Cross-Entropy Probabilistic Motion Planning

Marin Kobilarov
California Institute of Technology
marin@cds.caltech.edu

***Abstract*— This work, initially presented in [1], considers motion planning for nonlinear robotic systems operating in constrained environments. Motivated by recent developments in sampling-based motion planning and Monte Carlo optimization we propose a general randomized path planning method based on sampling in the space of trajectories. The idea is to construct a probability distribution over the set of feasible paths and to perform the search for an optimal trajectory through importance sampling. At the core of the approach lies the cross-entropy method for estimation of rare-event probabilities. The algorithm recursively approximates the optimal sampling distribution which guides the set of sampled trajectories towards regions of progressively lower cost until converging to a delta distribution at the optimum. Our main goal is to provide a framework for consistent adaptive sampling correlating the spatial structure of trajectories and their computed costs. We also illustrate how the method can be useful for constructing roadmap vertex sampling densities which are approximately optimal with respect to the resulting trajectory cost.**

I. INTRODUCTION

Consider a robotic vehicle with state $x \in \mathcal{X}$ controlled using actuator inputs $u \in \mathcal{U}$, where $\mathcal{X}$ is the state space and $\mathcal{U}$ denotes the set of controls. The vehicle dynamics is described by the function $f : \mathcal{X} \times \mathcal{U} \to T\mathcal{X}$ defined by

$$\dot{x}(t) = f(x(t), u(t)), \tag{1}$$

which is used to evolve the vehicle state forward in time. In addition, the vehicle is subject to constraints arising from actuator bounds and obstacles in the environment. These constraints are expressed through the vector of inequalities

$$F(x(t)) \geq 0, \tag{2}$$

for all $t \in [0, t_f]$, where $t_f > 0$ is the final time of the trajectory. The goal is to compute the optimal controls u^* and time t_f^* driving the system from its initial state $x_0 \in \mathcal{X}$ to a given goal region $\mathcal{X}_f \subset \mathcal{X}$, i.e.

$$\begin{aligned} (u^*, t_f^*) = \operatorname*{argmin}_{u, t_f} & \int_0^{t_f} C(u(t), x(t))\mathrm{dt}, \\ \text{subject to } & \dot{x}(t) = f(x(t), u(t)), \\ & F(x(t)) \geq 0,\ x(0) = x_0,\ x(t_f) \in \mathcal{X}_f \end{aligned} \tag{3}$$

for all $t \in [0, t_f]$ and where $C : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ is a given cost function.

II. PROBLEM FORMULATION

A trajectory recording the controls and states over the time interval $[0, t_f]$ is denoted by the function $\pi : [0, t_f] \to \mathcal{U} \times \mathcal{X}$, i.e. $\pi(t) = (u(t), x(t))$ for all $t \in [0, t_f]$. The space of all trajectories originating at point x_0 and satisfying the dynamics is denoted $\mathcal{P}$. Consider a finite-dimensional parametrization of trajectories in terms of vectors $z \in \mathcal{Z}$ where $\mathcal{Z} \subset \mathbb{R}^{n_z}$ is the *parameter space*. Assume that the parametrization is given by the function $\varphi : \mathcal{Z} \to \mathcal{P}$ according to $\pi = \varphi(z) \equiv \varphi_z$. The (control, state) tuples along a trajectory parametrized by z are written as $\pi(t) = \varphi_z(t)$. Define the *cost function* $J : \mathcal{Z} \to \mathbb{R}$ according to

$$J(z) = \int_0^{t_f} C(\varphi_z(t))dt.$$

Problem (3) can now be equivalently restated as finding the optimal t_f^* and $(x^*, u^*) = \varphi(z^*)$ such that

$$z^* = \operatorname*{arg\,min}_{z \in \mathcal{Z}_{\text{con}}} J(z), \tag{4}$$

where $\mathcal{Z}_{\text{con}} \subset \mathcal{Z}$ satisfies the constraints (2) and the final boundary conditions.

III. CROSS-ENTROPY OPTIMIZATION

A. Estimation of Rare-Event Probabilities

Our approach is based on importance sampling in the space of trajectories $\mathcal{Z}$. Assume that Z has a pdf $p(\cdot; \bar{v})$ belonging to some parametric family $\{p(\cdot; v), v \in \mathcal{V}\}$ where $\bar{v}$ is the true or nominal parameter. It is natural to consider an importance density q from the same family. Its optimal parameter v is found through the optimization $\min_v \mathfrak{D}(q^*, p(\cdot, v))$ where $\mathfrak{D}$ is the Kullback-Leibler (KL) or cross-entropy (CE) distance. Consider the estimation of the probability ℓ that a trajectory $z \in \mathcal{Z}_{con}$ sampled from $p(\cdot; \bar{v})$ has a cost $J(z)$ smaller than a given constant γ, i.e.

$$\ell = \mathbb{P}_{\bar{v}}(J(Z) \leq \gamma) = \mathbb{E}_{\bar{v}}[I_{\{J(Z) \leq \gamma\}}], \tag{5}$$

where $I_{\{\cdot\}}$ is the indicator function. It is approximated by

$$\hat{\ell} = \frac{1}{N} \sum_{i=1}^{N} I_{\{J(Z_i) \leq \gamma\}} \frac{p(Z_i; \bar{v})}{p(Z_i; v)},$$

where $Z_1, ..., Z_N$ are i.i.d. samples from $p(\cdot, v)$. In order to determine the optimal v the CE approach [2] leads to

$$\hat{v}^* = \operatorname*{argmax}_{v \in \mathcal{V}} \frac{1}{N} \sum_{i=1}^{N} I_{\{J(Z_i) \leq \gamma\}} \ln p(Z_i, v), \tag{6}$$

where $Z_1, ..., Z_N$ are i.i.d. samples from $p(\cdot, \bar{v})$. The problem is that when $\{J(Z) \leq \gamma\}$ is a rare event, $\hat{\ell}$ will be incorrectly estimated as zero!

The idea behind the CE method is to employ a multilevel approach using a sequence of parameters $\{v_j\}_{j \geq 0}$ and levels $\{\gamma_j\}_{j \geq 1}$. At the end the sequence converges to the optimal v^* which then can be used to estimate the integral $\hat{\ell}$ correctly. The procedure starts by drawing N samples $Z_1, ..., Z_N$ using

an initial parameter v_0, for instance $v_0 = \bar{v}$. Let ϱ be a small number, e.g. $10^{-2} \leq \varrho \leq 10^{-1}$. The value γ_1 is set to the ϱ–th quantile of $H(Z)$, i.e. γ_1 is the largest number for which $\mathbb{P}_{v_0}(H(Z) \leq \gamma_1) = \varrho$. The level γ_1 can be computed approximately by sorting the costs of the samples $J(Z_1), ..., J(Z_N)$ in an increasing order, say $J_1 \leq ... \leq J_N$, and setting $\hat{\gamma}_1 = J_{\lceil \varrho N \rceil}$. The optimal parameter v_1 for level $\hat{\gamma}_1$ is then estimated using (6) by replacing γ with $\hat{\gamma}_1$. In summary, each iteration of the algorithm perform two steps, starting with v_0,

1) *Sampling and updating of* γ_j: Sample $Z_1, ..., Z_n$ from $p(\cdot, \hat{v}_{i-1})$ and compute the ϱ-th quantile $\hat{\gamma}_t$.
2) *Adaptive Updating of* v_j: Compute $\hat{v}_j$ such that

$$\hat{v}_j = \underset{v \in \mathcal{V}}{\operatorname{argmin}} \frac{1}{|\mathcal{E}_j|} \sum_{Z_k \in \mathcal{E}_j} \ln p(Z_k; v), \qquad (7)$$

where $\mathcal{E}_j$ is the *elite* set, i.e. for which $J(Z_k) \leq \hat{\gamma}_j$.

IV. Motion Planning Algorithm

We now construct a general motion planning algorithm based on the cross-entropy method. We choose to represent the importance density using a Gaussian mixture model since it is a compact way to encode a rich set of trajectories across multiple homotopy classes.

The parameter space is $\mathcal{V} = (\mathbb{R}^{n_z} \times \mathbb{R}^{(n_z^2+n_z)/2})^K \times \mathbb{R}^K$ with elements $v = (\mu_1, \Sigma_1, ..., \mu_K, \Sigma_K, w_1, ..., w_K)$ corresponding to K mixture components with means μ_k, covariance matrices Σ_k and weights w_k. The number of mixture components K is chosen adaptively (see e.g. [3]). Even the simplest case with $K = 1$ is capable of solving complex multi-extremal problems globally. The complete algorithm is summarized below. [h]

Algorithm 4.1: CE Motion Planning

1) Choose initial trajectory samples $Z_1, ..., Z_N$ so that the set $\hat{\mathcal{P}} = \{\varphi_{Z_i}\}_{i=1}^N$ approximates (sparsely) the set of feasible trajectories over $\mathcal{X}$; Set $j = 0$ and $\hat{\gamma}_0 = \infty$
2) Update $\hat{v}_j$ using (7) (e.g. by EM) over the elite set $\mathcal{E}_j = \{Z_i \mid J(Z_i) \leq \hat{\gamma}_j\}$
3) Generate samples $Z_1, ..., Z_N$ from $p(\cdot, v_j)|_{\mathcal{Z}_{\text{con}}}$ and compute the ϱ–th quantile $\hat{\gamma}_{j+1} = J_{\lceil \varrho N \rceil}$
4) If a stopping criteria is met then finish, otherwise set $j = j + 1$ and goto step (2)

V. Links to Sampling-based Motion Planning

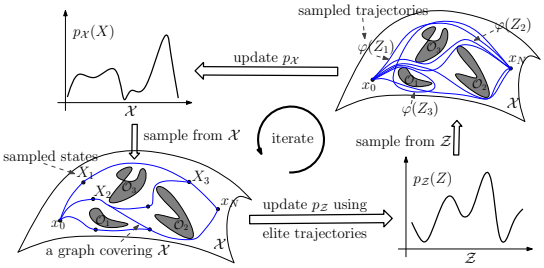
It is interesting to note that the CE approach effectively constructs a sequence of parametric models of sets of progressively "better" trajectories, i.e. with decreasing costs. Since a trajectory is a set of states this model can also be regarded as a density over the state space $\mathcal{X}$. A state $x \in \mathcal{X}$ has a high importance density, say $p_{\mathcal{X}}(x)$, if it belongs to a trajectory $\varphi_z \in \mathcal{P}$ with high density $p(z)$. Such a sampling density is defined as

$$p_{\mathcal{X}}(X; v) = \eta \cdot \max_{Z \in \mathcal{Z}_{\text{con}}} \{p(Z; v) \mid X \in \kappa(\varphi_Z)\}, \qquad (8)$$

where X is a random variable over $\mathcal{X}$ and $\eta > 0$ is a normalizing constant. For instance, the lower plots of Fig. 1 show the evolution of $p_{\mathcal{X}}$ restricted to the position space of a double-integrator vehicle example. The cost function in this example is $C(u, x) = \|v\| + \lambda \|u\|^2$, where v is velocity, u is acceleration input, and λ is a scaling factor.

Thus, *sampling-based motion planning (SMP)* can be combined with *cross-entropy importance sampling (IS)* to develop a consistent way for balancing between *exploration* and *exploitation*, respectively, dictated by the following properties:

- SMP: efficient exploration, but slow optimality,
- IS: fast convergence, but needs many feasible samples.



We conclude by proposing the following algorithm:

Algorithm: Cross-Entropy PRM

1) Sample states $\{X_i \sim p_{\mathcal{X}}\}_{i=1}^{N_{\mathcal{X}}}$
2) Connect $\{X_i\}_{i=1}^{N_{\mathcal{X}}}$ to graph G
3) Update $p_{\mathcal{Z}}$ using elite trajectory set from G
4) Sample trajectories $\{Z_i \sim p_{\mathcal{Z}}\}_{i=1}^{N_{\mathcal{Z}}}$
5) Update $p_{\mathcal{X}}$ using the set $\{\varphi(Z_i)\}_{i=1}^{N_{\mathcal{Z}}}$
6) Terminate, e.g. when $p_{\mathcal{Z}}$ is close to delta, otherwise goto (1).

References

[1] M. Kobilarov, "Cross-entropy randomized motion planning," in *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2011.

[2] R. Y. Rubenstein and D. P. Kroese, *Simulation and the Monte Carlo Method*. Wiley, 2008.

[3] M. A. F. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381–396, 2002.

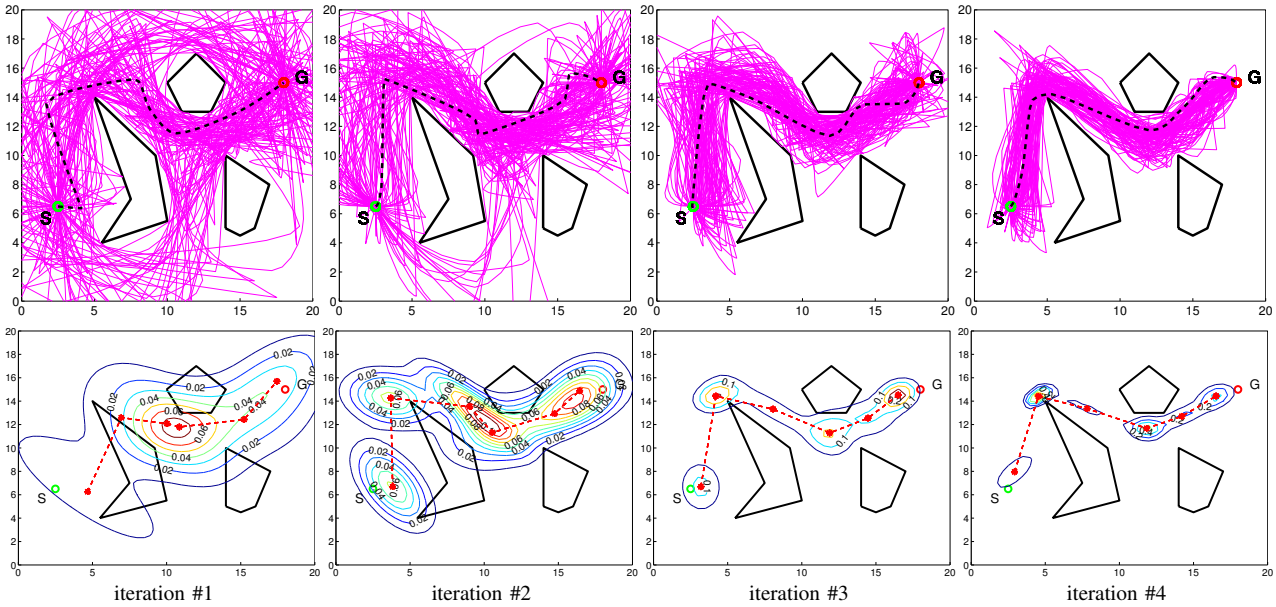


Fig. 1. The first four iterations of the CE algorithm 4.1, i.e. $j = 1, ..., 4$, applied to a double-integrator vehicle among obstacles. The upper plots show the sampled trajectories $\varphi(Z_1), ..., \varphi(Z_N)$ and the current optimal path $\varphi(z^*)$ (dashed). The lower plots visualize $p(\cdot, \hat{v}_j)$ as the level sets of another density over the (x, y)-position space. The density of each point corresponds to the density of the best trajectory passing through it. The algorithm starts from a nearly uniform state space exploration and converges towards a delta distribution at the global optimum.