АННОТАЦИЯ

Данная работа посвящена разработке и исследованию метода управления на основе Model Predictive Path Integral (MPPI), обеспечивающего точное выполнение линейных по ускорениям ограничений (Пфаффовых ограничений) вида $A(q, v, t)\,\dot{v} = b(q, v, t)$, где $q \in \mathbb{R}^n$ — обобщённые координаты, v — обобщённые скорости, $A \in \mathbb{R}^{m\times n}$ — матрица ограничений, $b \in \mathbb{R}^m$ — вектор правой части.

Предлагаемый подход (Gauss-MPPI) основан на интеграции принципа наименьшего принуждения Гаусса в процедуру семплирования траекторий MPPI, что позволяет гарантировать выполнение ограничений на каждом шаге интегрирования динамики системы. Для проекции свободного ускорения на множество допустимых ускорений используется аналитическое решение — уравнение Удвадия–Калабы.

Работа включает обзор существующих методов обработки ограничений в MPPI, теоретическое обоснование предложенного подхода, программную реализацию алгоритма с поддержкой параллельных вычислений, а также экспериментальную проверку на задачах робототехники: неголономный мобильный робот, робот-манипулятор с ограничениями в операционном пространстве и система с контактными взаимодействиями.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Задача управления механическими системами в условиях ограничений является одной из центральных в современной робототехнике. Многие практически важные системы — мобильные роботы, манипуляторы, шагающие роботы — подчиняются кинематическим и динамическим ограничениям, которые должны выполняться точно на каждом шаге управления.

Model Predictive Path Integral (MPPI) — метод модельно-предиктивного управления, основанный на стохастической выборке траекторий. MPPI не требует вычисления градиентов функции стоимости, работает с невыпуклыми и негладкими целевыми функциями и эффективно параллелизуется на GPU. Однако исходная формулировка MPPI не предусматривает явной обработки ограничений.

Существующие подходы к обработке ограничений в MPPI — штрафные методы, барьерные функции, вероятностные фильтры — ориентированы преимущественно на ограничения типа неравенств (безопасность, границы управления). Ограничения равенств на уровне ускорений (Пфаффовы ограничения) в контексте MPPI практически не исследованы.

Целью данной работы является разработка и исследование метода управления Gauss-MPPI, обеспечивающего точное выполнение линейных по ускорениям ограничений вида:

$$A(q, v, t)\, \dot{v} = b(q, v, t). \tag{1}$$

Предлагаемый подход основан на интеграции принципа наименьшего принуждения Гаусса в процедуру семплирования траекторий MPPI, что позволяет гарантировать выполнение ограничений на каждом шаге интегрирования динамики.

Для достижения поставленной цели необходимо решить следующие **задачи**:

1. Провести анализ существующих методов обработки ограничений в MPPI, выявить их преимущества и недостатки.

2. Разработать теоретические основы метода Gauss-MPPI: сформулировать задачу проекции свободного ускорения на множество допустимых ускорений как задачу квадратичного программирования.

3. Исследовать применимость принципа наименьшего принуждения Гаусса для стохастических дифференциальных уравнений.

4. Разработать алгоритм Gauss-MPPI и его программную реализацию с поддержкой параллельных вычислений на GPU.

5. Провести экспериментальную проверку на модельных задачах робототехники.

6. Выполнить сравнительный анализ с классическим MPPI, CBF-MPPI и штрафными методами.

1. ЛИТЕРАТУРНЫЙ ОБЗОР

1.1. Теоретические основы Path Integral Control

Теория управления на основе интегралов по траекториям берёт начало в работах Каппена [1], который показал связь между стохастическим оптимальным управлением и статистической механикой. Для определённого класса задач (линейная зависимость динамики от управления, квадратичная стоимость управления) уравнение Гамильтона–Якоби–Беллмана преобразуется в линейное уравнение, решение которого выражается через интеграл по траекториям:

$$u^*(x,t) = \mathbb{E}\left[u(x,t) \cdot \exp\left(-\frac{1}{\lambda}S(\tau)\right)\right], \tag{1.1}$$

где $S(\tau)$ — стоимость траектории τ, λ — параметр температуры.

Теодору и др. [2] обобщили этот подход на более широкий класс систем и предложили алгоритм PI^2 (Policy Improvement with Path Integrals) для обучения с подкреплением.

1.2. Model Predictive Path Integral Control

Метод MPPI был предложен Уильямсом и др. [3] как подход к модельно-предиктивному управлению, основанный на стохастической выборке траекторий. В отличие от градиентных методов, MPPI генерирует множество случайных траекторий и вычисляет оптимальное управление путём взвешенного усреднения:

$$u^* = \sum_{k=1}^{K} w_k \cdot u_k, \quad w_k = \frac{\exp(-\frac{1}{\lambda}S_k)}{\sum_{j=1}^{K}\exp(-\frac{1}{\lambda}S_j)}, \tag{1.2}$$

где K — число траекторий, S_k — стоимость k-й траектории.

Ключевые преимущества MPPI:

- Не требует вычисления градиентов функции стоимости;
- Работает с невыпуклыми и негладкими функциями стоимости;
- Эффективно параллелизуется на GPU [4].

MPPI успешно применяется в автономном вождении [4], управлении квадрокоптерами и манипуляторами.

1.2.1. Постановка задачи и алгоритм MPPI

MPPI решает задачу стохастического оптимального управления для систем вида:

$$\begin{aligned} dx &= f(x,u)dt + B(x)dw \\ x(t_0) &= x_0 \end{aligned} \tag{1.3}$$

где $x \in \mathbb{R}^n$ — вектор состояния, $u \in \mathbb{R}^m$ — вектор управления, f — функция динамики, B — матрица диффузии, dw — Винеровский процесс.

Цель — минимизация функционала стоимости на горизонте T:

$$J(u) = \phi(x_T) + \int_{t_0}^{t_0+T} \mathcal{L}(x_t, u_t)\, dt, \tag{1.4}$$

где $\phi(x_T)$ — терминальная стоимость, $\mathcal{L}(x,u)$ — мгновенная стоимость.

Параметр $\lambda > 0$ регулирует «температуру» распределения: при $\lambda \to 0$ алгоритм выбирает траекторию с минимальной стоимостью, при $\lambda \to \infty$ — равномерное усреднение.

1.3. Обработка ограничений в MPPI

Исходная формулировка MPPI не предусматривает явной обработки ограничений. Существующие подходы классифицируются следующим образом.

1. Штрафные методы. Наиболее простой подход — добавление штрафов за нарушение ограничений в функцию стоимости [3]. Недостатки: не гарантируют выполнение ограничений, требуют настройки весовых коэффициентов.

2. Методы на основе барьерных функций. Shield-MPPI [5] использует Control Barrier Functions (CBF) для фильтрации небезопасных управлений. GS-MPPI [6] применяет композитные CBF для систем с множественными ограничениями. Эти методы гарантируют выполнение ограничений типа неравенств, но увеличивают вычислительную сложность.

3. Вероятностные методы. SC-MPPI [7] встраивает фильтр безопасности в процесс выборки траекторий. BC-MPPI [8] присваивает вероятность допустимости каждой траектории и корректирует веса соответственно.

4. Проекционные методы. π-MPPI [9] использует проекцию для обеспечения ограничений на величину и производные управления.

1.4. Принцип наименьшего принуждения Гаусса

Принцип наименьшего принуждения Гаусса утверждает: из всех кинематически возможных движений системы реализуется то, для которого величина принуждения минимальна. Математически принцип формулируется как задача квадратичного программирования:

$$\begin{aligned} \min_{\dot v} \quad & [\dot v - a]^T M [\dot v - a] \\ \text{s.t.} \quad & A\dot v = b, \\ & a = M^{-1}Q, \end{aligned} \tag{1.5}$$

где M — матрица масс, Q — вектор обобщённых сил, $a = M^{-1}Q$ — ускорение системы без ограничений (свободное ускорение).

1.5. Уравнение Удвадия–Калабы

Удвадия и Калаба [10, 11] разработали универсальный метод получения уравнений движения для систем с ограничениями вида $A(q, v, t)\dot v = b(q, v, t)$. Решение задачи минимизации принуждения имеет аналитический вид:

$$\dot v = \dot v_{\text{free}} + M^{-1/2}(AM^{-1/2})^{+}(b - A\dot v_{\text{free}}), \tag{1.6}$$

где $(\cdot)^+$ — псевдообратная матрица Мура–Пенроуза.

Уравнение применимо к:

- Голономным ограничениям: $\phi(q, t) = 0$;
- Неголономным ограничениям: $\phi(q, v, t) = 0$;
- Ограничениям в операционном пространстве: $J(q)\dot v + \dot J v = \ddot x_d$.

1.6. Выводы из обзора

Анализ литературы позволяет сделать следующие выводы:

1. Существующие методы обработки ограничений в MPPI ориентированы преимущественно на ограничения типа неравенств (безопасность, границы управления).

2. Ограничения равенств на уровне ускорений (Пфаффовы ограничения) в контексте MPPI практически не исследованы.

3. Принцип наименьшего принуждения Гаусса и уравнение Удвадия–Калабы предоставляют аналитическое решение для проекции на множество допустимых ускорений.

4. Связь между MPPI и принципом Гаусса не исследована.

2. МЕТОДОЛОГИЯ

2.1. Численное интегрирование стохастических дифференциальных уравнений

Ключевым этапом алгоритма MPPI является моделирование траекторий системы, описываемой стохастическим дифференциальным уравнением (СДУ):

$$dx = f(x, u)\, dt + B(x)\, dW, \tag{2.1}$$

где dW — приращение Винеровского процесса.

Качество численных методов для СДУ характеризуется двумя типами сходимости:

- **Сильная сходимость** (порядок γ) — сходимость по траекториям:

$$\mathbb{E}\left[|x_N - x(T)|^2\right]^{1/2} \leq C \cdot \Delta t^{\gamma}. \tag{2.2}$$

  Важна, когда требуется точное воспроизведение отдельных реализаций процесса.

- **Слабая сходимость** (порядок β) — сходимость по распределениям:

$$|\mathbb{E}[g(x_N)] - \mathbb{E}[g(x(T))]| \leq C \cdot \Delta t^{\beta} \tag{2.3}$$

  для гладких функций g. Важна для вычисления математических ожиданий (как в MPPI).

2.1.1. Метод Эйлера–Маруямы

Простейшая схема численного интегрирования СДУ:

$$x_{k+1} = x_k + f(x_k, u_k)\, \Delta t + B(x_k)\, \Delta W_k, \tag{2.4}$$

где $\Delta W_k \sim \mathcal{N}(0, \Delta t \cdot I)$ — дискретное приращение Винеровского процесса. Метод имеет сильный порядок сходимости 0.5 и слабый порядок 1.0.

2.1.2. Метод Мильштейна

Схема более высокого порядка, учитывающая поправку Ито:

$$x_{k+1} = x_k + f(x_k, u_k)\,\Delta t + B(x_k)\,\Delta W_k + \frac{1}{2} B(x_k) \frac{\partial B}{\partial x}(x_k) \left[(\Delta W_k)^2 - \Delta t \right]. \quad (2.5)$$

Метод имеет сильный порядок сходимости 1.0, но требует вычисления производных матрицы диффузии.

2.1.3. Детерминированное приближение

В практических реализациях MPPI часто используется детерминированная модель с аддитивным шумом в управлении:

$$x_{k+1} = x_k + f(x_k, u_k + \epsilon_k)\,\Delta t, \quad \epsilon_k \sim \mathcal{N}(0, \Sigma), \quad (2.6)$$

что позволяет применять стандартные методы интегрирования ОДУ (Эйлера, Рунге–Кутты) и упрощает реализацию на GPU.

2.1.4. Симплектические интеграторы

Механические системы естественно описываются в гамильтоновой форме:

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}, \quad (2.7)$$

где q — обобщённые координаты, p — обобщённые импульсы, $H(q, p)$ — гамильтониан. Симплектические интеграторы (Штёрмера–Верле, RATTLE) сохраняют геометрическую структуру фазового потока, обеспечивая:

- Ограниченность ошибки энергии на бесконечном интервале;
- Качественно верное поведение траекторий;
- Долговременную устойчивость численного решения.

2.2. Алгоритм Gauss-MPPI

Предлагаемый метод Gauss-MPPI интегрирует принцип наименьшего принуждения Гаусса в процедуру семплирования траекторий MPPI. На каждом шаге интегрирования свободное ускорение проецируется на множество допустимых ускорений с помощью уравнения Удвадия–Калабы.

Для систем с Пфаффовыми ограничениями $A(q, v, t)\dot{v} = b(q, v, t)$ проекция свободного ускорения $\dot{v}_{\text{free}} = M^{-1}(Q + Bu)$ на множество допустимых ускорений имеет вид:

$$\dot{v} = \dot{v}_{\text{free}} + M^{-1/2}(AM^{-1/2})^{+}(b - A\dot{v}_{\text{free}}). \tag{2.8}$$

Данная проекция выполняется для каждой семплированной траектории на каждом временном шаге, что гарантирует выполнение ограничений во всех точках горизонта планирования.

2.3. Применимость принципа Гаусса к СДУ

Классический принцип Гаусса сформулирован для детерминированных систем. Применение принципа к стохастическим системам требует дополнительного исследования:

- Открытый вопрос: как корректно определить «свободное ускорение» при наличии стохастического члена $B(x)\,dW$;
- Рассматриваются подходы: проекция на каждом шаге интегрирования, модификация функции принуждения с учётом шума.

2.4. Вычислительная сложность

3. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ И ЭКСПЕРИМЕНТЫ

3.1. Реализация классического MPPI

Выполнена программная реализация классического алгоритма MPPI в среде симуляции MuJoCo. Реализация включает:

- Параллельную выборку траекторий на CPU;
- Интеграцию с физическим движком MuJoCo для моделирования динамики;
- Тестирование на стандартных задачах управления (CartPole, Pendulum и др.).

3.2. Интегрирование СДУ

Проведены эксперименты по численному интегрированию физических систем, представленных в виде стохастических дифференциальных уравнений:

- Реализованы методы Эйлера–Маруямы и детерминированное приближение;
- Исследовано влияние шага интегрирования на точность и устойчивость;
- Выявлены особенности работы с механическими системами (проверка сохранения энергии).

3.3. Реализация алгоритма Gauss-MPPI

3.4. Экспериментальные сценарии

Для валидации предложенного метода выбраны следующие экспериментальные сценарии.

3.4.1. Неголономный мобильный робот

Колёсный робот с ограничением отсутствия проскальзывания:

$$\dot{x}\sin\theta - \dot{y}\cos\theta = 0. \tag{3.1}$$

3.4.2. Робот-манипулятор с ограничениями в операционном пространстве

Задача следования по траектории с ограничениями на конечный эффектор:

$$J(q)\dot{v} + \dot{J}v = \ddot{x}_d. \tag{3.2}$$

3.4.3. Система с контактными взаимодействиями

4. РЕЗУЛЬТАТЫ И СРАВНИТЕЛЬНЫЙ АНАЛИЗ

4.1. Метрики оценки

Для сравнительного анализа предлагаемого метода используются следующие метрики:

- Точность выполнения ограничений (максимальная и средняя невязка);
- Качество управления (значение функции стоимости);
- Время вычислений на один шаг управления.

4.2. Сравнение с классическим MPPI

4.3. Сравнение с альтернативными методами

4.4. Анализ чувствительности к параметрам

4.5. Обсуждение результатов

ЗАКЛЮЧЕНИЕ

В данной работе рассмотрена задача управления механическими системами с линейными по ускорениям ограничениями методом MPPI. Предложен метод Gauss-MPPI, интегрирующий принцип наименьшего принуждения Гаусса в процедуру семплирования траекторий.

Основные результаты работы:

1. Проведён анализ существующих методов обработки ограничений в MPPI, выявлен пробел в обработке ограничений равенств на уровне ускорений.
2. Предложен метод Gauss-MPPI, использующий уравнение Удвадия–Калабы для проекции свободного ускорения на множество допустимых ускорений.
3. Выполнена программная реализация классического MPPI в среде MuJoCo.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Kappen H. J. Path integrals and symmetry breaking for optimal control theory // Journal of Statistical Mechanics: Theory and Experiment. 2005. Vol. 2005. No. 11. P. P11011.
2. Theodorou E., Buchli J., Schaal S. A generalized path integral control approach to reinforcement learning // Journal of Machine Learning Research. Vol. 11. 2010. P. 3137–3181.
3. Information theoretic MPC for model-based reinforcement learning / G. Williams, N. Wagener, B. Goldfain et al. // IEEE International Conference on Robotics and Automation (ICRA). 2017. P. 1714–1721.
4. Aggressive driving with model predictive path integral control / G. Williams, P. Drews, B. Goldfain et al. // IEEE International Conference on Robotics and Automation (ICRA). 2017. P. 1433–1440.
5. Ji J., Chen Z., Li Y. Shield-MPPI: Robust and safe model predictive path integral control via control barrier functions // IEEE Robotics and Automation Letters. 2024. To appear.
6. Yin S. et al. GS-MPPI: Generalized smooth MPPI with composite control barrier functions // arXiv preprint. 2024.
7. Zhu W. et al. SC-MPPI: Safe and constrained model predictive path integral control // arXiv preprint. 2023.
8. Nakamura M. et al. BC-MPPI: Barrier-certified model predictive path integral control // arXiv preprint. 2024.
9. Kim C. et al. π-MPPI: Projection-based model predictive path integral control // arXiv preprint. 2024.
10. Udwadia F. E., Kalaba R. E. A new perspective on constrained motion // Proceedings of the Royal Society of London. Series A. 1992. Vol. 439. No. 1906. P. 407–410.
11. Udwadia F. E., Kalaba R. E. Analytical Dynamics: A New Approach. Cambridge University Press. 2002. ISBN: 978-0521048330.

ПРИЛОЖЕНИЕ А