Combining Sampling- and Gradient-based Planning for Contact-rich Manipulation

Filippo Rozzi[1], Loris Roveda[2], and Kevin Haninger[3]

Abstract Planning over discontinuous dynamics is needed for robotics tasks like contact-rich manipulation, which presents challenges in the numerical stability and speed of planning methods when either neural network or analytical models are used. On the one hand, sampling-based planners require higher sample complexity in high-dimensional problems and cannot describe safety constraints such as force limits. On the other hand, gradient-based solvers can suffer from local optima and convergence issues when the Hessian is poorly conditioned. We propose a planning method with both sampling- and gradient-based elements, using the Cross-entropy Method to initialize a gradient-based solver, providing better search over local minima and the ability to handle explicit constraints. We show the approach allows smooth, stable contact-rich planning for an impedance-controlled robot making contact with a stiff environment, benchmarking against gradient-only MPC and CEM.

1 Introduction

Contact-rich manipulation tasks such as door opening and gear mating involve significant environmental dynamics and constraints, which discretely switch during the task as contact conditions change. Human manipulation can effectively handle these contact changes, realizing safe and robust behavior even with uncertain geometry and dynamics. Robotic manipulation methods which can maintain safety (e.g. limits on contact force and actuation) and robustness (variation in contact geometry or dynamics) are needed which are suited to discontinuous dynamics.

Gradient-based optimization methods are popular for contact planning, with a range of formulations from contact implicit [1–4] to hybrid contact modes [5, 6]. Sampling-based methods have been used for higher-level contact mode planning [7, 8] and continuous planning [9].

The tradeoffs between sampling and gradient-based planners has led to recent interest in combining them, such as adding gradient steps to the sampling-based Cross-Entropy Method (CEM) [10, 11]. In contact, recent work has compared the variance of gradient-based and gradient-free policy learning in contact [12]. This raises the question if a planning method can effectively use both parallelized sampling and gradients, where sampling may address the initialization sensitivity of gradient-based methods [13], while gradient-based methods can provide explicit safety constraints. This paper proposes a planning method which leverages CEM and interior-point optimization for contact-rich planning tasks.

1.1 Related works

Model-based planning for contact-rich robotic tasks consists of finding the action sequence that minimizes a cumulative cost function, according to learned, parametrized, or simulated dynamics. Usually, optimized actions are executed with model predictive control (MPC), replanning at each time step. The problem can be formulated using complimentarity conditions for contact constraints [1, 14], where sequential quadratic [2] or interior-point [3] solvers can be used. This contact-implicit trajectory optimization allows planning of contact mode and trajectory in one problem with multiple contacts and Coulomb constraints [3, 4, 13]. Alternatively, hybrid dynamics can be written with dynamics for each contact mode [5, 6].

The contact planning problem is challenging, typically needing relaxations [3, 9], linearization [4, 13], or quasi-static assumptions [15] to achieve acceptable speeds. The resulting problems can be sensitive to initialization [13], and in general gradient-based methods can suffer from local minima or poor

[1]Corresponding author, Politecnico di Milano, Department of Mechanical Engineering, Milano, Italy filippo.rozzi@mail.polimi.it
[2]Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA), Scuola Universitaria Professionale della Svizzera Italiana (SUPSI), Università della Svizzera Italiana (USI) IDSIA-SUPSI, Lugano, Switzerland loris.roveda@idsia.ch
[3]Department of Automation at Fraunhofer IPK, Berlin, Germany kevin.haninger@ipk.fraunhofer.de
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 101058521 — CONVERGING.

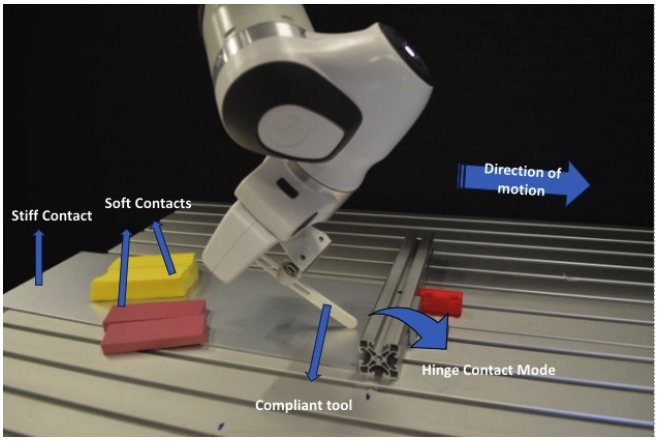


Figure 1: Experimental setup showing a robot making contact with a table, moving towards a hinge contact condition, providing a task with three modes: free space, vertical stiff contact, and hinge contact. The soft contacts are used for validating the effect of contact stiffness on performance.

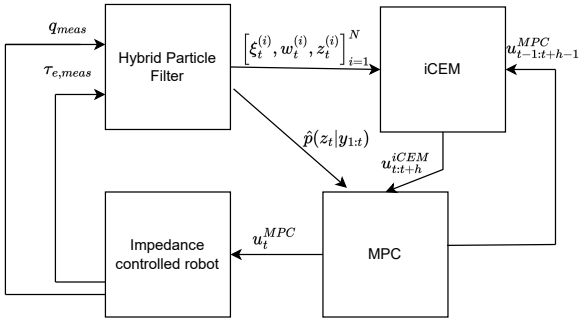


Figure 2: Block diagram of the proposed approach

convergence when the Hessian of the Lagrangian is poorly conditioned [16].

On the other hand, population-based sampling planners can parallelize well and do not need gradients, such as Random-Sampling Shooting [17], Model Predictive Path Integral Control [18], and the Cross-Entropy Method (CEM) [19]. Sampling-based have been used in contact planning, for higher-level mode initialization of a contact-implicit [7] and reachability planner [8]. Sampling can also be used for the continuous planning, where Rapidly-exploring random trees have been applied to relaxed contact models [9].

These zeroth-order optimizers maintain a sampling distribution action sequences [20], which is sampled from then updated at each iteration, assigning a higher probability to higher reward action sequences. In recent years, there has been a huge research effort on sampling based optimization in the context of model-based reinforcement learning. By adding temporally correlated actions and memory, CEM can be improved in both sample-efficiency and reward performance [21]. In [22], authors proposed a model-based reinforcement learning approach in which the interaction dynamics was directly learned through an ensemble of Artificial Neural Networks (ANNs) and CEM was used for online planning. However, CEM, like all sampling-based planners, suffers from scalability as the dimension of the action space or planning horizon grows, requiring more samples for convergence and having increasing reward variance [10,23]. Additionally, sampling-based methods cannot explicitly handle state constraints, raising questions about the ability to handle safety in contact.

An interesting idea is to combine zeroth-order and first-order planning methods in order to enjoy the benefits of both simultaneously. To this end, authors in [10, 11] proposed a combined planning framework that performs gradient updates on the sampled trajectories. Although able to exploit the advantages of both approaches, these methods still show scalability issues with very high dimensional planning problems, and don't handle constraints. For policy learning in contact, recent work [24] has shown that first order methods can have higher variance than zeroth order methods at higher stiffness and in longer planning horizons.

1.2 Paper contribution

To address poor convergence and local minima problems of gradient-based optimizers when dynamics are discontinuous, we propose a hybrid planner that uses both sampling- and gradient-based MPC. Contact is modelled with hybrid dynamics, where dynamics are switched according to contact mode. A particle filter is used to estimate the hybrid state online, providing a belief of current mode for the planner. CEM is used to initialize the action and state trajectory, modified to normalize sample cost by the weight and minimizing expected cost over the belief in mode. A gradient-based optimizer is then used, enforcing explicit constraints on the robot trajectory, as seen in Figure 2.

Compared to the work of [21], this CEM rolls out trajectories on a hybrid system, leveraging the contact mode belief provided by the filter to weight the computed samples costs. Compared to other combined sampling- and gradient-based planning approaches [10, 11], this method can enforce explicit constraints on the robot trajectory and handles hybrid dynamics. Our method is also validated experimentally, showing the feasibility of the approach with a real high-stiffness contact task.

The proposed approach is applied to contact-rich manipulation as seen in Figure 1, where the impedance controller reference trajectory is planned. The ability to estimate of contact mode online, as well as safely respond with changes in control is validated. This paper first introduces the dynamic models used, estimation of contact mode. Then, planning framework is introduced and studied in simulation. Finally, experiments are used to validate the approach, comparing the proposed planner to only MPC and only CEM. We find that the CEM initialization helps with discontinuous dynamics, reducing the mean and variance of the total

solve time, even from just one iteration, as compared to a pure gradient-based MPC for contact.

2 Problem Statement and Models

This section introduces the dynamic models used for contact and the hybrid dynamic model.

2.1 Stiffness Contact Model

We model environment contact as parallel 1-DOF stiffnesses to compromise between model identification efficiency, ease of deployment, generalizability, and differentiability [25]. The i^{th} stiffness element exerts contact force F_i at the point of contact of

$$F_i = K_i^T \left(x_i^o - (Rx_i + x)\right), \tag{1}$$

where stiffness $K_i \in \mathbb{R}^3$ has a rest pose at $x_i^o \in \mathbb{R}^3$ and $x_i \in \mathbb{R}^3$ is the contacting point in TCP frame, transformed into world coordinates with R and x , the orientation and position of the end-effector. The contact normal and stiffness are jointly described by the vector K_i, *i.e.*, the contact normal is $n_i = K_i/\|K_i\|$, and the stiffness is $\|K_i\|$. For a single point stiffness, the joint torque induced is $\tau_i = J_i^T F_i$, where $J_i = \partial(R(q)x_i + x(q))/\partial q$.

When N_c contacts are active, the total external joint torque is $\tau_e = \sum_{i=1}^{N_c} \tau_i(q)$. The parameters K_i, x_i^o, and x_i can be identified from least-squares fit on a dataset as $\min_{K,x^o,x} \sum_t \|\tau_t^m - \tau_e(q_t, K, x^o, x)\|$. These contact forces are composed of differentiable operations, and easily implemented in automatic differentiation (AD) frameworks to support their use in gradient-based planning.

2.2 Discontinuous Dynamics

The standard serial manipulator robot dynamics

$$M(q)\ddot{q} + C(q,\dot{q}) + B\dot{q} + G(q) = \tau_m + \tau_e \tag{2}$$

is assumed to be available in an AD framework, with joint position $q \in \mathbb{R}^n$, inertia matrix $M(q)$, Coriolis terms $C(q,\dot{q})$, viscous damping B, gravitational torque $G(q)$, input torque τ_m, and torques from contact τ_e.

We discriminate between different contact modes by letting the torque τ_e depend on the mode $z \in [1, \ldots, N_z]$, where each mode has different point contact models, resulting in a different τ_e. We suppose the mode changes during the task, denoting the value at t as z_t.

As typical in contact models, we use semi-implicit integration which handles stiff differential equations better [14]. For a time step Δ_t:

$$z_t \sim p(z_t|z_{t-1}) \tag{3}$$

$$\xi_t = \begin{bmatrix} q_t \\ \dot{q}_t \end{bmatrix} = \begin{bmatrix} q_{t-1} + \Delta_t \dot{q}_t \\ \dot{q}_{t-1} + \Delta_t M^{-1}(\tau_m + \tau_e(z_t) - B\dot{q}_t) \end{bmatrix} \tag{4}$$

where $p(z_t|z_{t-1})$ is a diagonal-dominant discrete transition matrix. As the external contact force $\tau_e(z_t)$ depends on the current discrete mode z_t, the dynamics can be written as:

$$\xi_t = \begin{bmatrix} q_t \\ \dot{q}_t \end{bmatrix} = f\left(\begin{bmatrix} q_{t-1} \\ \dot{q}_{t-1} \end{bmatrix}, z_{t-1}, \tau_{m,t-1}\right). \tag{5}$$

For an EKF, the dynamics in (5) would be linearized with respect to the state, adding a process noise contribution of $w \sim \mathcal{N}(0, Q)$. The observation model would comprise the external joint torques τ_e plus the joint positions as

$$y^z = C^z \xi + v, \tag{6}$$

where $C^z = [C_q^T, \frac{\partial \tau_e(z)}{\partial \xi}^T]^T$, $C_q = [I_n, 0_n]$, $v \sim \mathcal{N}(0, R)$ is the measurement noise, assumed to be independent and identically distributed, $\partial\tau_e(z)/\partial\xi$ is the Jacobian of the external contact torques $\tau_e(z)$ with respect to the joint states ξ, and I_n an identity matrix of dimension n.

2.3 Robot impedance dynamics

The robot is controlled with a Cartesian space impedance controller, where the 3 DoF impedance controlled robot dynamics are

$$\tau_m = J^T(q)\left[D\dot{x} + K\left(x - x_0\right)\right] + G(q) + C(q,\dot{q}) \tag{7}$$

where x, $\dot{x} \in \mathbb{R}^3$ are respectively the TCP Cartesian position and velocity, $J = \partial x/\partial q$ and the impedance parameters are $x_0 \in \mathbb{R}^3$ which represents the Cartesian impedance rest position, the damping $D \in \mathbb{R}^{3\times3}$, and stiffness $K \in \mathbb{R}^{3\times3}$ matrices, which are here all diagonal.

The dynamics in Eq.7 can then be updated to provide controlled discretized dynamics of

$$\begin{bmatrix} q_t \\ \dot{q}_t \end{bmatrix} = f\left(\begin{bmatrix} q_{t-1} \\ \dot{q}_{t-1} \end{bmatrix}, x_{0,t-1}, z_{t-1}\right). \tag{8}$$

3 Online Planning Method

This section introduces the belief estimation, cross-entropy and gradient-based planning method used.

3.1 Belief estimation

For estimating the hybrid state, we adopt a hybrid particle filter algorithm [26]. The proposed algorithm maintains an overall belief of the discrete mode $\hat{p}(z_t|y_{1:t})$, from the weighted belief of individual particles $\hat{p}(z^{(i)}|y_{1:t-1})$. In each step, each particle samples a discrete mode $\hat{z}_t^{(i)} \sim \hat{p}(z_t^{(i)}|y_{1:t-1})$, then propagates the mean $\mu_t^{(i)}$ and the covariance $\Sigma_t^{(i)}$ of continuous state ξ_t with the corresponding mode dynamics and observation equations via a standard Extended Kalman filter step [26]. Each particle $p^{(i)}$ is characterized by

the tuple $[\mu_t^{(i)}, \Sigma_t^{(i)}, \hat{p}(z_t^{(i)}|y_{1:t-1}), \hat{z}_t^{(i)}]$. An overview can be seen in Algorithm 1, where $\hat{z}_t^{(i)}$ is the sampled mode for particle i, $\hat{S}_t^{(i)} = \text{cov}(y_t|y_{1:t-1})$ the predicted observation covariance, $w_t^{(i)}$ the weight, $\hat{y}_{t|t-1}^{(i)} = \mathbb{E}(y_t|y_{1:t-1})$ is the predicted measurement, and δ is the Dirac delta function. The posterior belief $\hat{p}(z_t|y_{1:t})$ is used to detect changes in contact, and is used in the planning methods to weight the costs associated with different modes.

Algorithm 1 Hybrid filter

for all particles $p^{(i)}$ **do**
  $\hat{\mu}_{1|0}^{(i)} \leftarrow \mu_0$
  $\hat{\Sigma}_{1|0}^{(i)} \leftarrow \Sigma_0$
  $\hat{p}(z_1^{(i)}|y_1) \leftarrow p(z_0)$
end for
for all time step t **do**
  for all particles $p^{(i)}$ **do**
    $\hat{p}(z_t^{(i)}|, y_{1:t-1}) = \hat{p}(z_{t-1}^{(i)}|y_{1:t-1})p(z_t|z_{t-1})$
    $\hat{z}_t{}^{(i)} \sim \hat{p}(z_t^{(i)}|, y_{1:t-1})$
    $(\hat{y}_{t|t-1}^{(i)}, \hat{S}_t^{(i)}, \hat{\mu}_t^{(i)}, \hat{\Sigma}_t^{(i)})$
      $\leftarrow EKF(\mu_{t-1}^{(i)}, \Sigma_{t-1}^{(i)}, y_t, \hat{z}_t{}^{(i)})$
    $w_t^{(i)} \propto \mathcal{N}(y_t; \hat{y}_{t|t-1}^{(i)}, \hat{S}_t^{(i)})$
  end for
  Resample particles $\{p^{(i)}, w_t^{(i)}\}_{i=1}^N$
  Estimate marginal posterior belief as
    $\hat{p}(z_t|y_{1:t}) \propto \sum_{i=1}^N w_t^{(i)} \delta_{z_t}(\hat{z}_t^{(i)})$
end for

3.2 Proposed Cross-Entropy Method

The Cross-Entropy Method (CEM) samples actions from an action distribution, rolls out these trajectories and evaluates their resulting costs, adapting the action distribution to reduce loss. Sampling-based approaches may avoid local minima problems or convergence issues when gradients are not smooth due to the discontinuous dynamics, potentially reaching state-action pairs that would not be reached by gradient-based methods. Thus, we use CEM iterations before proceeding with a gradient-based non-linear solver.

The adopted CEM algorithm is based on the work of [21]. The CEM here is initialized from the hybrid filter at each time step t with the set of particles consisting of continuous joint states values $\xi_t^{(i)}$, their weight $w_t^{(i)}$ and discrete sampled state $z_t^{(i)}$,

$$\left[\xi_t^{(i)}, w_t^{(i)}, z_t^{(i)}\right]_{i=1}^N. \tag{9}$$

This list is used for calling the dynamics with the hybrid mode $z_t^{(i)}$ with initial continuous state $\xi_t^{(i)}$ along with the sampled actions $u_{t:t+h}$. As in typical CEM, actions are iteratively sampled and the sampling distribution shifted towards the elite samples which have the best loss. A detailed explanation of the steps is given in Algorithm 2. In Algorithm 2,

Algorithm 2 Proposed iCEM

Require: N: number of particles, h: planning horizon, d: action dimension, β: colored noise exponent, N_{iter}: CEM-iterations, $\left[\xi_t^{(i)}, w_t^{(i)}, z_t^{(i)}\right]_{i=1}^N$: list of particles, K: elite size
Initialize:
  $u_{t:t+h} = u_{t-1:t+h-1}^{MPC}$
  $\Sigma^u = \mathbf{1}$
Loop:
  for k in N_{iter} **do**
    noise samples $\leftarrow$ N samples from $C^\beta(d, h)$
    action samples $u_{t:t+h}^{(i)}$
      $\leftarrow$ N samples from $\texttt{clip}(u_t + C^\beta(d, h) \cdot \Sigma^u)$
    normalize weights $w_t^{(i)} \leftarrow \frac{w_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}}$
    trajectory rollouts $\xi_{t+1:t+1+h}^{(i)}$
      $\leftarrow f(\xi_t^{(i)}, u_{t:t+h}^{(i)}, z_t^{(i)})$
    costs $l^{(i)} \leftarrow$
      cost of each $\{\xi_{t+1:t+1+h}^{(i)}, u_{t:t+h}^{(i)}\}_{i=1}^N$
    normalize costs $l^{(i)} \leftarrow \frac{l^{(i)}}{w_t^{(i)}}$
    elite set $\leftarrow$ best K samples according to cost
    $u_{t:t+h}, \Sigma^u \leftarrow$ fit Gaussian distribution to elite set
  end for
return best action trajectory $u_{t:t+h}^{iCEM}$

$u_{t:t+h}$ and Σ^u represent the mean and the standard deviation of the action sequences, at the beginning of each time step t we initialise the mean with the optimal value found at the previous time step $t-1$ by the non linear MPC solver while the standard deviation of the sequences is re-initialised every time with the matrix $\mathbf{1} \in \mathbb{R}^{d \times h}$, whose elements are ones. $C^\beta(d, h)$ represents the colored noise distribution, built following the approach in [21]. Within the CEM inner loop, we keep updating the action sequences statistics by fitting at each iteration the best K trajectories; as long as the solver converges the mean $u_{t:t+h}$ is initialized at each time step with the optimal solution and the likelihood of sampling optimal solutions increases. The cost of each particle $l^{(i)}$ is divided by the particle weight, so a low likelihood particle results in a higher cost, to avoid actions which favor low-likelihood particles from being sampled more heavily. The cost function itself is the same as described in the following section for the MPC problem.

The CEM returns both the best action trajectory, along with the associated state trajectory, which are used to initialize the decision variables in a gradient-based MPC problem.

3.3 MPC problem

The MPC problem is formulated using multiple-shooting transcription with a generic problem statement of

$$\min_{u_{t:t+h}} \sum_{t}^{t+h} \sum_{z=1}^{N_z} \hat{p}(z|y_{1:t}) l^z(\xi_t^z, u_t) \tag{10}$$

$$\text{s.t. } \forall z \in [1, \dots, N_z], \quad \tau \in [t, \dots, t+h-1]: \tag{11}$$

$$\|\xi_{\tau+1}^z - f(\xi_\tau^z, u_\tau)\| \leq \rho \tag{12}$$

$$g(\xi_t^z, u_t) \geq 0 \tag{13}$$

$$u \in \mathcal{U} \tag{14}$$

where h is the planning horizon, $\mathcal{U}$ is the range of allowed inputs, ρ is the slack for continuity constraints (the inequality is applied element-wise), $z \in [1, \dots, N_z]$ represents the set of the different contact modes, and g represent general state and input dependent inequality constraints. A trajectory is rolled out for each mode, ξ_t^z, where continuity is imposed according to that mode's dynamics in f; $\hat{p}(z|y_{1:t})$ is the contact state belief information coming from the hybrid filter at time step t. The constraints shown in the optimization problem are nonlinear, so an interior-point nonlinear optimization solver is used. While nonlinear, the problem is written in an automatic differentiation frame- work, allowing calculation of the gradient and Hessian of the objective and constraints, significantly improving convergence.

The general stage cost function, which would be used also in the CEM loop, is defined as

$$\begin{aligned} l^z(\xi^z, u) =& (x_d - x)^T Q_x (x_d - x) \\ &+ (u - x)^T Q_u (u - x) + \dot{x}^T Q_{\dot{x}} \dot{x} \end{aligned} \tag{15}$$

where x_d is a desired position in world frame to be tracked, x and $\dot{x}$ are the Cartesian position and velocity of the robot TCP expressed with the forward kinematics function, directly depending on the joint configuration. In this scenario, the state $\xi^z = (q^z, \dot{q}^z)$ is represented by the robot joint variables for each mode z, while the adopted control input u is shared among the different modes and is impedance rest position x_0. The constraint is $g = \overline{F}_{imp} - \|K(x - x_0)\|_2$, limiting the virtual impedance force applied to $\overline{F}_{imp}$.

4 Simulation Study

To study the proposed approach, studies in simulation are used to investigate the solve time and resulting performance. To see the variance in the performance, each problem is solved repeatedly with the initial state ξ_t perturbed with $\mathcal{N}(0, 0.1)$ between solves, finding the mean and covariance.

4.1 CEM vs MPC in stiff contact

To compare qualitative properties of sampling- and gradient-based solvers, a simple tracking problem in stiff contact $K = 1e4$ N/m is set up with a cost function to track desired position. As can be seen in Figure 3, the CEM (solid line) has more variation in the control trajectory, and the state trajectory is slower to converge than the MPC (dash-dotted line).

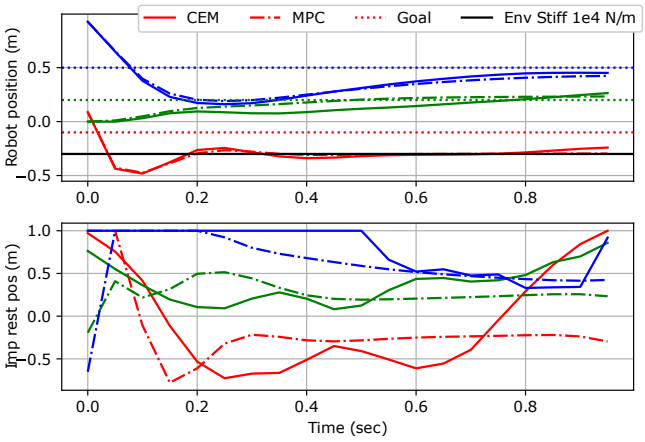


Figure 3: Comparing state and control trajectory from CEM vs MPC for stiff dynamics.

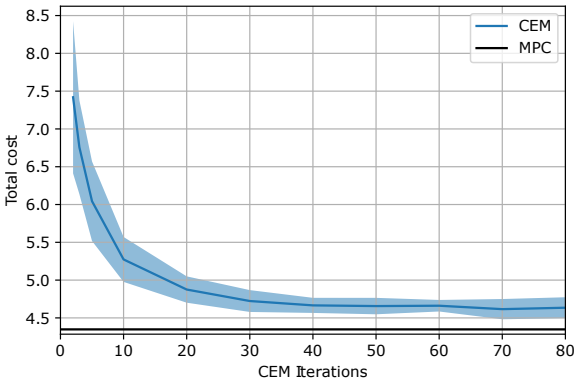


Figure 4: Comparing the resulting cost for a CEM- and MPC-only solve with stiff dynamics. While the CEM approaches the MPC performance at a high number of iterations, it does not fully converge.

The mean and variance as N_{iter} is varied can be seen in Figure 4, where the CEM performance saturates before approaching that of the MPC solver. The variance of the CEM performance decreases at higher iterations, but remains orders of magnitude larger than the variance of the MPC solver $(1e-9)$.

4.2 CEM Warmstart on total solve time

To study the relative performance of the gradient- and sampling-based methods in discontinuous dynamics, we examine two performance characteristics: time to solve and resulting expected cost.

First, the solver is used to solve a free-space problem according to (10) with horizon $h = 20$ steps and stepsize

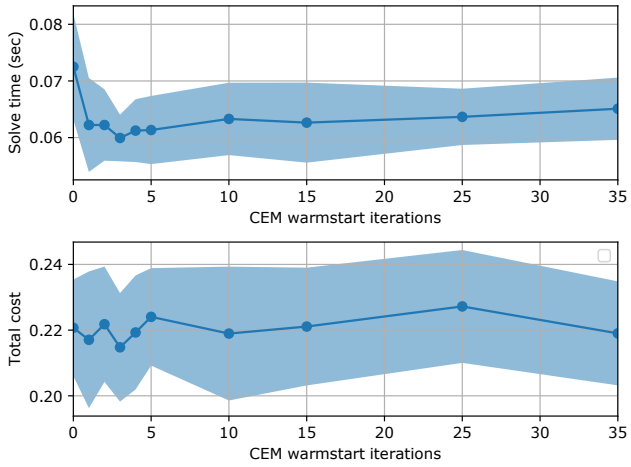
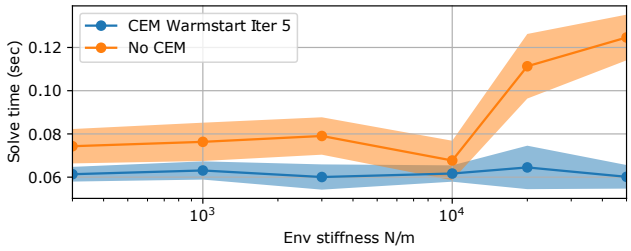
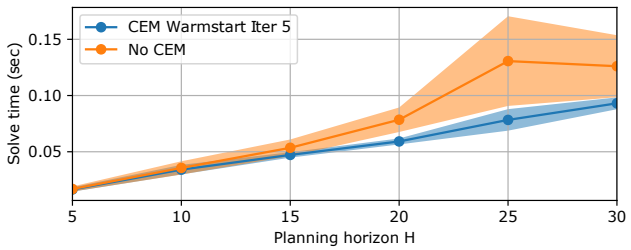


Figure 5: Effect of the number of CEM warmstart iterations on total solve time (CEM+MPC)



(a) Effect of stiffness



(b) Effect of planning horizon

Figure 6: Effect of problem parameters (stiffness and planning horizon) on solve time for only MPC (orange) and proposed approach (blue).

$\Delta_t = 0.05$ seconds. Then, the dynamics is changed to a single point contact model of stiffness $K = 5e3$ N/m and the previous solution used initialize the current step. This initialization is then used for N_{iter} iterations of the CEM before starting the next MPC solve.

The mean and covariance of the total solve time for the CEM iterations and MPC solve can be seen in Figure 5. As can be seen, the CEM helps to reduce both the mean and variance of the solve time, even from just one iteration. However, as the number of CEM iterations increases, the time required increases, indicating the additional time required to do more CEM iterations is not saving as much time in the MPC solver.

To analyze the impact of important process parameters, we study the improvement by $N_{iter} = 5$ warmstarts over a range of stiffnesses and planning horizons. The results can be seen in Figure 6, where as the stiffness and horizon increase, the CEM improves both the mean and variance of solve time. This trend matches the impact of stiffness and planning horizon in policy gradient estimation [12], where gradient-free methods performed better at long horizons and higher stiffnesses.

5 Experiments

To validate the performance on real-world contact tasks, we apply the proposed method to the two tasks seen in Figure 1. The code and data are available at https://gitlab.cc-asp.fraunhofer.de/hanikevi/contactmpc.

5.1 Vertical contact task

In the vertical contact task, the robot moves from free space into contact with an environment. The stiffness of the environment is adjusted by adding the yellow or purple foam seen in Figure 1, or direct contact with the aluminum table. In Figure 7 we can see how the stiffness parameter influences the contact mode detection responsiveness. Despite the developed CEM+MPC approach successfully worked for all the contact environments, from the plot we can see that with stiffer contacts the mode detection is more sharp, with the transition triggered right in correspondence of the vertical force peak. For lower stiffness contacts, mode detection is slightly delayed with respect to the force peak, probably because the lower measured interaction force makes it more difficult for the filter to assign different likelihood values.

5.2 Pivot task

In this task, the robot makes contact with the table, slides along the surface until making contact with the wall. The MPC problem is solved with $h = 13$, $\Delta_t = 0.04$, an impedance cost of $Q_u = 0.05$, maximum impedance force of $\overline{F}_{imp} = 25$, a desired pose of $x^d = [-, -, 0.01]$ in free space, and $x^d = [0.35, -, 0.01]$ in plane and hinge modes, where $-$ are omitted from the cost. In this experimental setup, the proposed approach was compared to a simple MPC approach without the CEM warm-start, in order to validate experimentally the results achieved in simulation. From the plots in Figure 8, it can be seen that without the CEM warm-start, the

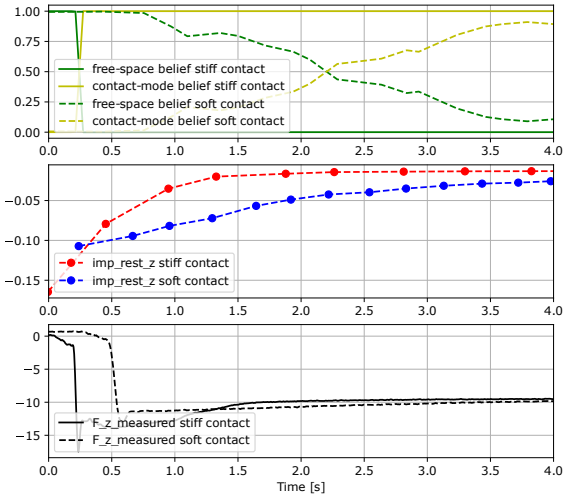


Figure 7: Effect of stiffness on contact detection

impedance rest position shows a small discontinuity in its trend, and this can be noticed especially when the first transition is triggered. This result is in line with what we have shown in the simulation study, the proposed CEM+MPC approach helps in decreasing the total solve time at contact transitions.

6 Conclusion

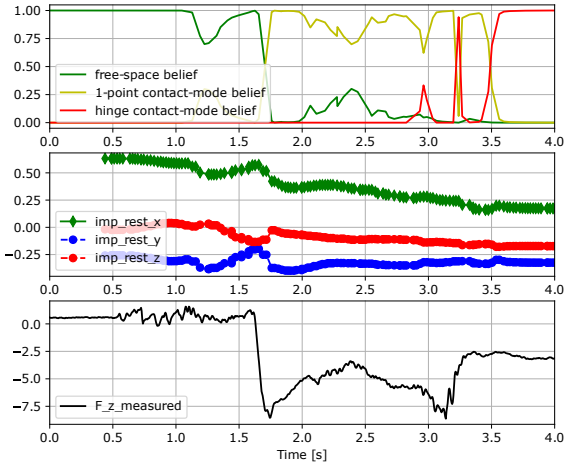
This paper proposed an approach to planning on discontinuous dynamics which uses a modified CEM to initialize a gradient-based planner. It was shown that such an approach can reduce the variation in solve time when discontinuities in dynamics occur. The approach was also vlidated on real contact tasks, showing that planning method results in safe robot trajectories which respond to the contact conditions.

References

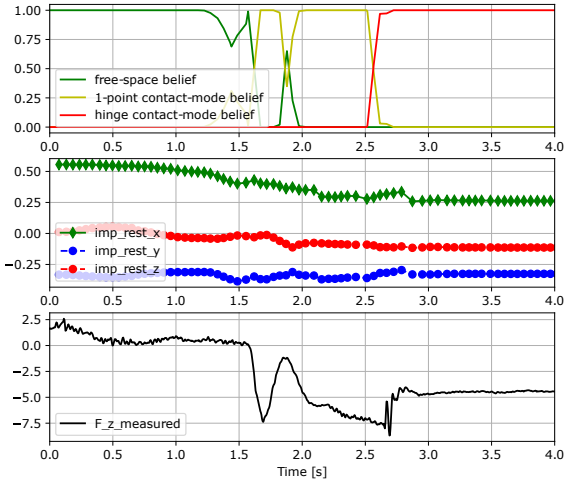
[1] M. Anitescu and F. A. Potra, "Formulating Dynamic Multi-Rigid-Body Contact Problems with Friction as Solvable Linear Complementarity Problems," p. 17.

[2] M. Posa, C. Cantu, and R. Tedrake, "A direct method for trajectory optimization of rigid bodies through contact," *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 69–81, 2014.

[3] S. L. Cleac'h, T. Howell, M. Schwager, and Z. Manchester, "Fast Contact-Implicit Model-Predictive Control," *arXiv:2107.05616 [cs, eess]*, Sep. 2021.



(a) Simple gradient-based MPC



(b) Proposed CEM+MPC approach

Figure 8: Comparison between the simple MPC and CEM+MPC approaches for the pivot task.

[4] A. Aydinoglu and M. Posa, "Real-Time Multi-Contact Model Predictive Control via ADMM," *arXiv:2109.07076 [cs]*, Mar. 2022.

[5] W. Jin and M. Posa, "Task-Driven Hybrid Model Reduction for Dexterous Manipulation," Nov. 2022.

[6] F. R. Hogan and A. Rodriguez, "Reactive planar non-prehensile manipulation with hybrid model predictive control," *The International Journal of Robotics Research*, vol. 39, no. 7, pp. 755–773, Jun. 2020.

[7] C. Chen, P. Culbertson, M. Lepert, M. Schwager, and J. Bohg, "TrajectoTree: Trajectory Optimization Meets Tree Search for Planning Multi-contact Dexterous Manipulation," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2021, pp. 8262–8268.

[8] K. Wu and G. Hao, "Design and nonlinear modeling of a novel planar compliant parallelogram mechanism with general tensural-compresural beams," *Mechanism and Machine Theory*, vol. 152, p. 103950, 2020.

[9] T. Pang, H. J. T. Suh, L. Yang, and R. Tedrake, "Global Planning for Contact-Rich Manipulation via Local Smoothing of Quasi-dynamic Contact Models," Jun. 2022.

[10] H. Bharadhwaj, K. Xie, and F. Shkurti, "Model-Predictive Control via Cross-Entropy and Gradient-Based Optimization," Apr. 2020.

[11] K. Huang, S. Lale, U. Rosolia, Y. Shi, and A. Anandkumar, "CEM-GD: Cross-Entropy Method with Gradient Descent Planner for Model-Based Reinforcement Learning," Dec. 2021.

[12] H. J. T. Suh, T. Pang, and R. Tedrake, "Bundled Gradients through Contact via Randomized Smoothing," *arXiv:2109.05143 [cs]*, Jan. 2022.

[13] A. Ö. Önol, R. Corcodel, P. Long, and T. Padır, "Tuning-Free Contact-Implicit Trajectory Optimization," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, May 2020, pp. 1183–1189.

[14] D. E. Stewart, "Rigid-Body Dynamics with Friction and Impact," *SIAM Review*, vol. 42, no. 1, pp. 3–39, Jan. 2000.

[15] F. R. Hogan, J. Ballester, S. Dong, and A. Rodriguez, "Tactile Dexterity: Manipulation Primitives with Tactile Feedback," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, May 2020, pp. 8863–8869.

[16] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model Predictive Control: Theory, Computation, and Design*. Nob Hill Publishing Madison, 2017, vol. 2.

[17] A. Nagabandi, G. Kahn, R. S. Fearing, and S. Levine, "Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning," *CoRR*, vol. abs/1708.02596, 2017. [Online]. Available: http://arxiv.org/abs/1708.02596

[18] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. M. Rehg, B. Boots, and E. A. Theodorou, "Information theoretic MPC for model-based reinforcement learning," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 1714–1721.

[19] R. Y. Rubinstein, "Optimization of computer simulation models with rare events," *European Journal of Operational Research*, vol. 99, no. 1, pp. 89–112, May 1997. [Online]. Available: https://ideas.repec.org/a/eee/ejores/v99y1997i1p89-112.html

[20] M. Okada and T. Taniguchi, "Variational inference MPC for bayesian model-based reinforcement learning," *CoRR*, vol. abs/1907.04202, 2019. [Online]. Available: http://arxiv.org/abs/1907.04202

[21] C. Pinneri, S. Sawant, S. Blaes, J. Achterhold, J. Stueckler, M. Rolinek, and G. Martius, "Sample-efficient Cross-Entropy Method for Real-time Planning," Aug. 2020.

[22] L. Roveda, J. Maskani, P. Franceschi, A. Abdi, F. Braghin, L. M. Tosatti, and N. Pedrocchi, "Model-based reinforcement learning variable impedance control for human-robot collaboration," *Journal of Intelligent & Robotic Systems*, pp. 1–17, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:216293188

[23] K. Chua, R. Calandra, R. McAllister, and S. Levine, "Deep Reinforcement Learning in a Handful of Trials using Probabilistic Dynamics Models," *arXiv:1805.12114 [cs, stat]*, May 2018.

[24] H. J. T. Suh, M. Simchowitz, K. Zhang, and R. Tedrake, "Do Differentiable Simulators Give Better Policy Gradients?" Aug. 2022.

[25] A. M. Castro, A. Qu, N. Kuppuswamy, A. Alspach, and M. Sherman, "A Transition-Aware Method for the Simulation of Compliant Contact With Regularized Friction," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1859–1866, Apr. 2020.

[26] S. Thrun, "Probabilistic robotics," *Communications of the ACM*, vol. 45, no. 3, pp. 52–57, 2002.