

MODEL PREDICTIVE PATH INTEGRAL CONTROL FOR SPACECRAFT RENDEZVOUS AND PROXIMITY OPERATIONS ON ELLIPTIC ORBITS

Tomohiro Sasaki*, Koki Ho[†], and E. Glenn Lightsey[‡]

This paper presents a nonlinear control framework for spacecraft rendezvous and proximity operations on elliptic orbits using Model Predictive Path Integral (MPPI) control. Path integral control is a sampling-based nonlinear stochastic optimal control algorithm that can avoid linear and quadratic approximations in both dynamics and cost functions. While this control method has gained popularity in the robotics community due to its algorithmic effectiveness, it remains unexplored in astrodynamics. This paper demonstrates a comprehensive closed-loop simulation of spacecraft rendezvous employing MPPI and evaluates its control performance through these simulations.

INTRODUCTION

Path integral control, rooted in sampling-based optimal control,^{1,2} has made notable strides in the field of robotics. Path integral control delivers a framework to solve a finite horizon nonlinear optimal control problem efficiently. Particularly, the path integral controller is extremely powerful when it incorporates nonlinear dynamics, stochastic dynamics, high-dimensional state space, and constraints due to its natural features of robustness to modeling errors and disturbances.

A standout application of path integral control is its use in predictive control, termed Model Predictive Path Integral Control (MPPI).^{3,4} Recent research reveals its potential across various areas, including autonomous vehicles,⁵ quadrotors,⁵ fixed-wing aircraft,⁶ vertical take-off and landing aircraft,⁷ robot manipulation,⁸ and space robotic manipulation.⁹ Given these successes, there is a reason to believe MPPI could also be valuable for spacecraft rendezvous and proximity operations.

MPPI's strength lies in its ability to optimize using samples without making assumptions about system dynamics or cost functions. The original idea of the path integral control is that the value function of the optimal control problem, which is the minimum value of the cost function, is transformed into an expectation over the possible trajectories known as path integral using the Feynman-Kac lemma.¹⁰ Then, the stochastic Hamilton-Bellman-Jacobi (HJB) equation of the optimal control problem, a stochastic partial differential equation, can be transformed into a stochastic differential equation. While the HJB equation can be solved backward in time, the stochastic differential equation can be solved forward in time with a forward sampling of stochastic diffusion processes. The classical path integral control assumes that the dynamics must be affine in control and satisfy

*Graduate Student, Daniel Guggenheim School of Aerospace Engineering, Atlanta, GA 30332 USA.

[†]Dutton-Duoffe Professor, Daniel Guggenheim School of Aerospace Engineering, Atlanta, GA 30332 USA.

[‡]David Lewis Professor of Space Systems Technology, Daniel Guggenheim School of Aerospace Engineering, Atlanta, GA 30332 USA.

a specific relationship between noise and controls. However, Theodorou and Todorov generalized the path integral framework by making a connection between classical path integral control and information theoretic notions of free energy and relative entropy, known as Kullback–Leibler (KL) divergence.² Furthermore, Williams et. al generalized the framework for a complete control non-affine system.⁵

Historically, relative motion control, a subclass of astrodynamics guidance, navigation, and control problems, has been achieved by open-loop control and error correction control each time. This type of control approach is sometimes undesirable regarding safe and robust spacecraft relative motion control. With increased computing capability, model predictive control (MPC) is becoming popular in space applications as a modern feedback control algorithm. The MPC controller solves a deterministic finite optimal control problem at every control update step and uses the first component of the generated control sequence. The MPC controller has advantages over traditional open-loop and feedback control algorithms in handling state and control constraints and effectively predicting the system’s dynamic behavior. Linear MPC with quadratic programming has already shown its effectiveness in spacecraft rendezvous and docking on circular orbits.^{11,12} Moreover, linear MPC has demonstrated simulation performance for spacecraft rendezvous on near-rectilinear Halo orbits.¹³ Nonlinear MPC with sequential quadratic programming has also been studied for rendezvous and proximity operations. Recent interest in rendezvous and proximity operations study is shifting towards robust nonlinear feedback control for achieving robust and safe maneuvers.

This paper presents a path integral control framework for spacecraft rendezvous and proximity operations under a predictive control setting and aims to explore how well sampling-based controllers like MPPI can be adapted to astrodynamics applications.

DYNAMICS MODEL

This paper investigates the high-fidelity relative dynamic model for achieving high-precision rendezvous and proximity operations. This paper considers the chief-deputy representation for modeling spacecraft rendezvous and proximity operations. The chief-deputy model defines the deputy position and velocity with respect to the chief state. Hence, the chief-deputy model does not have to deal with both chief and deputy absolute position and velocity but deals with chief absolute position and velocity and deputy relative position and velocity.

This paper leverages the nonlinear dynamics model developed by Xu and Wang,¹⁴ which provides the representation of nonlinear relative dynamics in the presence of J_2 perturbations without assuming specifics about the chief spacecraft orbits. Two distinct coordinate frames must be defined to derive the equations of motion for relative motion, as depicted in Figure 1. An Earth-centered inertial (ECI) coordinate frame is defined to locate the position of the chief spacecraft in an inertial frame. Here, the $\hat{\mathbf{X}}$ direction points toward the vernal equinox, the $\hat{\mathbf{Z}}$ direction toward the North pole, and the $\hat{\mathbf{Y}}$ direction completes the right-handed coordinate frame. Subsequently, a local-vertical-local-horizontal (LVLH) frame is defined to locate the position of the deputy spacecraft with respect to the chief spacecraft. In this frame, the $\hat{\mathbf{x}}$ direction aligns with the chief’s position vector pointing along the radial vector, and the $\hat{\mathbf{z}}$ aligns with the angular momentum vector of the chief spacecraft, and the $\hat{\mathbf{y}}$ direction completes the right-handed coordinate frame:

$$\hat{\mathbf{x}} = \mathbf{r}/r, \quad \hat{\mathbf{z}} = \mathbf{h}/h, \quad \hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{x}}$$

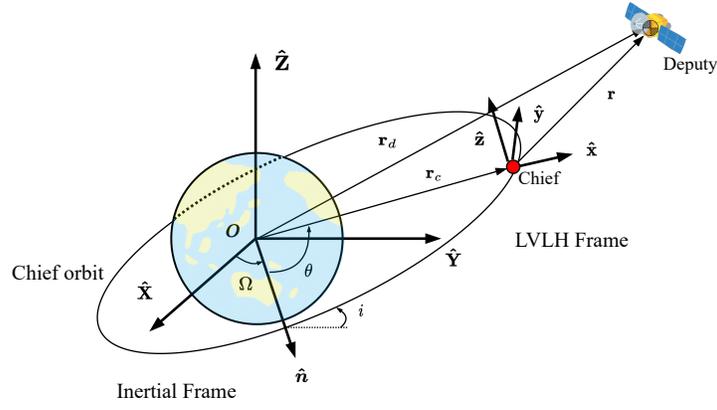


Figure 1. ECI and LVLH coordinate frames

Now, the nonlinear equations of motion of the chief orbit are expressed in terms of orbital elements as follows:

$$\begin{aligned}
 \dot{r} &= v_x \\
 \dot{v}_x &= -\frac{\mu}{r^2} + \frac{h^2}{r^3} - \frac{k_{J_2}}{r^4}(1 - 3 \sin^2 i \sin^2 \theta) \\
 \dot{h} &= -\frac{k_{J_2} \sin^2 i \sin 2\theta}{r^3} \\
 \dot{\theta} &= \frac{h}{r^2} + \frac{2k_{J_2} \cos^2 i \sin^2 \theta}{hr^3} \\
 \dot{i} &= -\frac{k_{J_2} \sin 2i \sin 2\theta}{2hr^3} \\
 \dot{\Omega} &= -\frac{2k_{J_2} \cos i \sin^2 \theta}{hr^3}
 \end{aligned} \tag{1}$$

where r is the magnitude of the position vector, v_x is the magnitude of the radial velocity vector, h_{ang} is the magnitude of the specific angular momentum, i is the inclination, Ω is the right ascension of the ascending node, and θ is the argument of latitude, $k_{J_2} = 3J_2\mu R_e^2/2$ is a constant parameter of J_2 perturbation. J_2 is the second zonal harmonic coefficient of the Earth, μ is the Earth's gravitational parameter, and R_e is the Earth's equatorial radius.

Next, the nonlinear equations of motion of relative motion for j -th deputy spacecraft with respect to the chief spacecraft are as follows:

$$\begin{aligned}
 \ddot{x}_j &= 2\dot{y}_j\omega_z - x_j(n_j^2 - \omega_z^2) + y_j\dot{\omega}_z - z_j\omega_x\omega_z - (\zeta_j - \zeta)\sin i \sin \theta - r(n_j^2 - n^2) + u_{j,x} \\
 \ddot{y}_j &= -2\dot{x}_j\dot{\omega}_z + 2\dot{z}_j\omega_x - x_j\dot{\omega}_z - y_j(n_j^2 - \omega_z^2 - \omega_x^2) + z_j\dot{\omega}_x - (\zeta_j - \zeta)\sin i \cos \theta + u_{j,y} \\
 \ddot{z}_j &= -2\dot{y}_j\omega_x - x_j\dot{\omega}_x\omega_z - y_j\dot{\omega}_x - z_j(n_j^2 - \omega_x^2) - (\zeta_j - \zeta)\cos i + u_{j,z}
 \end{aligned} \tag{2}$$

where (ω_x, ω_z) and $(\dot{\omega}_x, \dot{\omega}_z)$ are angular velocities and accelerations of the chief spacecraft, respectively given by:

$$\begin{aligned}
\omega_x &= -\frac{k_{J2}(\sin 2i \sin \theta)}{(hr^3)}, & \omega_z &= \frac{h}{r^2} \\
\dot{\omega}_x &= -\frac{k_{J2}(\sin 2i \cos \theta)}{r^5} + \frac{3v_x k_{J2}(\sin 2i \sin \theta)}{(r^4 h)} \\
&\quad - \frac{8k_{J2}^2 \sin^3 i \cos i \sin^2 \theta \cos \theta}{(r^6 h^2)} \\
\dot{\omega}_z &= -\frac{2hv_x}{r^3} - \frac{k_{J2}(\sin^2 i \sin 2\theta)}{r^5}
\end{aligned}$$

and n, n_j, ζ, ζ_j are given by:

$$\begin{aligned}
n^2 &= \frac{\mu}{r^3} + \frac{k_{J2}}{r^5} - \frac{5k_{J2} \sin^2 i \sin^2 \theta}{r^5} \\
n_j^2 &= \frac{\mu}{r_j^3} + \frac{k_{J2}}{r_j^5} - \frac{5k_{J2} \sin^2 i \sin^2 \theta}{r_j^7} \\
\zeta &= \frac{2k_{J2} \sin i \sin \theta}{r^4}, & \zeta_j &= \frac{2k_{J2} r_j Z}{r_j^5}
\end{aligned}$$

with

$$\begin{aligned}
r_j &= \sqrt{(r + x_j)^2 + y_j^2 + z_j^2} \\
r_j Z &= (r + x_j) \sin i \sin \theta + y_j \sin i \cos \theta + z_j \cos i.
\end{aligned}$$

This paper considers a single chief and single deputy system, and thus $j = 1$ for all time.

PATH INTEGRAL CONTROL

This section introduces information theoretic path integral control with importance sampling. The following discrete-time stochastic dynamical system is considered in this paper:

$$\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t, \mathbf{v}_t), \quad k = 0, 1, \dots, N - 1 \quad (3)$$

where $\mathbf{x}_t \in \mathbb{R}^n$ denotes the state of the dynamical system at time step t , $\mathbf{v}_t \in \mathbb{R}^m$ denotes a control input for the system, \mathbf{F} denotes the discrete nonlinear state-transition function of the system. It is assumed that the system does not have direct control over the input variable \mathbf{v}_t , but the system transforms control mean input \mathbf{u}_t into the actual control input \mathbf{v}_t by a white-noise process with the density function as

$$\mathbf{v}_t \sim \mathcal{N}(\mathbf{u}_t, \Sigma), \quad (4)$$

where $\Sigma \in \mathbb{R}^{m \times m}$ denotes the covariance of the control variable. This is a valid noise assumption for spacecraft rendezvous and proximity operations since the controller input has to go through a

lower-level control block before reaching the actual plant. The next step is to construct sequences of input variables and mean input variables as

$$\begin{aligned}\mathbf{V} &= [\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}] \in \mathbb{R}^{m \times N}, \\ \mathbf{U} &= [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1}] \in \mathbb{R}^{m \times N}.\end{aligned}$$

There are two different distributions to address the optimal control process. One distribution is the probability distribution \mathbb{P} of an input sequence \mathbf{V} in the uncontrolled system ($\mathbf{U} \equiv 0$):

$$\mathbf{p}(\mathbf{V}) = Z^{-1} \prod_{t=0}^{N-1} \exp\left(-\frac{1}{2} \mathbf{v}_t^T \Sigma \mathbf{v}_t\right), \quad (5)$$

and the other distribution is the probability distribution \mathbb{Q} of an open-loop input sequence in the controlled system:

$$\mathbf{q}(\mathbf{V}) = Z^{-1} \prod_{t=0}^{N-1} \exp\left(-\frac{1}{2} (\mathbf{v}_t - \mathbf{u}_t)^T \Sigma (\mathbf{v}_t - \mathbf{u}_t)\right), \quad (6)$$

where $Z = ((2\pi)^m |\Sigma|)^{\frac{1}{2}}$. Here, the discrete-time optimal control problem is assumed to be of the form of:

$$\min_{\mathbf{U} \in \mathcal{U}} \mathbb{E}_{\mathbb{Q}} \left[\phi(\mathbf{x}_N) + \sum_{t=0}^{N-1} \ell(\mathbf{x}_t, \mathbf{u}_t) \right], \quad (7)$$

where \mathcal{U} is the set of admissible control sequences, $\ell(\cdot, \cdot)$ is the instantaneous cost, and $\phi(\cdot)$ is the terminal cost. The running cost is further defined with a state-dependent cost and a control cost as

$$\ell(\mathbf{x}_t, \mathbf{u}_t) = q(\mathbf{x}_t) + \frac{\lambda}{2} \mathbf{u}_t^T \Sigma^{-1} \mathbf{u}_t, \quad (8)$$

where $\lambda \in \mathbb{R}^+$ denotes the inverse temperature of the free energy of the control system and is the hyper-parameter of this control algorithm. Note that the state violation cost, such as safe constraints, can be taken into account in the state-dependent cost $q(\cdot)$. Let C be the total state-dependent cost of the trajectory given by:

$$C(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N) = \phi(\mathbf{x}_N) + \sum_{t=0}^{N-1} q(\mathbf{x}_t). \quad (9)$$

Let \mathcal{H} be a mapping from the input sequence along with an initial condition to a resulting trajectory. Then, the state-dependent cost given an input sequence is defined as

$$S(\mathbf{V}) = C(\mathcal{H}(\mathbf{V}; \mathbf{x}_0)) \quad (10)$$

The free energy of the control system is now defined as:

$$\mathcal{F}(S, \mathbf{p}, \mathbf{x}_0, \lambda) = -\lambda \log \left(\mathbb{E}_{\mathbb{P}} \left[-\frac{1}{\lambda} S(\mathbf{V}) \right] \right), \quad (11)$$

where the expectation is taken with respect to \mathbb{P} . Applying Jensen's inequality, we can find the upper bound of the free energy as:

$$\mathcal{F}(S, \mathbf{p}, \mathbf{x}_0, \lambda) \leq \mathbb{E}_{\mathbb{Q}} [S(\mathbf{V})] + \lambda \mathbb{D}_{\text{KL}}(\mathbb{Q}|\mathbb{P}), \quad (12)$$

where $\mathbb{D}_{\text{KL}}(\mathbb{Q}|\mathbb{P})$ denotes the Kullback-Leibler (KL) divergence defined as follows:

$$\mathbb{D}_{\text{KL}}(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}} \left[\log \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right].$$

The KL divergence is a measure for comparing distances between two probability distributions. In this case, the KL divergence is a distance between uncontrolled and controlled distributions. Let \mathbb{Q}^* be optimal distribution, then the free energy is minimized computed with S and \mathbb{Q}^* . Now Equation (12) implies the optimization problem for the given dynamical system and control input where free energy is the lower bound of the inequality. Practically speaking, optimal distribution \mathbb{Q}^* cannot directly optimized. Instead, the controlled distribution \mathbb{Q} is shifted as close as possible to the optimal distribution \mathbb{Q}^* . The minimization problem that has to be solved can be transformed into:

$$\mathbf{U}^* = \arg \min_{\mathbf{U} \in \mathcal{U}} \mathbb{D}_{\text{KL}}(\mathbb{Q}^*|\mathbb{Q}). \quad (13)$$

In the literature,⁵ the authors showed that the solution to $\min_{\mathbf{U}} \mathbb{D}_{\text{KL}}(\mathbb{Q}^*|\mathbb{Q})$ as

$$\mathbf{u}_t = \int \mathbf{q}^*(\mathbf{V}) \mathbf{v}_t d\mathbf{V}. \quad (14)$$

where $\mathbf{q}^*(\mathbf{V})$ is the density function of optimal distribution \mathbb{Q} given by:

$$\mathbf{q}^*(\mathbf{V}) = \frac{1}{\eta} \exp \left(-\frac{1}{\lambda} S(\mathbf{V}) \right) \mathbf{p}(\mathbf{V}). \quad (15)$$

Typical information theoretic MPPI utilizes iterative importance sampling to estimate the optimal control solution given a current control distribution. Rewrite Equation (14) as:

$$\int \mathbf{q}(\mathbf{V}) \frac{\mathbf{q}^*(\mathbf{V}) \mathbf{p}(\mathbf{V})}{\mathbf{p}(\mathbf{V}) \mathbf{q}(\mathbf{V})} \mathbf{v}_k d\mathbf{V}. \quad (16)$$

Thus, the optimal control input can be expressed in terms of an expectation with respect to \mathbb{Q} as:

$$\mathbf{u}_t^* = \mathbb{E}_{\mathbb{Q}} [w(\mathbf{V}) \mathbf{v}_t], \quad (17)$$

where \mathbf{w} is the importance sampling weight given by:

$$w(\mathbf{V}) = \frac{1}{\eta} \exp \left(-\frac{1}{\lambda} S(\mathbf{V}) + \sum_{t=0}^{N-1} -\mathbf{v}_t^T \Sigma^{-1} \mathbf{v}_t + \frac{1}{2} \mathbf{u}_t^T \Sigma^{-1} \mathbf{u}_t \right) \quad (18)$$

Algorithm 1: Model Predictive Path Integral Control Simulation

Input: f : Transition model
 K : Number of samples
 N : Number of timesteps
 \mathbf{U} : Initial control sequence
 Σ, ϕ, q, λ : Control hyper-paramter
 \mathbf{X}_{ref} : Reference trajectory

```
1 while Goal not reached do
2    $\mathbf{x}_0 \leftarrow \text{StateEstimate}()$ ;
3   for  $k \leftarrow 0$  to  $K - 1$  do
4      $\mathbf{x} \leftarrow \mathbf{x}_0$ ; Sample  $\mathbf{V}^k = \{\mathbf{v}_0^k, \mathbf{v}_1^k, \dots, \mathbf{v}_{N-1}^k\}$ ;
5     for  $t \rightarrow 0$  to  $N - 1$  do
6        $\mathbf{x}_t \leftarrow \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$ ;
7        $S(\mathbf{V})_+ = q(\mathbf{x}_t) + \lambda \mathbf{u}_{t-1}^T \Sigma^{-1} (\mathbf{v}_{t-1} - \mathbf{u}_{t-1})$ ;
8      $S(\mathbf{V})_+ = \phi(\mathbf{x}_N)$ ;
9    $\beta \leftarrow \min_k [S(\mathbf{V})]$ ;
10   $\eta \leftarrow \sum_{k=1}^{K-1} \exp(-\frac{1}{\lambda}(S(\mathbf{V}) - \beta))$ ;
11  for  $k \leftarrow 0$  to  $K - 1$  do
12     $w(\mathbf{V}) = \frac{1}{\eta} \exp(-\frac{1}{\lambda}S(\mathbf{V}) + \sum_{k=0}^{N-1} -\mathbf{v}_k^T \Sigma^{-1} \mathbf{v}_k + \frac{1}{2} \mathbf{u}_k^T \Sigma^{-1} \mathbf{u}_k)$ ;
13  for  $t \leftarrow 0$  to  $N - 1$  do
14     $\mathbf{u}_t = \sum_{k=0}^{K-1} w(\mathbf{V}) (\mathbf{v}_t - \mathbf{u}_t)$ ;
15  ExecuteCommand( $\mathbf{u}_0$ );
16  for  $t \leftarrow 0$  to  $N - 1$  do
17     $\mathbf{u}_{t-1} \leftarrow \mathbf{u}_t$ ;
18   $\mathbf{u}_{N-1} \leftarrow \text{Sample}(\mathbf{u}_{N-1})$ ;
```

NUMERICAL SIMULATION

This paper demonstrates a spacecraft rendezvous on a geosynchronous transfer orbit. Information theoretic MPPI using importance sampling is implemented for the maneuver controller. The controller considers collision avoidance and control constraints in the optimal control problem, ensuring a safe and reliable operation. The control commands are executed in a predictive control setting, with an extended Kalman filter estimating the current states of relative position and velocity vectors at each time step. MPPI is implemented as a tracking controller for a given reference trajectory. The open-loop reference trajectory is generated through constrained differential dynamic programming (CDDP) algorithm.¹⁵ The simulation duration is set to $t_f = 37000$ s with time step of $dt = 1$ s. The prediction length of MPPI control is 50 with its sampling time of $T_s = 10$ s. The control update occurs at every 10 s. The EKF estimates the deputy spacecraft state given the measurements at every 1 s. The initial mean orbital elements of the chief spacecraft are:

$$\begin{aligned} a = 24439 \text{ km}, \quad e = 0.719, \quad i = 0.05 \text{ deg}, \quad \Omega = 0 \text{ deg} \\ \omega = 20 \text{ deg}, \quad f = 0 \text{ deg} \end{aligned} \tag{19}$$

The initial and final states are chosen to satisfy relatively stable deputy states given by:¹⁶

$$\begin{aligned} x_{j,0} = 100 \text{ m}, \quad y_{j,0} = 0 \text{ m}, \quad z_{j,0} = 200 \text{ m}, \quad \dot{x}_{j,0} = 0 \text{ m/s} \\ \dot{y}_{j,0} = -0.23 \text{ m/s}, \quad \dot{z}_{j,0} = 0 \text{ m/s} \end{aligned} \quad (20)$$

$$\begin{aligned} x_{j,f} = 20 \text{ m}, \quad y_{j,f} = 0 \text{ m}, \quad z_{j,f} = 40 \text{ m}, \quad \dot{x}_{j,f} = 0 \text{ m/s} \\ \dot{y}_{j,f} = -0.0459 \text{ m/s}, \quad \dot{z}_{j,f} = 0 \text{ m/s} \end{aligned} \quad (21)$$

The state-dependent cost for tracking control with a collision avoidance sphere constraint is given by:

$$q(\mathbf{x}_t) = (\mathbf{x}_t - \mathbf{x}_{\text{ref},t})^T \mathbf{Q} (\mathbf{x}_t - \mathbf{x}_{\text{ref},t}) + C_{CA} \mathbb{1}_{CA}(\mathbf{x}_t), \quad (22)$$

where $\mathbf{Q} = 10\mathbf{I}_6$ and $c_{CA} = 10^8$. $\mathbb{1}_{CA}(\mathbf{x})$ is a collision avoidance indicator function to imitate state constraints for collision avoidance in this framework given by:

$$\mathbb{1}_{CA}(\mathbf{x}) = \begin{cases} 1 & \text{if } \sqrt{x^2 + y^2 + z^2} \leq 20 \text{ m}, \\ 0 & \text{else.} \end{cases}$$

The terminal cost is given by:

$$\phi(\mathbf{x}, \mathbf{x}_{\text{ref}}) = (\mathbf{x} - \mathbf{x}_{\text{ref}})^T \mathbf{Q}_f (\mathbf{x} - \mathbf{x}_{\text{ref}}). \quad (23)$$

where $\mathbf{Q}_f = 1000\mathbf{I}_6$. In the process of MPPI optimization, control-bound constraints are taken into account as,

$$\mathbf{u}_{\min} \leq \mathbf{u}_t \leq \mathbf{u}_{\max}, \quad (24)$$

with

$$\mathbf{u}_{\max} = -\mathbf{u}_{\min} = 6 \times 10^{-6}$$

Finally, this paper uses the following measurement model for estimating the position and velocity of deputy spacecraft:

$$\mathbf{y} = \mathbf{I}_6 \mathbf{x} + \mathbf{v}_{\text{meas}} \quad (25)$$

where $\mathbf{v}_{\text{meas}} \sim \mathcal{N}(0, \mathbf{R}_{\text{meas}})$ and $\mathbf{R}_{\text{meas}} = \text{diag}(10^{-2}, 10^{-2}, 10^{-2}, 10^{-4}, 10^{-4}, 10^{-4})$. Additive process noise matrix of the dynamics is set to $\mathbf{Q}_{\text{proc}} = 10^{-11}\mathbf{I}_6$. The initial covariance matrix of state is $\mathbf{P} = \text{diag}(5 \times 10^{-7}, 5 \times 10^{-7}, 5 \times 10^{-7}, 10^{-4}, 10^{-4}, 10^{-4})$.

Figure 2 shows the three-dimensional relative trajectory of the deputy spacecraft with respect to the chief spacecraft. The number of samples is $K = 3000$. The covariance of the control variables

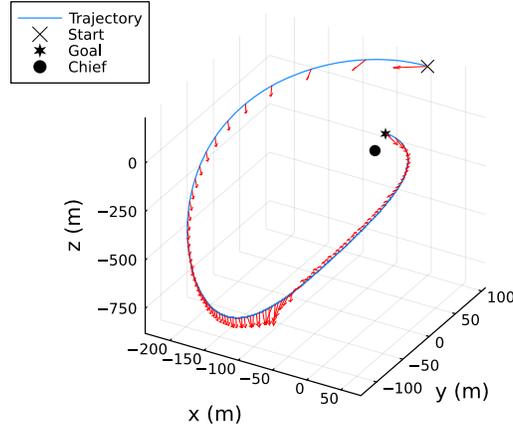


Figure 2. Three-dimensional MPPI closed-loop simulation trajectory

is $\Sigma = 10^{-11} \mathbf{I}_3$. The inverse temperature is $\lambda = 1$. The MPPI controller uses an estimated state by EKF and output control command to the spacecraft. The filter converges rapidly after the start of simulation. The maneuver involves both in-plane and out-of-plane relative motion control. Larger portions of the control effort are taken after passing the relative apoapsis, which coincides with the inertial apoapsis.

Figure 18 shows two two-dimensional relative trajectories of the deputy spacecraft with respect to the chief spacecraft. Both in-plane (left figure in Figure 18) and out-of-plane (right figure in Figure 18) trajectories have satisfied the sphere collision avoidance constraint.

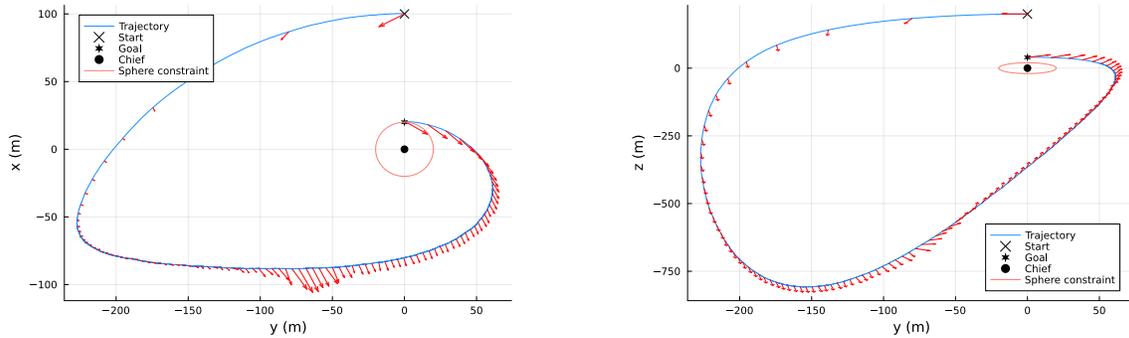


Figure 3. Two-dimensional MPPI closed-loop simulation trajectories with a sphere constraint. The left figure is the plane of the radial direction (\hat{x}) and the along-track direction. The right figure is the plane of the along-track direction and the cross-track direction.

Figure 4 shows the control input history of MPPI spacecraft rendezvous. Throughout the simulation, MPPI found controls within the given control bounds. The out-of-plane control effort fluctuation is larger than the in-plane ones. As control input may flutter due to the nature of sampling in sampling-based control, the resulting control sequence has a small variation on the control trajectory. This flutter phenomenon would be fixed by tuning the covariance of control variables and the number of samples. However, MPPI still finds a feasible control input during its operation. The

total velocity change, which is calculated by summing up the two norms of velocity changes, is $\Delta V = 0.139$ m/s.

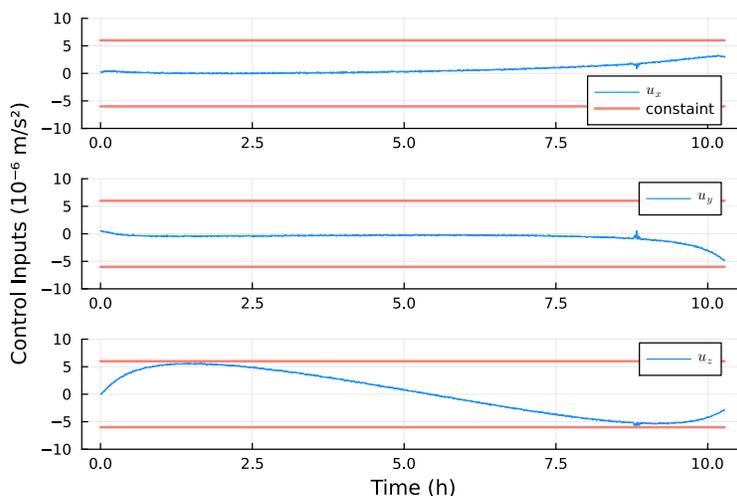


Figure 4. Control Input History

Table 1 shows the state-dependent trajectory cost and fuel usage ΔV comparison given the number of samples and the inverse temperature. The table implies that the increase in the number of samples does not directly improve the controller performance for both state-dependent trajectory cost and fuel usage, which contradicts to natural expectation. In addition, the state-dependent trajectory cost varies with the inverse temperature λ . If the inverse temperature is high and the number of samples is not enough, the controller performance will be dropped significantly.

Table 1. Trajectory and ΔV cost comparison given the number of samples and the inverse temperature

| Inverse Temperature λ | 0.1 | | 1 | | 10 | |
|-------------------------------|------------|------------------|----------------|------------------|------------|------------------|
| Number of Samples, K | State Cost | ΔV , m/s | State Cost | ΔV , m/s | State Cost | ΔV , m/s |
| 1000 | 0.0559 | 0.139 | 1.09 | 0.139 | 1.71 | 0.139 |
| 2000 | 0.00976 | 0.139 | 0.0232 | 0.139 | 0.00608 | 0.139 |
| 3000 | 0.0389 | 0.139 | 0.00208 | 0.139 | 0.0299 | 0.139 |

CONCLUSION AND FUTURE WORK

This paper demonstrated a sampling-based nonlinear model predictive control (MPC) for spacecraft rendezvous and proximity operations. The model predictive path integral (MPPI) controller successfully completed a relative motion maneuver with state and control constraints and an extended Kalman filter. The MPPI controller can also be used for highly nonlinear spacecraft relative motion control with a realistic number of samples, as general robotic applications consider. Future work includes line-of-sight constraints, range and angle measurements, and attitude control for the full spacecraft rendezvous and docking.

ACKNOWLEDGMENT

The first author has received support from the Nakajima Foundation.

REFERENCES

- [1] H. J. Kappen, “A linear theory for control of non-linear stochastic systems,” *arXiv*, 2004, 10.1103/phys-revlett.95.200201.
- [2] E. A. Theodorou, “Nonlinear Stochastic Control and Information Theoretic Dualities: Connections, Interdependencies and Thermodynamic Interpretations,” *Entropy*, Vol. 17, No. 5, 2015, pp. 3352–3375, 10.3390/e17053352.
- [3] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, “Information-Theoretic Model Predictive Control: Theory and Applications to Autonomous Driving,” *IEEE Transactions on Robotics*, Vol. 34, No. 6, 2017, pp. 1603–1622, 10.1109/tro.2018.2865891.
- [4] G. Williams, A. Aldrich, and E. A. Theodorou, “Model Predictive Path Integral Control: From Theory to Parallel Computation,” *Journal of Guidance, Control, and Dynamics*, Vol. 40, No. 2, 2017, pp. 1–14, 10.2514/1.g001921.
- [5] 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,” *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 1714–1721, 10.1109/icra.2017.7989202.
- [6] J. Pravitra, E. Theodorou, and E. N. Johnson, “Flying Complex Maneuvers with Model Predictive Path Integral Control,” *AIAA Scitech 2021 Forum*, 2021, 10.2514/6.2021-1957.
- [7] M. D. Houghton, A. B. Oshin, M. J. Acheson, E. A. Theodorou, and I. M. Gregory, “Path Planning: Differential Dynamic Programming and Model Predictive Path Integral Control on VTOL Aircraft,” *AIAA SCITECH 2022 Forum*, 2022, 10.2514/6.2022-0624.
- [8] L. Hou, H. Wang, H. Zou, and Y. Zhou, “Robotic Manipulation Planning for Automatic Peeling of Glass Substrate Based on Online Learning Model Predictive Path Integral,” *Sensors (Basel, Switzerland)*, Vol. 22, No. 3, 2022, p. 1292, 10.3390/s22031292.
- [9] M. Raisi, A. Noohian, and S. Fallah, “A fault-tolerant and robust controller using model predictive path integral control for free-flying space robots,” *Frontiers in Robotics and AI*, Vol. 9, 2022, p. 1027918, 10.3389/frobt.2022.1027918.
- [10] M. Kac, “On Distributions of Certain Wiener Functionals,” *Transactions of the American Mathematical Society*, Vol. 65, No. 1, 1949, p. 1, 10.2307/1990512.
- [11] H. Park, S. D. Cairano, and I. Kolmanovsky, “Model Predictive Control of Spacecraft Docking with a Non-rotating Platform,” *IFAC Proceedings Volumes*, Vol. 44, No. 1, 2011, pp. 8485–8490, 10.3182/20110828-6-it-1002.00031.
- [12] A. Weiss, M. Baldwin, R. S. Erwin, and I. Kolmanovsky, “Model Predictive Control for Spacecraft Rendezvous and Docking: Strategies for Handling Constraints and Case Studies,” *IEEE Transactions on Control Systems Technology*, Vol. 23, No. 4, 2015, pp. 1638–1647, 10.1109/tcst.2014.2379639.
- [13] D. A. Marsillach, S. D. Cairano, U. Kalabić, and A. Weiss, “Fail-Safe Spacecraft Rendezvous on Near-Rectilinear Halo Orbits,” *2021 American Control Conference (ACC)*, Vol. 00, 2021, pp. 2980–2985, 10.23919/acc50511.2021.9483328.
- [14] G. Xu and D. Wang, “Nonlinear Dynamic Equations of Satellite Relative Motion Around an Oblate Earth,” *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 5, 2012, pp. 1521–1524, 10.2514/1.33616.
- [15] T. Sasaki, K. Ho, and E. G. Lightsey, “Nonlinear Spacecraft Formation Flying using Constrained Differential Dynamic Programming,” *2022 AAS/AIAA Astrodynamics Specialist Conference*, 2022.
- [16] D. Wang, B. Wu, and E. K. Poh, *Satellite Formation Flying: Relative Dynamics, Formation Design, Fuel Optimal Maneuvers and Formation Maintenance*. Springer Singapore, 2017, 10.1007/978-981-10-2383-5.