

Stochastic Hybrid Processes as Solutions of Stochastic Differential Equations

Estimation of Rare Event Probability in Stochastic Hybrid Systems



Jaroslav Krystul, Henk A.P. Blom and Arunabha Bagchi

Outline

- Introduction
- Classical SDE models
- Stochastic Hybrid Systems
 - Model by Ghosh and Bagchi
 - Models by Krystul and Blom
- Estimation of rare event probability in stochastic hybrid systems

Stochastic Differential Equation

$$dX_t = a(X_t) dt + \sigma(X_t) dW_t, \quad t > 0$$

$$X(0) = X_0$$

$\{W_t, t \geq 0\}$ is the standard Brownian motion.

$a(\cdot)$: drift coefficient and $b(\cdot)$: diffusion coefficient

Solution : existence and uniqueness; under “Lipschitz” and “Growth” condition

$$(L) \quad |a(x) - a(y)|^2 + |b(x) - b(y)|^2 \leq K|x - y|^2$$

$$(G) \quad |a(x)|^2 + |b(x)|^2 \leq K|x|^2$$

SDE: Strong solution: Properties

I. Continuous path

2. Semimartingale : $X(t) = X_0 + A(t) + M(t)$

$A(t)$ is of bounded variation, $M(t)$ is a **Martingale**.

– Important for mathematical analysis (well developed theory)

3. Markov : $\mathcal{D}(X(t)|\mathcal{F}_s) = \mathcal{D}(X(t)|X(s))$, $t > s$.

Transition prob. : $p(s, x, t, A) = P(X(t) \in A | X(s) = x)$

Time homogeneous : $p(s, x, t, A) = p(0, x, t - s, A)$

$$\Rightarrow \mathcal{D}(X(t+s)|\mathcal{F}_s) = \mathcal{D}_{X(s)}(X(t)).$$

Strong Markov :

$\mathcal{D}(X(t+\tau)|\mathcal{F}_\tau) = \mathcal{D}_{X(\tau)}(X(t))$, τ stopping time

– Helpful to analyse processes stopped at random times

Jump Diffusion Process

- Extension while keeping Semimartingale property?
- Discontinuous paths?

Jump Diffusion Process :

$$dX_t = a(X_t) dt + b(X_t) dW_t + h(X_t) dN_t, \quad t > 0$$

$\{N(t)\}$: (pure) jump process.

Alternative representation:

$$dX_t = a(X_t) dt + b(X_t) dW_t + \int_{\mathbb{R}} g(X_{t-}, u) p(dt, du)$$

$p(\cdot, \cdot)$ Poisson random measure associated with the jump process.

Solution: existence, uniqueness under Growth and Lipschitz condition on a , b and g .

Properties : Semimartingale, Markov, but *Discontinuous*

Stochastic Hybrid Systems

Further extension : Not only the process but also the governing model/equation jumps.

— *regime switch* or *Hybrid* model.

Model I (Ghosh & Bagchi):

$$(X_t, \theta_t) \in \mathbb{R}^d \times \mathbb{M}, \quad \mathbb{M} = \{1, \dots, N\}, \quad t \geq 0$$

$$dX_t = a(X_t, \theta_t) dt + b(X_t, \theta_t) dW_t + \int_{\mathbb{R}} g(X_{t-}, \theta_{t-}, u) p(dt, du)$$

$$P(\theta_{t+\delta t} = j | \theta_t = i, X_s, \theta_s, s \leq t) = \lambda_{ij}(X_t) \delta t + o(\delta t), \quad i \neq j$$

$$X(0) = X_0 \quad \theta(0) = \theta_0$$

$p(\cdot, \cdot)$ — Poisson random measure with intensity $dt \times l(du)$

$$\lambda_{ij}(\cdot) \geq 0, \quad i \neq j, \quad i, j = 1, 2, \dots, N \quad \text{and} \quad \sum_{j=1}^N \lambda_{ij}(\cdot) = 0$$

Hybrid Systems - Model I

Assumptions :

- $a(\cdot, i), b(\cdot, i)$ are bounded and Lipschitz continuous
- $\lambda_{ij}(\cdot)$ are bounded and measurable
- Support of $g(x, \theta, u)$ w.r.t. 'u' (i.e., the proj. on \mathbb{R}) is bdd.

Existence of Solution: (Put in Ito-Skorohod's framework)

Identify i with e_i , i th unit vector in \mathbb{R}^N and embed \mathbb{M} into \mathbb{R}^N . For $x \in \mathbb{R}^d$, define $\Delta_{ij}(x)$ to be the consecutive (with lexicographic ordering on $(i, j) \in \mathbb{M} \times \mathbb{M}$) left closed right open intervals on \mathbb{R}_+ with length $\lambda_{ij}(x)$.

Define $c : \mathbb{R}^d \times \mathbb{M} \times \mathbb{R} \rightarrow \mathbb{R}^N$ as

$$c(x, i, u) = e_j - e_i \quad \text{if } u \in \Delta_{ij}(x), \quad 0 \text{ otherwise.}$$

Hybrid Systems - Model I

Model can be expressed as

$$dX_t = a(X_t, \theta_t) dt + b(X_t, \theta_t) dW_t + \int_{\mathbb{R}} g(X_{t-}, \theta_{t-}, u) p(dt, du)$$

$$d\theta_t = \int_{\mathbb{R}} c(X_{t-}, \theta_{t-}, u) p(dt, du), \quad t > 0$$

$$X(0) = X_0 \quad \theta(0) = \theta_0$$

and the existence of unique strong solution can be shown.

Remarks :

- Conditions on the coefficients can be relaxed
- $\{X(t)\}$ is not Markov
- Augmented process $(X(t), \theta(t))$ is Markov

Hybrid Systems - Model I

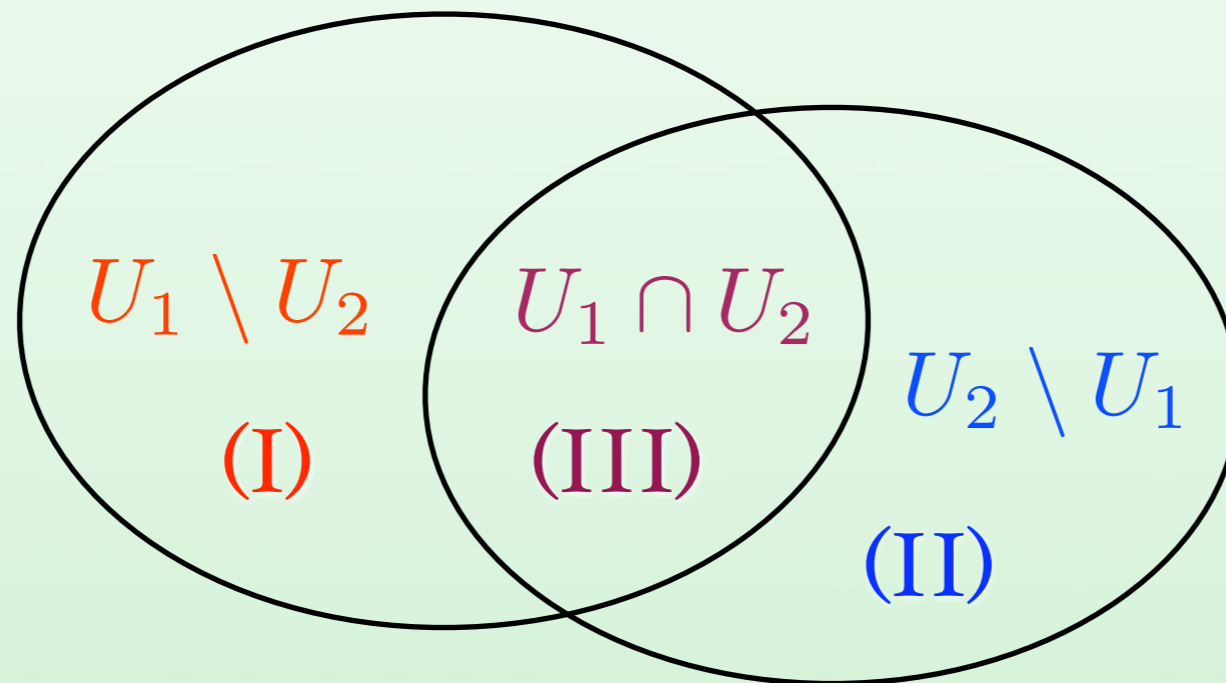
Let U_1 and U_2 denote respectively the supports of $g(x, \theta, u)$ and $c(x, \theta, u)$ w.r.t. 'u' (i.e. the projections on \mathbb{R}).

Different situations can arise:

I. **Process jumps but no switching:** $u \in U_1 \setminus U_2$

II. **No jump in process but switching occurs:** $u \in U_2 \setminus U_1$

III. **Simultaneous switching and jump (Hybrid Jump):** $u \in U_1 \cap U_2$



Hybrid Systems - Model 2 (Krystul and Blom)

$$dX_t = a(X_t, \theta_t) dt + b(X_t, \theta_t) dW_t + \int_{\mathbb{R}^n} g_1(X_{t-}, \theta_{t-}, u) q_1(dt, du) \\ + \int_{\mathbb{R}^n} g_2(X_{t-}, \theta_{t-}, u) p_2(dt, du)$$

$$d\theta_t = \int_{\mathbb{R}^n} c(X_{t-}, \theta_{t-}, u) p_2(dt, du), \quad t > 0$$

- W is an m -dimensional standard Wiener process.
- $q_1(dt, du)$ is a martingale random measure associated to a Poisson random measure p_1 with intensity $dt \times m_1(du)$.
- $p_2(dt, du)$ is a Poisson random measure with intensity $dt \times m_2(du) = dt \times du_1 \times \mu(\underline{u})$, μ is a probability measure on \mathbb{R}^{n-1} , $u_1 \in \mathbb{R}$.

Hybrid Systems - Model 2

Assumptions :

- Lipschitz condition. For all $i = 1, 2, \dots, N$

$$|a(x, e_i)|^2 + |b(x, e_i)|^2 + \int_{\mathbb{R}^n} |g_1(x, e_i, u)|^2 m_1(du) \leq K(1 + |x|^2).$$

- Growth condition:

$$\begin{aligned} & |a(x, e_i) - a(y, e_i)|^2 + |b(x, e_i) - b(y, e_i)|^2 \\ & + \int_{\mathbb{R}^n} |g_1(x, e_i, u) - g_1(y, e_i, u)|^2 m_1(du) \leq K_r |x - y|^2. \end{aligned}$$

for $|x| \leq r, |y| \leq r$.

- $\lambda_{ij}(\cdot)$ are bounded and measurable
- Support of $g_2(x, \theta, u)$ w.r.t. ' u_1 ' (i.e., the proj. on \mathbb{R}) is bdd.

Hybrid Systems - Model 3 (Krystul, Blom and Bagchi)

Idea:

X_t starts in some set E^i and (X_t, θ_t) evolves according to some Ito-Skorohod SDE.

If X_t reaches a boundary of that set, the process immediately jumps to a new set E^j (possibly including jump in θ)

Process continues again according to the Ito-Skorohod SDE in a new set until it reaches the boundary of the current set.

$$E = \{x \mid x \in E^i, \text{ for some } i = 1, \dots, L\} = \bigcup_{i=1}^L E^i$$

$$\partial E = \{x \mid x \in \partial E^i, \text{ for some } i = 1, \dots, L\} = \bigcup_{i=1}^L \partial E^i$$

Hybrid Systems - Model 3

$(X_t, \theta_t) \in \mathbb{R}^d \times \mathbb{M}$, $\mathbb{M} = \{e_1, e_2, \dots, e_N\}$.

Suppose $\tau_1^E < \tau_2^E < \dots < \tau_m^E < \dots$ are the successive hitting times of ∂E .

Between the boundary jumps $\{X_t, \theta_t\}$ is a switching jump-diffusion process. For $\tau_m^E < t < \tau_{m+1}^E$,

$$\begin{aligned} X_t &= X_{\tau_m^E} + \int_{\tau_m^E}^t a(X_s, \theta_s) ds + \int_{\tau_m^E}^t b(X_s, \theta_s) dW_s \\ &\quad + \int_{\tau_m^E}^t \int_{\mathbb{R}^n} g_2(X_{s-}, \theta_{s-}, u) p_2(ds, du), \\ \theta_t &= \theta_{\tau_m^E} + \int_{\tau_m^E}^t \int_{\mathbb{R}^n} c(X_{s-}, \theta_{s-}, u) p_2(ds, du). \end{aligned}$$

Hybrid Systems - Model 3

At jump points ($t = \tau_m$),

$$X_{\tau_m^E} = f^x(X_{\tau_m^E-}, \theta_{\tau_m^E-}, Z),$$

$$\theta_{\tau_m^E} = f^\theta(X_{\tau_m^E-}, \theta_{\tau_m^E-}, Z),$$

where Z is some V -valued random variable,

$$f^x : \partial E \times \mathbb{M} \times V \longrightarrow E,$$

$$f^\theta : \partial E \times \mathbb{M} \times V \longrightarrow \mathbb{M}.$$

Hybrid Systems - Model 3

Assumptions :

- a and b satisfy Lipschitz and Growth conditions
- $\lambda_{ij}(\cdot)$ are bounded and measurable
- Support of $g_2(x, \theta, u)$ w.r.t. ' u_1 ' (i.e., the proj. on \mathbb{R}) is bdd.
- Function g_2 has the following property: $(x + g_2(x, \theta, u)) \in E$ for each $x \in E^i$, $\theta \in \mathbb{M}$, $u \in \mathbb{R}^n$, $i = 1, \dots, L$
- $d(\partial E, f^x(\partial E, \mathbb{M}, V)) > 0$,
i.e. when $\{X_t\}$ has reached the boundary ∂E it always jumps inside of open set E .

Estimation of rare event probability in stochastic hybrid systems

$$\{X_t, \theta_t\} \in \mathbb{R}^n \times \mathbb{M}, \quad \mathbb{M} = \{e_1, \dots, e_S\}$$

switching diffusion process:

$$dX_t = a(\theta_t, X_t)dt + b(\theta_t, X_t)dW_t,$$

$$P(\theta_{t+\delta} = \theta | \theta_t = \eta, X_t = x) = \lambda_{\eta\theta}(x)\delta + o(\delta), \eta \neq \theta.$$

(X_t, θ_t) starts at $t = 0$ in $D_0 \subset \mathbb{R}^n \times \mathbb{M}$ with known initial probability distribution $P_{X_0, \theta_0}(\cdot)$.

$\tau_A = \inf\{t \geq 0 : X_t \in A\}$: hitting time of $A \subset \mathbb{R}^n$, $A \cap D_0 = \emptyset$.

To calculate $P(\tau_A < T)$: probability that X_t will reach the set A in the time interval $(0, T)$.

Using crude Monte Carlo to estimate $P(\tau_A < T)$ may require a huge number of simulations. For events with probability of the order 10^{-10} , often 10^{11} or more Monte Carlo simulated trajectories are needed.

Factorization Approach

Idea : Identify intermediate states that are (sequentially) visited much more often than the rare target set.

D_0 initial set;

$$D_1 \supset \cdots \supset D_{m-1} \supset D_m = A \subset \mathbb{R}^n; \quad D_0 \cap D_1 = \emptyset.$$

Define for $k = 1, \dots, m$, stopping times $\tau_k \triangleq \inf\{t \geq 0 : X_t \in D_k\}$

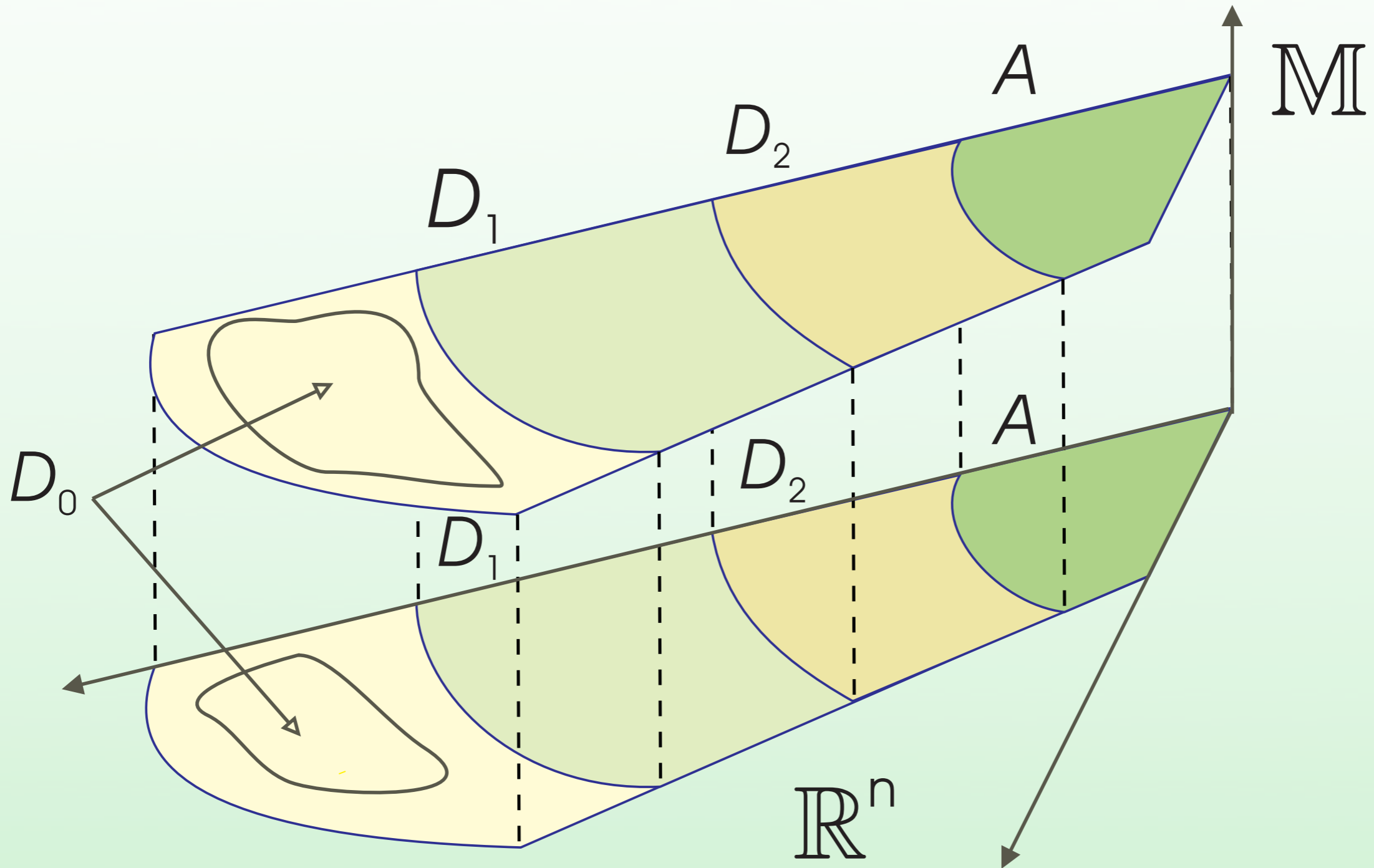
$$\begin{aligned} P(\tau_A < T) &= P(\tau_m < T) \\ &= P(\tau_m < T, \tau_{m-1} < T, \dots, \tau_0 < T) \\ &= \prod_{k=1}^m P(\tau_k < T | \tau_{k-1} < T). \end{aligned}$$

Individual probabilities on r.h.s. are not very small.

Nested sequence of level sets

D_0 initial set; $D_1 \supset \dots \supset D_{m-1} \supset D_m = A$; $D_0 \cap D_1 = \emptyset$.

$$P(\tau_A < T) = \prod_{k=1}^m P(\tau_k < T | \tau_{k-1} < T).$$



Factorization Approach (continued)

$$P(\tau_A < T) = \prod_{k=1}^m P(\tau_k < T | \tau_{k-1} < T) \triangleq \prod_{k=1}^m \gamma_k$$

To calculate γ_k 's we proceed as follows:

Define $\xi_k = (X_{\tau_k \wedge T}, \theta_{\tau_k \wedge T})$,

$$\pi_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_k < T), \quad B \subset \mathbb{R}^n \times \mathbb{M},$$

$$p_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_{k-1} < T)$$

$\{(X_t, \theta_t)\}$ **strong Markov** \Rightarrow $\{\xi_k\}$ is a **Markov (chain)**

Factorization Approach (continued)

$$P(\tau_A < T) = \prod_{k=1}^m P(\tau_k < T | \tau_{k-1} < T) \triangleq \prod_{k=1}^m \gamma_k$$

To calculate γ_k 's we proceed as follows:

Define $\xi_k = (X_{\tau_k \wedge T}, \theta_{\tau_k \wedge T})$,

$$\pi_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_k < T), \quad B \subset \mathbb{R}^n \times \mathbb{M},$$

$$p_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_{k-1} < T)$$

$\{(X_t, \theta_t)\}$ **strong Markov** \Rightarrow $\{\xi_k\}$ is a **Markov** (chain)

$$p_k(B) = \int_{\mathbb{R}^n \times \mathbb{M}} P_{\xi_k | \xi_{k-1}}(B | \xi) \pi_{k-1}(d\xi),$$

$$\pi_k(B) = \frac{P(\xi_k \in B, \tau_0 < T, \dots, \tau_k < T)}{P(\tau_0 < T, \dots, \tau_k < T)} = \frac{P(\xi_k \in B \cap \bar{D}_k, \tau_0 < T, \dots, \tau_{k-1} < T)}{P(\tau_0 < T, \dots, \tau_k < T)}$$

$$= \frac{P(\xi_k \in B \cap \bar{D}_k | \tau_0 < T, \dots, \tau_{k-1} < T)}{P(\tau_k < T | \tau_0 < T, \dots, \tau_{k-1} < T)} = \frac{\int_B \mathbf{I}_{\{\xi \in \bar{D}_k\}} p_k(d\xi)}{\int_{\mathbb{R}^n \times \mathbb{M}} \mathbf{I}_{\{\xi \in \bar{D}_k\}} p_k(d\xi)}$$

$$\bar{D}_k \triangleq D_k \times \mathbb{M}$$

$$\gamma_k = \int_{\mathbb{R}^n \times \mathbb{M}} \mathbf{I}_{\{\xi \in \bar{D}_k\}} p_k(d\xi)$$

Factorization Approach (continued)

$$P(\tau_A < T) = \prod_{k=1}^m P(\tau_k < T | \tau_{k-1} < T) \triangleq \prod_{k=1}^m \gamma_k$$

To calculate γ_k 's we proceed as follows:

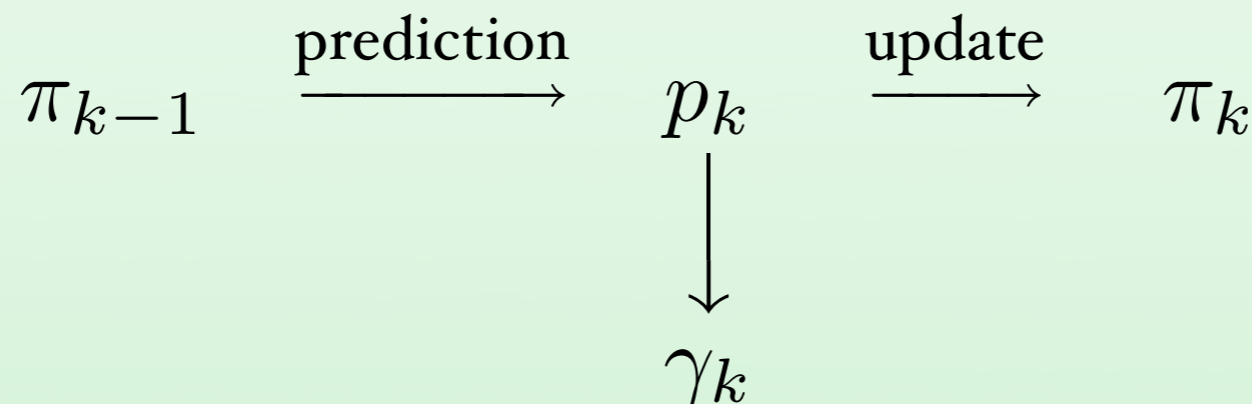
Define $\xi_k = (X_{\tau_k \wedge T}, \theta_{\tau_k \wedge T})$,

$$\pi_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_k < T), \quad B \subset \mathbb{R}^n \times \mathbb{M},$$

$$p_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_{k-1} < T)$$

$\{(X_t, \theta_t)\}$ **strong Markov** \Rightarrow $\{\xi_k\}$ is a **Markov** (chain)

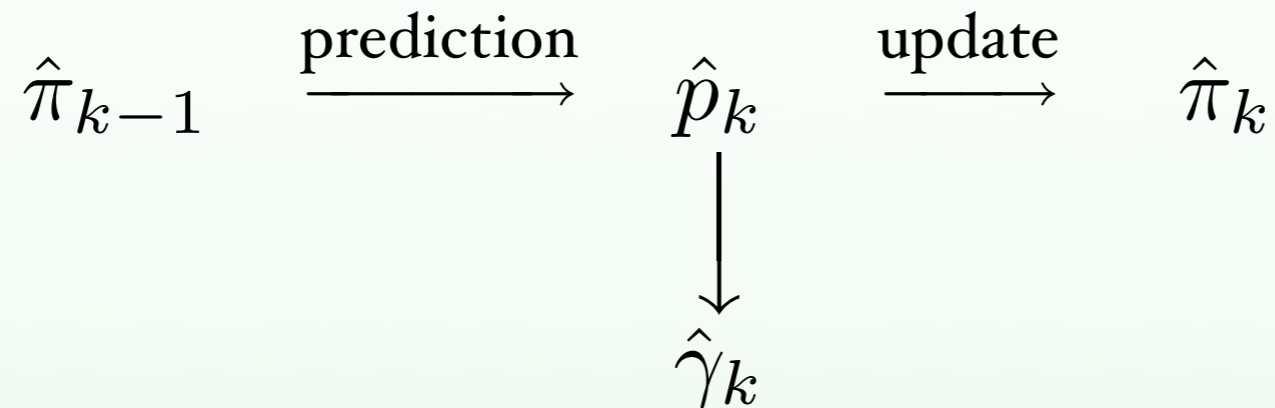
Evolution of the flow $\{\pi_k, p_k, \gamma_k; k = 0, 1, \dots, m\}$ is described by the following diagram:



with initial condition $\pi_0(d\xi) = P_{\xi_0}(d\xi) = P(\xi_0 \in d\xi)$.

Interacting Particle System (IPS) approach

An approximating sequence $\{\hat{\pi}_k, \hat{p}_k, \hat{\gamma}_k; k = 0, 1, \dots, m\}$:



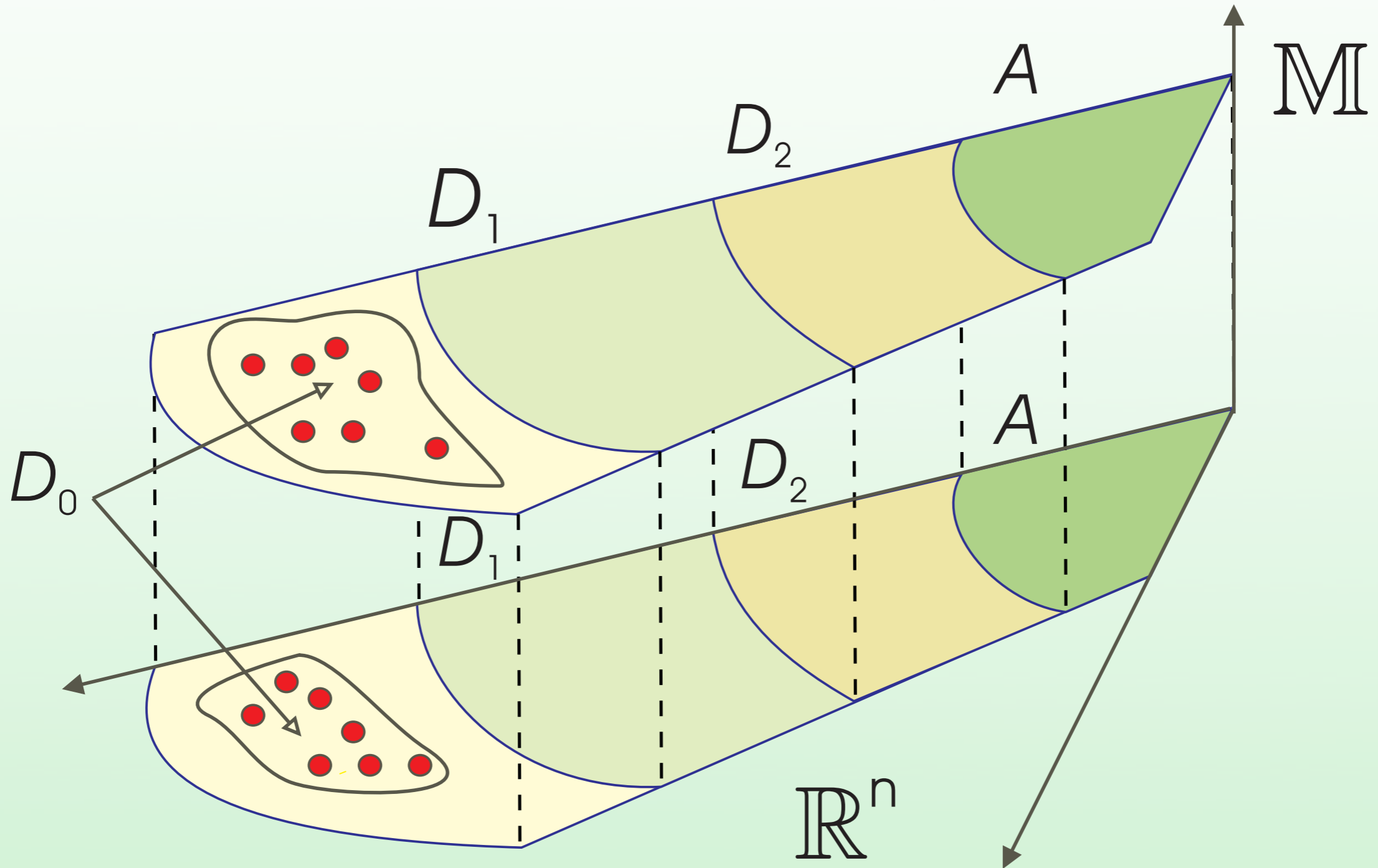
Approximations in the form of weighted empirical distributions associated with the particle system $\{\xi_k^i, \omega_k^i\}_{i=1}^N$:

$$\begin{aligned}
 \pi_0 \approx \hat{\pi}_0 &= \sum_{i=1}^N \omega_0^i \delta_{\{\xi_0^i\}}, & \pi_k \approx \hat{\pi}_k &= \sum_{i=1}^N \frac{\mathbf{I}_{\{\xi_k^i \in \bar{D}_k\}}}{\sum_{j=1}^N \omega_k^j \mathbf{I}_{\{\xi_k^j \in \bar{D}_k\}}} \delta_{\{\xi_k^i\}}, \\
 p_k \approx \hat{p}_k &= \sum_{i=1}^N \omega_k^i \delta_{\{\xi_k^i\}}, & \gamma_k \approx \hat{\gamma}_k &= \sum_{i=1}^N \omega_k^i \mathbf{I}_{\{\xi_k^i \in \bar{D}_k\}}.
 \end{aligned}$$

IPS algorithm

Step 1: At $t = 0$ choose N particles $\{\xi_0^i, i = 1, \dots, N\}$ from

$P_{X_0, \theta_0} \equiv \pi_0$, equivalently, $\hat{\pi}_0 \triangleq \sum_{i=1}^N w_0^i \delta_{\{\xi_0^i\}}$, $w_0^i = \frac{1}{N}$.



IPS algorithm (continued)

Step 2. Prediction step: $\pi_{k-1} \rightarrow p_k$

Given $\hat{\pi}_{k-1} \triangleq \sum_{i=1}^N w_{k-1}^i \delta_{\{\xi_{k-1}^i\}}$, i.e., a weighted particle cloud $\{(\xi_{k-1}^i, w_{k-1}^i), i = 1, \dots, N\}$ with $\xi_{k-1}^i \in \bar{D}_{k-1}$,

Let particles evolve following the hybrid system equations until \bar{D}_k or the final time T is hit.

$\hat{\xi}_k^i$: Value of the i -th particle at the end of the step

$\hat{p}_k = \sum_{i=1}^N w_{k-1}^i \delta_{\{\hat{\xi}_k^i\}}$ is the approximation of p_k .

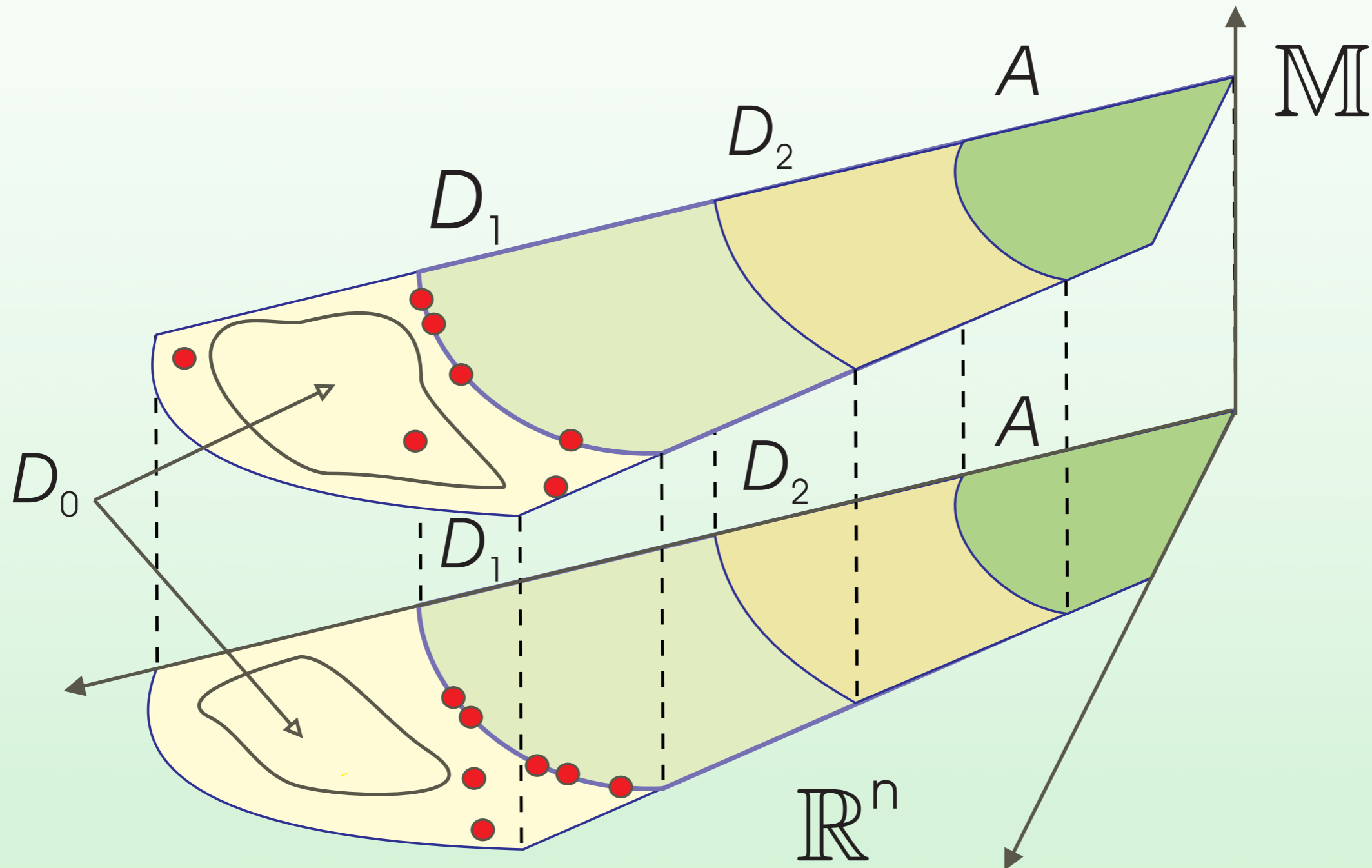
Recall that:

$$p_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_{k-1} < T), B \subset \mathbb{R}^n \times \mathbb{M}.$$

IPS algorithm (continued)

$\hat{\xi}_k^i$: Value of the i -th particle at the end of the prediction step

$\hat{p}_k = \sum_{i=1}^N w_{k-1}^i \delta_{\{\hat{\xi}_k^i\}}$ is the approximation of p_k .

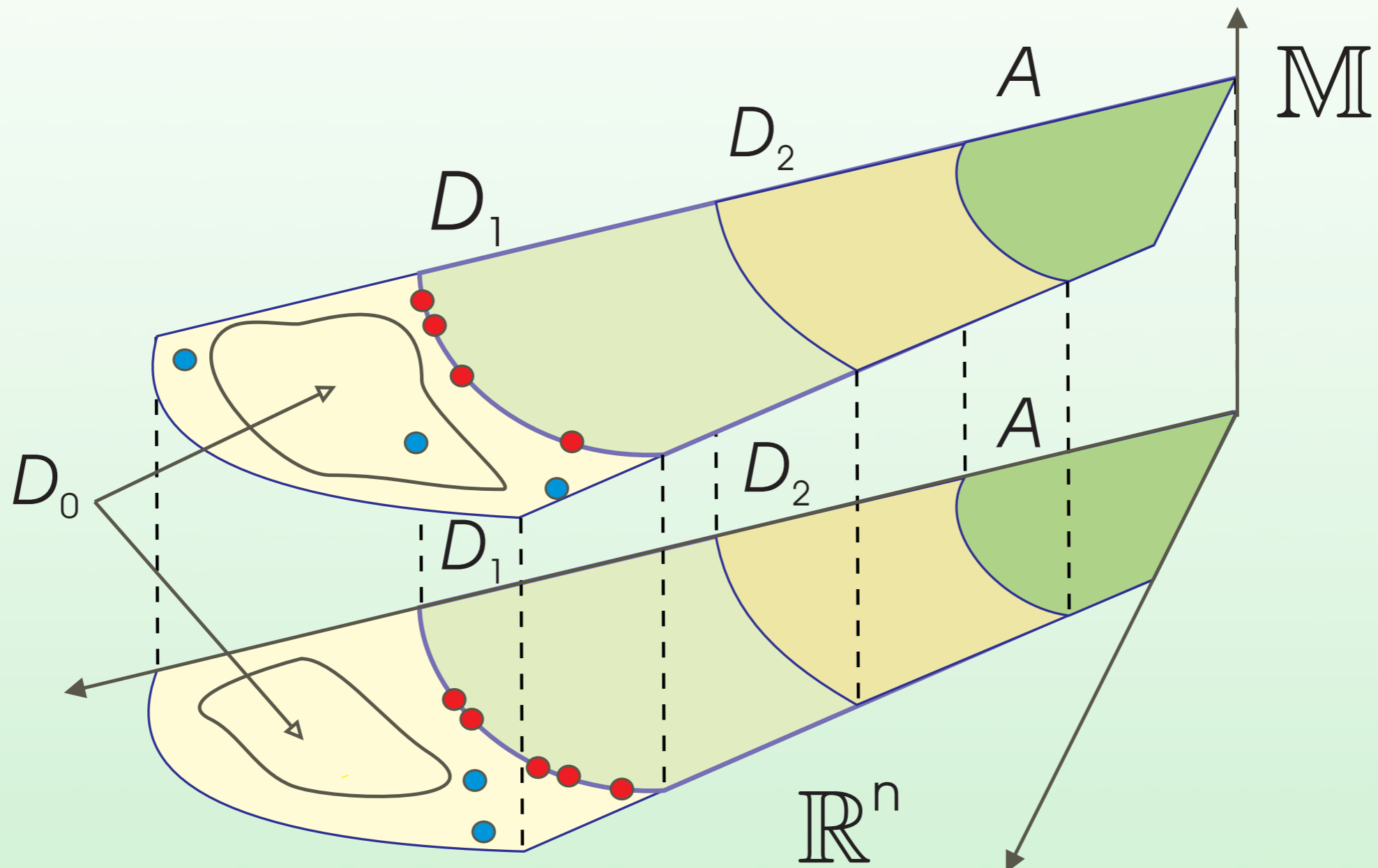


IPS algorithm (continued)

Set $\hat{w}_k^i = 0$, if $X_{\tau_k \wedge T}^i \notin D_k$, otherwise $\hat{w}_k^i = w_{k-1}^i$.

Approximate $\gamma_k = P(\tau_k < T | \tau_{k-1} < T)$ by $\hat{\gamma}_k = \sum_{i=1}^N \hat{w}_k^i$.

Stop algorithm if $\hat{\gamma}_k = 0$, and $P(\tau_A < T) \approx \prod_{k=1}^m \hat{\gamma}_k = 0$.



IPS algorithm (continued)

Step 3. Updating step: $p_k \rightarrow \pi_k$

Approximate π_k by $\hat{\pi}_k \triangleq \sum_{i=1}^N \tilde{w}_k^i \delta_{\{\hat{\xi}_k^i\}}$, with $\tilde{w}_k^i = \hat{w}_k^i / \sum_{i=1}^N \hat{w}_k^i$.

Recall that: $\pi_k(B) = P(\xi_k \in B | \tau_0 < T, \dots, \tau_k < T)$, $B \subset \mathbb{R}^n \times \mathbb{M}$.

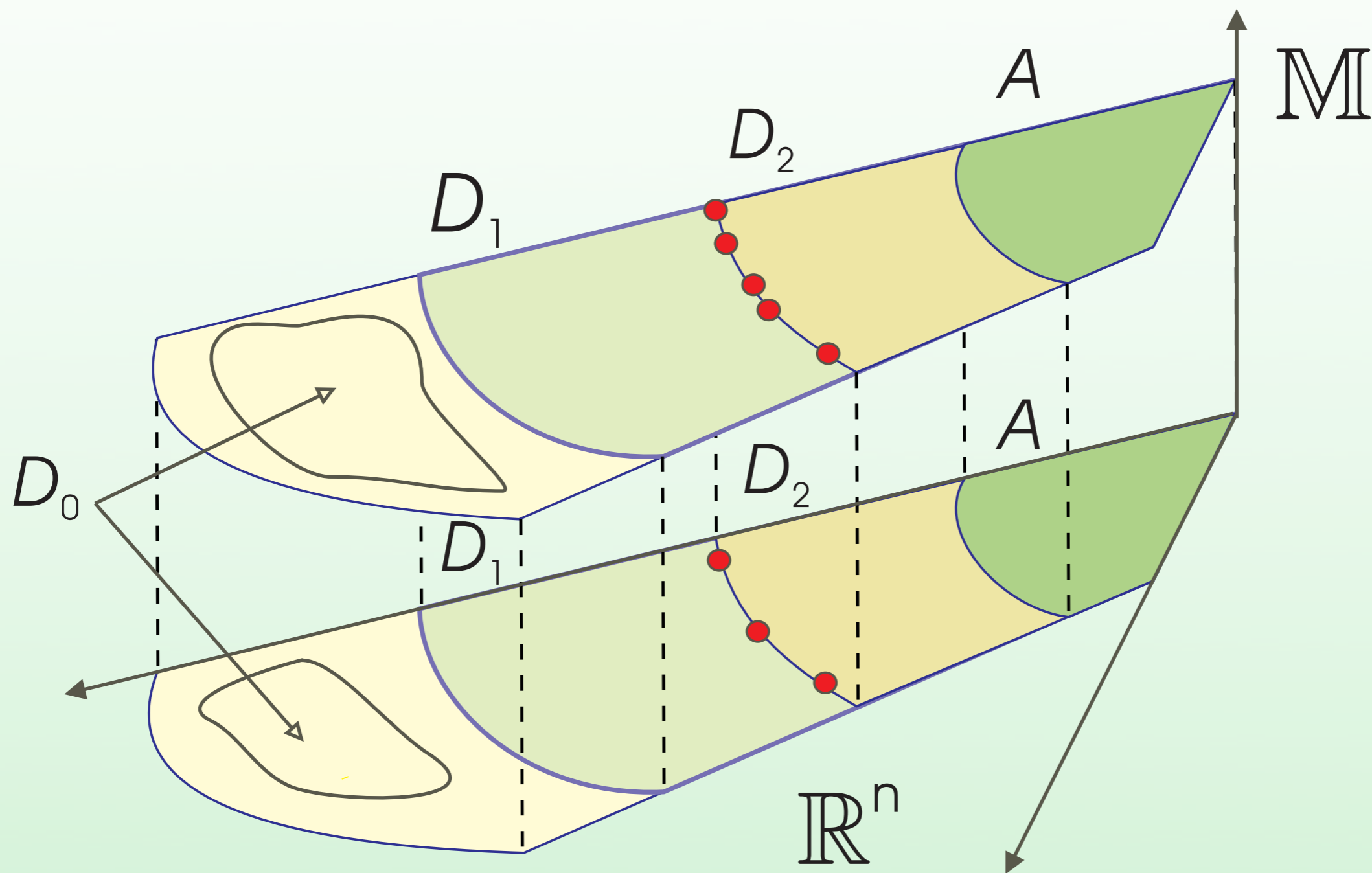
To avoid carrying the particles with no weight, resample N particles from $\hat{\pi}_k$.

The new set of particles is $\{\xi_k^i, \omega_k^i\}_{i=1}^N$ with $\omega_k^i = \frac{1}{N}$.

Go to Step 2.

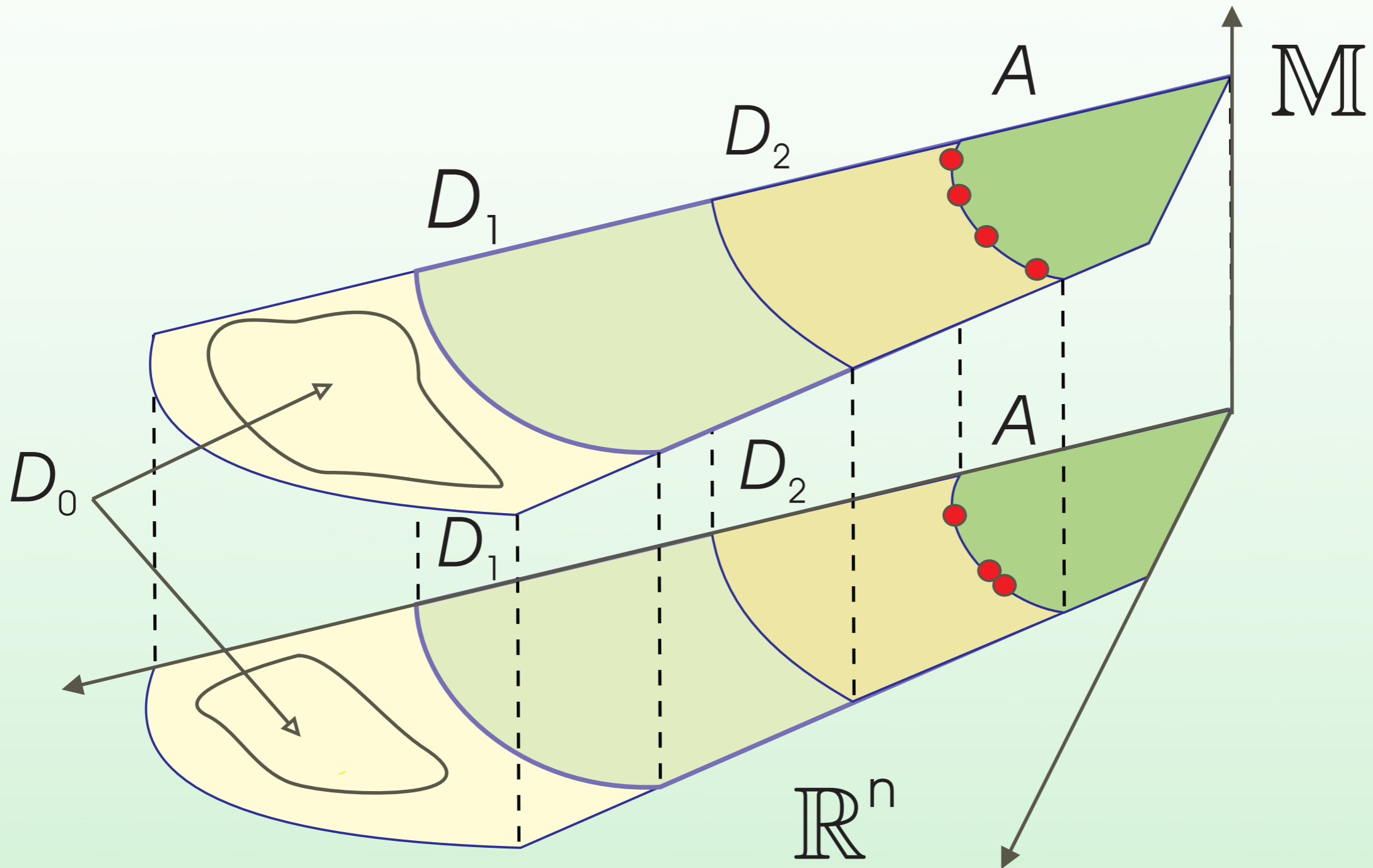
IPS algorithm (continued)

$$k := k + 1$$



IPS algorithm (continued)

At stage $k = m$ we get $P(\tau_A < T) \approx \prod_{k=1}^m \hat{\gamma}_k$.



Theorem (Cerou, Del Moral, LeGland and Lezaud, 2002)

IPS estimator is unbiased, i.e.

$$\mathbb{E} \left[\prod_{k=1}^m \gamma_k^N \right] = P(\tau_A < T) = P_{hit}(0, T)$$

and

$$\left(\mathbb{E} \left(\prod_{k=1}^m \gamma_k^N - \prod_{k=1}^m \gamma_k \right)^p \right)^{\frac{1}{p}} \leq \frac{a_p b_m}{\sqrt{N}},$$

for some finite constant a_p which depends only on the parameter p , and for some finite constant b_m which depends only on the parameter m .

Diffusion Example: Geometric Brownian Motion

$$dX_t = \left(\mu + \frac{\sigma^2}{2}\right)X_t dt + \sigma X_t dW_t, \quad X_0 = x$$

Probability of hitting level d before time T :

$$P(\tau_d \leq T), \quad \tau_d \triangleq \inf\{t > 0 : X_t \geq d\}.$$

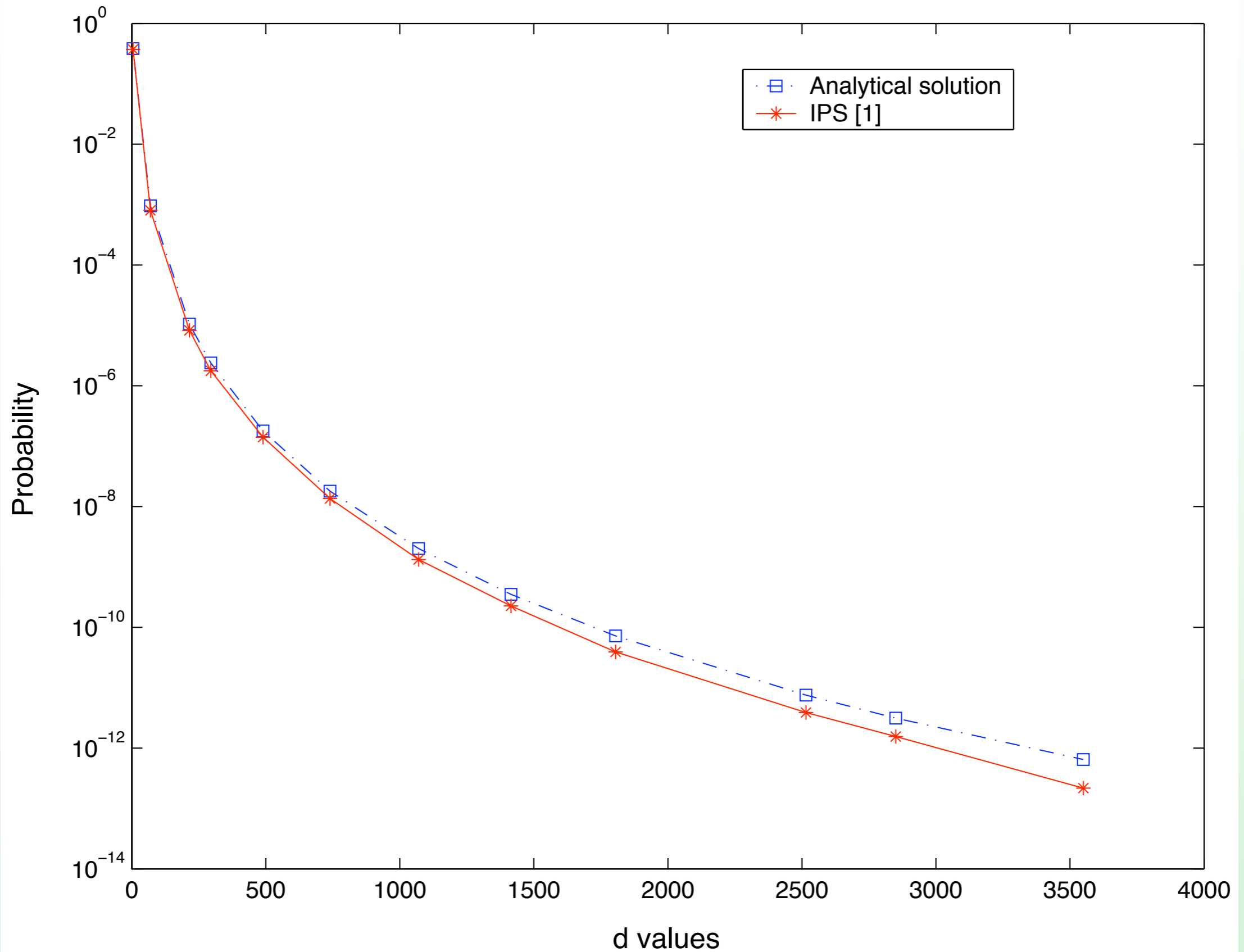
Analytical solution:
$$\int_0^t \frac{\ln(d/x)}{\sqrt{2\pi\sigma^3 s^3}} \exp\left\{\frac{-(\ln(d/x) - \mu s)^2}{2\sigma^2 s}\right\} ds$$

IPS :

- $\mu = 1, \sigma = 1, x_0 = 1,$ varying $d \leq 3550$
- intermediate levels d_j 's are chosen experimentally; 40%-55% of particles starting at D_{j-1} reaches D_j .
- 1000 simulations of 1000 particles each

$$P_{hit}(0, T) \approx \frac{1}{1000} \sum_{i=1}^{1000} \left(\prod_{k=1}^m \hat{\gamma}_k\right)^{(i)}$$

Probability to hit level d before time $T = 1$: diffusion



References

- Krystul, J. (2006). Modeling of Stochastic Hybrid Systems with Applications to Accident Risk Assessment, PhD Thesis.
- Ghosh, M.K. and Bagchi, A. (2005). Modeling stochastic hybrid systems. In J. Cagnol and J.P. Zolesio (Eds.), System Modeling and Optimization (pp. 269-280). Kluwer Academic Publishers.
- Blom, H.A.P. (2003). Stochastic hybrid processes with hybrid jumps, Proc. IFAC Conf. Analysis and Design of Hybrid Systems Saint-Malo, Brittany, France (pp. 361-366).