot collides at T=1s. We compare SLQ-MPC only, CAS-MPC only, and the hybrid approach, with the mode switch (orange dot in Figure 5) triggered at $\tau_{\text{sw}} = 10$.

Figure 5 shows the complementary advantages of the hybrid approach: SLQ-MPC (red) exhibits initial instability with residuals exceeding 100, indicating local linearization struggles when initialized far from the optimum; CAS-MPC (blue) converges stably but oscillates around 1 due to its intrinsic stochasticity; the hybrid method (green) leverages CAS-MPC's exploration to navigate the non-convex landscape, then switches to SLQ-MPC for millimeter-level refinement, yielding faster convergence without sacrificing accuracy. Table 1 summarizes performance across five SE(3) tasks: the hybrid method attains the fastest average convergence (3.40 s vs. 6.36 s for SLQ and 4.05 s for CAS-MPC) with the lowest variance, and the lowest average position error (0.055 m) while maintaining competitive rotation accuracy.

4.4 Hardware Results

For hardware experiments, we test all three methods in a fixed-base configuration due to lack of torque control API capability for the robot wheels, and safety concerns. Results shown in Figure 6 exhibit similar trends to the simulation. Since the target is closer to the robot than in the simulation scenario, SLQ-MPC does not exhibit the same initial deviation as seen in previous simulation results, but still converges slower than its sampling-based and hybrid counterparts. Once near the optimum, CAS-MPC fails to achieve stable convergence due to its inherent stochasticity, whereas SLQ-MPC and the hybrid method remain stable at the optimum.

5. CONCLUSION AND FUTURE WORK

This work introduced constrained annealing-based sampling MPC with primal-dual weight reshaping, enabling safe collision avoidance while maintaining sample availability during exploration. The hybrid refinement strategy successfully demonstrated seamless transitions from fast sampling-based exploration to precise gradient-based control, achieving faster convergence than pure SLQ-MPC while attaining millimeter-level accuracy. Experiments on the Unitree Go2 quadruped validated safe navigation around obstacles, while the Fetch manipulator confirmed the framework's ability to achieve precision required for SE(3) reaching tasks. Future directions include integrating control barrier functions to provide formal safety guarantees and exploring adaptive constraint tube tightening strategies for dynamic environments.

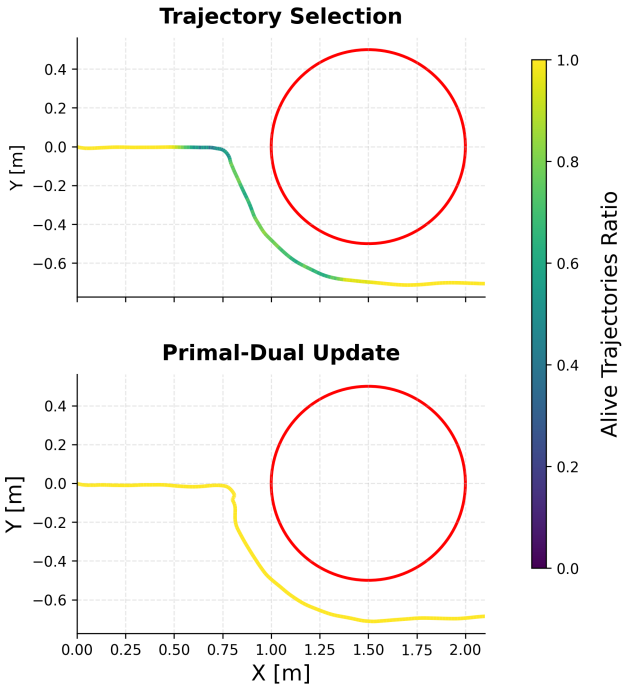

Fig. 4. Comparison of center-of-mass trajectories and alive trajectory ratios between trajectory selection and primal-dual methods.

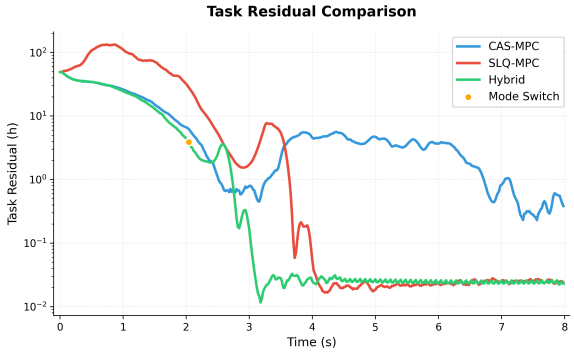

Fig. 5. Comparison of task residual convergence in simulation across three control strategies.

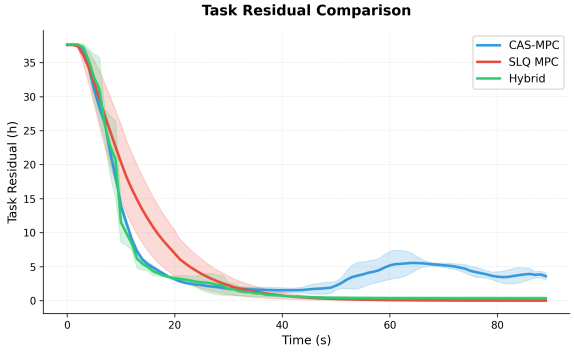

Fig. 6. Comparison of task residual convergence in hardware experiments.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used ChatGPT (enterprise version) in order to brainstorm ideas, debug code, and check writing and grammar. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

REFERENCES

Askari, I., Vaziri, A., Tu, X., Zeng, S., and Fang, H. (2025). Model Predictive Inferential Control of Neural State-Space Models for Autonomous Vehicle Motion Planning. *IEEE Transactions on Robotics*, 41, 3202–3222.

Dai, M., Lu, Z., Li, N., and Wang, Y. (2025). Enhanced Agility and Safety in Mobile Manipulators through Centroidal Momentum-Based Motion Planning. In *2025 European Control Conference (ECC)*, 2733–2739.

Dong, Z., Papalia, A., Jung, L., Spiro, A., Osteen, P.R., Robison, C.S., and Everett, M. (2025). Learning Smooth State-Dependent Traversability from Dense Point Clouds.

Grandia, R., Farshidian, F., Ranftl, R., and Hutter, M. (2019). Feedback MPC for Torque-Controlled Legged Robots. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4730–4737. 53-0866.

Howell, T., Gileadi, N., Tunyasuvunakool, S., Zakka, K., Erez, T., and Tassa, Y. (2022). Predictive Sampling: Real-time Behaviour Synthesis with MuJoCo. 2.00541 [cs].

Neunert, M., de Crousaz, C., Furrer, F., Kamel, M., Farshidian, F., Siegwart, R., and Buchli, J. (2016). Fast nonlinear Model Predictive Control for unified trajectory optimization and tracking. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 1398–1404.

Pankert, J. and Hutter, M. (2020). Perceptive Model Predictive Control for Continuous Mobile Manipulation. *IEEE Robotics and Automation Letters*, 5(4), 6177–6184.

Trevisan, E. and Alonso-Mora, J. (2024). Biased-MPPI: Informing Sampling-Based Model Predictive Control by Fusing Ancillary Controllers. *IEEE Robotics and Automation Letters*, 9(6), 5871–5878.

Williams, G., Drews, P., Goldfain, B., Rehg, J.M., and Theodorou, E.A. (2016). Aggressive driving with model predictive path integral control. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 1433–1440.

Williams, G., Goldfain, B., Drews, P., Saigol, K., Rehg, J., and Theodorou, E. (2018). Robust Sampling Based Model Predictive Control with Sparse Objective Information. In *Robotics: Science and Systems XIV*. Robotics: Science and Systems Foundation.

Xue, H., Pan, C., Yi, Z., Qu, G., and Shi, G. (2024). Full-Order Sampling-Based MPC for Torque-Level Locomotion Control via Diffusion-Style Annealing. 5610 [cs].

Yi, Z., Pan, C., He, G., Qu, G., and Shi, G. (2024). CoVO-MPC: Theoretical Analysis of Sampling-based MPC and Optimal Covariance Design. .07369 [cs].

Yin, J., Zhang, Z., Theodorou, E., and Tsiotras, P. (2022). Trajectory Distribution Control for Model Predictive Path Integral Control using Covariance Steering. 09.12147 [cs].