arXiv:2303.00873v1 [math.OC] 2 Mar 2023

用于控制的状态估计：一种输出反馈随机模型预测控制方法

Mohammad S. Ramadan [a, b], Robert R. Bitmead [a]和 Ke Huang [b]

[a]机械与航空航天工程系，加利福尼亚大学圣迭戈分校，美国加利福尼亚州拉霍亚，邮编 *92093-0411*。

[b]电气与计算机工程系，圣迭戈州立大学，美国加利福尼亚州圣迭戈，邮编 *92182* .

摘要

本文提出了一种新方法，用于为随机输出反馈问题确定单个状态值。该方法借鉴了模型预测控制的范式，尤其关注开环控制与闭环控制的区别，以及确定性最优控制与随机最优控制的区别。文中介绍了状态选择算法，该算法基于给定的动力学和约束、名义确定性状态反馈控制器，以及一种基于采样的方法；它通过在粒子滤波器提供的可用候选状态值中进行选择，优化预先设定的有限时域性能函数，从而选出最佳状态值。算法使用名义控制器和所选状态，在整个时域内最小化代价函数。因此，最小化并非针对控制的选择本身，而只是通过选择要使用的状态值间接实现。该算法普遍适用于非线性随机系统，并依赖蒙特卡洛采样与求平均。不过，对于线性二次多面体约束情形，该技术可简化为关于状态值的二次规划。本文通过一系列计算示例评估该算法，展示其有效性与局限性，并讨论数值方面及并行化的可能性。示例展示了该算法如何在长时域内与其配套的粒子滤波器一起以闭环方式运行。

关键词：状态估计、随机模型预测控制、机会约束、部分可观测马尔可夫决策过程。

1 引言

我们研究非线性随机系统输出反馈控制中的状态选择。给定状态的滤波后粒子密度和名义控制律，本文提出一种算法，从该密度中选取单个状态值，并将其用于控制。该算法反复选取单个粒子，随后进入一个闭环阶段，生成一组候选控制序列，再根据这些序列在所有粒子上的平均开环性能进行评估。开环与闭环计算并存这一现象，凸显了确定性最优反馈控制与随机最优反馈控制之间的重要区别；单一估计值的传播与整体密度的传播也体现了这一点。我们将在下文简要讨论这两方面。这也凸显了输出反馈随机模型预测控制形式化中的一个难点。该问题是一个典型的控制问题，有助于理解相关方法及其部分发展脉络。

电子邮箱地址：msramada@eng.ucsd.edu (Mohammad S. Ramadan), rbitmead@eng.ucsd.edu (Robert R. Bitmead), khuang@sdsu.edu (Ke Huang)。

背景与模型预测控制

自20世纪80年代末以来，人们对模型预测控制（模型预测控制）的兴趣与研究激增，这在很大程度上归因于其处理约束 [1]的能力。模型预测控制是在滚动时域最优控制框架下实现这一点的：它从假定可用的当前状态出发，反复求解一个带约束的开环最优控制问题。也就是说，模型预测控制被表述为全状态反馈；反馈通过滚动时域机制实现，即在测量下一个状态并重新求解开环问题之前，仅应用开环控制序列中的第一个元素，而该序列的每个元素都是初始状态的函数。由于这些优化问题是开环的，每个带约束优化问题的复杂度都易于管理。

使用部分观测状态的模型预测控制进入了随机最优控制的领域，而众所周知，除最简单的情况外，随机最优控制在计算上都难以处理 [2,3],，无论是否存在约束。此外，对于这些问题，等价的开环最优解序列不复存在；最优控制必然只能是反馈控制

[4]。我们的研究旨在随机最优控制的计算难处理性（以及最优性）与开环方法的便利性之间采取折中路径。给定一个特定的状态反馈控制律 $u_k = \kappa(x_k)$（例如由显式模型预测控制提供），我们提出一种启发式（显然是次优的）方法，根据状态粒子滤波条件密度选择单个状态值。由此实现的控制律 κ 仅通过选择过程依赖于滤波密度。对于无约束线性二次问题，选择粒子密度的条件均值将是最优的 [5]。然而，对于受约束和/或非线性问题，状态选择取决于给定的控制律，并需要以开环方式、基于粒子集合，在整个预测时域内评估约束满足情况和性能指标。候选状态向量的评估建立在模型预测控制迭代的有限时域开环特性之上，并结合预测时域内的两个方面：约束违反的概率，以及预测性能函数的计算；二者均根据粒子密度进行平均，具体说明如下。

部分可观测状态下的随机最优控制（以下简称随机最优控制）通过随机动态规划求解 [4,6] ，这需要在每个时刻传播完整的条件密度。最优控制律必然只能采用反馈形式；它无法沿时域预先计算，因为它取决于预期的测量值及相关的条件密度传播。将其与确定性最优控制（以及隐含其中的模型预测控制方法）相比较：在确定性最优控制中，可以通过庞特里亚金最大值原理等方法计算开环最优控制，而且从时间函数的角度看，它与动态规划中的反馈最优控制一致。正是确定性最优控制的这一特性，使滚动时域模型预测控制能够从开环约束优化中生成反馈控制律。由于优化是开环的，因此在计算上是可行的。开环与闭环的这些概念构成了随机最优控制中计算问题的基础。

由于随机最优控制所带来的沉重负担，人们已经开发出若干针对这类问题的简化方法；这些方法虽不那么繁琐，但只是近似且非最优的。

- ✻ 广义方法通过初始状态不确定性的前两个中心矩的近似值来表示状态不确定性，并利用线性化或泰勒展开，近似描述这些矩在输入和预期测量值作用下的传播动力学 [3]。✻ 显式探测，即在随机问题的确定性版本的代价中显式加入基于状态方差的表达式，如 [7]。✻ 随机模型预测控制方法通常不考虑主动学习或依赖未来测量值的机制，从而可以获得开环解。旨在处理初始状态不确定性的随机模型预测控制方法示例包括：

用于处理初始状态不确定性的示例包括：

- ✯ 一种面向初始状态不确定性的、经修改的基于场景的方法 [8],✯ 用于对非线性但无约束问题的输入空间进行采样的序贯蒙特卡罗方法 [9],✯ 随机管方法 [10,11,12],其表述方式使其适用性仅限于线性系统。

关于随机模型预测控制方法的更完整综述可见于 [13]。这里，我们提出了一种类似于随机模型预测控制方法（或等价地，类似于开环反馈控制 [6]）的算法。不同于这些方法，我们寻求选择初始状态值，而非设计控制器。与随机管方法不同，我们的表述可自然扩展至受约束非线性动力学。这得益于我们所采用的采样方法，与 [9],相比，该方法可在各样本间并行化。

正如 Striebel 在 [14],中所讨论的，目的或功能应当指导数据约简工作。我们的方法侧重于从表示密度的粒子集合中选择单个状态向量值。这一约简受到既定控制目标的影响，并采用一个给定的——尚不一定具有稳定性、最优性或约束可行性的——名义控制律。因此，在控制目标的指引下，控制值由名义控制律和选择过程间接确定。这不同于其他直接选择控制值的方法。该方法基于以下信息：具有白过程噪声的系统状态方程、名义状态反馈控制律、约束、噪声密度、经滤波的粒子状态密度、预测时域上的代价函数、状态与输入约束，以及其违反率容差。与该控制计算并行进行的是经滤波状态粒子密度的传播，其性质也会影响闭环行为。

在已有文献中，针对具有全状态反馈的线性动力学，曾将初始状态纳入模型预测控制的决策变量。在 [15], 中，初始状态被增广为基于管的模型预测控制问题的决策变量，以提高问题可行的可能性。此外， [16],采用了先前提出的增广概念，但目的是在原点周围的一个不变集上实现恒定的价值函数。我们的方法不假设初始状态已知，并且可以自然地推广到非线性动力学。

对于具有线性标称反馈控制、二次代价和多面体约束的线性问题，状态选择方法得到的解等价于一个关于初始状态的二次规划。

提纲

精确的问题表述和我们提出的算法见第2节。第3节介绍了一些基础

性质，并将我们的方法与随机模型预测控制联系起来。第4节解释了计算工作量，包括状态粒子滤波器的传播，以及所提算法进行并行计算的可能性。第5节讨论了该算法的特殊情形：具有多面体约束的线性系统，并给出一个结果，表明在无约束情形下，该算法可得到条件均值。该算法在线性、二次、多面体情形下的这些性质，是支持该算法的最有力技术依据。第6节列出了非线性情形和线性特殊情形的数值示例。

2 问题表述

在当前时刻 $k = 0$，我们掌握以下信息。

I. 离散时间状态动力学

$$x_{k+1} = f(x_k, u_k, w_k), \tag{1}$$

其中，状态为 $x_k \in \mathbb{R}^{r_x}$，控制输入为 $u_k \in \mathbb{R}^{r_u}$，外生扰动为 $w_k \in \mathbb{R}^{r_w}$。

II. 随机扰动过程 $\{w_k\}$，假设其为独立同分布，且具有已知密度 $\mathcal{W}$。对于所有 k，状态 x_0 与 w_k 独立。

III. 状态约束集和输入约束集分别为 $\mathbb{X}$ 和 $\mathbb{U}$，以及 $\epsilon \in [0, 1)$，即可接受的概率约束违反率。

IV. 初始状态 x_0密度 p_0，以粒子集合 $\Xi = \{\xi_0^i \in \mathbb{R}^{r_x}, j = 1, \ldots, L\}$的形式给出。V. 名义全状态反馈控制律 $u_k = \kappa(x_k)$，该控制律目前既未假设相对于 $\mathbb{X}$ 和 $\mathbb{U}$ 具有递归可行性，也未假设相对于以下代价（事实上也未假设相对于任何代价）是最优的。[1]

VI. 一个 N阶段的有限时域轨迹代价函数 J

$$J = \sum_{k=0}^{N} \ell_k(x_k, u_k).$$

由于初始状态 x_0 并非完全已知（除非p_0 是点质量函数），因此我们试图选择一个候选状态值 $x_0^\star$，该值由初始密度支持，并在有限时域内实现：可行性（可能是在仿真中以概率或统计方式实现）；以及对随后以粒子集平均值评估轨迹代价 J 产生有利影响。

如果式 (1) 中的动力学具有 $w_k s$ 零和 x_0 已知，我们称之为确定性情形。

[1] 虽然名义控制律不必具备可行性或最优性，但其选择对于状态选择的性能和可行性至关重要。第6节通过使用多个控制器评估状态选择算法对此进行了展示。

接下来，我们提出一种算法，即状态选择算法，该算法返回一个候选状态 $x_0^\star$，并将其输入控制$u_0 = \kappa(x_0^\star)$，然后应用于系统。系统输出用于更新粒子滤波密度，并在新的当前时刻 $k = 0$ 重复这一过程。第6节中的示例将对此进行说明。

状态选择算法

(1) 选择样本重复次数 M 和统计可行性容差 $\alpha \in [0, \epsilon)$，作为算法参数。(2) 对每个 $i \in \{1, 2, \ldots, L\}$，选择 $x_0' = \xi^i \in \Xi$。(a) 使用 (Nr_w)-向量

$$W_j' = \begin{pmatrix} w_{0,j}'^T & \cdots & w_{N-1,j}'^T \end{pmatrix}^T,$$
$$W_j'' = \begin{pmatrix} w_{0,j}''^T & \cdots & w_{N-1,j}''^T \end{pmatrix}^T,$$

样本，从 $\mathcal{W}^N$ 和 Ξ中抽取 M 个独立实现，组成 $(2Nr_w + r_x)$-向量序列

$$\left\{ \begin{pmatrix} W_j' \\ W_j'' \\ x_{0,j}'' \end{pmatrix} : j = 1, \ldots, M \right\}.$$

(b) 对每个 j 和 k，根据 x_0'计算 κ闭环状态序列，即 $x_{0,\ j}' = x_0'$。

$$x_{k+1,j}' = f(x_{k,j}', \kappa(x_{k,j}'), w_{k,j}'). \tag{2}$$

这定义了当前初始状态候选值 x_0'下，长度为 N 的闭环控制序列 $\{\kappa(x_{k,\ j}')\}_{k=0}^{N-1}$ 的 M 个样本。

(c) 对每个 $j \in \{1, \ldots, M\}$ 和 $k \in \{0, N-1\}$，根据 $x_{0,\ j}''$计算开环控制状态序列，

$$x_{k+1,j}'' = f(x_{k,j}'', \kappa(x_{k,j}'), w_{k,j}''). \tag{3}$$

(d) 对每个 k 计算样本平均闭环控制违规率，[2]

$$\hat{\beta}_k(x_0') = \frac{1}{M} \sum_{j=1}^{M} \mathbb{1}(\kappa(x_{k,j}') \in \mathbb{U}). \tag{4}$$

(e) 对每个 k 计算样本平均开环控制状态违规率，

$$\hat{\lambda}_k(x_0') = \frac{1}{M} \sum_{j=1}^{M} \mathbb{1}(x_{k,j}'' \in \mathbb{X}). \tag{5}$$

[2] $\mathbb{1}(\cdot)$ 是某事件的指示函数。

ˆ ˆ (f) 如果对于所有 k， $\beta_k(x'_0) \geq 1-\alpha$ 且 $\lambda_k(x'_0) \geq 1-\alpha$，则将该候选状态 x'_0 判定为可行并继续。否则，返回步骤 2。(g) 如果 x'_0 可行，则计算其样本平均性能，

$$J_c^M(x'_0) = \frac{1}{M}\sum_{j=1}^{M}\sum_{k=0}^{N}\ell_k(x''_{k,j}, \kappa(x'_{k,j})).$$

(3) 选择 $x_0^\star$，使其成为最小化 $J_c^M(\cdot)$ 的可行 x'_0，前提是可行集非空。

在该算法中，序列 $\{x'_{k,j}\}_k$ 和 $\{x''_{k,j}\}_k$ 是由递推式（2）和（3）生成的随机过程的实现，这些递推式由构造的相互独立的白噪声序列 W'_j 和 W''_j 驱动。

这些过程是 κ 和 x'_0 的函数。为使记号简洁，我们省略这种依赖关系。当实现次数 M 很大时，在正则条件 [17,18], 下，平均代价 $J_c^M(\cdot)$ 收敛

$$\begin{aligned} J_c(x'_0) &= \mathbb{E}\left(\sum_{k=0}^{N}\ell_k\left(x''_k, \kappa\left(x'_k\right)\right)\right), \\ &= \mathbb{E}_{x''_0}\mathbb{E}_{W'}\mathbb{E}_{W''}\left(\sum_{k=0}^{N}\ell_k\left(x''_k, \kappa\left(x'_k\right)\right)\right). \end{aligned} \tag{6}$$

由于元素 x''_0, W', W'' [17] 之间具有相互独立性，期望可以进行因式分解。此外，（4）和（5）中的指示函数样本均值收敛于约束满足的概率。

我们将该概率分布下的可行状态值集合 x'_0 记为

$$\begin{aligned} \mathbb{X}_0^\epsilon = \{x'_0 \in \Xi |\, &\mathbb{P}(x''_k \in \mathbb{X}) \geq 1-\epsilon,\, k=0,1,\ldots,N, \\ &\mathbb{P}(\kappa(x'_k) \in \mathbb{U}) \geq 1-\epsilon,\, k=0,1,\ldots,N-1\}. \end{aligned} \tag{7}$$

我们将候选状态 $x_0^\star$ 定义为

$$x_0^\star = \arg\min_{x'_0 \in \mathbb{X}_0^\epsilon} J_c(x'_0). \tag{8}$$

$x_0^\star$ 的相应代价为 $J_c^\star$

$$J_c^\star = \min_{x'_0 \in \mathbb{X}_0^\epsilon} J_c(x'_0) = J_c(x_0^\star). \tag{9}$$

尽管这些定义是以无限的 M为前提的，但后文将证明，$\mathcal{O}$的 M (log L)足以保证可行性和最优性，并使其概率/可靠性至少达到 $1-\delta$。例如，第6节的数值示例中使用了 $M = 135$。证明将在接下来的两节中给出，相关结果将以期望和概率的形式推导。

对核心递推式 (2) 和 (3) 的一种解释是：对于每个粒子 $x'_0 \in \Xi$，式 (2) 会生成随机激励闭环系统的 M 控制序列$\{u_{k,j} = \kappa(x'_{k,\ j})\}$。随后，递推式 (3) 使用这些闭环序列自身的开环代价的样本平均值，再对粒子状态密度取平均。从某种意义上说，搜索针对的是这些随机化闭环序列作为 x'_0 的函数时，其性能沿时域 N 取平均、并在粒子密度 Ξ上取平均后的结果。

状态选择算法的正则条件

上述步骤 (2b) 会生成一组 M 闭环控制序列 $\{\kappa(x'_{k,\ j})\}_{k=0}^{N-1}$，这些序列是 $x'_{0,\ j}$ 以及$w'_{k,\ j}$s 的函数；之后的步骤会考察相对于 ($\mathbb{X}$, $\mathbb{U}$) 的可行性以及与 J 相关的性能。为了使这一阶段有助于状态选择，状态和控制律必须能够通过不同的x'_0 以及 w'_k得到充分激励。考虑系统

$$x'_{k+1} = f(x'_k, \kappa(x'_k), w'_k), \quad x'_0, \tag{10}$$
$$u'_k = \kappa(x'_k), \tag{11}$$

该算法要求（10）能够从过程噪声 w'_k 到达，并要求（10）—（11）这一对具有可观测性和可重构性。这是为了确保控制序列对初始状态值和噪声过程的敏感性。这些是关于系统和控制律 κ 的条件。若缺乏用于选择 x'_0 的多样化控制序列，就无法保证 $\mathbb{X}_0^\epsilon$ 非空，最小化也无法有效进行。

与随机最优控制一样，闭环性能既取决于控制律 κ，也取决于由粒子密度 Ξ近似表示的信息状态 [4] 。粒子滤波器的传播应避免粒子耗尽问题，并应正确反映条件状态密度。我们假设其中包含条件均值等信息，并且覆盖范围足够广，能够产生丰富的可行控制序列。

除上述说明外，我们不再深入探讨。尽管如此，在计算示例中，我们指出了可行状态的可达性问题，以及粒子滤波器过于局部化时的性能限制。

3 候选状态的性质

借鉴确定性等价控制，其中将最小二乘最佳状态估计，即 Ξ 的条件均值 $\xi_{0|0}$，选作候选状态，我们比较 $x_0^\star$ 的代价 ˆJ_c 与 $\xi_{0|0}$ 的代价。

ˆ命题1。 假设 $\xi_{0|0}$（ Ξ 中ˆ粒子的样本均值）是可行的，即 $\xi_{0|0} \in \mathbb{X}_0^\epsilon$。则ˆ$J_c(x_0^\star) = J_c^\star \leq J_c(\xi_{0|0})$。 □

这是对状态选择算法最后阶段中关于x_0' 进行最小化的直接推论。

接下来，我们定义并比较设计代价函数 J_{des}，以便与代价 J_c 进行比较。

定义（设计代价）。假设在（*1*）中采用全状态反馈，则状态值 x_0' 的设计代价及其对应的最优因果控制律 $\kappa^\star$ 定义为

$$\begin{aligned} J_{des}(x_0') &= \min_{\kappa} \mathbb{E}_{W'} \left(\sum_{k=0}^{N} \ell_k(x_k', \kappa(x_k')) \right), \\ &= \mathbb{E}_{W'} \left(\sum_{k=0}^{N} \ell_k(x_k', \kappa^*(x_k')) \right), \end{aligned} \tag{12}$$

其中， κ^* 是使代价最小的因果控制律。在（12）的第一行中， $\{x_k'\}_{k\geq 0}$ 使用名义控制律 κ构造；而在第二行中，则使用 $\kappa^\star$构造。假设极小值点存在，否则将 min 替换为 inf。

命题2。若初始状态已知且可行，即 p_0 是位于$z \in \mathbb{R}^{r_x}$ 处的点质量，且$\mathbb{X}_0^\epsilon = \{z\}$，则 $J_{des}(z) \leq J_c^*$。

证明。

$$\begin{aligned} J_c^* &= J_c(z), \\ &= \mathbb{E}_{W'} \mathbb{E}_{W''} \left(\sum_{k=0}^{N} \ell_k \left(x_k'', \kappa^* \left(x_k' \right) \right) \right), \\ &\geq \mathbb{E}_{W'} \mathbb{E}_{W''} \left(\sum_{k=0}^{N} \ell_k \left(x_k', \kappa^* \left(x_k' \right) \right) \right), \\ &= \mathbb{E}_{W'} \left(\sum_{k=0}^{N} \ell_k \left(x_k', \kappa^* \left(x_k' \right) \right) \right), \\ &= J_{des}(z), \end{aligned}$$

其中，不等式成立是因为 $\{x_k'\}_k$ 是从 $x_0' = x_0'' = z$ 出发应用最优控制律 κ^* 后得到的最优序列。 □

注意，对于 J_{des} ，假设对所有 $k \geq 0$ 均可获得全状态反馈，而命题2的假设仅要求在 J_c^* 中的时刻0获得全状态反馈。

命题3。 在确定性情形下，即 $\Xi = \{z\}$和 $w_k = 0$，对于所有 k，若 z 可行，则 $J_{des}^*(z) = J_c^*$。

证明。 由 $W' = W'' = 0$和 $x_0' = x_0'' = z$可知，对于所有 k，根据这两个状态序列的定义 (2) 和 (3)，有$x_k' = x_k''$。因此，

$$\begin{aligned} J_c^* &= J_c(z), \\ &= \mathbb{E}_{W'} \mathbb{E}_{W''} \left(\sum_{k=0}^{T} \ell_k \left(x_k'', \kappa^* \left(x_k' \right) \right) \right), \\ &= \left(\sum_{k=0}^{N} \ell_k \left(x_k', \kappa^* \left(x_k' \right) \right) \right), \\ &= J_{des}(z), \end{aligned}$$

□

命题3重述了以下结论：对于确定性情形，开环控制与闭环控制等价；对于 $k \geq 0$ 的全状态反馈等价于仅在 $k = 0$ 时获得全状态反馈。在这种情况下，最优控制 κ^* 的候选状态 就是真实状态本身。

4 状态选择算法的计算复杂度

首先，求出 M 的下界，以保证具有裕度 δ 的 ϵ-概率可行性。然后，讨论应用状态选择算法所需的计算时间。

4.1 $M = \mathcal{O}(\log L)$

根据第2节，初始状态密度表示为粒子集 $\Xi = \{\xi^i | \, i = 1, 2, \ldots\ldots, L\}$上的有限粒子质量函数。

$$p_0(\cdot) = \sum_{i=1}^{L} \delta(\cdot - \xi^i), \tag{13}$$

其中 δ 是定义在 $\mathbb{R}^{r_x}$ 上的狄拉克δ函数。由于 Ξ 是有限的，因此式（7）中的 $\mathbb{X}_0^\epsilon$ 也是有限的，可写为

$$\begin{aligned} \mathbb{X}_0^\epsilon = &\bigcap_{k=1}^{N} \{x_0' \in \Xi | \, \mathbb{P}(x_k'' \in \mathbb{X}) \geq 1 - \epsilon\} \bigcap \\ &\bigcap_{k=0}^{N-1} \{x_0' \in \Xi | \, \mathbb{P}(\kappa(x_k') \in \mathbb{U}) \geq 1 - \epsilon\}. \end{aligned}$$

根据随机向量 x_k'' 和 x_k'的构造，条件 $\{x_k'' \in \mathbb{X}\}$ 取决于 x_0'、 x_0''、 W'和 W''，而$\{\kappa(x_k') \in \mathbb{U}\}$ 取决于 x_0' 和 W'。因此，两者都是 x_0'和 $\Lambda = (x_0'', \ W', \ W'')$的函数。它们可以改写为

$$G_k(x_0', \Lambda) = \{x_k'' \in \mathbb{X}\}, \, k = 1, \ldots, N,$$

and

$$G_{k+N+1}(x_0', \Lambda) = \{\kappa(x_k') \in \mathbb{U}\},\ k = 0, \ldots, N-1.$$

因此，更简洁地说，

$$\mathbb{X}_0^{\epsilon} = \bigcap_{k=1}^{2N} \{x_0' \in \Xi |\, \mathbb{P}(G_k(x_0', \Lambda)) \geq 1 - \epsilon\}.$$

算法第 1 步的统计可行性容差 α 满足 $\alpha \in [0, \epsilon)$，且在第 2a 步，$j = 1$，2，……，M 是 Λ的独立蒙特卡罗样本。定义集合

$$\mathbb{X}_0^{\alpha, M} = \bigcap_{k=1}^{2N} \{x_0' \in \Xi |\, \frac{1}{M} \sum_{j=1}^{M} \mathbb{1}(G_k(x_0', \Lambda^j)) \geq 1 - \alpha\}. \tag{14}$$

该集合以依赖于样本数 M 的概率，是 $\mathbb{X}_0^{\epsilon}$ 的子集。对于给定的可靠性 $\delta \in (0, 1)$，可以找到足够大的 M，使计算得到的样本均值 $\mathbb{X}^{\alpha}{}_{,\ 0}{}^{M}$ 能够恰当地近似 $\mathbb{X}_0^{\epsilon}$。

下面的定理是对 [19] 中定理 5 的一个版本，经过修改以适应该论文的问题表述。

定理 1。对于任意 $\delta \in (0, 1)$，若

$$M \geq \frac{1}{2(\epsilon - \alpha)^2} \log\left(\frac{L}{\delta}\right), \tag{15}$$

then

$$\mathbb{P}(\mathbb{X}_0^{\alpha, M} \subseteq \mathbb{X}_0^{\epsilon}) \geq 1 - \delta. \tag{16}$$

引理 1。*(霍夫丁不等式 [20])*。对于独立的随机变量 Z_q、$q = 1$、……、M、$\mathbb{P}(Z_q \in [a_q, b_q]) = 1$、$a_q \leq b_q$，对所有 $t \geq 0$

$$\mathbb{P}\left(\sum_{q=1}^{\bar{M}} (Z_q - \mathbb{E}\ Z_q) \geq t\bar{M}\right) \leq \exp\left(-\frac{2\bar{M}^2 t^2}{\sum_{q=1}^{\bar{M}} (b_q - a_q)^2}\right)$$

□

证明。（定理 1 的证明）。设 $x \in \Xi \backslash \mathbb{X}_0^{\epsilon}$。那么存在 l，它是 $\mathbb{P}(G_l(x,\ \Lambda))$ 的一个极小值点，$l \in \{1, 2, \ldots, N\}$。根据 x 的定义，我们有 $\mathbb{P}(G_l(x,\ \Lambda)) < 1 - \epsilon$。定义随机变量 Y_j，对于 $j = 1$，2，…，M，使得 $Y_j = \mathbb{1}\{G_l(x,\ \Lambda^j)\}$。因此，

$$\mathbb{E}\, Y_j = \mathbb{P}\left(G_l(x, \Lambda^j)\right) < 1 - \epsilon. \tag{17}$$

因此，

$$\begin{aligned}
&\mathbb{P}(x \in \mathbb{X}_0^{\alpha, M}) \\
&\leq^{①} \mathbb{P}\left(x \in \{x_0' \in \Xi |\, \frac{1}{M} \sum_{j=1}^{M} \mathbb{1}(G_l(x_0', \Lambda^j)) \geq 1 - \alpha\}\right), \\
&=^{②} \mathbb{P}\left(\frac{1}{M} \sum_{j=1}^{M} Y_j \geq 1 - \alpha\right), \\
&\leq^{③} \mathbb{P}\left(\frac{1}{M} \sum_{j=1}^{M} (Y_j - \mathbb{E}\, Y_j) \geq -1 + \epsilon + 1 - \alpha\right), \\
&\leq \mathbb{P}\left(\sum_{j=1}^{M} (Y_j - \mathbb{E}\, Y_j) \geq M(\epsilon - \alpha)\right), \\
&\leq^{④} \exp\left(-2M(\epsilon - \alpha)^2\right),
\end{aligned}$$

其中，带编号的不等式分别由以下原因得出：①- 事件交集的概率不大于其中任意单个事件的概率；②, ③- (17) 中 Y_j 的定义；④- Hoeffding 不等式（引理 1），其中 $a_j = 0$，$b_j = 1$ 对所有$j \in \{1, 2, \ldots, M\}$均成立。因此，

$$\begin{aligned}
\mathbb{P}\left(\mathbb{X}_0^{\alpha, M} \not\subseteq \mathbb{X}_0^{\epsilon}\right) &= \mathbb{P}\left\{\text{there exists } x \in \mathbb{X}_0^{\alpha, M} \textit{ s.t. } x \notin \mathbb{X}_0^{\epsilon}\right\}, \\
&\leq \sum_{x \in \Xi \backslash \mathbb{X}_0^{\epsilon}} \mathbb{P}(x \in \mathbb{X}_0^{\alpha, M}), \\
&\leq L \exp\left(-2M(\epsilon - \alpha)^2\right).
\end{aligned}$$

若左侧的值应为 $\delta \in (0, 1)$，则对两边取 log(·)，可得 (15)。 □

若 $x_0' \in \mathbb{X}^{\alpha}{}_{,\ 0}{}^{M}$，则以至少 $1 - \delta$ 的概率有 $x \in \mathbb{X}_0^{\epsilon}$。对于实现最优性，也可以得到 $\mathcal{O}(\log L)$ 的下界 M，其结果类似：算法第 3 步中的 x_0' 满足 (8) 的概率接近于 1，这里使用了对 [19] 的推论 6 的扩展。

选择参数 $\alpha \in (0,\ \epsilon)$ 和 $\delta \in (0,\ 1)$ 时固有的近似，以及不等式 (15) 中的不精确性，构成了状态选择算法量化结果内在保守性的来源。在第6节中使用样本重复次数 $M = 135$ 和粒子数 $L = 400$ 进行计算示例时，所达到的约束违反率将清楚地体现这一点。

4.2 计算时间

状态选择算法每个时间步的计算复杂度为 $\mathcal{O}(NL \log\ L)$。它可在以下方面并行化：第二步；在选择 x_0' 时跨越 i；以及在样本中跨越 j。使用图形处理器（GPU）还能进一步缩短所需的计算时间。

随着状态维度 r_x 增大，这种方法的复杂度也会随之增加，因为粒子滤波器所需的样本数 L 可能会大幅增加。不过，一旦 L 固定，针对给定的 M、ϵ、α 和 δ 的计算便会遵循既定规律。

5 带约束的随机线性系统与二次代价

本节将说明，在满足以下假设时：线性动力学、线性状态反馈控制律、二次阶段成本以及多面体状态与输入约束，候选状态是基于密度二阶矩、在 $\mathbb{R}^{r_x}$上求解二次规划得到的解。这样就无需计算 M 样本序列及其约束违反和代价的样本均值，也无需像前文那样引入裕度参数 α 和 δ。我们假设滤波条件状态密度由其前两阶矩描述，而这两个矩可由卡尔曼滤波器或粒子滤波器提供。

At time $k = 0$, we have:

I. 离散时间线性状态动力学

$$x_{k+1} = Fx_k + Gu_k + w_k. \tag{18}$$

II. $\{w_k\}$ 是独立同分布的，$\mathbb{E}\, w_k = 0$ 且 $\mathsf{cov}(w_k) = \Sigma_w$，对所有 k均成立。III. 多面体状态和输入约束集的最小约束违反概率 $\epsilon \in (0,\ 1)$,

$$\begin{aligned} \mathbb{X} &= \{x_k \in \mathbb{R}^{r_x} | Tx_k \le \bar{x}\},\ \bar{x} \in \mathbb{R}^t, \\ \mathbb{U} &= \{u_k \in \mathbb{R}^{r_u} | Su_k \le \bar{u}\},\ \bar{u} \in \mathbb{R}^m, \end{aligned} \tag{19}$$

其中 $T \in \mathbb{R}^{t\times r_x}$ 和 $S \in \mathbb{R}^{m\times r_u}$ 具有满行秩。IV. 初始状态 x_0的密度的前两阶矩 $\mathbb{E}\, x_0 = \hat{}x_0$ 和 $\mathsf{cov}(x_0) = \Sigma_0$，，且该初始状态与 w_k 相互独立，对所有 k均成立。

V. 线性全状态反馈控制律 $u_k = Kx_k$，但既未假设其关于 $\mathbb{X}$ 和 $\mathbb{U}$ 递归可行，也未假设其关于某个代价最优，实际上也未假设其具有稳定性。

VI. 有限时域二次轨迹成本函数 J

$$J = x_k^T Q_N x_k + \sum_{k=0}^{N-1} \left[x_k^T Q x_k + u_k^T R u_k\right].$$

针对这种情况，我们推导出一种状态选择算法的变体，无需对仿真进行样本均值计算，即可直接刻画状态的期望、代价和约束违反情况。最终得到一个用于确定所选状态 $x_0^\star$的二次规划。这相当于在非常大的M 下运行状态选择算法，且无需承担随之而来的计算开销。

我们将在后面的数值示例中看到，对于这些受约束线性二次问题，当约束不起作用时，状态选择算法会返回条件均值。

(2)和(3)中的状态序列变为

$$x'_{k+1} = (F + GK)x'_k + w'_k,\ x'_0 \in \mathbb{R}^{r_x}, \tag{20}$$
$$x''_{k+1} = Fx''_k + GKx'_k + w''_k,\ x''_0 \sim p_0. \tag{21}$$

状态选择算法中的这些递推关系给出:

- x'_k affine in x'_0 and w'_{k-j},
- x''_k affine in x'_0, x''_0, w'_{k-j} and w''_{k-j}.

进而，这意味着(6)中的期望成本函数 J_c 关于 x'_0是二次的。约束的分析更为繁琐，但会得出关于 x'_0的新的、但更保守的多面体约束。

5.1 代价函数 J_c

对于 $x'_0 \in \mathbb{X}_0^\epsilon$，代价函数 J_c 由 (6) 给出。

$$\begin{aligned} J_c(x'_0) &= \mathbb{E}_{x''_0}\mathbb{E}_{W'}\mathbb{E}_{W''}\Bigg(\sum_{k=0}^{N-1}\Big[(x''_k)^T Q x''_k \\ &\qquad + (x'_k)^T K^T R K x'_k\Big] + (x''_N)^T Q_N x''_N\Bigg), \quad (22) \\ &= \mathbb{E}_{x''_0}\mathbb{E}_{W'}\mathbb{E}_{W''}\Bigg(\sum_{k=0}^{N-1}\Big[tr(Q x''_k (x''_k)^T) \\ &\quad + tr(K^T R K x'_k (x'_k)^T)\Big] + tr(Q_N x''_N (x''_N)^T)\Bigg). \quad (23) \end{aligned}$$

由 (20) 和 (21),

$$x'_k = F_K^k x'_0 + \sum_{p=0}^{k-1} F_K^p w'_{k-p-1}, \tag{24}$$

$$\begin{aligned} x''_k &= F^k x''_0 + \sum_{j=0}^{k-1} F^j w''_{k-1-j} \\ &\quad + \Psi_{k-1} x'_0 + \sum_{h=0}^{k-1} F^h GK \sum_{p=0}^{k-h-2} F_K^p w'_{k-h-2-p}, \end{aligned} \tag{25}$$

其中 $F_K =$(F+GK) 且 $\Psi_{k-1} = \sum_{h=0}^{k-1} F^h GK F_K^{k-h-1}$。也就是说，这些状态序列是 x''_0和x'_0的仿射函数。因此，代价 J_c 关于 x'_0 和 x''_0是二次的。

记均值为 $\hat{}x'_k = \mathbb{E}_{W'} x'_k$和 $\hat{}x''_k = \mathbb{E}_{x''_0}\mathbb{E}_{W'}\mathbb{E}_{W''} x''_k$，则有

$$\hat{x}'_k = F_K^k x'_0, \quad \hat{x}''_k = F^k \hat{x}''_0 + \Psi_{k-1} x'_0.$$

因此，由于 w'_k 和 w''_k 的均值为零，误差 $\tilde{x}'_k = x'_k - \hat{}x'_k$ 和 $\tilde{x}''_k = x''_k - \hat{}x''_k$ 均为零。

利用独立性假设，并忽略加性常数，

$$J_c(x_0') = (x_0')^T \mathcal{A}_1 x_0' + (\hat{x}_0'')^T \mathcal{A}_2 x_0'. \tag{26}$$

$\mathcal{A}_1$ 和 $\mathcal{A}_2$ 的公式及推导见附录。

推论1。 假设 K_k 是无约束*N*时域线性二次问题的最优时变反馈增益，其中 Q、 R 和 Q_N。那么，候选状态就是条件均值。 □

这一结果可直接由 [21, 第 221]页给出的论证的离散时间等价形式，以及无约束情况下控制性能对状态估计协方差迹的单调依赖关系推出。

5.2 约束

状态序列和输入序列 $\{x_k''\}$、 $\{u_k\} = \{Kx_k'\}$ 按构造均为随机的。 x_k' 的预测协方差Σ_k'可表示为

$$\Sigma_{k+1}' = (F+GK)\Sigma_k'(F+GK) + \Sigma_w,\ \Sigma_0' = 0, \tag{27}$$

其中，初始协方差为零，因为x_0' 是确定性的。对于 Σ_k'', x_k''的协方差为

$$\Sigma_{k+1}'' = F\Sigma_k''F^T + GK\Sigma_k'K^TG^T + \Sigma_w,\ \Sigma_0'' = \Sigma_0. \tag{28}$$

线性问题陈述中的第 IV 项给出了初始协方差 Σ_0。

输入序列的协方差可由式 (27) 得出。

$$\text{cov}(u_k) = \text{cov}(Kx_k') = K\Sigma_k'K^T. \tag{29}$$

接下来，我们说明如何将式 (19) 中的概率约束转化为关于状态序列均值 (26) 和协方差 (27)、(28) 的确定性线性约束。

命题4。 若满足以下确定性多面体约束，则式 *(19)* 中的概率多面体状态约束也得到满足

$$T\hat{x}_k'' \le \bar{x} - \sqrt{\frac{t-\epsilon}{\epsilon}}\sqrt{diag(T\Sigma_k T^T)}, \tag{30}$$

其中： t 是 T_k的行数，函数 *diag*$(\cdot)$返回方阵对角线元素组成的列向量，第二个平方根按元素计算。

引理 2. ***(Cantelli*不等式*)*** 对于均值为 $\hat{\gamma}$、方差为 Γ 的标量随机变量 γ，

$$\mathbb{P}(\gamma - \hat{\gamma} \ge \eta) \le \frac{\Gamma}{\Gamma + \eta^2},\ \eta \ge 0. \tag{31}$$

□

引理 3. 对于 $j = 1, \ldots, t$，令 $T(j)$ 为 T 的第 j^{th} 行，$\bar{x}(j)$ 为 $\bar{x}$ 的第 j^{th} 个元素。概率约束

$$\mathbb{P}\Big(T(j)x_k'' \le \bar{x}(j)\Big) \le 1 - \frac{\epsilon}{t}, \tag{32}$$

若以下线性不等式成立，则这些约束得到满足

$$T(j)\hat{x}_k'' \le \bar{x}(j) - \sqrt{\frac{t-\epsilon}{\epsilon}}\sqrt{T(j)\Sigma_k''T(j)^T}. \tag{33}$$

证明。 与 [22], 中的工作类似，假设存在$\rho \ge 0$ 使得

$$T(j)\hat{x}'' \le \bar{x} - \rho, \tag{34}$$

因此，条件 $T(j)x'' \ge T(j)\hat{}x'' + \rho$ 可由$T(j)x'' \ge \bar{x}(j)$推出，从而，

$$\begin{aligned}\mathbb{P}\Big(T(j)x'' \ge \bar{x}\Big) &\le \mathbb{P}\Big(T(j)x'' \ge T(j)\hat{x}'' + \rho\Big),\\ &= \mathbb{P}\Big(T(j)x'' - T(j)\hat{x}'' \ge \rho\Big),\\ &\le \frac{T(j)\Sigma''T(j)^T}{T(j)\Sigma''T(j)^T + \rho^2},\end{aligned}$$

其中，最后一个不等式由 Cantelli不等式 (31) 得出。如果 ρ 足够大，使得最后一项的上界为 ϵ/t,

$$\frac{T(j)\Sigma''T(j)^T}{T(j)\Sigma''T(j)^T + \rho^2} \le \frac{\epsilon}{t},$$

或等价地

$$\rho \ge \sqrt{\frac{t-\epsilon}{\epsilon}}\sqrt{T(j)\Sigma''T(j)^T}.$$

将 ρ的这一下界代入 (34)，即可得到 (33)。 □

引理 4. 设$(\Omega\ \mathcal{B}, \mathbb{P})$ 为一个概率空间，且 $E_i \in \mathcal{B}$对于 $i = 1, \ldots, n$ 成立。若对所有 $i = 1, \ldots, n$，均有 $\mathbb{P}(E_i) \ge 1 - \epsilon/n$，则 $\mathbb{P}(\bigcap_{i=1}^n E_i) \ge 1 - \epsilon$。 □

证明。 **（命题4的证明）**。注意，(19)中的状态约束集可以写成集合交集

$$\begin{aligned}\mathbb{X}_k &= \{x_k \in \mathbb{R}^{r_x} | Tx_k \le \bar{x}_k\},\\ &= \bigcap_{j=1}^{t} \{x_k \in \mathbb{R}^{r_x} | T(j)x_k \le \bar{x}_k(j)\},\end{aligned}$$

由于所有行都必须同时执行 [19]。根据引理 4， $\mathbb{P}(x_k \in \mathbb{X}_k)$ $\ge 1 - \epsilon$ 由 $\mathbb{P}(\{x_k \in$推出

$\mathbb{R}^{r_x}|T(j)x_k \le \bar{x}_k(j)\}) \ge 1-\epsilon/t$。后者由引理 3 中的 (33) 推出。将 (33) 中针对 T 的所有 t 行不等式叠加起来，我们得到 (30)。请注意，随着行数 t 的增加，约束会变得更加严格，近似也会更加保守。 □

采用类似的论证，可得到与上述结果类似的概率输入约束结果。最后，概率约束 (19) 关于 x'_0 仍为多面体形式，如下所示，其中 K 为名义状态反馈增益， S 为控制约束 (19) 中的矩阵。

$$\begin{aligned} T\hat{x}''_k &\le \bar{x} - \sqrt{\frac{t-\epsilon}{\epsilon}}\sqrt{diag(T\Sigma''_k T^T)}, \\ &\quad for\ all\ k = 1,2,\dots,N, \\ SK\hat{x}'_k &\le \bar{u} - \sqrt{\frac{m-\epsilon}{\epsilon}}\sqrt{diag(SK\Sigma'_k K^T S^T)}, \\ &\quad for\ all\ k = 1,2,\dots,N-1, \\ SK\hat{x}'_0 &\le \bar{u}, \end{aligned} \tag{35}$$

其中，$\hat{}x'_k$ 和 $\hat{}x''_k$ 仅为 x'_0的函数，与 (26) 中相同。协方差矩阵 Σ'_k 和 Σ''_k 可离线计算，且与 x'_0无关。最后一个不等式是因为 $\Sigma'_0 = 0$。

我们沿用 [10]在 (35) 中使用“闭环协方差”的做法。也就是说，我们将协方差矩阵 Σ'_k 和 Σ''_k 替换为它们的一步预测

$$\Sigma'_1 = \Sigma_w, \quad \Sigma''_1 = F\Sigma''_0 F^T + \Sigma_w, \tag{36}$$

其中 $\Sigma''_0 = \Sigma_0$ 由式（1）给出。这考虑到了在下一个时刻可以获得系统的新测量值这一事实。因此，这放宽了约束，并避免了当 F 不稳定时， Σ''_k 的 k 出现无界增长。

以紧凑形式表示的求解 $x_0^\star$的最小化问题，是一个关于 x'_0的二次规划。

$$\begin{aligned} \min_{x'_0} \quad & (x'_0)^T \mathcal{A}_1 x'_0 + (\hat{x}''_0)^T \mathcal{A}_2 x'_0, \\ & subject\ to \\ T\Psi_{k-1}x'_0 &\le \bar{x} - \sqrt{\frac{t-\epsilon}{\epsilon}}\sqrt{diag(T\Sigma''_1 T^T)} - TF^k\hat{x}''_0, \\ & \quad for\ all\ k = 1,2,\dots,N \\ SKF_K^k x'_0 &\le \bar{u} - \sqrt{\frac{m-\epsilon}{\epsilon}}\sqrt{diag(SK\Sigma_w K^T S^T)}, \\ & \quad for\ all\ k = 1,2,\dots,N-1 \\ SKx'_0 &\le \bar{u}, \end{aligned} \tag{37}$$

对于具有多面体约束的线性二次随机问题，状态选择算法可简化为针对 x'_0 的二次规划，从而能够在不进行采样、也不考虑样本均值收敛要求的情况下评估该方法。

进而，这使得我们能够将其与确定性情形 [23] 下的已知解以及通过卡尔曼滤波获得的状态条件均值估计进行比较。

6 数值示例

本节介绍应用第4节状态选择算法的计算示例。我们还给出一个线性系统示例，该系统具有二次代价和多面体约束，适用于第5节介绍的二次规划方法。此外，我们还评估了该算法的计算开销，包括自助粒子滤波器的传播。

6.1 非线性系统

我们沿用第2节的问题表述框架。

I. 状态动力学描述如下:

$$\begin{aligned} z_{k+1} &= 0.9z_k + 0.2h_k + w_k^1, \\ h_{k+1} &= -0.15z_k + 0.9h_k + 0.05z_k h_k + u_k + w_k^2, \\ y_k &= z_k + v_k. \end{aligned}$$

其中， $x_k = (z_k\ h_k)^T \in \mathbb{R}^2$ 是状态向量， u_k 是标量控制量， y_k 是标量测量值， $w_k = (w_k^1\ w_k^2)^T \in \mathbb{R}^2$ 是过程噪声， v_k 是标量测量噪声。II. 噪声 $w_k \sim \mathcal{N}(0_2, 0.3\mathbb{I}_2)$ 和 $v_k \sim \mathcal{N}(0, 0.3)$ [3]。III. 输入约束集为 $\mathbb{U} = [-3,3]$。状态约束集 $\mathbb{X}$ 是 $\mathbb{R}^2$ 中 $\mathcal{L}$ 的补集，其中$\mathcal{L} = [3,5]\times[-4,2]\cup[-2,5]\times[-7,-4]$。下图展示了需要避开的这个 L 形集合。约束违反率为 $\epsilon = 0.3$。IV. x_0 状态密度由 $\mathbb{R}^2$ 中的一组 $L = 400$ 粒子给出。V. 我们考虑两个连续名义控制器：(a) 稳定反馈线性化控制律 $u_k = \kappa_1(x_k) = -0.05z_k h_k$；然后是 (b) 一个可行最优控制器 $\kappa_2(x_k)$，稍后将详细介绍。

VI. 运行成本为 $\ell_k(x_k, u_k) = x_k^T x_k + u_k^2$。时域 $N = 6$。

对于状态选择算法，选择 $\alpha = .1, \delta = .01$。定理 1 随后给出 $M = 135$。接下来执行以下序列，从 Ξ 开始，其中 $L = 400$ 个粒子从 $\mathcal{N}((\ 7.5\ -7.5)^T, .5\mathbb{I}_2)$ 中采样得到。

i) The selected state, $x_0^\star$, is applied in the control $u_k = \kappa(x_0^\star)$,

[3] $\mathcal{N}(\mu,\ \Sigma)$表示均值向量为 μ、协方差矩阵为 Σ的高斯密度。

ii) 测量输出 y_{k+1} ； iii) 自助粒子滤波器计算出一组更新后的 Ξ，个用于 x_{k+1}的 L 滤波粒子；iv) 重新执行状态选择。

通过这次迭代，可以联合评估状态选择算法、与之配套的粒子滤波器以及名义控制律的控制性能和约束处理能力。由于本文旨在研究用于控制的状态估计，因此这是一项关键评估。

为了更好地直观展示开环动力学，图 1 展示了输入和状态扰动均为零时状态动力学的流线。

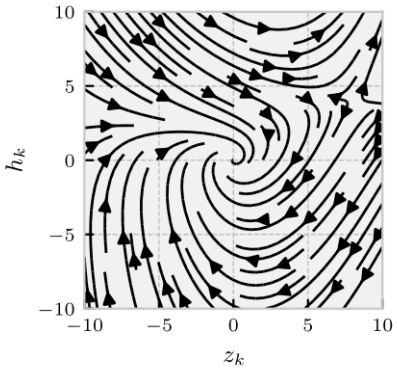


图 1. 将 u_k 和 w_k 设为零时非线性系统状态的流线。

6.1.1 示例 *1*：稳定控制器

我们对受控非线性系统进行了两组对比仿真：一组采用上述状态选择算法，即 $u_k = \kappa_1(x_0^\star)$；另一组采用可称为确定性等价控制器的方法，$u_k = \kappa_1(\hat{x}_{0|0})$，其中 $\hat{x}_{0|0}$ 是相应粒子滤波器的条件均值。两种情况下，粒子滤波器都根据相应的测量输出演化，而测量输出又取决于施加的控制。图 2 展示了采用状态选择算法反馈时受控状态的滤波后粒子密度。图 3 展示了相应的条件均值反馈情形。图 4 展示了每个时间步中违反状态约束并落入集合 $\mathcal{L}$ 的粒子百分比。这表明状态选择算法采取了更谨慎的策略。

关于这两次仿真，我们有以下观察。

- 闭环粒子密度的演变有所不同，因为控制信号不同，因此状态和测量值也不同。
- 状态选择算法避免了状态约束违反，但可能过于保守，而条件均值控制

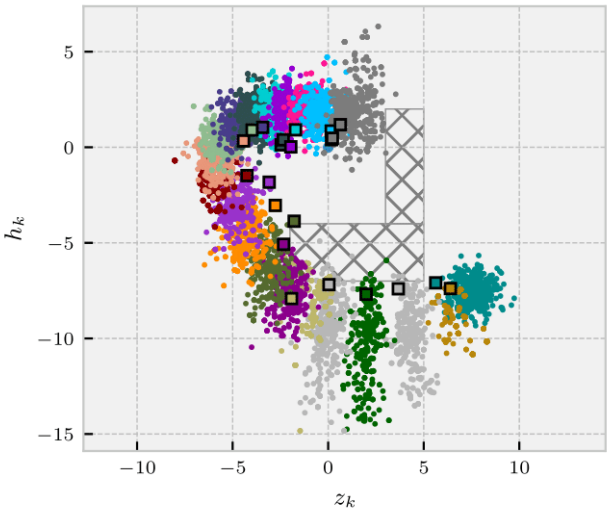


图 2. **状态选择算法：** 在使用 $u_k = \kappa_1(x_0^\star)$ 时，粒子滤波密度的演变，即控制器在每个时间步使用的候选状态 $x_0^\star$ 。黑色方块表示从粒子密度中选出的状态的位置。交叉阴影区域表示状态约束集 $\mathbb{X}$ 的补集。

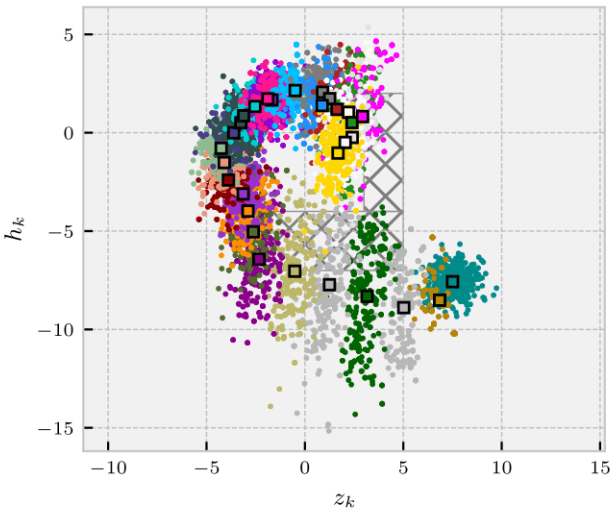


图 3. **条件均值：** 在使用 $u_k = \kappa_1(\hat{x}_{0|0})$ 时，粒子滤波密度的演变，其中 $\hat{x}_{0|0}$ 是粒子滤波密度的条件均值，每个时刻均以黑色方块标出。

违反了状态约束要求（以粒子密度衡量）。

- 显然，所选状态通过选择接近约束边界的 $x_0^\star$，使后续粒子密度更加谨慎。这一特性取决于名义控制律 $\kappa_1(\cdot)$。
- 状态选择模拟在时间步 k =22 停止，因为粒子密度中的可行状态集 $\mathbb{X}_0^{\alpha,M}$ 为空。也就是说，没有任何状态选择

x_0' 得出了可行解。同样，这是由于受控动态系统在原点附近缺乏可达性所导致的性质。

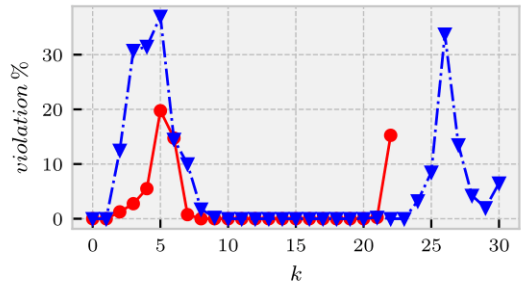


图 4. 蓝色三角形表示使用粒子密度的当前条件均值控制 $u_k = \kappa_1(\hat{x}_{0|0})$ 时非线性系统的状态约束违反率。红色圆圈表示基于状态选择算法的控制 $u_k = \kappa_1(x_0^\star)$ 的相同违反率。

仿真使用 PYTHON（一种未经编译的解释型程序）在配备 16.00 GB 随机存取存储器的 M1 芯片 2021 款 MacBook Pro 上进行。图 2 中仿真每个时间步的平均运行时间约为 4.9 秒，其中包括状态选择算法（主要计算负载）和粒子滤波器。该算法可完全并行化，既可针对状态选择 x_0'，也可针对相应的样本进行并行化。虽然本文未实现这一点，但它有望缩短计算时间。

6.1.2 示例 *2*：可行最优控制器

接下来，我们使用状态反馈控制器 $\kappa_2(\cdot)$ 重新进行先前的实验：这是一个由采样和一次值迭代计算得到的 $(\mathbb{U}, \mathbb{X})$-可行、无限时域折扣成本最优控制器。其构造过程详见附录。

针对控制律 $\kappa_2(\cdot)$，进行了两次闭环仿真。图 5 展示了在控制器中使用状态选择算法的 $x_0^\star$ 所得到的结果。图 6 展示了使用粒子滤波器的条件均值 $\hat{x}_{0|0}$ 时的相应行为。图 7 展示了每个时间步违反状态约束的粒子所占百分比。ϵ 的值为 0.3。

注： 如第 *2* 节所述，信息状态/粒子密度 Ξ 在状态选择算法中发挥着核心作用；它使闭环控制序列 $\{\kappa(x'_{k,\ j})\}_{k=0}^{N}$ 多样化，因为 x_0' 是从 Ξ 中的粒子里选取的。因此，状态选择算法对粒子滤波器的要求可分为两部分。首先，粒子滤波器必须足够准确地近似贝叶斯滤波器。这是构建该算法的基本假设。其次，必须避免粒子耗尽，以保留粒子变异性。

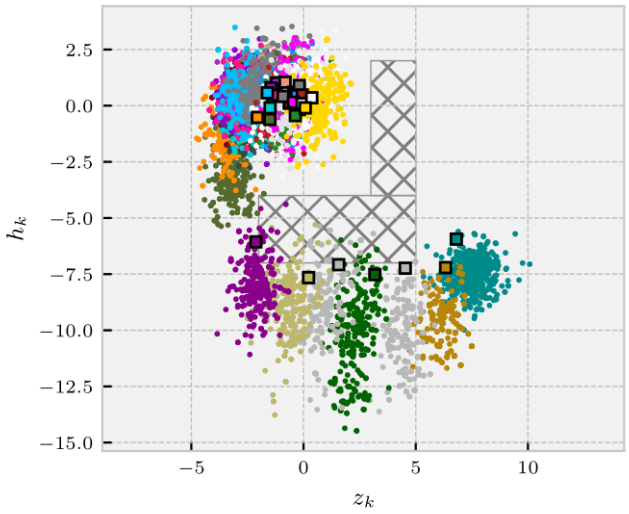


图 5. **状态选择算法：** 粒子滤波密度随 $u_k = \kappa_2(x_0^\star)$ 的演变。也就是说，控制器在每个当前时间步使用由黑色方块标出的候选状态 $x_0^\star$。黑色方块表示选定的状态。

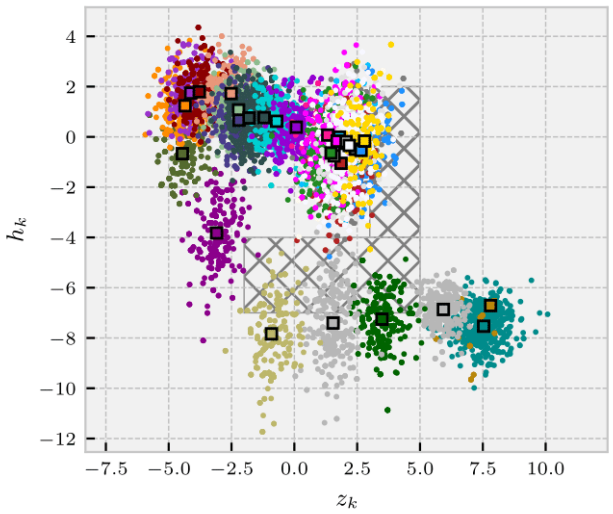


图 6. **条件均值：** 粒子滤波密度随 $u_k = \kappa_2(\hat{x}_{0|0})$ 的演变，也就是说，控制器在每个当前时间步使用粒子滤波条件均值。

在上述示例中，我们使用了方差为0.3 的测量扰动 v_k，以及协方差为 $0.3\mathbb{I}_2$ 的状态扰动 w_k。这有助于保持粒子滤波密度的多样性，如图所示。例如，若将测量扰动的方差设为0.1，并将粒子数固定为400，则会导致粒子耗竭，进而引发状态选择算法的可行性问题，并且使贝叶斯滤波密度的表示效果变差。

与粒子滤波器的要求相反但又相辅相成的是对控制律 κ 以及状态扰

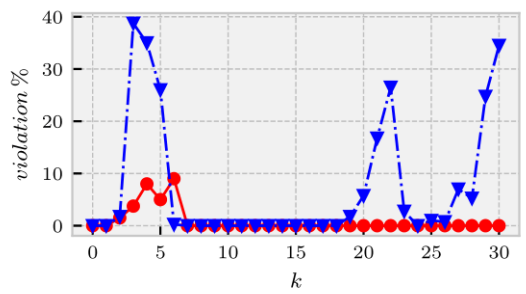


图 7. 蓝色三角形表示采用控制 $u_k = \kappa_2(\hat{x}_{0|0})$（即使用粒子密度的条件均值）时非线性系统的状态约束违反率。红色圆圈表示基于状态选择算法的控制 $u_k = \kappa_2(x_0^\star)$ 对应的相同指标。ϵ 的值为 30%。

动 w_k 的要求。第2节将这一主题作为正则性条件进行了讨论。对于状态选择算法而言，从 w_k 获取状态和控制信号的可达性会影响该算法的闭环可行集 $\mathbb{X}_0^\epsilon$。在采用反馈线性化控制器 κ_1 的非线性示例中，经过22步后可行集为空——这一结果在每次运行时都会发生变化。这是由于原点附近控制增益减小，导致激励不足。

6.1.3 *Example 3: Linear dynamics with polyhedral constraints*

在本例中，考虑一个标准的 DC-DC变换器调节问题。它是随机模型预测控制文献中的一个基准问题，并用于 [15,24,25]。

- 其动力学描述如下：

$$\begin{pmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{pmatrix} = x_{k+1} = Fx_k + Gu_k + w_k,$$
$$y_k = Hx_k + v_k,$$

其中

$$F = \begin{bmatrix} 1 & 0.0075 \\ -0.143 & 0.996 \end{bmatrix}, \quad G = \begin{bmatrix} 4.798 \\ 0.115 \end{bmatrix},$$
$$H = \mathbb{I}_{2\times 2}.$$

- 噪声信号 w_k 和 v_k 均为零均值白噪声，且彼此独立，也与 x_0独立。

$$\mathrm{cov}(w_k) = \Sigma_w = 0.1\mathbb{I}_{2\times 2},\ \mathrm{cov}(v_k) = \mathrm{diag}(0.5, 0.4),$$
$$\mathbb{E}\, x_0 = (0.6455, 1.3751)^T,\ \mathrm{cov}(x_0) = \Sigma_0 = 0.1\mathbb{I}_{2\times 2}.$$

在仿真中，我们将密度设为高斯分布。

- 概率约束为

$$\mathbb{P}(x_k^1 \le 2) \ge 1 - \epsilon, \text{ for all } k = 1, \ldots, N$$

其中，与 [15] 的 40% 不同，$\epsilon = 10\%$。根据命题4和（37），这意味着

$$T\Psi_{k-1}\hat{x}_0 \le \bar{x} - \sqrt{\frac{t-\epsilon}{\epsilon}}\sqrt{\mathrm{diag}(T\Sigma_1'' T^T)} - TF^k\hat{x}_0'',$$

对于所有 $k = 1, \ldots, N$。其中 $T = (1, 0)$、$\bar{x} = 2$、$t = 1$ 和 Σ_1'' 是预测协方差，并在 (36) 中定义。

- 线性状态反馈控制器 $u_k = Kx_k$ 选为具有如下权重矩阵的无限时域LQR控制器，与 [15] 中的权重矩阵类似：Q =diag(1,10) 和 $R = 10$

$$K = \begin{bmatrix} -0.2409 & 0.3930 \end{bmatrix},$$

这些权重矩阵也用于构建优化问题 (37)，且 $Q_N = Q$。

- 对于 (37)，预测时域选为 $N = 8$。相应的矩阵 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 为

$$\mathcal{A}_1 = \begin{bmatrix} 47.23 & -43.76 \\ -43.76 & 45.51 \end{bmatrix}, \quad \mathcal{A}_2 = \begin{bmatrix} -93.98 & 87.18 \\ 85.33 & -89.45 \end{bmatrix},$$

每个时刻都求解二次规划 (37)，以找到候选状态 $x_0^\star$，并将其用于线性状态反馈控制器。使用卡尔曼滤波器更新状态条件均值和协方差；该滤波器仍是最小二乘最优无偏估计器。100 次闭环迭代的结果如图8所示。使用卡尔曼滤波器得到的状态条件均值搭配线性控制器进行相应仿真的结果如图9所示。

请注意，当 x_k^2 低于某条特定直线时，优化问题返回的候选状态近似为条件均值。然而，超过该直线后，候选状态会有所不同。这是约束所产生的影响，该约束存在于状态分段过程中，但在线性二次型调节器或卡尔曼滤波器中不存在。

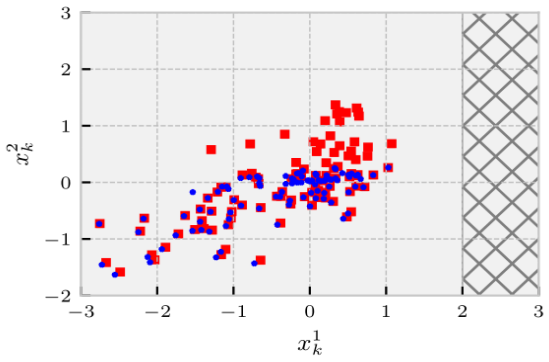


图 8. 控制 $u_k = Kx_0^\star$ 时卡尔曼滤波器条件均值（红色方块）的取值。蓝色圆点表示控制器采用的候选状态取值，$x_0^\star$，。

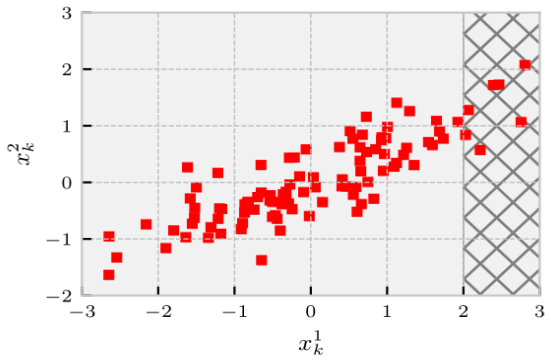


图 9. 卡尔曼滤波器条件均值 $\hat{x}_{0|0}$ 的取值，其中 $u_k = K\hat{x}_{0|0}$。

图10展示了：图8中 x_k^1的条件均值；由卡尔曼滤波器传播的其双西格玛区间（条件协方差第(1,1)项的平方根）；以及真实状态。阴影区域表示两个标准差，即95%的置信区间。由于本示例算法中使用的容差为 $\epsilon = 10\%$，因此该解是保守的。

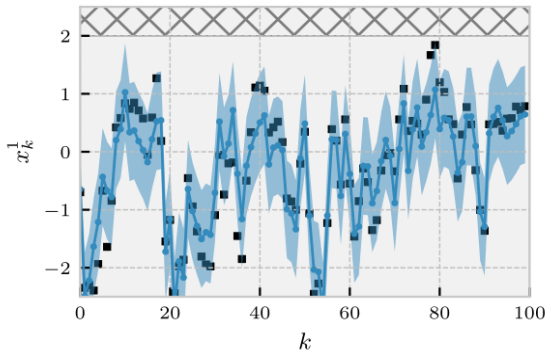


图 10. 深蓝色曲线表示 x_k^1 的条件均值，阴影区域表示条件均值上下两个标准差的范围，真实状态以黑色方块表示。

7 结论

状态选择算法需要一个非空的初始状态集合 $\mathbb{X}_0^\epsilon$ 才能运行。正如前文所述，估计器（此处为粒子滤波器）和固定控制器 κ 对算法的可行性也起着各自的作用。容易耗尽粒子的粒子滤波器会限制 Ξ 的变化性，因此无法很好地近似贝叶斯滤波器。另一方面，某个控制器（例如 $\kappa = 0$）会消除控制对初始状态和 w_k' 的所有敏感性，如第6节的第一个示例所示。这种控制器也会消除平均成本中的变化。

虽然用软概率约束替代硬概率约束至少可以从算法层面避免不可行性问题，但仍需更具体地理解该问题。因此，我们未来的工作将致力于

- 通过为 w_k''添加更多选项来丰富所提出的控制序列分布，例如使其具有不同于 w_k''的统计特性。 w_k'
- 明确控制器所需的更多实用特性。 κ
- 探索控制中的参数估计。可以通过简单的状态增广来实现这一点，所提供的框架无论采用粒子滤波器方法还是状态选择算法，都可以自然地适应。

参考文献

[1] D. Q. Mayne，《模型预测控制：近期进展与未来前景》，Automatica，第 50 卷，第 12 期，第 2967–2986 页，2014 年。[2] A. A. Fel'dbaum，《最优控制系统》。Elsevier，1966 年。[3] Y. Bar-Shalom，《随机动态规划：谨慎与探索》，IEEE Transactions on Automatic Control，第 26 卷，第 5 期，第 1184–1195 页，1981 年。[4] P. R. Kumar 和 P. Varaiya，《随机系统：估计、辨识与自适应控制》。SIAM，2015 年。[5] K. J. ˚Astr¨om，《随机控制理论导论》。Courier Corporation，2012 年。[6] D. Bertsekas，《动态规划与最优控制：第一卷》。Athena scientific，2012 年，第 1 卷。[7] H. La、A. Potschka、J. Schl¨oder 和 H. Bock，《不确定过程控制中的双重控制与信息增益》，IFAC-PapersOnLine，第 49 卷，第 7 期，第 139–144 页，2016 年。[8] M. A. Sehr 和 R. R. Bitmead，《粒子模型预测控制：可处理的随机非线性输出反馈模型预测控制》，IFAC- PapersOnLine，第 50 卷，第 1 期，第 15 361–15 366 页，2017 年。[9] N. Kantas、J. Maciejowski 和 A. Lecchini-Visintini，《用于模型预测控制的序贯蒙特卡罗》，载于《非线性模型预测控制》。Springer，2009 年，第 263–273 页。[10] J. Yan 和 R. R. Bitmead，《将状态估计纳入模型预测控制及其在网络流量控制中的应用》，Automatica，第 41 卷，第 4 期，第 595–604 页，2005 年。[11]M. Cannon、Q. Cheng、B. Kouvaritakis 和 S. V. Rakovi´c，《带状态估计的随机管模型预测控制》，Automatica，第 48 卷，第 3 期，第 536–541 页，2012 年。[12]T. A. N. Heirung、J. A. Paulson、J. O'Leary 和 A. Mesbah，《随机模型预测控制——它如何工作？》，Computers & Chemical Engineering，第 114 卷，第 158–170 页，2018 年。[13]A. Mesbah，《随机模型预测控制：概述与未来研究展望》，IEEE Control Systems Magazine，第 36 卷，第 6 期，第 30–44 页，2016 年。[14] C. Striebel，《随机系统最优控制中的充分统计量》，Journal of Mathematical Analysis and Applications，第 12 卷，第 3 期，第 576–592 页，1965 年。[15]H. Schl¨uter 和 F. Allg¨ower，《使用初始状态优化的随机模型预测控制》，arXiv 预印本 arXiv:2203 01844，2022 年。[16] D. Q. Mayne、M. M. Seron 和 S. Rakovi´c，《有界扰动下约束线性系统的鲁棒模型预测控制》，Automatica，第 41 卷，第 2 期，第 219–224 页，2005 年。[17] S. Resnick，《概率之路》。Springer，2019 年。

[18] A. Doucet、S. Godsill 和 C. Andrieu，《用于贝叶斯滤波的序贯蒙特卡洛采样方法》，*Statistics and computing*，第 10 卷，第 3 期，第 197–208 页，2000 年。

[19]J. Luedtke 和 S. Ahmed，“用于概率约束优化的样本近似方法”，*SIAM* 优化期刊，第 19 卷，第 2 期，第 674–699 页，2008 年。

[20] W. Hoeffding，“有界随机变量之和的概率不等式”，载于 *Wassily Hoeffding* 文集。Springer，1994 年，第 409–426 页。

[21] B. D. Anderson 和 J. B. Moore，最优控制：线性二次方法。Prentice-Hall，1989 年。

[22] M. Farina、L. Giulioni、L. Magni 和 R. Scattolini，“模型预测控制的概率方法”，载于第*52*届 *IEEE* 决策与控制会议。IEEE，2013 年，第 7734–7739 页。

[23] A. Bemporad、M. Morari、V. Dua 和 E. Pitsikopoulos，“受约束系统的显式线性二次调节器”，*Automatica*，第 38 卷，第 1 期，第 3–20 页，2002 年。

[24]M. Cannon、B. Kouvaritakis、S. V. Rakovi´c 和 Q. Cheng，“具有概率约束的模型预测控制中的随机管束”，*IEEE Transactionson Automatic Control*，第 56 卷，第 1 期，第 194–200 页，2010 年。

[25] M. Lorenzen、F. Dabbene、R. Tempo 和 F. Allg¨ower，“随机模型预测控制中的约束收紧与稳定性”，*IEEE Transactions on Automatic Control*，第 62 卷，第 7 期，第 3165–3177 页，2016 年。[26] D. Bertsekas，“动态规划中离散化过程的收敛”，*IEEE Transactions on Automatic Control*，第 20 卷，第 3 期，第 415–419 页，1975 年。[27] J. Rust，“利用随机化打破维数灾难”，*Econometrica*：计量经济学会期刊，第 487–516 页，1997 年。

A 附录

推导 $\mathcal{A}_1$ 和 $\mathcal{A}_2$

根据式 (26)，将状态表示为其均值和误差，并代入式 (23)

$$
\begin{aligned}
J_c(x_0') = \mathbb{E}_{x_0''}\mathbb{E}_{W'}\mathbb{E}_{W''}\Bigg(\sum_{k=0}^{N-1} \Big[& tr(Q\hat{x}_k''(\hat{x}_k'')^T) \\
& + tr(K^T RK\hat{x}_k'(\hat{x}_k')^T) + tr(Q\tilde{x}_k''(\tilde{x}_k'')^T) + \\
& + tr(K^T RK\tilde{x}_k'(\tilde{x}_k')^T)\Big] + tr(Q_N\hat{x}_N''(\hat{x}_N'')^T) \\
& + tr(Q_N\tilde{x}_N''(\tilde{x}_N'')^T)\Bigg),
\end{aligned} \tag{A.1}
$$

由于状态误差的均值为零，交叉项（均值和误差）在取期望时最终被忽略。

误差协方差不是 x_0' 的函数，因此可以在不改变极小值点的情况下，用某个常数 C 替换式 (A.1) 中的代价。因此，除一个加法常数外

$$
\begin{aligned}
J_c(x_0') = \sum_{k=0}^{N-1} \Big[& (\hat{x}_k'')^T Q\hat{x}_k'' + (\hat{x}_k')^T K^T RK\hat{x}_k' \\
& + (\hat{x}_N'')^T Q_N\hat{x}_N''\Big] + C,
\end{aligned} \tag{A.2}
$$

其中，迹被还原为其二次型。将式 (26) 中的 $\hat{}x_k'$ 和 $\hat{}x_k''$ 代入 (A.2)，得到

$$
\begin{aligned}
J_c(x_0') = \sum_{k=0}^{N-1} \Big[& (\hat{x}_0'')^T (F^k)^T QF^k\hat{x}_0'' \\
& + 2(\hat{x}_0'')^T (F^k)^T Q\Psi_{k-1}x_0' \\
& + (x_0')^T \Psi_{k-1}^T Q\Psi_{k-1}x_0' \\
& + (x_0')^T (F_K^k)^T K^T RKF_K^k x_0' \\
& + (\hat{x}_0'')^T (F^N)^T QF^N\hat{x}_0'' \\
& + 2(\hat{x}_0'')^T (F^N)^T Q\Psi_{N-1}x_0' \\
& + (x_0')^T \Psi_{N-1}^T Q\Psi_{N-1}x_0'\Big] + C.
\end{aligned} \tag{A.3}
$$

所有相对于 x_0' 为常数的项都可以加到 C 上，使其成为 C_1,

$$
\begin{aligned}
J_c(x_0') = \sum_{k=0}^{N-1} \Big[& 2(\hat{x}_0'')^T (F^k)^T Q\Psi_{k-1}x_0' \\
& + (x_0')^T \Psi_{k-1}^T Q\Psi_{k-1}x_0' \\
& + (x_0')^T (F_K^k)^T K^T RKF_K^k x_0' \\
& + 2(\hat{x}_0'')^T (F^N)^T Q_N\Psi_{N-1}x_0' \\
& + (x_0')^T \Psi_{N-1}^T Q_N\Psi_{N-1}x_0'\Big] + C_1,
\end{aligned} \tag{A.4}
$$

或者写成紧凑形式，忽略加法常数

$$
J_c(x_0') = (x_0')^T \mathcal{A}_1 x_0' + (\hat{x}_0'')^T \mathcal{A}_2 x_0', \tag{A.5}
$$

其中

$$
\mathcal{A}_2 = \sum_{k=0}^{N-1} \Big[2(F^k)^T Q\Psi_{k-1}\Big] + 2(F^N)^T Q_N\Psi_{N-1}, \tag{A.6}
$$

$$
\begin{aligned}
\mathcal{A}_1 = \sum_{k=0}^{N-1} & \Big[\Psi_{k-1}^T Q\Psi_{k-1} + (F_K^k)^T K^T RKF_K^k\Big] + \\
& \Psi_{N-1}^T Q_N\Psi_{N-1}.
\end{aligned} \tag{A.7}
$$

计算 κ_2

令 V为无限时域折扣成本问题的价值函数，即最优成本；其阶段成本为 $\ell_k(x_k,\ u_k) =$

$x_k^T x_k + u_k^2$，且折扣因子为 $\gamma = 0.9$ [6]。我们将 κ_2 定义为相应的最优控制。也就是说，动态规划方程为

$$V(x_k) = x_k^T x_k + \min_{u_k \in \bar{\mathbb{U}}(x_k)} \left\{ u_k^2 + \gamma \mathbb{E}_{w_k} V(f(x_k, u_k, w_k)) \right\}, \tag{A.8}$$

其中，$\bar{\mathbb{U}}(x_k)$ 是状态相关输入约束集，定义为 $\bar{\mathbb{U}}(x_k) = \{u \in \mathbb{U} \mid \mathbb{P}(f(x_k, u_k, w_k) \in \mathcal{L}) < \epsilon\}$。

控制器 κ_2 可以通过在有限网格化状态空间和输入空间上的马尔可夫决策过程执行值迭代来计算；当这些空间的网格尺寸足够大时，κ_2 相对于原始的无限状态空间问题 [26]，该控制器是最优的。

按照 [27],中的动态规划方法，在状态空间中的 $[-10, 10] \times [-5, 15]$ 上均匀生成包含 4000 个点的随机化网格，并在 $[-3, 3]$ 控制空间 $\mathbb{U}$ 中生成包含 50 个点的网格。在 $\bar{\mathbb{U}}(x)$ 的定义中，连续概率被替换为网格中位于 $\mathcal{L}$ 内部各点上的离散概率，并将 (A.8) 中的期望替换为网格上的样本均值。在该网格上执行值迭代，直至收敛，该收敛性是有保证的。所得的最优控制

$$\kappa_2(x_k) = \arg\min_{u_k \in \bar{\mathbb{U}}(x_k)} \left\{ u_k^2 + \gamma \mathbb{E}_{w_k} V(f(x_k, u_k, w_k)) \right\}, \tag{A.9}$$

可通过在网格上用样本均值替代期望来计算。将网格上的点 ζ 纳入定义 $\mathcal{L}$ 的网格中，前提是 $\kappa(\zeta)$ 不可行。

计算出 κ_2 后，将其拟合为状态空间区域 $[-10, 10] \times [-15, 5]$ 上的分段线性曲面，然后通过二维线性插值获得 $\kappa_2(x)$。