

Monte Carlo Localization

Andreas Munk Jose Armenteros Ramiro Mata Bjarne Grossmann

Contents

1	Introduction	2
2	Problem statement	2
3	Derivation of Sequential Bayes filter	3
3.1	Bayesian Network of the Hidden Markov Model (HMM)	3
3.2	Grid-based approximation	4
3.3	Particle filters	5
3.4	Algorithmic implementation	6
4	Simulation	6
4.1	World model	6
4.2	Spacecraft model	7
4.3	Sensor model	7
4.3.1	Motion Sensor	7
4.3.2	Vision Sensor	8
5	Motion model	8
6	Observation model	9
7	Representations	10
7.1	Grid-based approximation	10
7.2	Particle-based approximation	11
7.3	Resampling Methods	11
7.3.1	Naive Sampling	11
7.3.2	Stratified Sampling	12
7.3.3	Systematic Sampling	12

1 Introduction

Monte Carlo localization is a method common in mobile robotics that enables a robot to estimate its location within a known environment using a recursive Bayesian estimation approach. The robot is equipped with different sensors to sense and observe the environment as it moves around the landscape. Using a given map of that landscape, it then uses the observations to sequentially infer its location within that map.

In Bayesian terms, the recursive Bayesian estimation (or Bayesian filter) is a general method to estimate an unknown probability density function over time by incorporating new observation through a processing and observation model. The probability density function represents the belief over the states of a dynamic system and can therefore be seen as a filtering process that eliminates less probable states and eventually estimates the true underlying state. The computation of the posterior distribution is often untractable due to the complexity of the observation model and the dynamically changing state, and is therefore often approximated by grid-based or particle-based Monte Carlo methods.

In this project, we developed a demonstration application that shows the mechanism behind the recursive Bayesian estimation for grid and particle-based Monte Carlo methods in a localization scenario. In this simulation, we send an autonomous robot (or spacecraft) to a remote area (or planet). The spacecraft knows a topological map of the landscape. The spacecraft is also equipped with a vision sensor observing the elevation of the terrain nearby, and a motion sensor measuring the motion of the spacecraft. Using these noisy observations, the task is to estimate our location with respect to the given topological map.

2 Problem statement

The goal of the Monte Carlo localization is to estimate the state of the spacecraft - in this particular case, we are interested only in the location (x, y) on the topological map \mathcal{M} with size (x_{max}, y_{max}) . The state space Ω can be represented by all possible states

$$\theta = (x, y) \in \Omega \quad \text{with } 0 < x < x_{max} \text{ and } 0 < y < y_{max}$$

Our goal is to find a good estimate $\hat{\theta}_t$ of the true state θ_t at time t . We can formulate the belief $Bel(\theta_t)$ of possible states at a time t given the independent observations of the terrain $z_{0:t}$ and motions $u_{0:t}$ (note that we do not explicitly state the conditioning on $u_{0:t}$ for readability) in a probabilistic manner as

$$Bel(\theta_t) = p(\theta_t | z_{0:t}) = \frac{p(z_t | \theta_t) p(\theta_t | z_{0:t-1})}{p(z_t | z_{0:t-1})} \quad (1)$$

with the normalization term

$$p(z_t | z_{0:t-1}) = \int p(z_t | \theta_t) p(\theta_t | z_{0:t-1}) d\theta_t \quad (2)$$

and the predictive distribution

$$p(\theta_t | z_{0:t-1}) = \int p(\theta_t | \theta_{t-1}) p(\theta_{t-1} | z_{0:t-1}) d\theta_{t-1}. \quad (3)$$

The system can be fully determined by defining the **observation model** $p(z_t|\theta_t) = p(z_t|\theta_t, \mathcal{M})$ to compute the likelihood of the observation z_t at location θ_t given the map model \mathcal{M} , and the **motion model** given by the transition distribution $p(\theta_t|\theta_{t-1}) = p(\theta_t|\theta_{t-1}, u_t)$ given the last estimated location θ_{t-1} and the observed motion u_t .

However, the integrals of the normalization term and the predictive distribution cannot be determined analytically and have to be solved using the Monte Carlo integration method. With the Monte Carlo representation of the posterior distribution, we are able to estimate the true state of the spacecraft with respect to any criterion. Typical estimates are the Minimum Mean Square Estimate (or the expected value)

$$\hat{\theta}_t^{\text{MMS}} = E[\theta_t|z_{0:t}] = \int \theta_t p(\theta_t|z_{0:t}) d\theta_t \quad (4)$$

or the Maximum A-Posteriori estimate (or the maximum mode)

$$\hat{\theta}_t^{\text{MAP}} = \arg \max_{\theta_t} p(\theta_t|z_{0:t}) \quad (5)$$

3 Derivation of Sequential Bayes filter

In this section, we will derive and formulate the *Recursive Bayesian Estimation* or Bayes Filter. The purpose is to lay the foundation for the application known as particle filtering. Specifically, we will use the *Sequential Importance Resampling* (SIR). Of those algorithms associated with SIR we utilize the special case known as *Bootstrap filter*. We generalize the notation for the derivation of SIR.

3.1 Bayesian Network of the Hidden Markov Model (HMM)

Using HMM and denoting a time-step as k , we will infer the probability of a current true state, denoted x_k , conditioned on all previous hidden and (including the current) observed states, which we denote $x_{1:k-1}$ and $z_{1:k}$ respectively. For HMM, we make the following assumption about any time sequence:

$$p(x_k|x_{1:k-1}) = p(x_k|x_{k-1}) \quad (6)$$

$$p(z_k|x_{1:k}) = p(z_k|x_k). \quad (7)$$

We further assume, that given the current state, the current observation will be independent of all previous observations,

$$p(z_{1:k}|x_k) = p(z_k|x_k)p(z_{1:k-1}|x_k). \quad (8)$$

For completion, we can then write the entire joint probability for a time sequence as follows,

$$p(x_{1:k}, z_{1:k}) = p(x_1) \prod_{i=2}^k p(z_i|x_i)p(x_i|x_{i-1}). \quad (9)$$

However, for predictive purposes, we wish to find the distribution of the current state x_k conditioned on all previous states. Since we assume any hidden state is independent of all previous hidden states, except the immediate previous one, we wish to infer,

$$\begin{aligned}
p(x_k|z_{1:k}) &= \frac{p(z_{1:k}|x_k)p(x_k)}{p(z_{1:k})} \\
&= \frac{p(z_k|x_k)p(z_{1:k-1}|x_k)p(x_k)}{p(z_{1:k})} \\
&= \frac{p(z_k|x_k)p(x_k|z_{1:k-1})p(z_{1:k-1})}{p(z_{1:k})} \\
&= \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}.
\end{aligned} \tag{10}$$

As we shall see later, we explicitly model $p(z_k|x_k)$, thus leaving us to reformulate $p(x_k|z_{1:k-1})$ as a marginalization,

$$\begin{aligned}
p(x_k|z_{1:k-1}) &= \int \frac{p(x_k, x_{k-1}, z_{1:k-1})}{p(z_{1:k})} dx_{k-1} \\
&= \int \frac{p(x_k|x_{k-1})p(z_{1:k-1}|x_{k-1})p(x_{k-1})}{p(z_{1:k})} dx_{k-1} \\
&= \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1}) dx_{k-1},
\end{aligned} \tag{11}$$

where we further utilized the following assumption

$$p(x_k, z_{1:k-1}|x_{k-1}) = p(x_k|x_{k-1})p(z_{1:k-1}|x_{k-1}). \tag{12}$$

3.2 Grid-based approximation

We then identify the previous equation as an average of $p(x_k|x_{k-1})$ over the conditional probability density of x_{k-1} . Thus, referring to Chap. 10 (*A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, Bayesian Data Analysis, Third Edition (Chapman & Hall/CRC Texts in Statistical Science) (Chapman and Hall/CRC), 3rd edn. (2014).*), we can estimate the aforementioned expected value through importance sampling, as sampling x_{k-1} in its current form is non-trivial. Hence, by denoting a proposal distribution as $\pi(\theta)$, where θ denotes all relevant parameters associated with the HMM model, we write,

$$p(x_k|z_{1:k-1}) \simeq \frac{1}{S} \sum_{s=1}^S p(x_k|x_{k-1}^s) \tilde{w}(\theta^s), \tag{13}$$

where θ^s denotes all relevant parameters but explicitly a sample x_{k-1}^s , and

$$\tilde{w}(\theta^s) = \frac{w(\theta^s)}{\sum_{s=1}^S w(\theta^s)}, \quad \text{with} \quad w(\theta^s) = \frac{p(x_{k-1}^s|z_{1:k-1})}{\pi(\theta^s)}. \tag{14}$$

It turns out that a proper choice of proposal distribution from which x_{k-1} is easily sampled, significantly simplifies the posterior estimation, as we will now show,

$$p(x_{k-1}|z_{1:k-1}) \propto p(z_\alpha|x_\alpha)p(x_\alpha|z_{\alpha-1}) = \sum_{s=1}^S p(z_\alpha|x_\alpha)p(x_\alpha|x_{\alpha-1}^s)\tilde{w}_\alpha(\theta^s), \quad (15)$$

where $\alpha = k - 1$ and any approximate equalities have been substituted with strict equalities for brevity. Using the above equation and dropping the parameter dependencies for the weights, we find,

$$w_k^s = \frac{p(x_\alpha^s|z_{1:\alpha})}{\pi(\theta^s)} \propto \frac{p(z_\alpha|x_\alpha^s) \sum_{l=1}^L p(x_\alpha^s|x_{\alpha-1}^l)\tilde{w}_\alpha^l}{\pi(\theta^s)}. \quad (16)$$

Clearly, weights at subsequent time-steps depend sequentially on immediate previous weights. However, it would perhaps seem computationally disadvantageous to sample S samples, for each we need to calculate probabilities through a sum of previous L samples. To make further simplifications, we turn to particle filtering.

3.3 Particle filters

Particle filters rely on the simple concept assigning each weight to a unique particle. We formulate this mathematically as the set of S samples,

$$\{x_k^s, \tilde{w}_k^s \mid s = 1, 2, \dots, S\}. \quad (17)$$

The fundamental idea is, that for each time step we keep our samples/particles, but update their weights. We formulate this as,

$$p(x_k^s|x_\alpha^l) = \begin{cases} p(x_k^s|x_\alpha^s) & s = l \\ 0 & \text{otherwise} \end{cases}. \quad (18)$$

Consequently, we still draw samples, but the net effect is that each sample is simply updated, and one can keep track of each sample/particle's weights development through a sequence. Armed with the particle formulation, we simplify the weights' update as follows:

$$w_k^s = \frac{p(x_\alpha^s|z_{1:\alpha})}{\pi(\theta^s)} \propto \frac{p(z_\alpha|x_\alpha^s)p(x_\alpha^s|x_{\alpha-1}^s)\tilde{w}_\alpha^s}{\pi(\theta^s)}. \quad (19)$$

By choosing $\pi(\theta^s) = p(x_\alpha^s|x_{\alpha-1}^s)$, we arrive at the bootstrap filter, where each weight is updated as:

$$w_k^s = p(z_{k-1}|x_{k-1}^s)\tilde{w}_{k-1}^s. \quad (20)$$

This formulation completely allows us to ignore the exact form of the original distribution $p(x_{k-1}|z_{1:k-1})$, making the particle an intuitive sampling algorithm. We should note, that the particle filters are highly prone to include weights that negatively influence the distribution. Thus, we keep track of the effective sample size (S_{eff}) - see Chap. 10 (*A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, Bayesian Data Analysis, Third Edition (Chapman & Hall/CRC Texts in Statistical Science) (Chapman and Hall/CRC), 3rd edn. (2014).*),

$$S_{eff} = \frac{1}{\sum_{s=1}^S (w(\tilde{\theta}^s))^2}. \quad (21)$$

3.4 Algorithmic implementation

We now provide algorithmic steps to calculate and update the particle weights. Assume $p(z_k|x_k)$, $p(z_k|x_k^s)$ and $p(x_k|x_{k-1})$ are given, then at each time-step,

- For $s = 1, 2, \dots, S$ draw x_k (by propagating x_{k-1} with the transition kernel),

$$x_k^s \sim p(x_k^s|x_{k-1}^s). \quad (22)$$

- For $s = 1, 2, \dots, S$ update weights w_k^s ,

$$w_k^s = p(z_k|x_k^s)\tilde{w}_{k-1}^s. \quad (23)$$

- For $s = 1, 2, \dots, S$ normalize the weights,

$$\tilde{w}_k^s = \frac{w_k^s}{\sum_{s=1}^S w_k^s} \quad (24)$$

- Compute the effective sample size

$$S_{eff} = \frac{1}{\sum_{s=1}^S (\tilde{w}(\theta^s))^2}. \quad (25)$$

- If $S_{eff} < S_{eff}^{threshold}$ perform resampling:
 - Choose a resampling method (we choose a systematic resampling strategy, as we will touch upon in the next sections).
 - Draw S new particles/samples from the sample population according to their weights \tilde{w}_k^s .
 - Reset all weights as $\tilde{w}_k^s = 1/S$.

4 Simulation

In order to demonstrate the Monte Carlo localization method, a model of a real scenario can be created as a setting for the robot localization simulation. Therefore, we need to model

- the environment given as a topological map
- the spacecraft which is able to navigate around the map
- the sensors used by the spacecraft to make observations

4.1 World model

The world model is generated automatically and is represented by an elevation map, i.e. a rasterized image is used to represent the elevation at discrete positions. The elevation values are created using a composition of perlin or simplex noise which is often used in computer graphics to create natural landscapes. The elevation values are then mapped to colors for visual purpose only. The world map is implemented in the file [World.py](#)

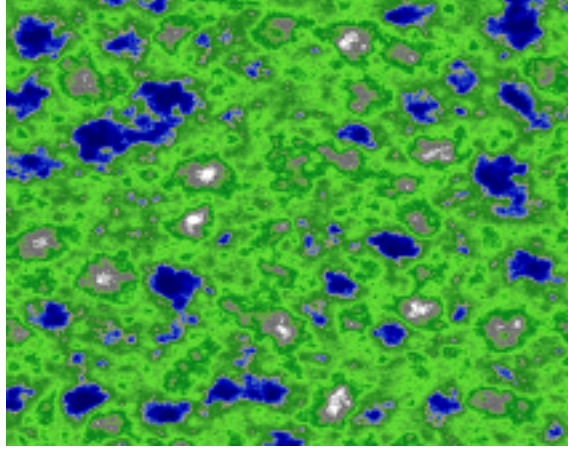


Figure 1: A typical map generated by our algorithm.

4.2 Spacecraft model

The spacecraft is basically just a location which can be moved around manually using the arrow keys. The motion is determined by its directional and angular velocity and an arbitrary friction value for each. The spacecraft is implemented in the file [Robot.py](#)

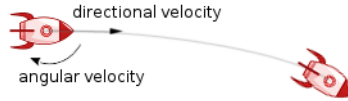


Figure 2: The spacecraft and its motion.

4.3 Sensor model

Sensors can be described by their observation method and their observation rate. The observations made by the sensor underly an unknown noise model. The abstract sensor and noise model are implemented in the file [Sensor.py](#).

In our case, we simulate two kinds of sensors with different noise models:

- Motion sensor: Captures motion of the spacecraft (Odometry/Position)
- Vision sensor: Captures terrain's elevation beneath the spacecraft

4.3.1 Motion Sensor

The motion sensor is used to capture the movement of the spacecraft since the last measurement and is basically just the motion vector between the last and current position. The sensor is capable of making observations at a given framerate. The implementation of the Motion sensor can be found in the file [Motion.py](#).

We simulate two different measurement models: (a) using the motion vector measurements or (b) using the odometry, i.e. independent measurement of the angle and distance of the motion. This is reflected by the two noise models:

- Vector: The motion error is given by a bivariate normal distribution
- Odometry: The error for rotation and distance are both normal distributed, but considered independently

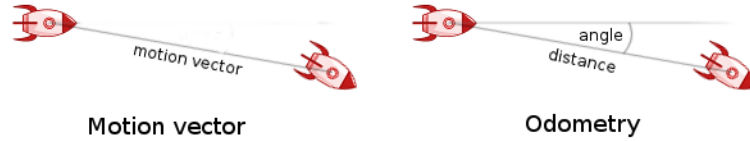


Figure 3: Different motion sensor models for the spacecraft.

The implementation of the Motion sensor noise can also be found in the file [Motion.py](#).

4.3.2 Vision Sensor

The vision sensor is able to capture the elevation of the terrain in its view. The view is determined by the position of the spacecraft and is able to get information about the terrain within a certain range given by a rectangle with a predefined size. The sensor is capable of making observations at a given framerate. The implementation of the Vision sensor can be found in the file [Vision.py](#).

Additionally, we implemented various methods for simulating sensor noise:

- Gaussian: Simple pixelwise Gaussian noise
- Salt: Some pixels are changed to the maximum possible value
- Pepper: Some pixels are changed to the minimum possible value
- Salt & Pepper: Some pixels are changed randomly to the maximum or minimum possible value
- Speckle: Additive Gaussian noise based on observed pixels

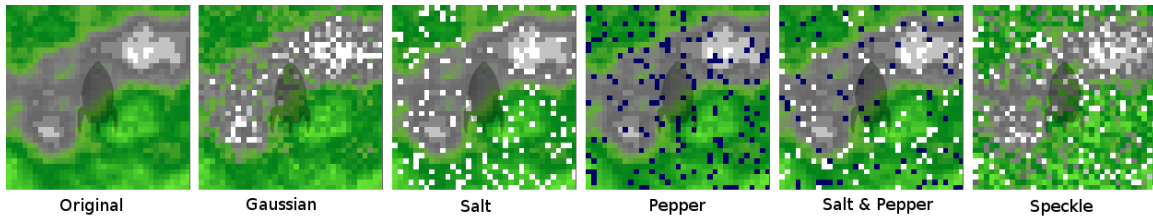


Figure 4: Different noise models for the vision sensor.

The implementation of the Vision sensor noise can also be found in the file [Vision.py](#).

5 Motion model

Motion model is used to predict the new position of the density or particles depending on the commands given to the robot. In particular, the motion model is given by the transition distribution

$$p(\theta_t|\theta_{t-1}) = p(\theta_t|\theta_{t-1}, u_t). \quad (26)$$

We implemented two motion models which are based on the motion sensor and its noise:

- The first model uses the motion vector itself to propagate the density or particles. The uncertainty of the new position is modelled by adding Gaussian noise to its final position.
- The second model uses the odometry measurements to compute a new position. Here, the density or particles are moved based on the rotation and translation of the robot and the uncertainty is modeled independently for the rotation and translation by a Gaussian noise.

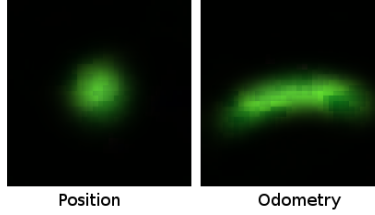


Figure 5: Distribution of the likely positions for different noise models of the motion sensor. The position after the true motion would be in the center of the image.

The implementation of the motion model can be found in the file [Motion.py](#).

6 Observation model

The observation model is used to compute the similarity between an observation z of the vision sensor and a patch M (of the same size as z) in the reference map. In particular, we compute the likelihood given as

$$p(z_t|\theta_t) = p(z_t|\theta_t, \mathcal{M}). \quad (27)$$

In this project, we implemented four different methods to compute the likelihood:

- Absolute mean difference: it is more suited for the particle-based method as it introduces more noise in comparison to the other methods. This metric thus produces a flatter distribution which helps capturing lower-weight particles that might otherwise be excluded and could lead to sample degeneracy or impoverishment.

$$R(x, y) = \sum_{x', y'} |Z(x', y') - M(x + x', y + y')| \quad (28)$$

- Normalized cross-correlation coefficient: it is better suited for the grid-based method as the metric is usually creates a distribution with steep modes which are prone to be “overlooked” by the particle-based method.

$$\begin{aligned}
R(x, y) &= \frac{\sum_{x', y'} (Z(x', y') \cdot M(x + x', y + y'))}{\sqrt{\sum_{x', y'} Z(x', y')^2 \cdot \sum_{x', y'} M(x + x', y + y')^2}} \\
Z'(x', y') &= Z(x', y') - 1/(w \cdot h) \cdot \sum_{x'', y''} Z(x'', y'') \\
M'(x + x', y + y') &= M(x + x', y + y') - 1/(w \cdot h) \cdot \sum_{x'', y''} M(x + x'', y + y'') \quad (29)
\end{aligned}$$

- Normalized cross-correlation: it is similar to the previous method; the main difference is that is not centered and it therefore dependent on the absolute values themselves, i.e. higher values result in a higher probability.

$$R(x, y) = \frac{\sum_{x', y'} (Z'(x', y') \cdot M'(x + x', y + y'))}{\sqrt{\sum_{x', y'} Z'(x', y')^2 \cdot \sum_{x', y'} M'(x + x', y + y')^2}} \quad (30)$$

- Mean squared difference: it is a good compromise in terms of performance for the grid-based and particle-based method similar to the absolute mean difference.

$$R(x, y) = \sum_{x', y'} (Z(x', y') - M(x + x', y + y'))^2 \quad (31)$$

The implementation of the observation models can be found in the file [Vision.py](#).

7 Representations

We approximate the posterior distribution $p(\theta_t | z_{0:t})$ using two methods: grid-based method and particle filter. The implementation for both methods can be found in the file [Representation.py](#).

7.1 Grid-based approximation

The grid-based localization approximates the posterior distribution by dividing the continuous state space into a discrete grid. Each cell in the grid has a probability (or weight) assigned to it that represents the belief of being the true locations. Based on the spacecraft's movement, these probabilities have to be propagated on the grid using our motion model and then corrected according to the observation model.

An advantage of the grid-based approach is that it can accurately model the distribution given a high resolution of the grid. However, since there is only one true state, the posterior tends to be very sparse, i.e. most of the cells in the grid only have a very low probability assigned to it. Especially in higher dimensional space, this leads to a system where most of the computational power is wasted by comparing unlikely map patches with the observation.

7.2 Particle-based approximation

In order to reduce the computational overhead created by the grid-based method, the particle-based filter has become popular. Similar to the grid-based methods, the posterior distribution is approximated at certain locations, in this case given by a set of particles which represent possible locations of the spacecraft. Instead of having an evenly divided grid, the particles are drawn from the posterior distribution itself such that they accumulate at the modes of the distribution (or to be more precise: they are propagated towards the modes). We therefore only compare the observation with relevant map patches, thus representing the distribution with less particles and consequently reducing the computational cost.

Even though the particle filter makes the computation and estimation much more efficient, it comes with the trade-off of less accurate approximations compared to the grid-based method.

7.3 Resampling Methods

We implement three different sampling methods for the robot localization problem: naive sampling, stratified sampling, and systematic sampling. Ideally, the goal of the resampling method should be to sample those particles that are representative of our unknown probability distribution ($p(\theta_t|z_{0:t})$, the probability of the robot's location). In other words, to filter out unlikely particles and choose those with a higher probability (higher weights). However, sampling randomly can result in systematically picking more low weight particles and discarding large weight particles, and can therefore not only result in an unrepresentative sample, but it can also gradually lead to degeneracy as the sampling sequence continues. We employ these resampling methods to combat degeneracy (i.e. to not discard highly weighted particles during the resampling phase). We briefly expose the three methods and their properties. All three methods share the same initial structure in the sense that we first compute a normalized cumulative sum of the weights and then we sample with replacement from it. The following function implements the normalized cumulative sum and the draws from it with the index supplied by the implemented resampling method.

```
def _resample_indices(self, positions):
    indices = np.searchsorted(np.cumsum(self._density), positions, side='right')
    self._locations = self._locations[indices]
    self._density = self._density[indices]
    self._density /= np.sum(self._density)
```

7.3.1 Naive Sampling

In naive sampling, we simply draw uniformly from this cumulative sum of the weights. We implement this in python by cross-referencing two arrays: one array which stores N uniform draws, and another storing the cumulative sum of the weights (see function above). We then use the index of the drawn particles from the cumulative sum array as the basis of our new sample distribution. In this manner, particles with larger weights are more likely to be drawn, and hence we prevent the problem of picking unrepresentative particles.

```
def resample_naive(self):
    positions = np.random.random(self.N)
    self._resample_indices(positions)
```

7.3.2 Stratified Sampling

Similar to naive sampling, we compute the cumulative sum of the weights and normalize it. However, it differs from it in that we divide our sampling space into N equal groups (i.e. the stratification), then we draw uniformly from each group. This results in drawing the same amount of times from each subdivision, and hence we ensure that we sample more evenly across the sampling space.

```
def resample_stratified(self):
    positions = (np.arange(self.N) + np.random.random(self.N)) / self.N
    self._resample_indices(positions)
```

However, this introduces a new problem - sample impoverishment - whereby with each iteration our sample becomes more and more highly concentrated with particles originating from large weight particles. That is, since large weight particles are more likely to be drawn, it can be seen that with each resampling step, we gradually reduce the diversity of the particles. In a worst case scenario (in the limit) of sample impoverishment, we would end up with a sample whose particles originated from one particle of large weight, which would result in (poorly) approximating our probability distribution with only one point estimate. We combat this by resampling only when the effective sample size (S_{eff} is set to 0.5 here) threshold is reached. In this manner, we slow down the process of sample impoverishment.

7.3.3 Systematic Sampling

Similar to stratified sampling, we divide the sampling space into N equal groups, except that instead of each group's draw being independent of each other, in systematic sampling the position of a draw is the same in each group. Therefore, the draws in each group are equally spaced.

```
def resample_systematic(self):
    positions = (np.arange(self.N) + np.random.random()) / self.N
    self._resample_indices(positions)
```