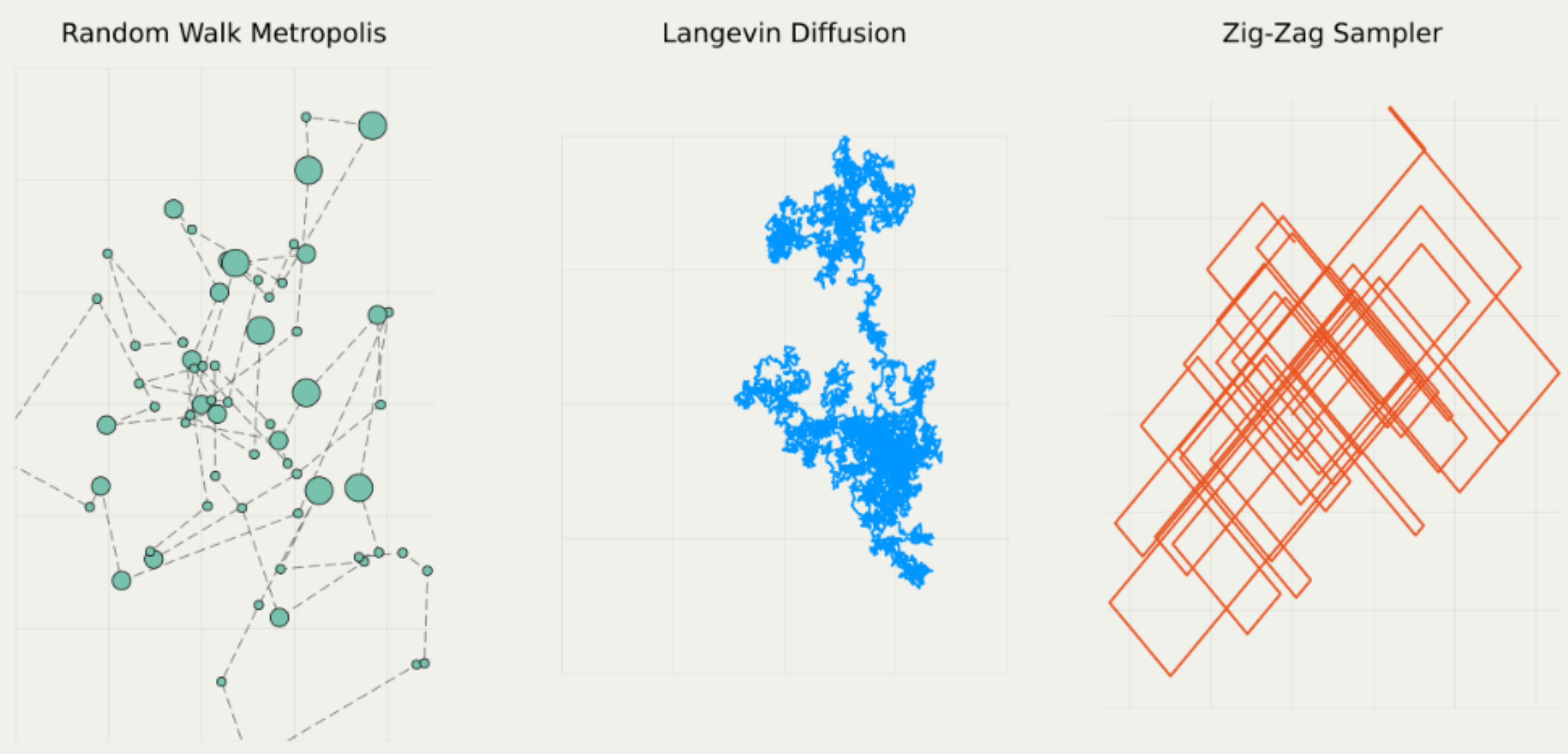
大学共同利用機関法人 情報・システム研究機構
統計数理研究所
The Institute of Statistical Mathematics

ベイズ変数選択の計算的解決

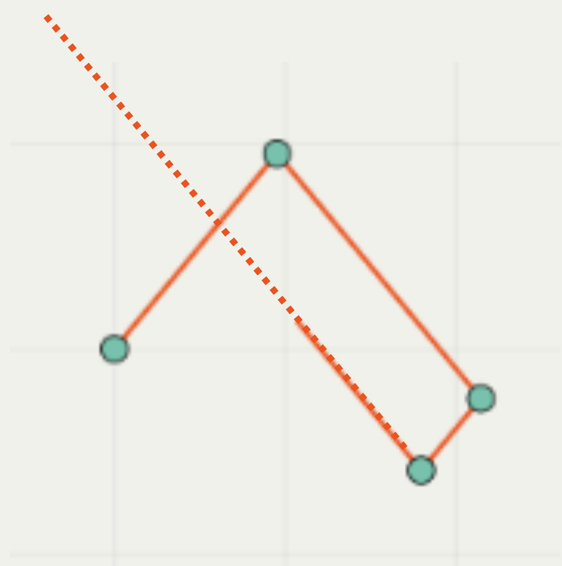
総合研究大学院大学　司馬博文

新たなサンプリング法：区分確定的モンテカルロ



Markov 連鎖　　拡散過程　　PDMP

MCMC とは全く異なるアルゴリズムを持つ



① 強度関数

$$m^{(i)}(t) = \left(v \cdot U(x_{i-1} + tv)\right)$$

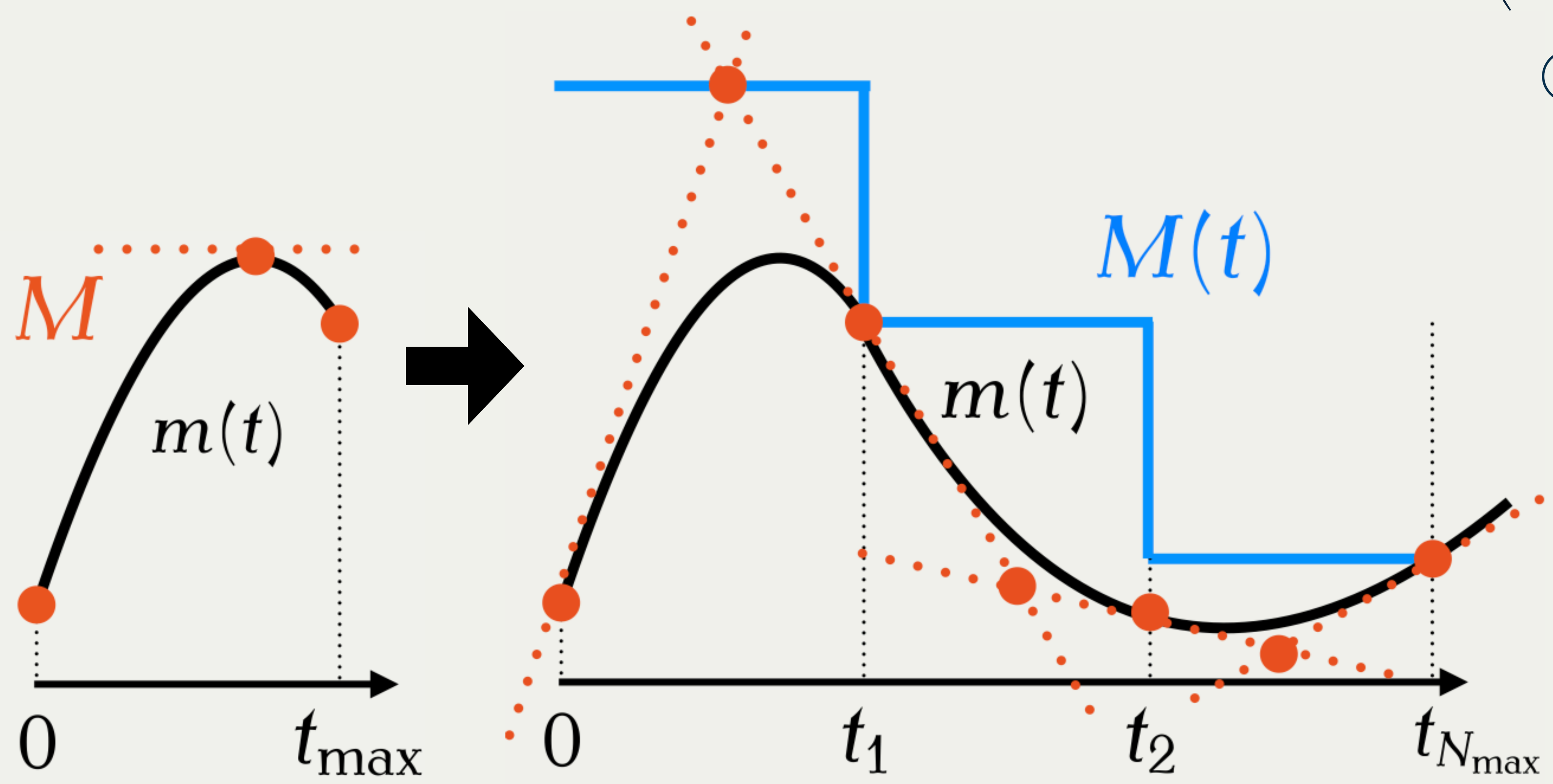
を持つ非一様 Poisson 点過程の最初の到着時刻 $T_1^{(i)}$ をシミュレート

$U(x) = -\log \pi(x)$:負の対数尤度 / ポテンシャル

Poisson 剪定の自動化

$$\mathrm{P}[T_1^{(i)} \geq s] = \exp\left(-\int_0^s m^{(i)}(t)\,dt\right)$$

⬅ この生存関数を持つ乱数を生成したい



② 上界 $M^{(i)}$ を構成する：

$$m^{(i)}(t) \leq M^{(i)}(t), \quad t \in [0, t_{\max}]$$

Poisson 剪定 (Lewis & Shedler, 1979)

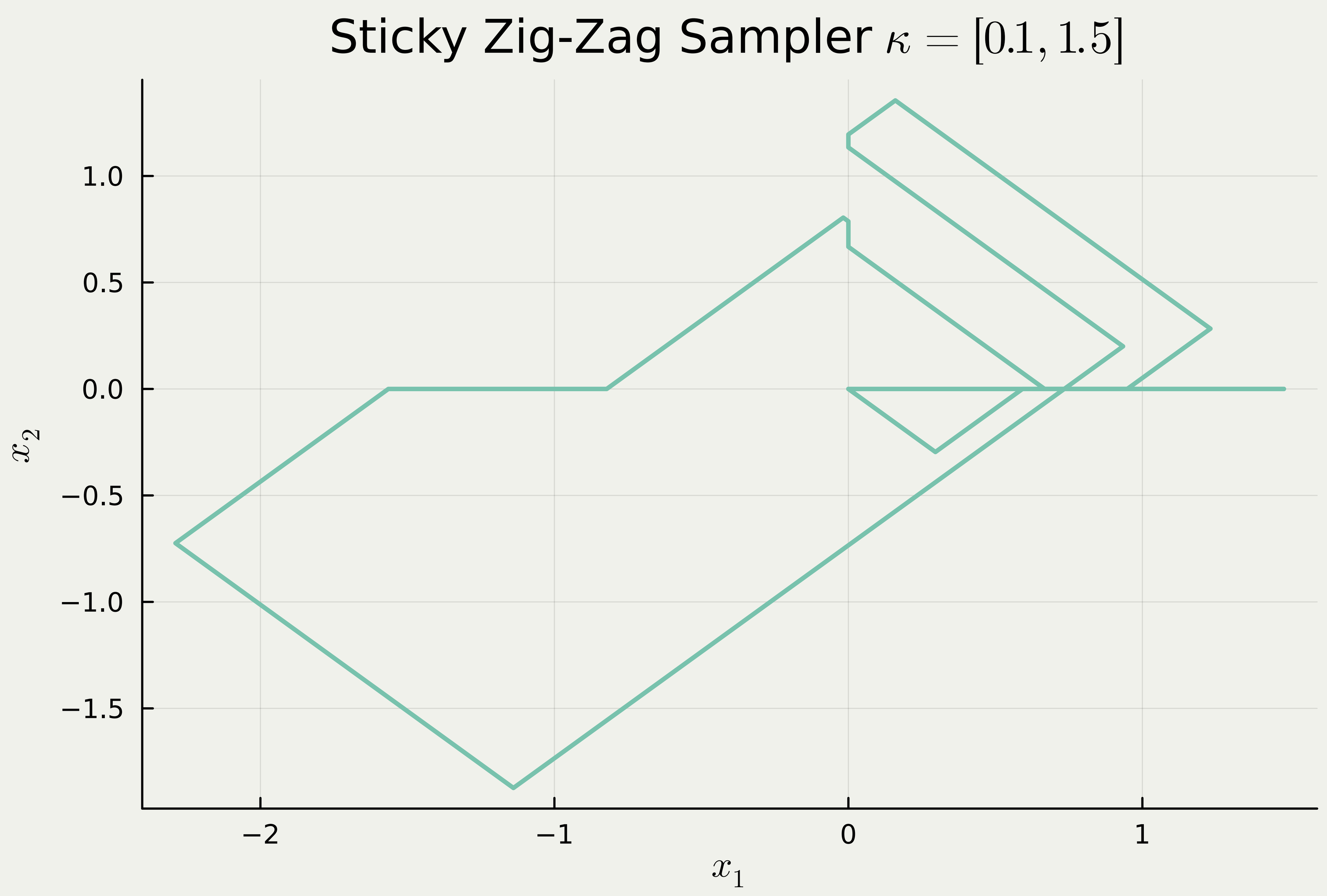
$$T_1, T_2, T_3, \cdots \sim \mathrm{Pois}(M^{(i)})$$

$$\frac{m(T_1)}{M(T_1)}, \frac{m(T_2)}{M(T_2)}, \frac{m(T_3)}{M(T_3)}, \cdots$$

の確率で採択

$$\cancel{T_1}, T_2, \cancel{T_3}, \cdots \sim \mathrm{Pois}(m^{(i)})$$

非絶対連続分布からのサンプリング



Bierkens, Grazzi, van der Meulen & Schauer (2023)

Spike-and-slab prior (Mitchell & Beauchamp, 1988)

$$p(dx) = \prod_{i=1}^d \left(\omega_i p_i(x_i)\,dx_i + (1-\omega_i)\delta_0(dx_i)\right)$$

という形の事前分布を用いると，事後包含確率 PIP

$$\mathrm{P}[\ i\ \text{番目の変数がモデルに入る}\ ] = \mathrm{P}[X_i = 0]$$

が求まる．

$$p(x|y)\,dx \propto Ce^{\ell(x)} \prod_{i=1}^d \left(p_i(x_i)\,dx_i + \kappa_i^{-1}\delta_0(dx_i)\right)$$

③ $x_i = 0$ となった瞬間，i 番目の変数を凍結する

➡ 部分空間を脱出するまでの時刻 $\tau \sim \mathrm{Exp}(\kappa)$ を計算