

The continuous-time equations of motion of the manipulator's end-effector frame are known by

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{v}, \\ \dot{\mathbf{R}} &= \mathbf{R}\boldsymbol{\omega}^\times,\end{aligned}$$

where \mathbf{x} is the position vector, \mathbf{R} is the rotation matrix representation of the orientation, \mathbf{v} is the linear spatial velocity, and $\boldsymbol{\omega}$ is the angular body velocity.

Let ΔT be the sampling time period. Then, in discrete time, the equations of motion of the manipulator's end-effector frame can be expressed as

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{x}_k + \Delta T \mathbf{v}_k, \\ \mathbf{R}_{k+1} &= \mathbf{R}_k \exp(\Delta T \boldsymbol{\omega}_k^\times),\end{aligned}$$

at any time step $k = 0, 1, 2, \dots$, where \mathbf{x}_* and \mathbf{R}_* denote the position vectors and rotation matrices at the time steps $* = k$ and $* = k + 1$, and \mathbf{v}_k and $\boldsymbol{\omega}_k$ denote the linear spatial and the angular body velocities at the time step k , respectively.

With sufficiently small ΔT , the first-order approximation

$$\exp(\Delta T \boldsymbol{\omega}_k^\times) \approx \mathbf{I} + \Delta T \boldsymbol{\omega}_k^\times$$

then linearizes the discrete-time equations of motion of the manipulator in the Cartesian space by

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{x}_k + \Delta T \mathbf{v}_k, \\ \mathbf{R}_{k+1} &\approx \mathbf{R}_k + \Delta T \mathbf{R}_k \boldsymbol{\omega}_k^\times.\end{aligned}$$

Similarly, one can predict the position vectors \mathbf{x}_* and rotation matrices \mathbf{R}_* at the following time steps $* = k + 2, \dots, k + N$ by

$$\begin{aligned}\mathbf{x}_{k+2} &= \mathbf{x}_k + \Delta T \mathbf{v}_k + \Delta T \mathbf{v}_{k+1}, \\ \mathbf{R}_{k+2} &\approx \mathbf{R}_k + \Delta T \mathbf{R}_k \boldsymbol{\omega}_k^\times + \Delta T \mathbf{R}_{k+1} \boldsymbol{\omega}_{k+1}^\times, \\ &\vdots \\ \mathbf{x}_{k+N} &= \mathbf{x}_k + \Delta T \mathbf{v}_k + \dots + \Delta T \mathbf{v}_{k+N-1}, \\ \mathbf{R}_{k+N} &\approx \mathbf{R}_k + \Delta T \mathbf{R}_k \boldsymbol{\omega}_k^\times + \dots + \Delta T \mathbf{R}_{k+N-1} \boldsymbol{\omega}_{k+N-1}^\times,\end{aligned}$$

where \mathbf{v}_{*-1} and $\boldsymbol{\omega}_{*-1}$ are the linear and angular velocities at the time steps $k + 1, \dots, k + N - 1$, respectively.

For the ease of notation, let \mathbf{s}_n be an $N \times 1$ vector of which the first $n = 1, \dots, N$ elements are 1 and the remaining ones are 0. Then, the position \mathbf{x}_{k+n} and orientation \mathbf{R}_{k+n} of the manipulator's end-effector frame at any time step $k + n$ in the prediction horizon can be expressed as follows

$$\begin{aligned}\mathbf{x}_{k+n} &= \mathbf{x}_k + \Delta T (\mathbf{s}_n \otimes \mathbf{I})^\top \mathbf{v}_{k,N}, \\ \mathbf{R}_{k+n} &= \mathbf{R}_k + \Delta T \mathbf{R}_k [(\mathbf{s}_n \otimes \mathbf{I})^\top \boldsymbol{\omega}_{k,N}]^\times,\end{aligned}$$

where the linear spatial $\mathbf{v}_{k,N}$ and angular body $\boldsymbol{\omega}_{k,N}$ velocity sequences in the prediction horizon $[k, k + N]$ are defined by

$$\mathbf{v}_{k,N} = \begin{bmatrix} \mathbf{v}_k \\ \mathbf{v}_{k+1} \\ \vdots \\ \mathbf{v}_{k+N-2} \\ \mathbf{v}_{k+N-1} \end{bmatrix}, \quad \boldsymbol{\omega}_{k,N} = \begin{bmatrix} \boldsymbol{\omega}_k \\ \boldsymbol{\omega}_{k+1} \\ \vdots \\ \boldsymbol{\omega}_{k+N-2} \\ \boldsymbol{\omega}_{k+N-1} \end{bmatrix}.$$

Let \mathbf{x}_d and \mathbf{R}_d be the desired position and orientation of the manipulator's end-effector frame. The control objective in the prediction horizon $[k, k + N]$ is then formulated as

$$\underset{\mathbf{v}_{k,N}, \boldsymbol{\omega}_{k,N}}{\text{minimize}} \quad V = \sum_{n=1}^N V_{k+n},$$

where the cost functions V_{k+n} at the time steps $k+n$ are defined by

$$\begin{aligned}
V_{k+n} &= \kappa_v (\mathbf{x}_{k+n} - \mathbf{x}_d)^\top (\mathbf{x}_{k+n} - \mathbf{x}_d) + (\Delta T)^2 \alpha_v \mathbf{v}_{k+n-1}^\top \mathbf{v}_{k+n-1} \\
&\quad + \kappa_\omega \text{tr}(\mathbf{I} - \mathbf{R}_{k+n}^\top \mathbf{R}_d) + (\Delta T)^2 \alpha_\omega \boldsymbol{\omega}_{k+n-1}^\top \boldsymbol{\omega}_{k+n-1} \\
&\approx \kappa_v [\mathbf{x}_k + \Delta T(\mathbf{s}_n \otimes \mathbf{I})^\top \mathbf{v}_{k,N} - \mathbf{x}_d]^\top [\mathbf{x}_k + \Delta T(\mathbf{s}_n \otimes \mathbf{I})^\top \mathbf{v}_{k,N} - \mathbf{x}_d] + (\Delta T)^2 \mathbf{v}_{k,N}^\top (\alpha_v \mathbb{I}_{N,n} \otimes \mathbf{I}) \mathbf{v}_{k,N} \\
&\quad + \kappa_\omega \text{tr}(\mathbf{I} - \mathbf{R}_k^\top \mathbf{R}_d) + \Delta T \kappa_\omega \text{tr}([(s_n \otimes \mathbf{I})^\top \boldsymbol{\omega}_{k,N}]^\times \mathbf{R}_k^\top \mathbf{R}_d) + (\Delta T)^2 \boldsymbol{\omega}_{k,N}^\top (\alpha_\omega \mathbb{I}_{N,n} \otimes \mathbf{I}) \boldsymbol{\omega}_{k,N} \\
&= \kappa_v (\mathbf{x}_k - \mathbf{x}_d)^\top (\mathbf{x}_k - \mathbf{x}_d) + 2\Delta T \kappa_v (\mathbf{x}_k - \mathbf{x}_d)^\top (\mathbf{s}_n \otimes \mathbf{I})^\top \mathbf{v}_{k,N} + (\Delta T)^2 \mathbf{v}_{k,N}^\top [(\kappa_v \mathbf{s}_n \mathbf{s}_n^\top + \alpha_v \mathbb{I}_{N,n}) \otimes \mathbf{I}] \mathbf{v}_{k,N} \\
&\quad + \kappa_\omega \text{tr}(\mathbf{I} - \mathbf{R}_k^\top \mathbf{R}_d) + 2\Delta T \kappa_\omega [\text{sk}(\mathbf{R}_d^\top \mathbf{R}_k)^\vee]^\top (\mathbf{s}_n \otimes \mathbf{I})^\top \mathbf{v}_{k,N} + (\Delta T)^2 \boldsymbol{\omega}_{k,N}^\top (\alpha_\omega \mathbb{I}_{N,n} \otimes \mathbf{I}) \boldsymbol{\omega}_{k,N},
\end{aligned}$$

where $\alpha > 0$ is the damping gain, and $\mathbb{I}_{N,n}$ is an $N \times N$ matrix of which all elements are zero except the n -th diagonal element being 1. Because $\mathbf{x}_k - \mathbf{x}_d$ and $\mathbf{R}_k \mathbf{R}_d^\top$ are constant in the prediction horizon $[k, k+N]$, one can thus simplify the objective of the model predictive control problem to

$$\underset{\mathbf{v}_{k,N}, \boldsymbol{\omega}_{k,N}}{\text{minimize}} \quad \Delta T \mathbf{v}_{k,N}^\top \mathbf{A}_v \mathbf{v}_{k,N} + 2\mathbf{b}_v^\top \mathbf{v}_{k,N} + \Delta T \boldsymbol{\omega}_{k,N}^\top \mathbf{A}_\omega \boldsymbol{\omega}_{k,N} + 2\mathbf{b}_\omega^\top \boldsymbol{\omega}_{k,N},$$

where the matrices \mathbf{A}_* and vectors \mathbf{b}_* with $*$ = v, ω are defined by

$$\mathbf{A}_v = \kappa_v \sum_{n=1}^N (\mathbf{s}_n \mathbf{s}_n^\top) \otimes \mathbf{I} + \alpha_v \mathbf{I}, \quad \mathbf{b}_v = \kappa_v \sum_{n=1}^N (\mathbf{s}_n \otimes \mathbf{I})(\mathbf{x}_k - \mathbf{x}_d), \quad \mathbf{A}_\omega = \alpha_\omega \mathbf{I}, \quad \mathbf{b}_\omega = \kappa_\omega \sum_{n=1}^N (\mathbf{s}_n \otimes \mathbf{I}) \text{sk}(\mathbf{R}_d^\top \mathbf{R}_k)^\vee.$$

Let \mathbf{x}_{eo} and \mathbf{R}_{eo} denote the position and orientation of the object frame relative to the end-effector frame of the manipulator. The position and orientation of the object frame in the world frame is then expressed as

$$\begin{aligned}
\mathbf{x}_o &= \mathbf{x} + \mathbf{R} \mathbf{x}_{eo}, \\
\mathbf{R}_o &= \mathbf{R} \mathbf{R}_{eo}.
\end{aligned}$$

Let the forces exerted by the tray on the object via contact points $i = 1, \dots, 4$ be expressed in the object's body frame by $\mathbf{f}_i = [f_{xi} \ f_{yi} \ f_{zi}]^\top$. The rigid-body dynamics of the object are described by

$$\begin{aligned}
m\dot{\mathbf{v}}_o + m\mathbf{g} &= \sum_{i=1}^4 \mathbf{R}_o \mathbf{f}_i, \\
\mathcal{I}\dot{\boldsymbol{\omega}}_o + \boldsymbol{\omega}_o^\times \mathcal{I} \boldsymbol{\omega}_o &= \sum_{i=1}^4 \mathbf{r}_i^\times \mathbf{f}_i,
\end{aligned}$$

where \mathbf{v}_o and $\boldsymbol{\omega}_o$ are the linear spatial and angular body velocities of the object, i.e.,

$$\begin{aligned}
\mathbf{v}_o &= \dot{\mathbf{x}}_o = \dot{\mathbf{x}} + \dot{\mathbf{R}} \mathbf{x}_{eo} = \mathbf{v} + \mathbf{R} \boldsymbol{\omega}^\times \mathbf{x}_{eo} = \mathbf{v} - \mathbf{R} \mathbf{x}_{eo}^\times \boldsymbol{\omega}, \\
\boldsymbol{\omega}_o^\times &= \mathbf{R}_o^\top \dot{\mathbf{R}}_o = \mathbf{R}_o^\top \dot{\mathbf{R}} \mathbf{R}_{eo} = \mathbf{R}_o^\top \mathbf{R} \boldsymbol{\omega}^\times \mathbf{R}_{eo} = (\mathbf{R}_{eo}^\top \boldsymbol{\omega})^\times.
\end{aligned}$$

One can thus derive that, at the acceleration level,

$$\begin{aligned}
\dot{\mathbf{v}}_o &= \dot{\mathbf{v}} + \mathbf{R}(\boldsymbol{\omega}^\times)^2 \mathbf{x}_{eo} - \mathbf{R} \mathbf{x}_{eo}^\times \dot{\boldsymbol{\omega}}, \\
\dot{\boldsymbol{\omega}}_o &= \mathbf{R}_{eo}^\top \dot{\boldsymbol{\omega}}.
\end{aligned}$$

Let $\mathbf{f} = [\mathbf{f}_1^\top \ \mathbf{f}_2^\top \ \mathbf{f}_3^\top \ \mathbf{f}_4^\top]^\top$ convert the rigid-body dynamics of the object into

$$\begin{aligned}
m\dot{\mathbf{v}} - m\mathbf{R} \mathbf{x}_{eo}^\times \dot{\boldsymbol{\omega}} + m\mathbf{R}(\boldsymbol{\omega}^\times)^2 \mathbf{x}_{eo} + m\mathbf{g} &= \mathbf{R} \mathbf{R}_{eo} \mathbf{G}_l \mathbf{f}, \\
\mathcal{I} \mathbf{R}_{eo}^\top \dot{\boldsymbol{\omega}} + (\mathbf{R}_{eo}^\top \boldsymbol{\omega})^\times \mathcal{I} (\mathbf{R}_{eo}^\top \boldsymbol{\omega}) &= \mathbf{G}_a \mathbf{f},
\end{aligned}$$

where $\mathbf{G}_l = [\mathbf{I} \ \mathbf{I} \ \mathbf{I} \ \mathbf{I}]$ and $\mathbf{G}_a = [\mathbf{r}_1^\times \ \mathbf{r}_2^\times \ \mathbf{r}_3^\times \ \mathbf{r}_4^\times]$.

Thus, at the k -th time step, the robot manipulator can approximate the motions of the object at the following steps $k+n$ by

$$m(\mathbf{v}_{k+n} - \mathbf{v}_{k+n-1}) - m\mathbf{R}_k \mathbf{x}_{eo}^\times (\boldsymbol{\omega}_{k+n} - \boldsymbol{\omega}_{k+n-1}) + \Delta T m \mathbf{R}_k (\boldsymbol{\omega}_k^\times)^2 \mathbf{x}_{eo} + \Delta T m \mathbf{g} = \Delta T \mathbf{R}_k \mathbf{R}_{eo} \mathbf{G}_l \mathbf{f}_{k+n-1},$$

$$\mathcal{I} \mathbf{R}_{eo}^\top (\boldsymbol{\omega}_{k+n} - \boldsymbol{\omega}_{k+n-1}) + \Delta T (\mathbf{R}_{eo}^\top \boldsymbol{\omega}_k)^\times \mathcal{I} (\mathbf{R}_{eo}^\top \boldsymbol{\omega}_k) = \Delta T \mathbf{G}_a \mathbf{f}_{k+n-1},$$

where $n = 1, 2, \dots, N$. In other words, the predicted contact forces $\mathbf{f}_{k,N}$ constrain the linear spatial $\mathbf{v}_{k,N}$ and angular body $\boldsymbol{\omega}_{k,N}$ velocities of the manipulator as per

$$m(\mathbf{\Gamma} \otimes \mathbf{I}) \mathbf{v}_{k,N} - m[\mathbf{\Gamma} \otimes (\mathbf{R}_k \mathbf{x}_{eo}^\times)] \boldsymbol{\omega}_{k,N} - \Delta T [\mathbf{I} \otimes (\mathbf{R}_k \mathbf{R}_{eo} \mathbf{G}_l)] \mathbf{f}_{k,N} = m \boldsymbol{\eta}_k - \Delta T m (\mathbf{1} \otimes [\mathbf{R}_k (\boldsymbol{\omega}_k^\times)^2 \mathbf{x}_{eo} + \mathbf{g}]),$$

$$[\mathbf{\Gamma} \otimes (\mathcal{I} \mathbf{R}_{eo}^\top)] \boldsymbol{\omega}_{k,N} - \Delta T (\mathbf{I} \otimes \mathbf{G}_a) \mathbf{f}_{k,N} = \boldsymbol{\lambda}_k - \Delta T (\mathbf{1} \otimes [(\mathbf{R}_{eo}^\top \boldsymbol{\omega}_k)^\times \mathcal{I} (\mathbf{R}_{eo}^\top \boldsymbol{\omega}_k)]),$$

where $\mathbf{1}$ is a vector of all ones, and

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}, \quad \mathbf{f}_{k,N} = \begin{bmatrix} \mathbf{f}_k \\ \mathbf{f}_{k+1} \\ \vdots \\ \mathbf{f}_{k+N-2} \\ \mathbf{f}_{k+N-1} \end{bmatrix}, \quad \boldsymbol{\eta}_k = \begin{bmatrix} \mathbf{v}_k - \mathbf{R}_k \mathbf{x}_{eo}^\times \boldsymbol{\omega}_k \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad \boldsymbol{\lambda}_k = \begin{bmatrix} \mathcal{I} \mathbf{R}_{eo}^\top \boldsymbol{\omega}_k \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

To further ease the presentation, let

$$\mathbf{H}_v = \begin{bmatrix} m(\mathbf{\Gamma} \otimes \mathbf{I}) \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{H}_\omega = \begin{bmatrix} m[\mathbf{\Gamma} \otimes (\mathbf{R}_k \mathbf{x}_{eo}^\times)] \\ -\mathbf{\Gamma} \otimes (\mathcal{I} \mathbf{R}_{eo}^\top) \end{bmatrix}, \quad \mathbf{H}_f = \begin{bmatrix} \mathbf{I} \otimes (\mathbf{R}_k \mathbf{R}_{eo} \mathbf{G}_l) \\ \mathbf{I} \otimes \mathbf{G}_a \end{bmatrix},$$

$$\mathbf{h} = \begin{bmatrix} m \boldsymbol{\eta}_k - \Delta T m (\mathbf{1} \otimes [\mathbf{R}_k (\boldsymbol{\omega}_k^\times)^2 \mathbf{x}_{eo} + \mathbf{g}]) \\ \boldsymbol{\lambda}_k - \Delta T (\mathbf{1} \otimes [(\mathbf{R}_{eo}^\top \boldsymbol{\omega}_k)^\times \mathcal{I} (\mathbf{R}_{eo}^\top \boldsymbol{\omega}_k)]) \end{bmatrix}$$

The constraints then become

$$\mathbf{H}_v \mathbf{v}_{k,N} - \mathbf{H}_\omega \boldsymbol{\omega}_{k,N} - \Delta T \mathbf{H}_f \mathbf{f}_{k,N} = \mathbf{h}.$$

To prevent the object from sliding, the contact forces \mathbf{f}_i need to respect the friction-cone constraints

$$f_{xi}^2 + f_{yi}^2 - \mu^2 f_{zi}^2 \leq 0$$

for the contact points $i = 1, \dots, 4$, where $\mu > 0$ is the friction coefficient. Also, the object keeps in contact with the tray by making their contact forces unilateral in the normal direction, i.e.,

$$f_{zi} \geq \epsilon$$

for some $\epsilon > 0$, where $i = 1, \dots, 4$. At the k -th time step, the friction cone constraints turn into

$$\mathbf{f}_{k,N}^\top (\mathbb{I}_{N,n} \otimes \mathbb{I}_{4,i} \otimes \boldsymbol{\Lambda}) \mathbf{f}_{k,N} \leq 0,$$

$$(\mathbf{1} \otimes \mathbf{I} \otimes [0 \ 0 \ 1]) \mathbf{f}_{k,N} \geq \epsilon \mathbf{1},$$

for the contact points $i = 1, \dots, 4$ in the prediction horizon $n = 1, \dots, N$, where $\mathbb{I}_{4,i}$ is a 4×4 matrix of which all elements are zero except the i -th diagonal element being 1, and

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\mu^2 \end{bmatrix}.$$

The problem then becomes a second-order cone program as follows

$$\begin{aligned} & \underset{\mathbf{v}_{k,N}, \boldsymbol{\omega}_{k,N}}{\text{minimize}} \quad \Delta T \begin{bmatrix} \mathbf{v}_{k,N} \\ \boldsymbol{\omega}_{k,N} \end{bmatrix}^\top \begin{bmatrix} \mathbf{A}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_\omega \end{bmatrix} \begin{bmatrix} \mathbf{v}_{k,N} \\ \boldsymbol{\omega}_{k,N} \end{bmatrix} + 2 \begin{bmatrix} \mathbf{b}_v \\ \mathbf{b}_\omega \end{bmatrix}^\top \begin{bmatrix} \mathbf{v}_{k,N} \\ \boldsymbol{\omega}_{k,N} \end{bmatrix}, \\ & \text{subject to} \quad \mathbf{1} \otimes \begin{bmatrix} \tilde{\mathbf{v}} \\ \tilde{\boldsymbol{\omega}} \end{bmatrix} \leq \begin{bmatrix} \mathbf{v}_{k,N} \\ \boldsymbol{\omega}_{k,N} \end{bmatrix} \leq \mathbf{1} \otimes \begin{bmatrix} \hat{\mathbf{v}} \\ \hat{\boldsymbol{\omega}} \end{bmatrix}, \\ & \quad \mathbf{H}_v \mathbf{v}_{k,N} - \mathbf{H}_\omega \boldsymbol{\omega}_{k,N} - \Delta T \mathbf{H}_f \mathbf{f}_{k,N} = \mathbf{h}, \\ & \quad \mathbf{f}_{k,N}^\top (\mathbb{I}_{N,n} \otimes \mathbb{I}_{4,i} \otimes \boldsymbol{\Lambda}) \mathbf{f}_{k,N} \leq 0, \\ & \quad (\mathbf{1} \otimes \mathbf{I} \otimes [0 \ 0 \ 1]) \mathbf{f}_{k,N} \geq \epsilon \mathbf{1}. \end{aligned}$$