Manifold-valued Dirichlet Processes

Hyunwoo J. Kim, Jia Xu, Baba Vemuri, Vikas Singh

http://pages.cs.wisc.edu/~hwkim/projects/dp-mglm/

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Motivation

Regression

$$f : \mathbf{R}^d \to \mathcal{M}$$

$$\mathcal{M} = \mathrm{SPD}(n)$$

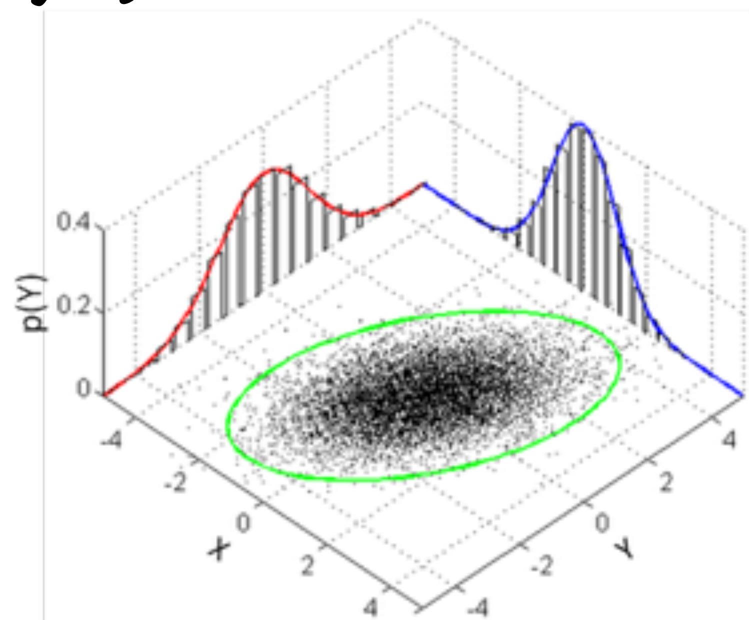
Motivation

Regression

$$f : \mathbf{R}^d \to \mathcal{M}$$

$$\mathcal{M} = \mathrm{SPD}(n)$$

- Covariance matrix

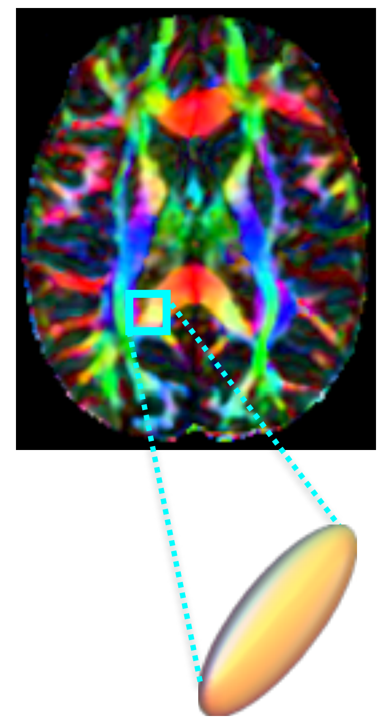


Motivation

Regression

$$f : \mathbf{R}^d \to \mathcal{M}$$

DTI



$$\mathcal{M} = \mathrm{SPD}(n)$$

- Covariance matrix
- Diffusion tensor imaging

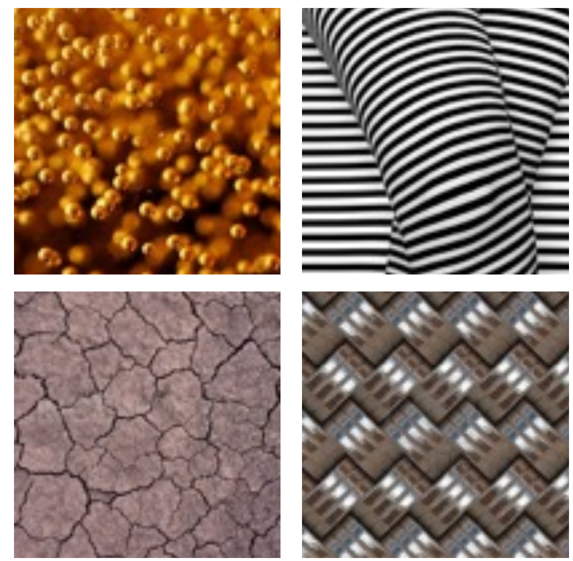
Motivation

Regression

$$f : \mathbf{R}^d \to \mathcal{M}$$

$$\mathcal{M} = \mathrm{SPD}(n)$$

- Covariance matrix
- Diffusion tensor imaging
- Region covariance
- …



M.Cimpoi et al., CVPR 2014

Linear regression on manifolds

$$f : R \to \mathcal{M} \qquad y_i = \mathrm{EXP}(\mathrm{EXP}(B, vx_i), \epsilon)$$

GR — Fletcher, IJCV 2013.

$$f : R^n \to \mathcal{M} \qquad y_i = \mathrm{EXP}(\mathrm{EXP}(B, \sum_{j=1}^{d} V^j x_i^j), \epsilon)$$

MGLM — Kim et al., CVPR 2014.

IR — Zhu et al., JASA, 2009.

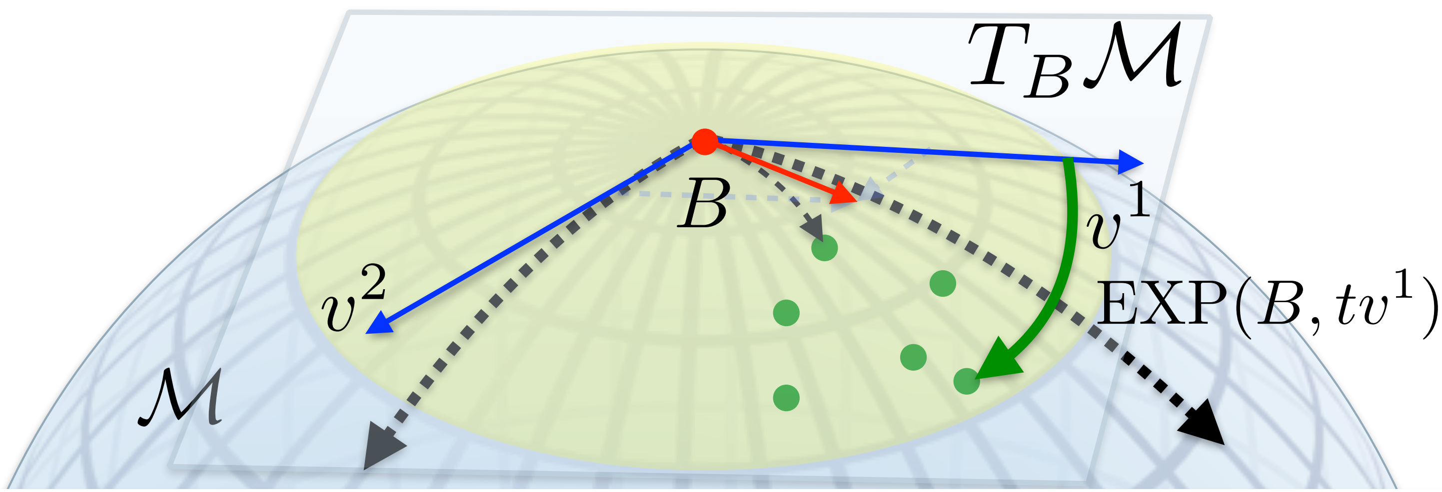
Motivations

Build models on manifolds

Locally defined parametric models

PGA (Fletcher et al., 2004), GR (Fletcher, IJCV 2013),

MGLM (Kim et al., CVPR 2014), RCCA (Kim et al., ECCV 2014)



Euclidean models

- General Linear Model

$$\boldsymbol{y}_i = \boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d + \epsilon$$

What if data have nonlinear correlation?

Euclidean models

- General Linear Model

$$\boldsymbol{y}_i = \boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d + \epsilon$$

- Generalized Linear Model (fixed link function)

$$\boldsymbol{y}_i = g^{-1}(\boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d) + \epsilon$$

- Single Index Model (searching link function)

$$\boldsymbol{y}_i = g^{-1}(\beta^0 + \beta^1 x_i^1 + \ldots + \beta^d x_i^d) + \epsilon$$

Euclidean models

- General Linear Model

$$\boldsymbol{y}_i = \boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d + \epsilon$$

- Generalized Linear Model (fixed link function)

$$\boldsymbol{y}_i = g^{-1}(\boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d) + \epsilon$$

- Single Index Model (searching link function)

$$\boldsymbol{y}_i = g^{-1}(\beta^0 + \beta^1 x_i^1 + \ldots + \beta^d x_i^d) + \epsilon$$

Link functions for manifold-valued response?

Euclidean models

- GLM

$$\boldsymbol{y}_i = \boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d + \epsilon$$

- ?-GLM

$$\boldsymbol{y}_i = \underline{\boldsymbol{\beta}_{\color{red}i}}^0 + \underline{\boldsymbol{\beta}_{\color{red}i}}^1 x_i^1 + \ldots + \underline{\boldsymbol{\beta}_{\color{red}i}}^d x_i^d + \epsilon$$

Euclidean models

- GLM

$$\boldsymbol{y}_i = \boldsymbol{\beta}^0 + \boldsymbol{\beta}^1 x_i^1 + \ldots + \boldsymbol{\beta}^d x_i^d + \epsilon$$

- DP-GLM

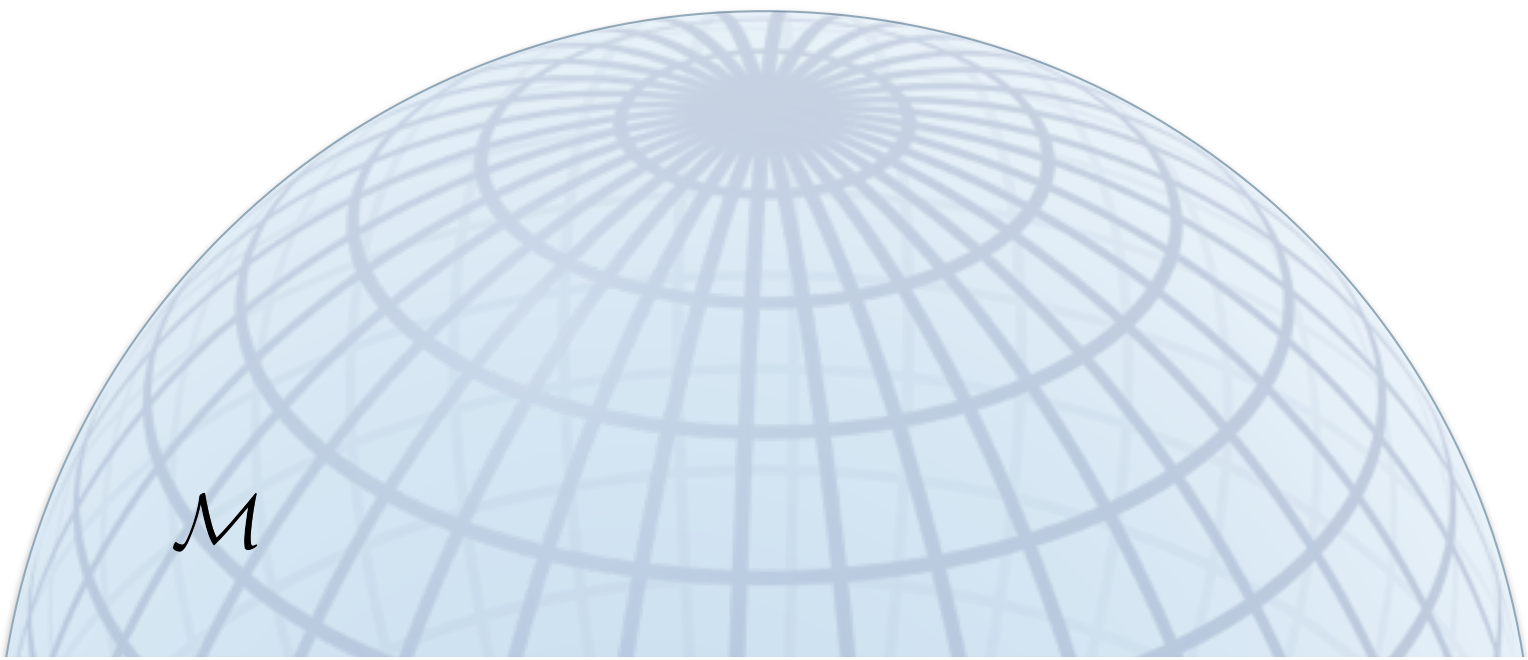
$$\boldsymbol{y}_i = \underline{\boldsymbol{\beta}_{\boldsymbol{i}}}^0 + \underline{\boldsymbol{\beta}_{\boldsymbol{i}}}^1 x_i^1 + \ldots + \underline{\boldsymbol{\beta}_{\boldsymbol{i}}}^d x_i^d + \epsilon$$

$$(x_i, y_i)|\theta_i \sim F(\theta_i), \theta_i|G \sim G, G \sim DP(G_0, \nu)$$

Hannah et al., JMLR 2011

MGLM on manifolds

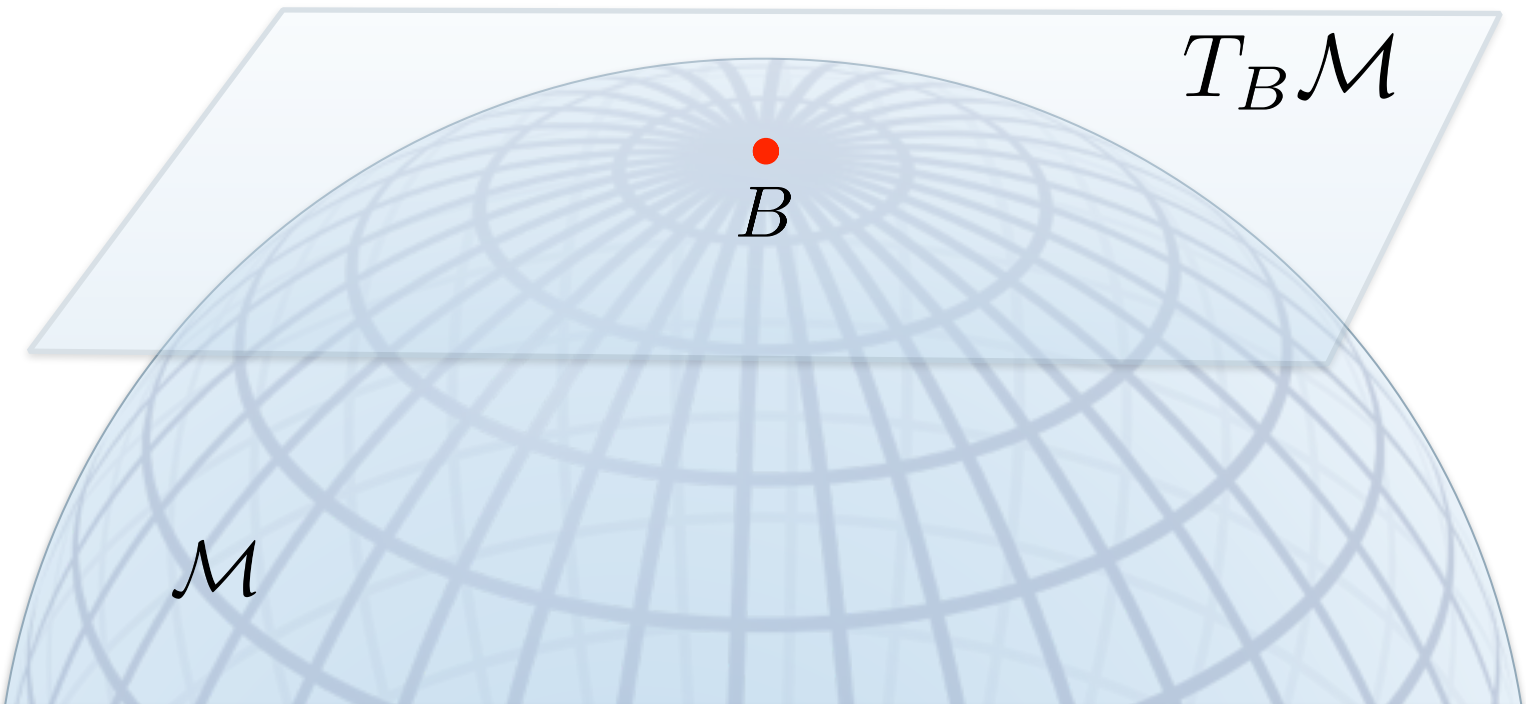
$$y_i = \mathrm{EXP}(\mathrm{EXP}(B, \sum_{j=1}^{d} V^j x_i^j), \epsilon)$$



Kim. et al., CVPR 2014

MGLM on manifolds

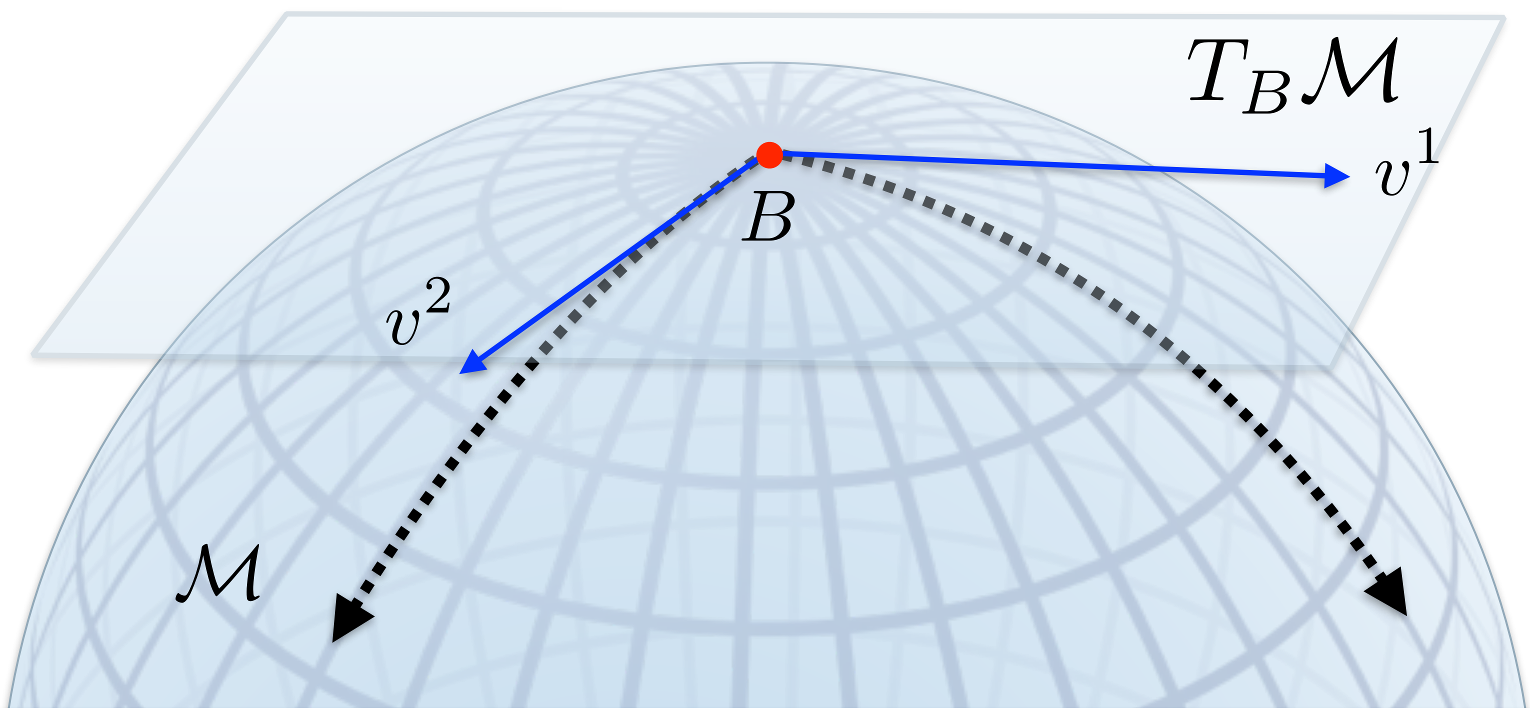
$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B}, \sum_{j=1}^{d} V^j x_i^j), \epsilon)$$



Kim. et al., CVPR 2014

MGLM on manifolds

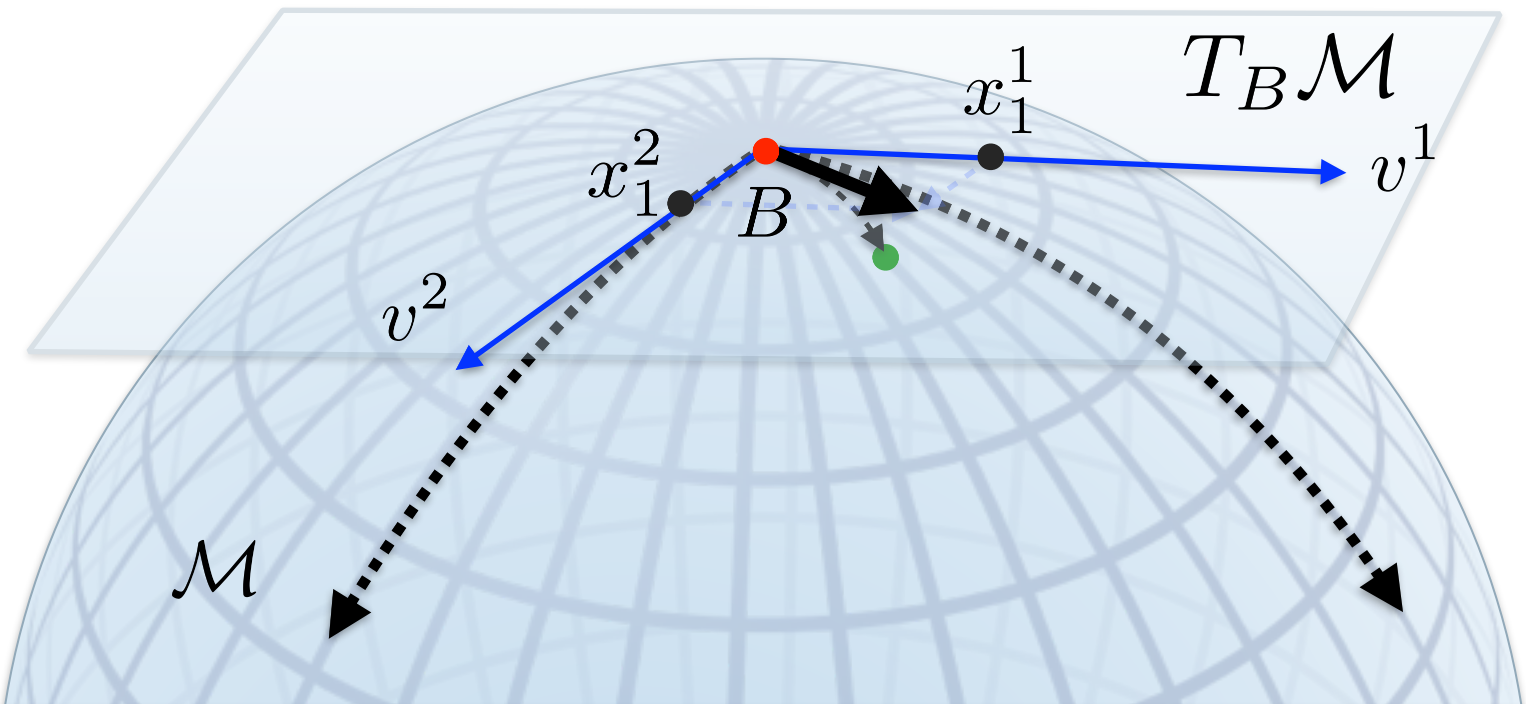
$$y_i = \text{EXP}(\text{EXP}(\underline{B}, \sum_{j=1}^{d} \underline{V^j} x_i^j), \epsilon)$$



Kim. et al., CVPR 2014

MGLM on manifolds

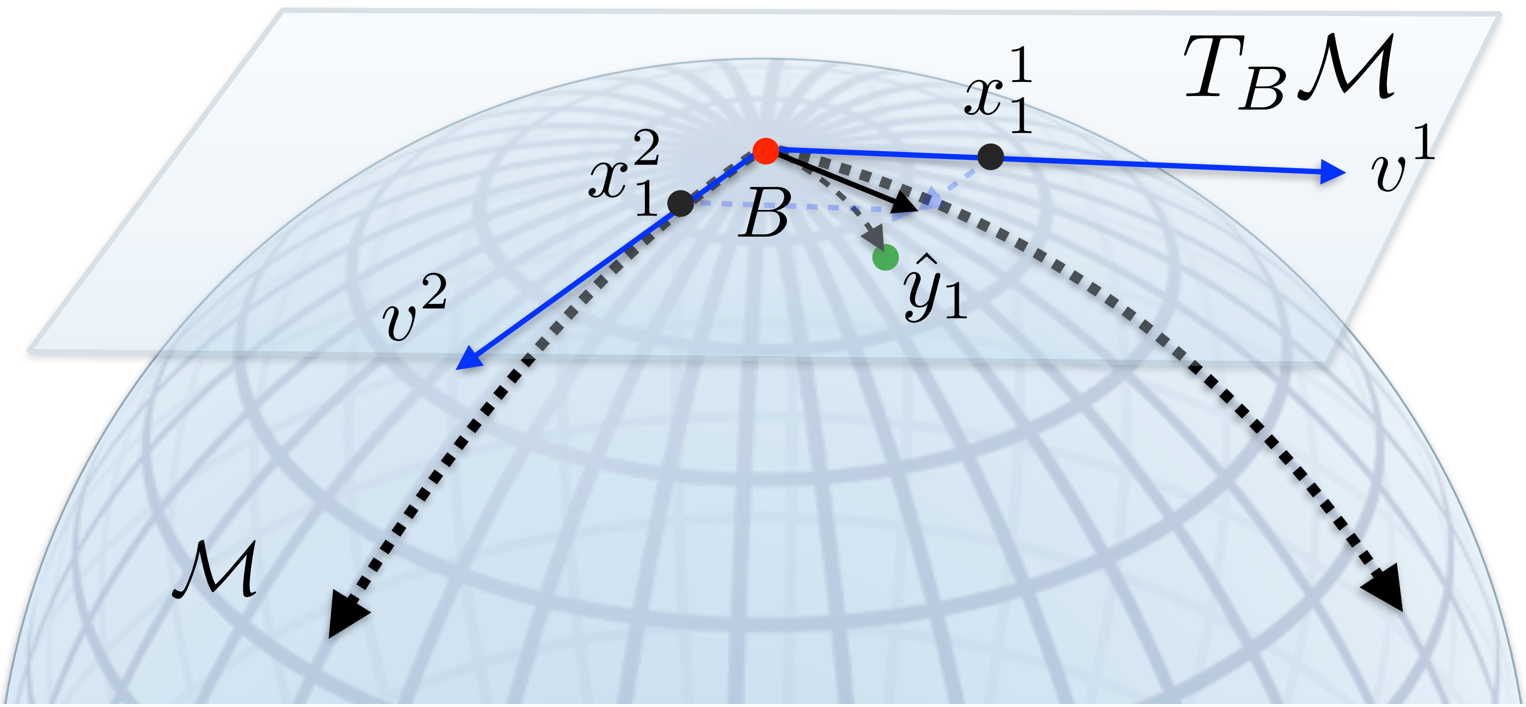
$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B}, \underline{\sum_{j=1}^{d} V^j x_i^j}), \epsilon)$$



Kim. et al., CVPR 2014

MGLM on manifolds

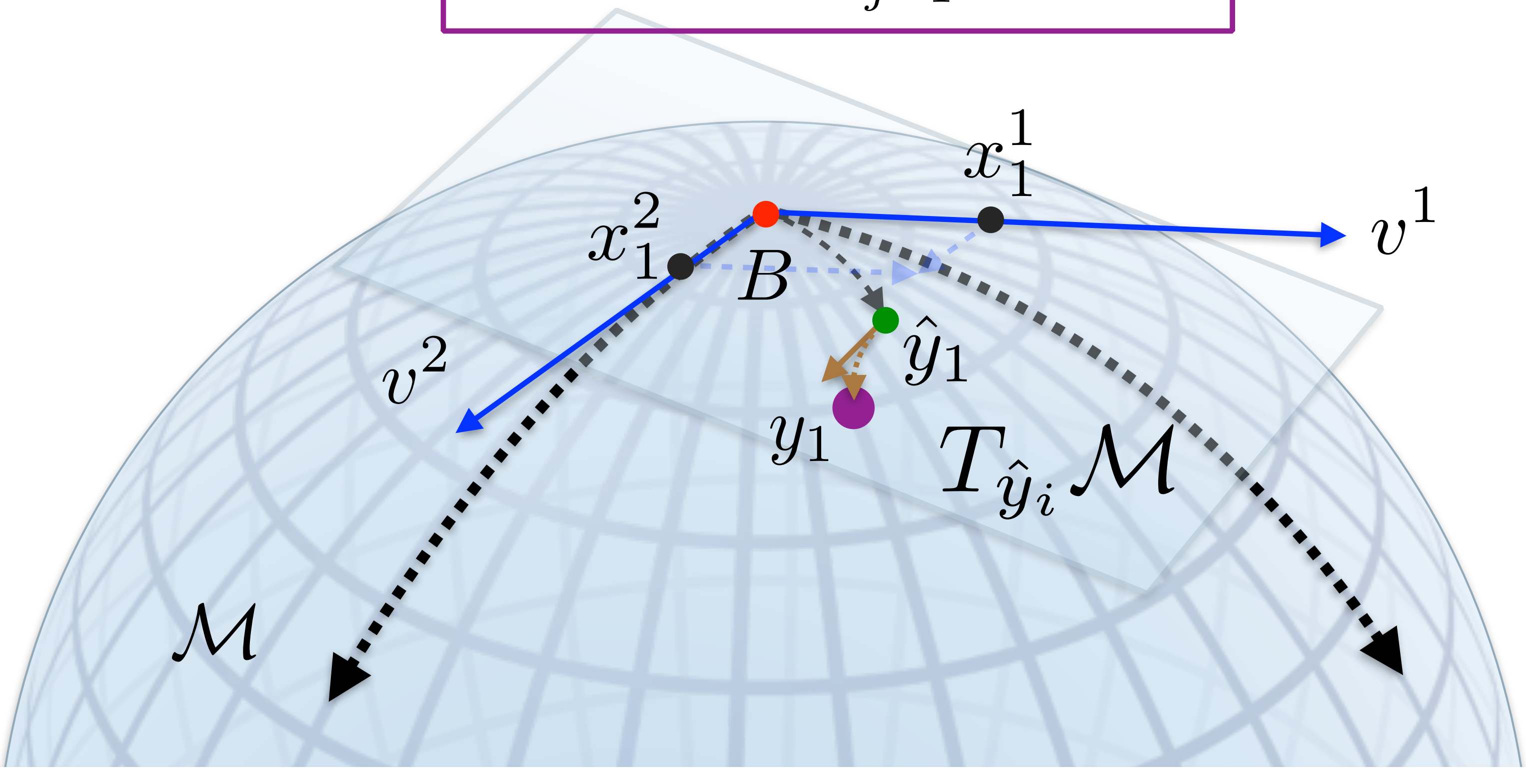
$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B}, \sum_{j=1}^{d} V^j x_i^j), \epsilon)$$



Kim. et al., CVPR 2014

MGLM on manifolds

$$y_i = \boxed{\mathrm{EXP}(\mathrm{EXP}(B, \sum_{j=1}^{d} V^j x_i^j), \epsilon)}$$

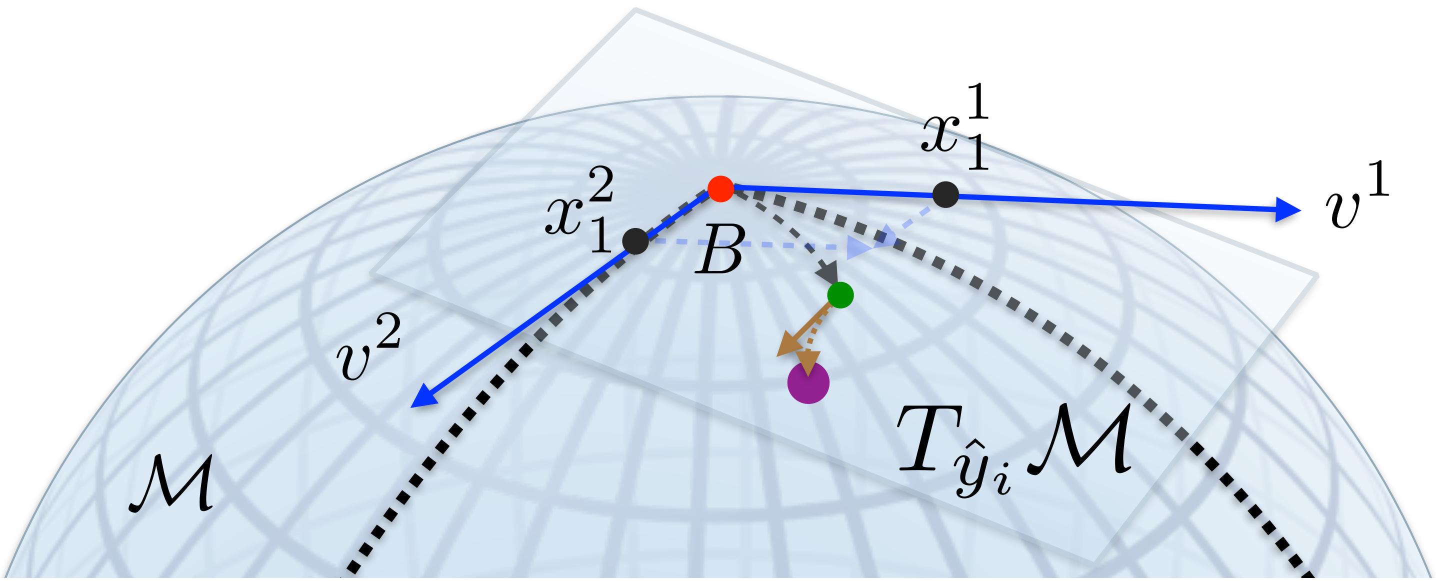


Kim. et al., CVPR 2014

MGLM on manifolds

$$\underline{y_i} = \text{EXP}(\text{EXP}(\underline{B}, \sum_{j=1}^{d} \underline{V^j} x_i^j), \underline{\epsilon})$$

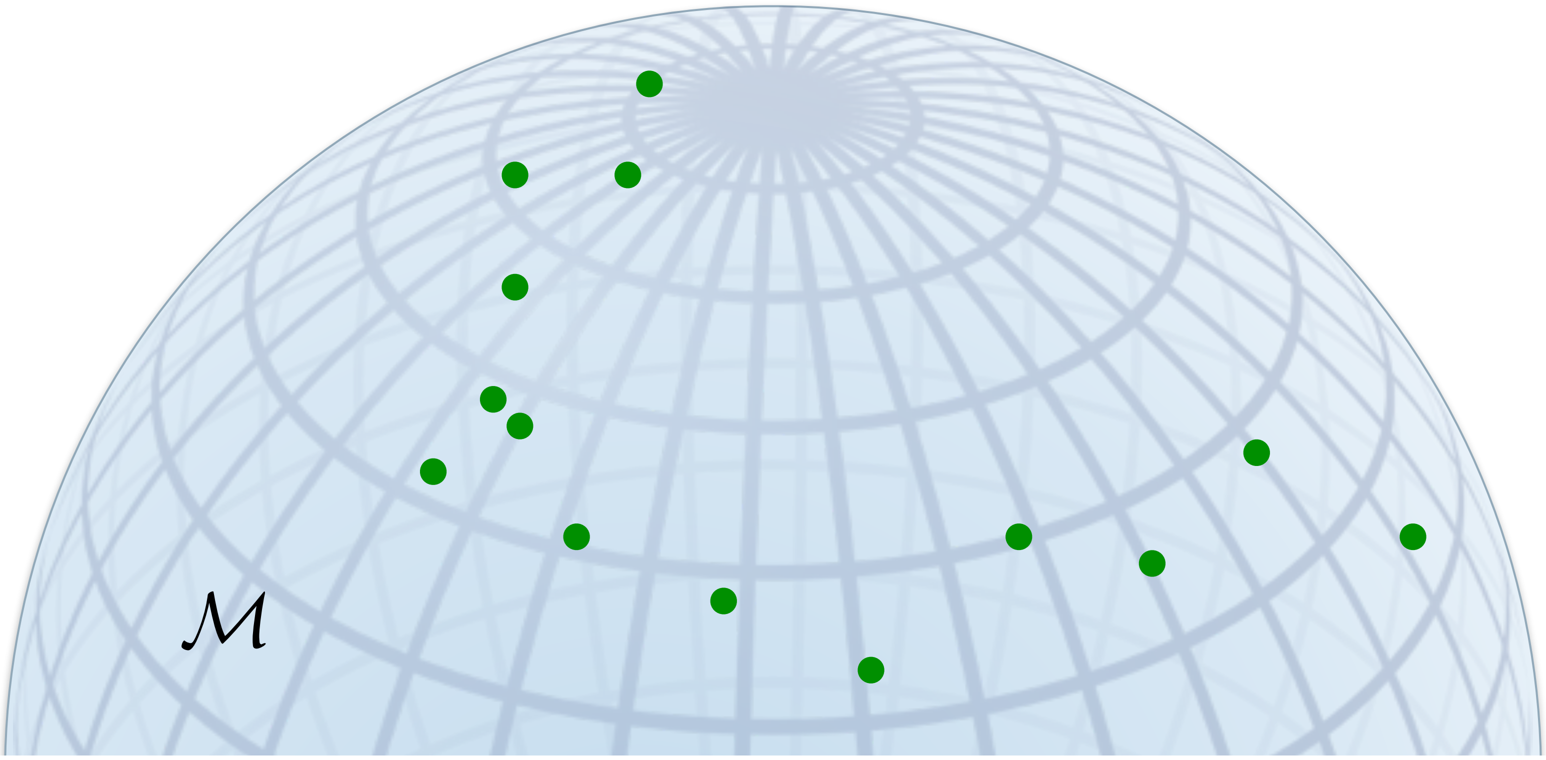
$$\underline{y_i} = \underline{\beta^0} + \underline{\beta^1} x_i^1 + \underline{\beta^2} x_i^2 + \ldots + \underline{\beta^d} x_i^d + \underline{\epsilon}$$



Kim. et al., CVPR 2014

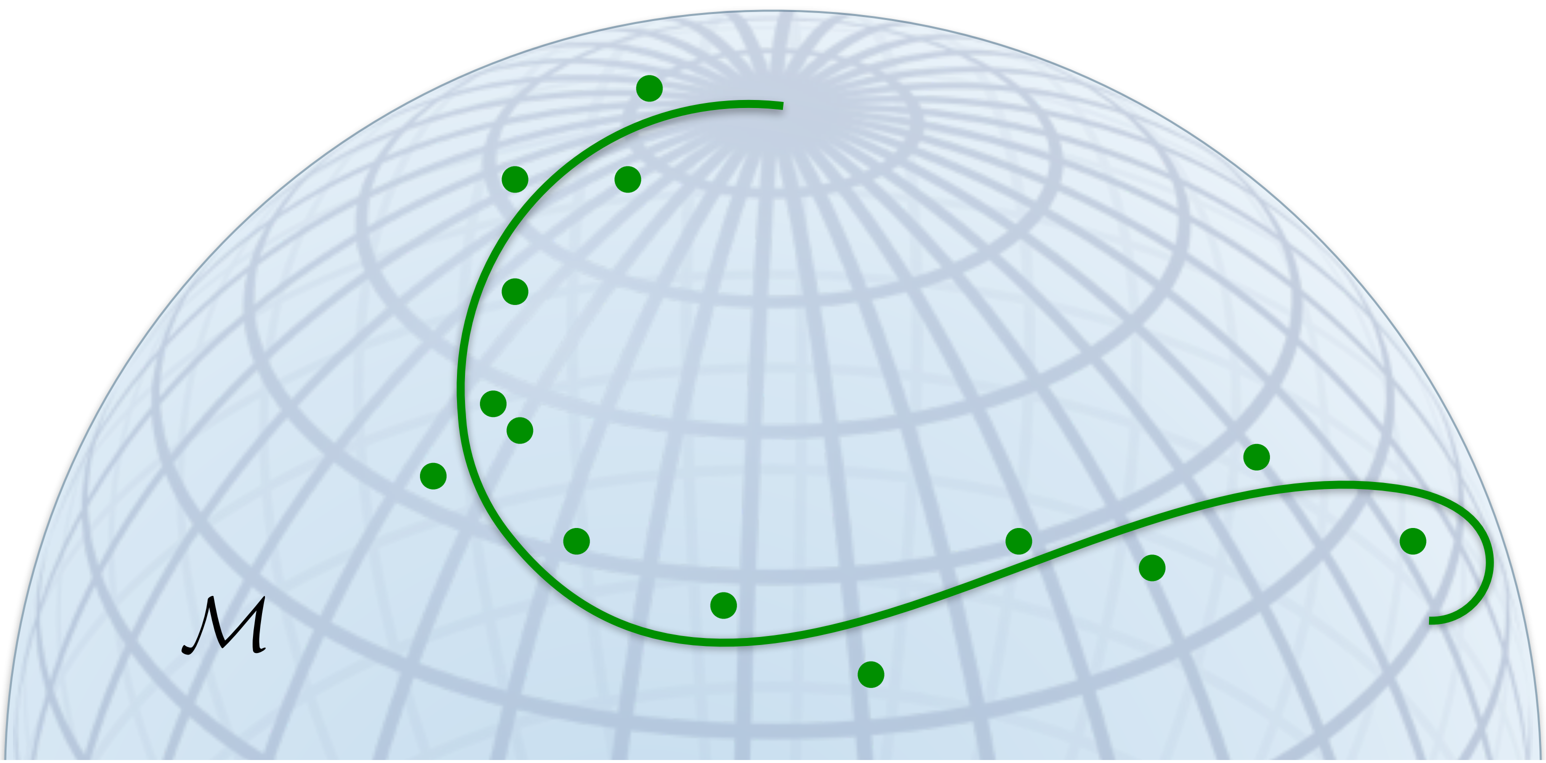
DP-MGLM

$$y_i = \mathrm{EXP}(\mathrm{EXP}(B_i, \sum_{j=1}^{d} V_i^j x_i^j), \epsilon)$$



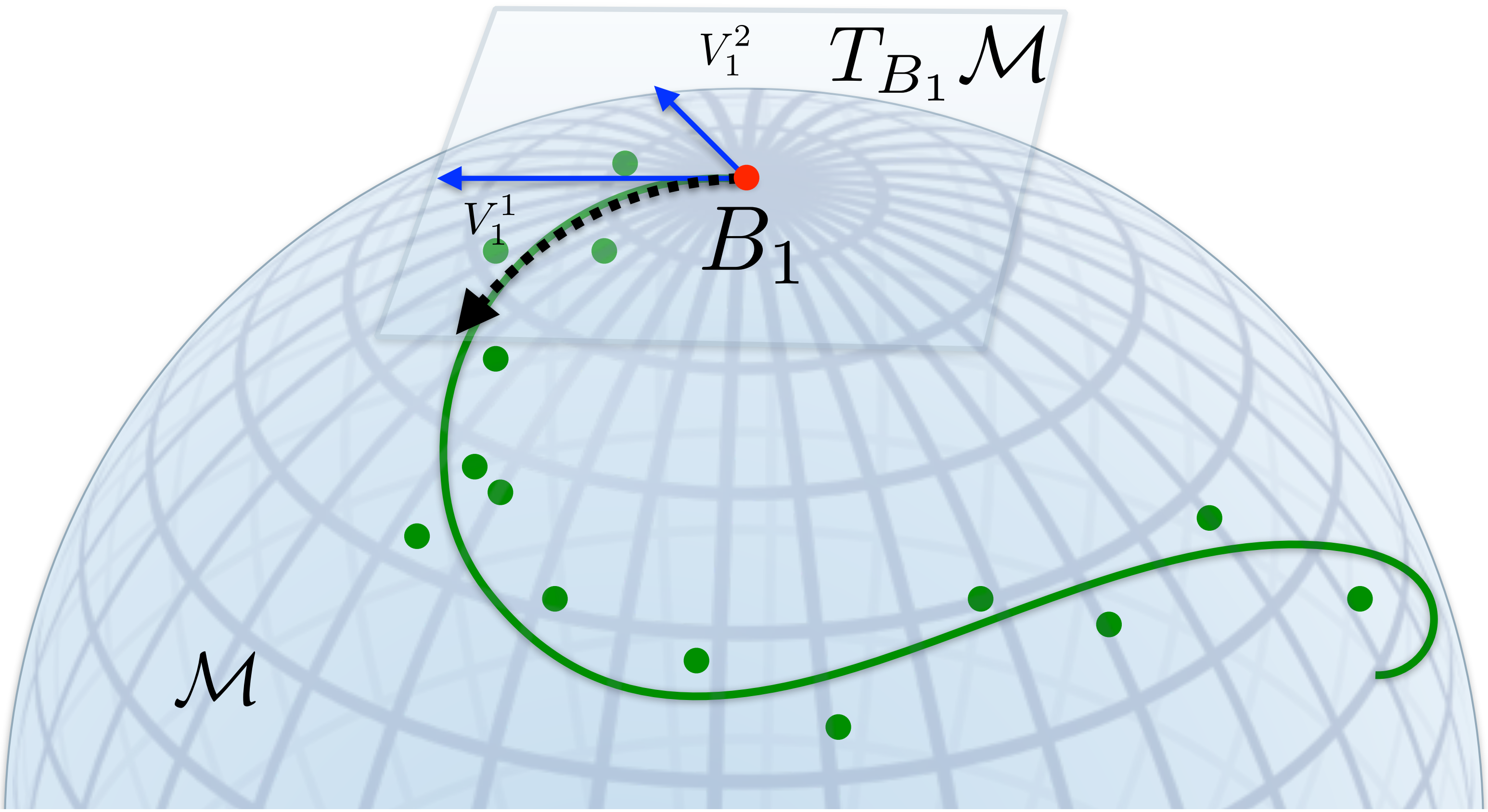
DP-MGLM

$$y_i = \mathrm{EXP}(\mathrm{EXP}(B_i, \sum_{j=1}^{d} V_i^j x_i^j), \epsilon)$$



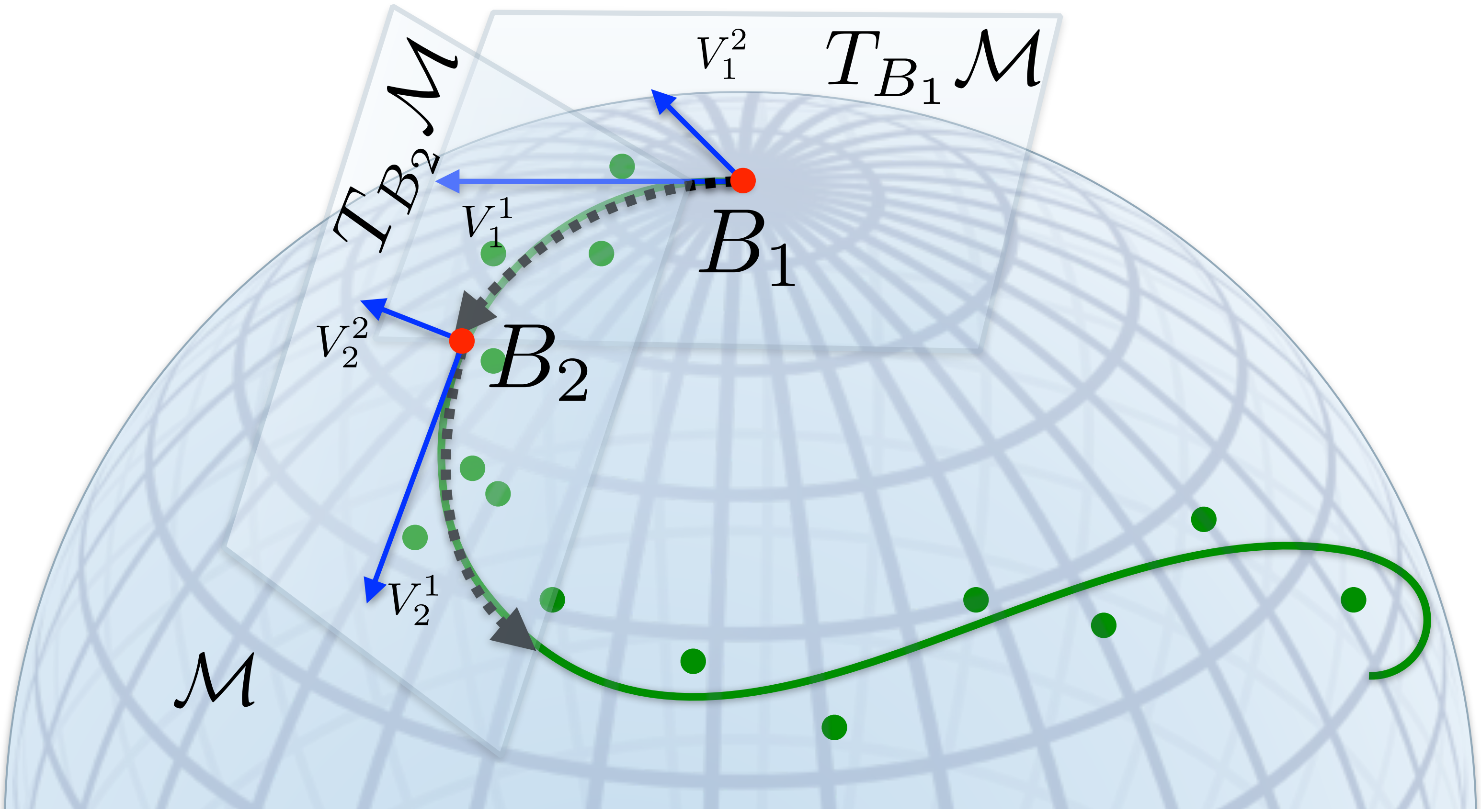
DP-MGLM

$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B_i}, \sum_{j=1}^{d} \underline{V_i^j} x_i^j), \epsilon)$$



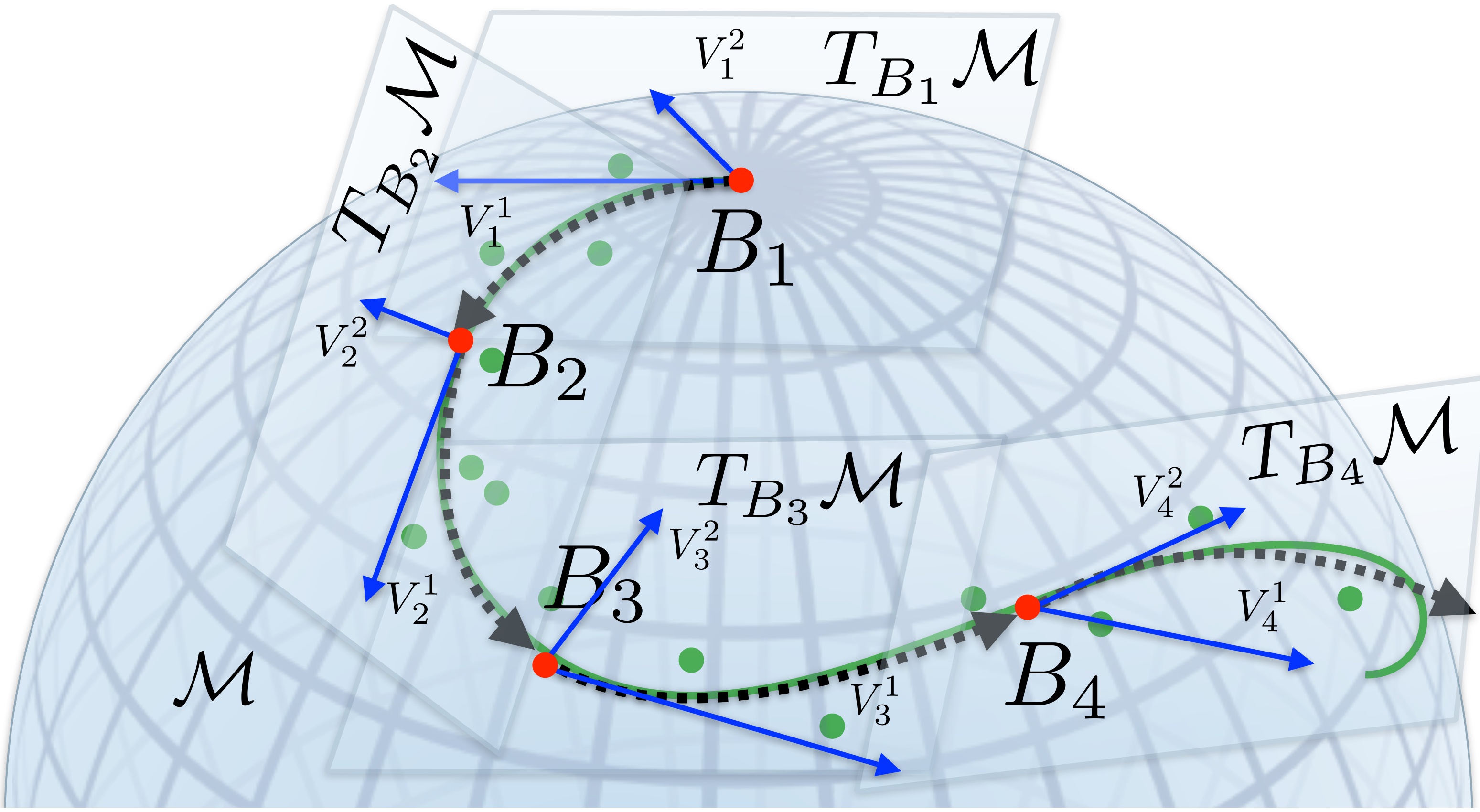
DP-MGLM

$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B_i}, \sum_{j=1}^{d} \underline{V_i^j} x_i^j), \epsilon)$$



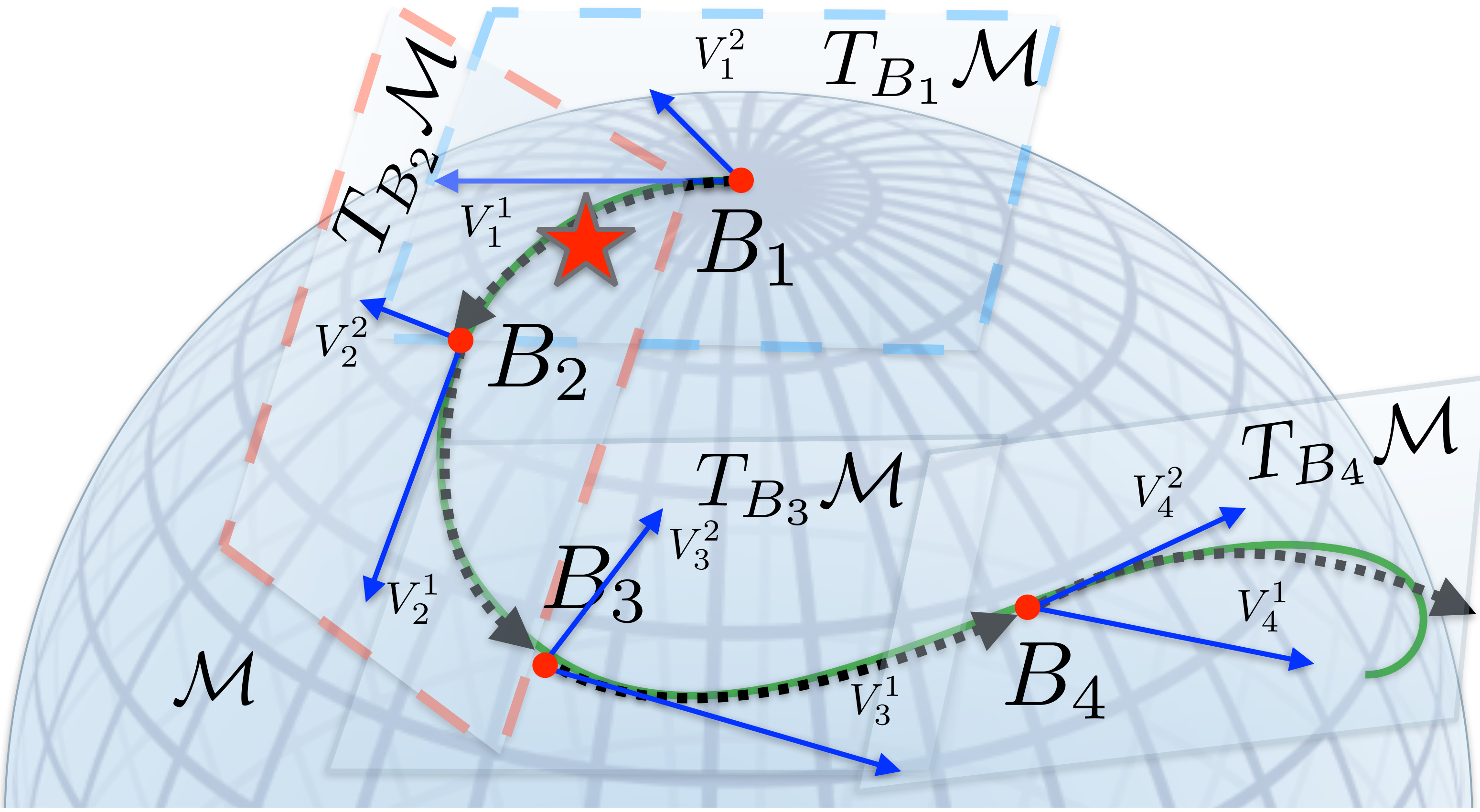
DP-MGLM

$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B_i}, \sum_{j=1}^{d} \underline{V_i^j} x_i^j), \epsilon)$$



DP-MGLM

$$y_i = \mathrm{EXP}(\mathrm{EXP}(\underline{B_i}, \sum_{j=1}^{d} \underline{V_i^j} x_i^j), \epsilon)$$



DP-GLM

$$
\begin{aligned}
G &\sim DP(G_0, \nu) \\
\theta_i = (\theta_{x_i}, \theta_{y_i})|G &\sim G, \\
x_i|\theta_{x_i} &\sim f_x(\theta_{x_i}), \\
y_i|x_i, \theta_{y_i} &\sim GLM(x_i, \theta_{y_i}),
\end{aligned}
$$

Hannah et al. JMLR 2011

DP-GLM

Manifold-valued?

$y_i \in \mathcal{M}$

$$G \sim DP(G_0, \nu)$$

$$\theta_i = (\theta_{x_i}, \theta_{y_i})|G \sim G,$$

$$x_i|\theta_{x_i} \sim f_x(\theta_{x_i}),$$

$$y_i|x_i, \theta_{y_i} \sim GLM(x_i, \theta_{y_i}),$$

Hannah et al. JMLR 2011

DP-MGLM

$$G \sim DP(G_0, \nu)$$

$$\theta_i = (\theta_{x_i}, \theta_{y_i})|G \sim G,$$

MGLM on manifolds

$$x_i|\theta_{x_i} \sim f_x(\theta_{x_i}),$$

$$y_i|x_i, \theta_{y_i} \sim MGLM(x_i, \theta_{y_i}),$$

DP-MGLM

$$G \sim DP(G_0, \nu)$$

$$\theta_i = (\theta_{x_i}, \theta_{y_i})|G \sim G,$$

$$x_i|\theta_{x_i} \sim f_x(\theta_{x_i}),$$

$$y_i|x_i, \theta_{y_i} \sim MGLM(x_i, \theta_{y_i}),$$

manifold-valued parameters?

$$\theta_{y_i} \in \mathrm{SPD}(n) \times \mathrm{Sym}(n)^d$$

DP-MGLM

- Distribution on manifolds
- Intrinsic metric

$$G \sim DP(G_0, \nu)$$

$$\theta_i = (\theta_{x_i}, \theta_{y_i}) | G \sim G,$$

$$x_i | \theta_{x_i} \sim f_x(\theta_{x_i}),$$

$$y_i | x_i, \theta_{y_i} \sim MGLM(x_i, \theta_{y_i}),$$

$$\theta_{y_i} \in \mathrm{SPD}(n) \times \mathrm{Sym}(n)^d$$

DP-MGLM

Hamiltonian Monte Carlo (HMC) sampling

$$G \sim DP(G_0, \nu)$$

$$\theta_i = (\theta_{x_i}, \theta_{y_i})|G \sim G,$$

$$x_i|\theta_{x_i} \sim f_x(\theta_{x_i}),$$

$$y_i|x_i, \theta_{y_i} \sim MGLM(x_i, \theta_{y_i}),$$

$$\theta_{y_i} \in \mathrm{SPD}(n) \times \mathrm{Sym}(n)^d$$

HMC
Hamiltonian/Hybrid Monte Carlo

$$H(q, p) = U(q) + K(p)$$

Duane, S., et al, “Hybrid Monte Carlo”, 1987.

Neal, R. “MCMC using Hamiltonian dynamics”, 2011

HMC
Hamiltonian/Hybrid Monte Carlo

$$H(q,p) = U(q) + K(p)$$

$$q = \theta \in \boldsymbol{R}^d$$

$$p = \dot{\theta} \in \boldsymbol{R}^d$$

Duane, S., et al, “Hybrid Monte Carlo”, 1987.

Neal, R. “MCMC using Hamiltonian dynamics”, 2011

HMC
Hamiltonian/Hybrid Monte Carlo

$$H(q,p) = U(q) + K(p)$$

$$U(q) := -\log f(q)$$

$$K(p) := \frac{1}{2}p^T M^{-1} p$$

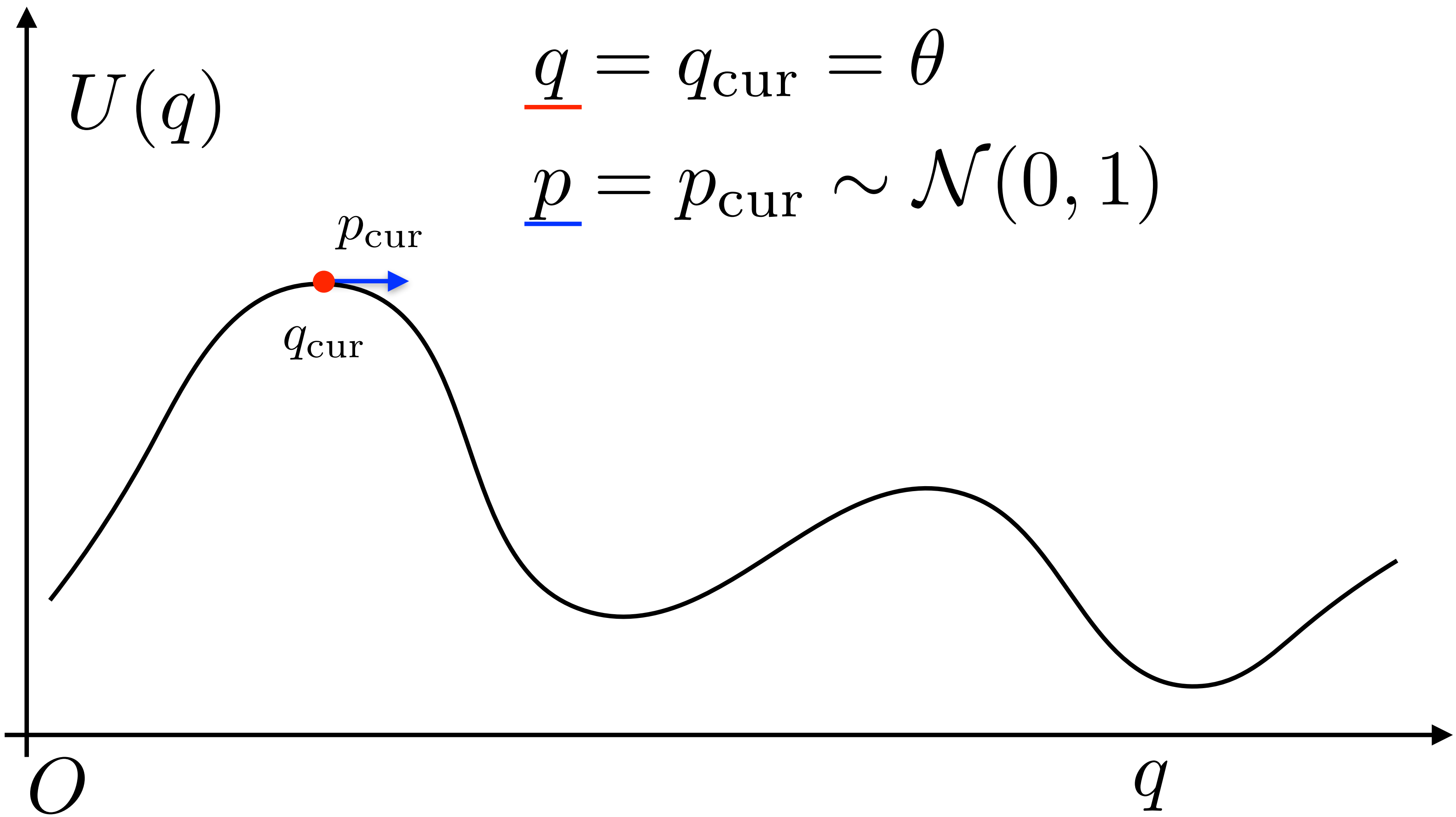
Duane, S., et al, “Hybrid Monte Carlo”, 1987.

Neal, R. “MCMC using Hamiltonian dynamics”, 2011

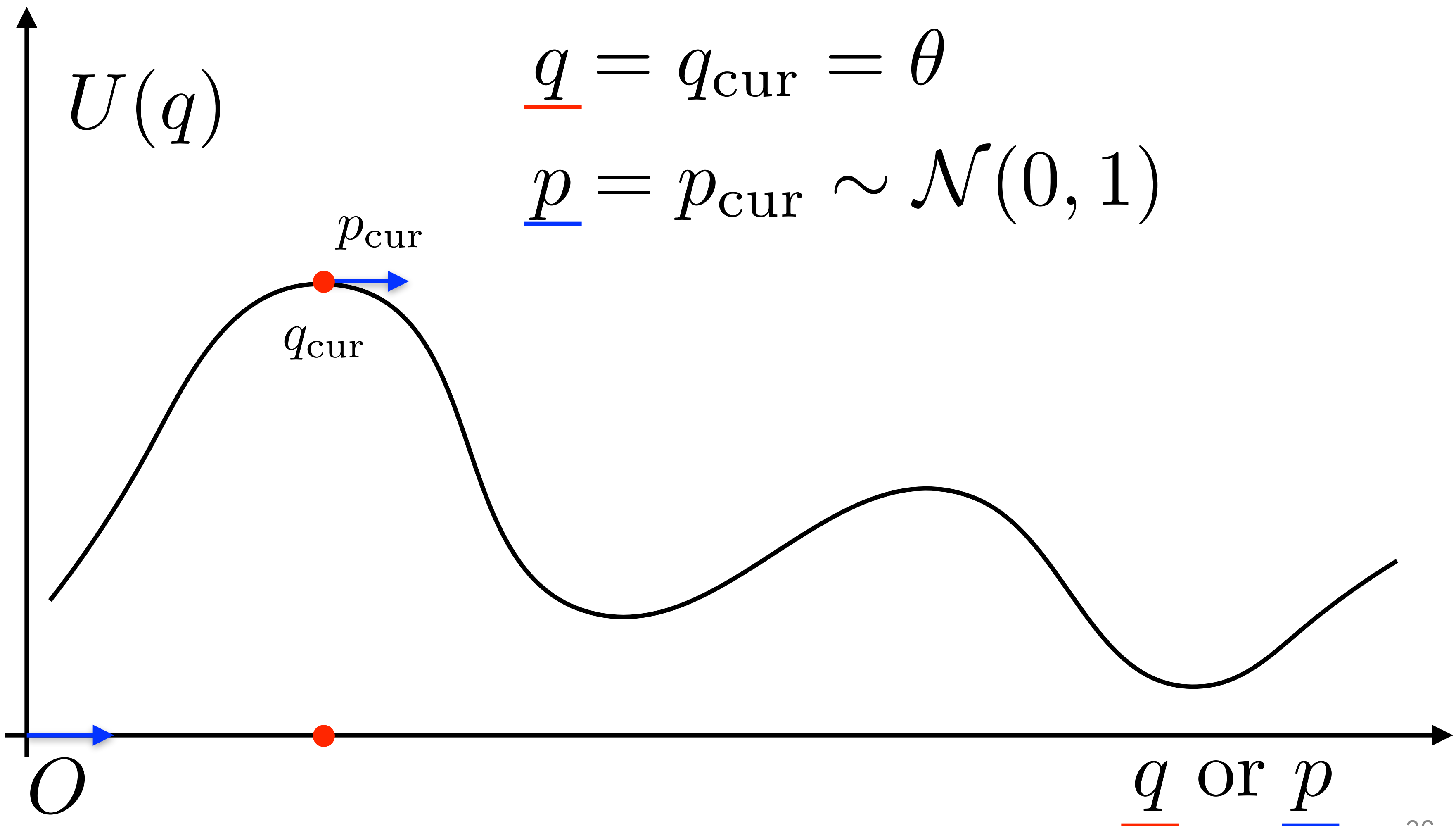
HMC
Hamiltonian/Hybrid Monte Carlo

$$q = q_{\rm cur} = \theta$$

$$p = p_{\rm cur} \sim \mathcal{N}(0, 1)$$

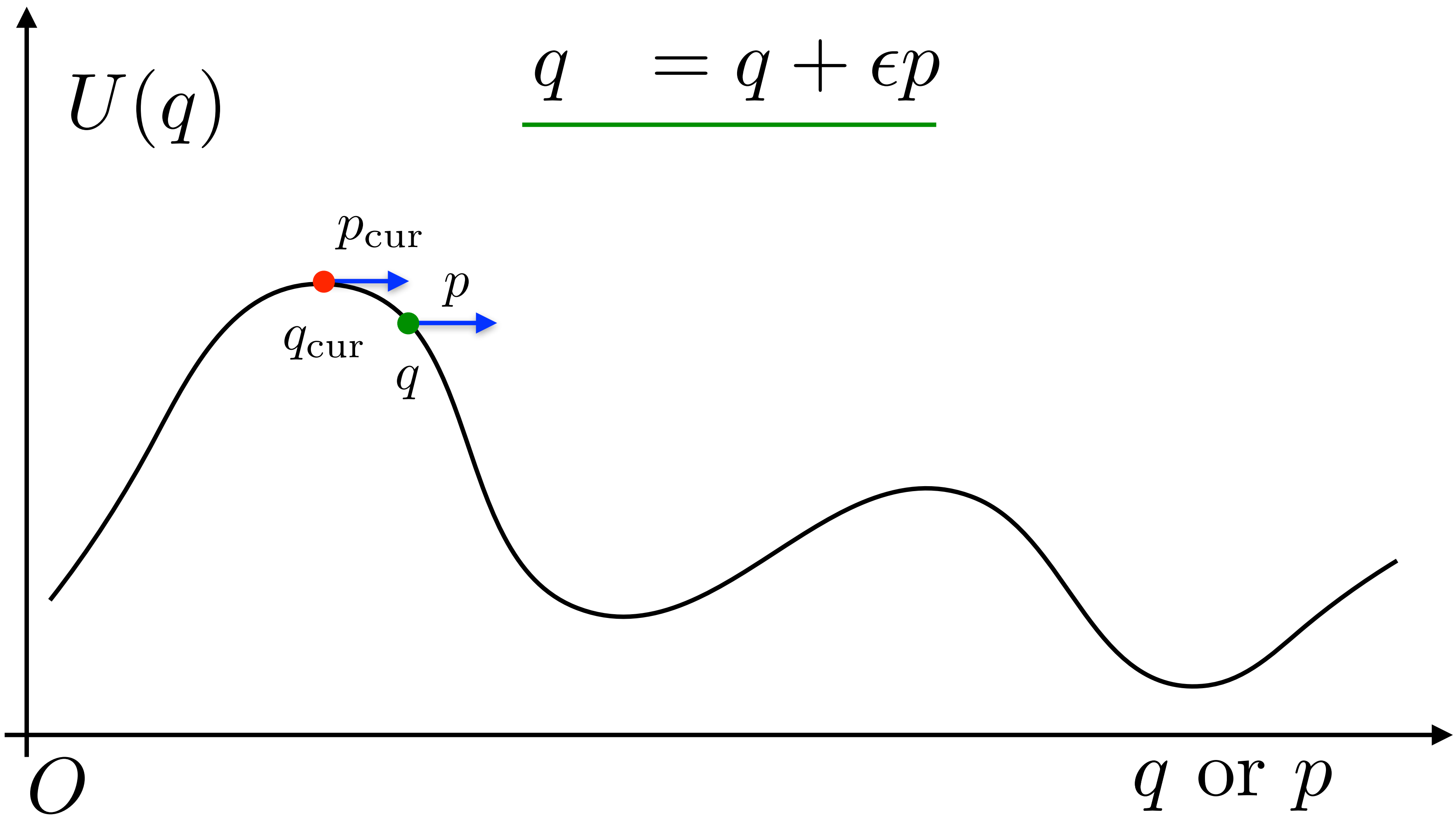


HMC
Hamiltonian/Hybrid Monte Carlo



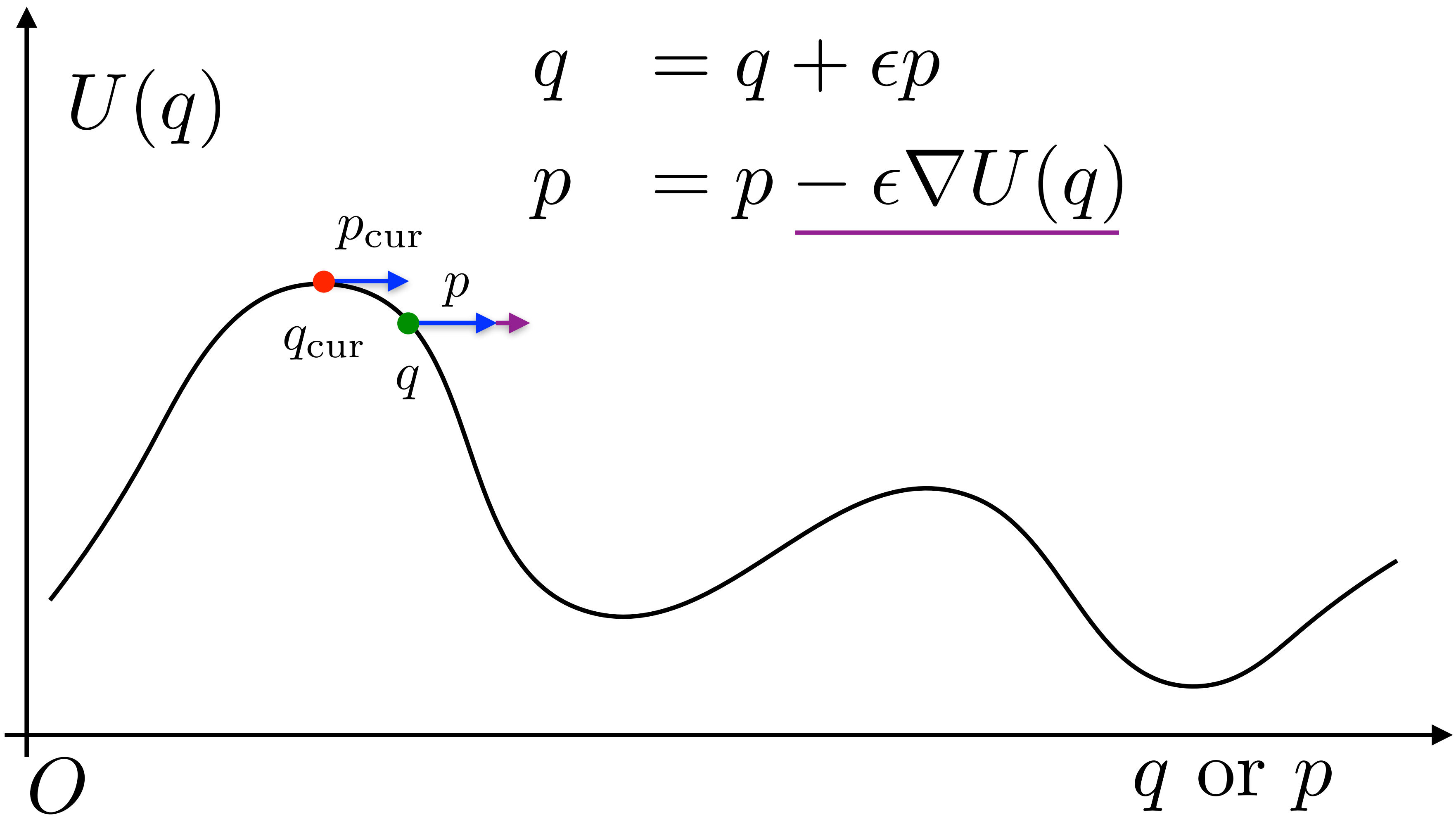
HMC
Hamiltonian/Hybrid Monte Carlo

$$q \quad = q + \epsilon p$$



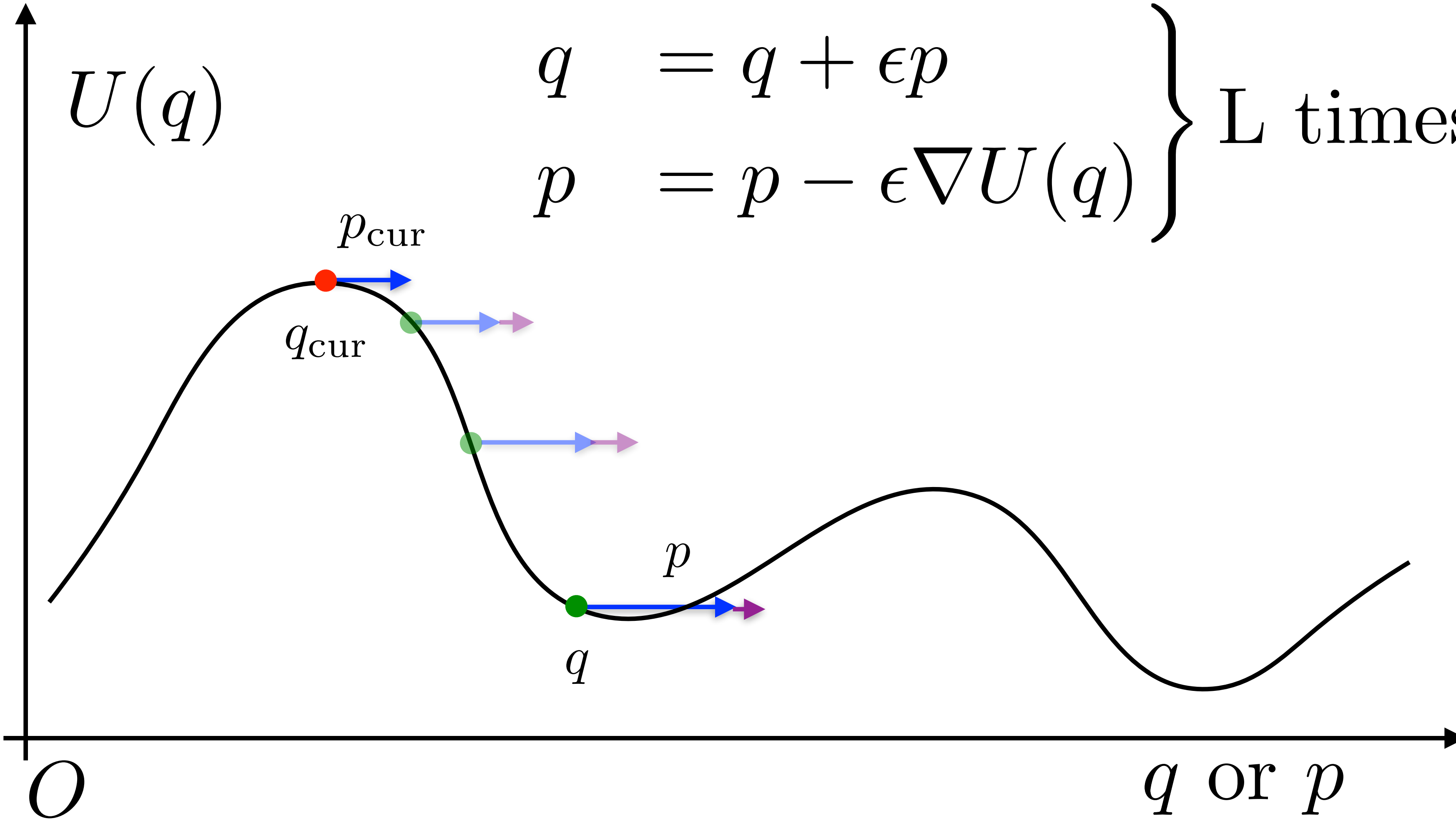
HMC
Hamiltonian/Hybrid Monte Carlo

$$q = q + \epsilon p$$
$$p = p - \epsilon \nabla U(q)$$

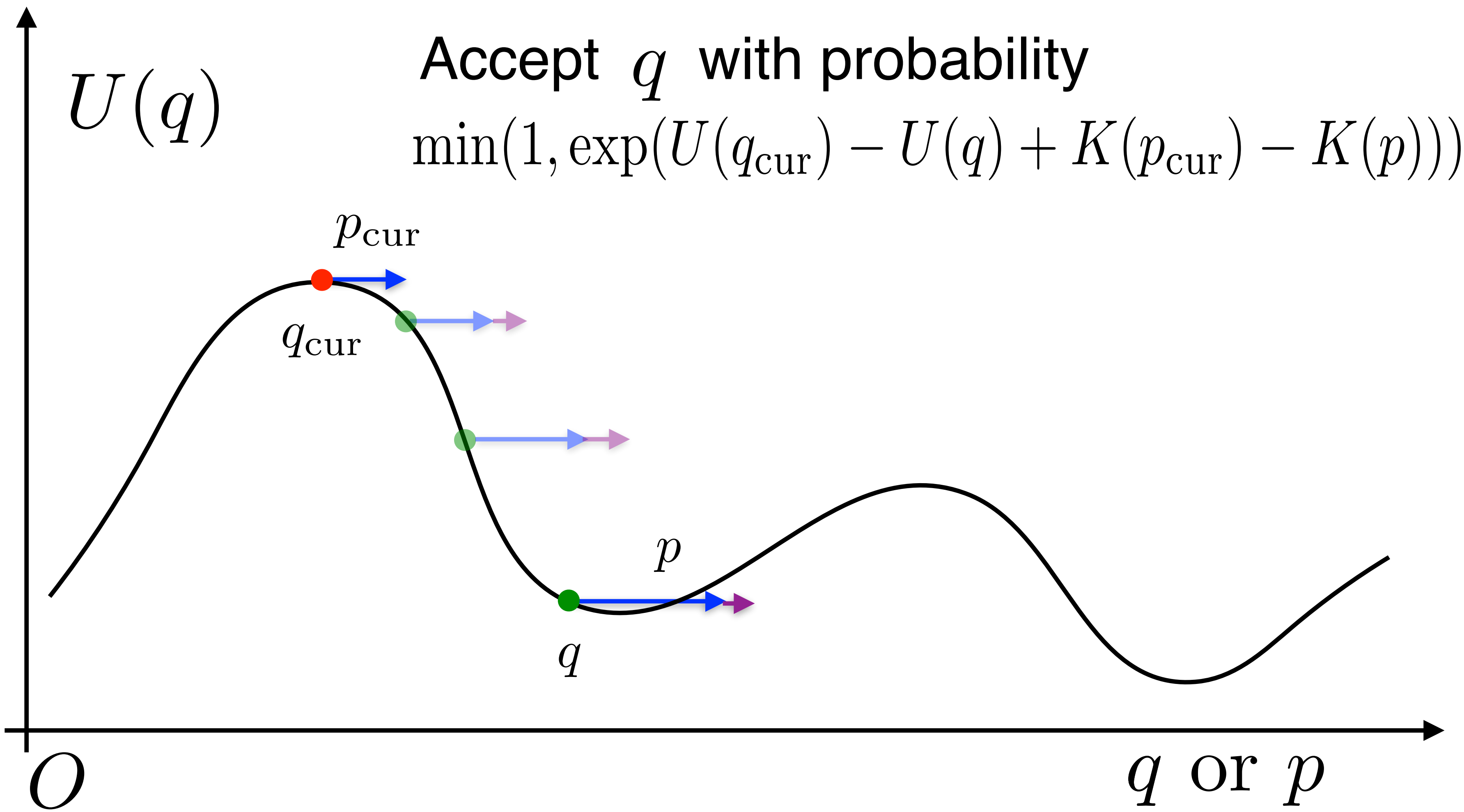


HMC
Hamiltonian/Hybrid Monte Carlo

$$\left.\begin{aligned} q &= q + \epsilon p \\ p &= p - \epsilon \nabla U(q) \end{aligned}\right\} \text{L times}$$

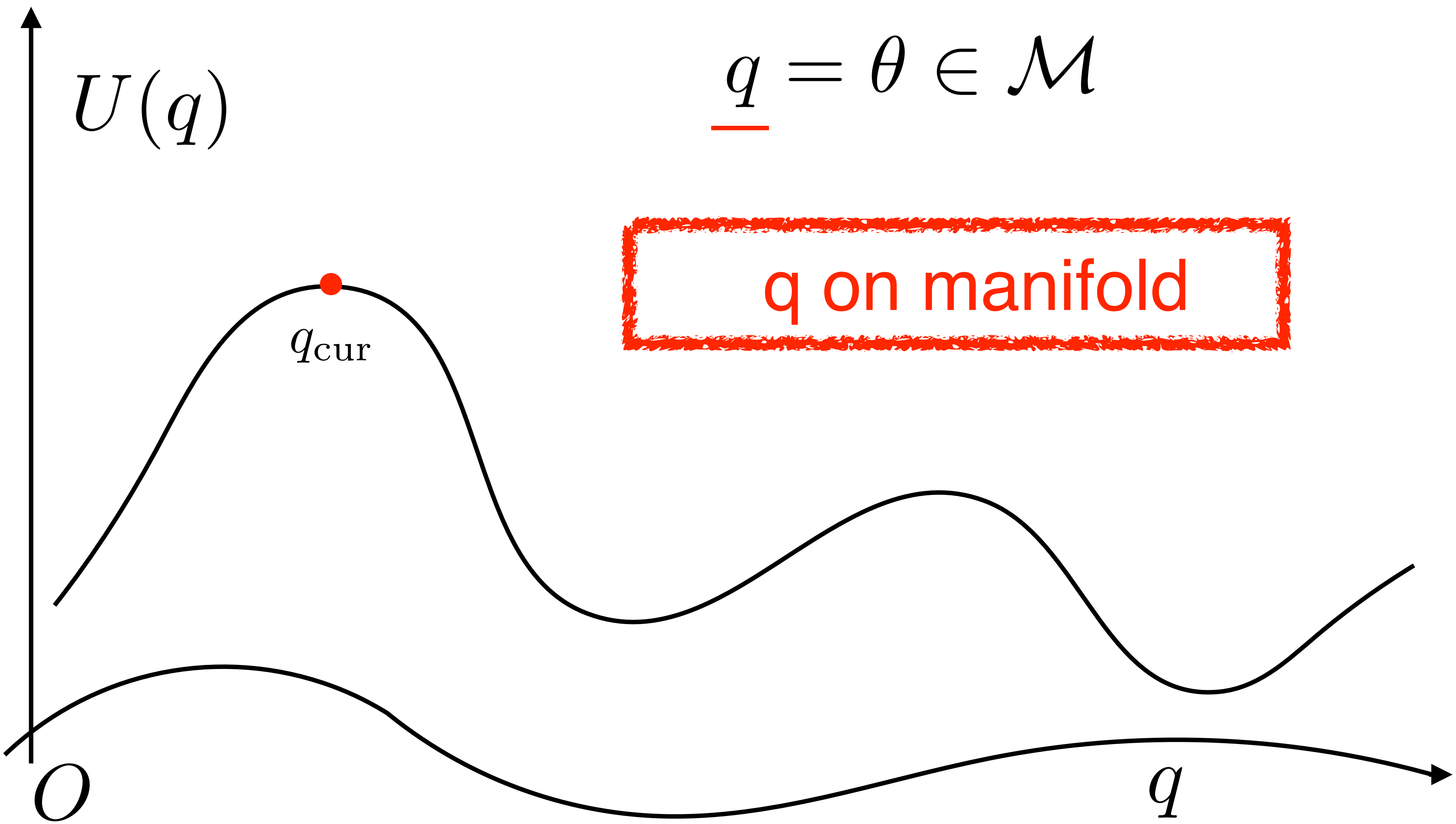


HMC
Hamiltonian/Hybrid Monte Carlo



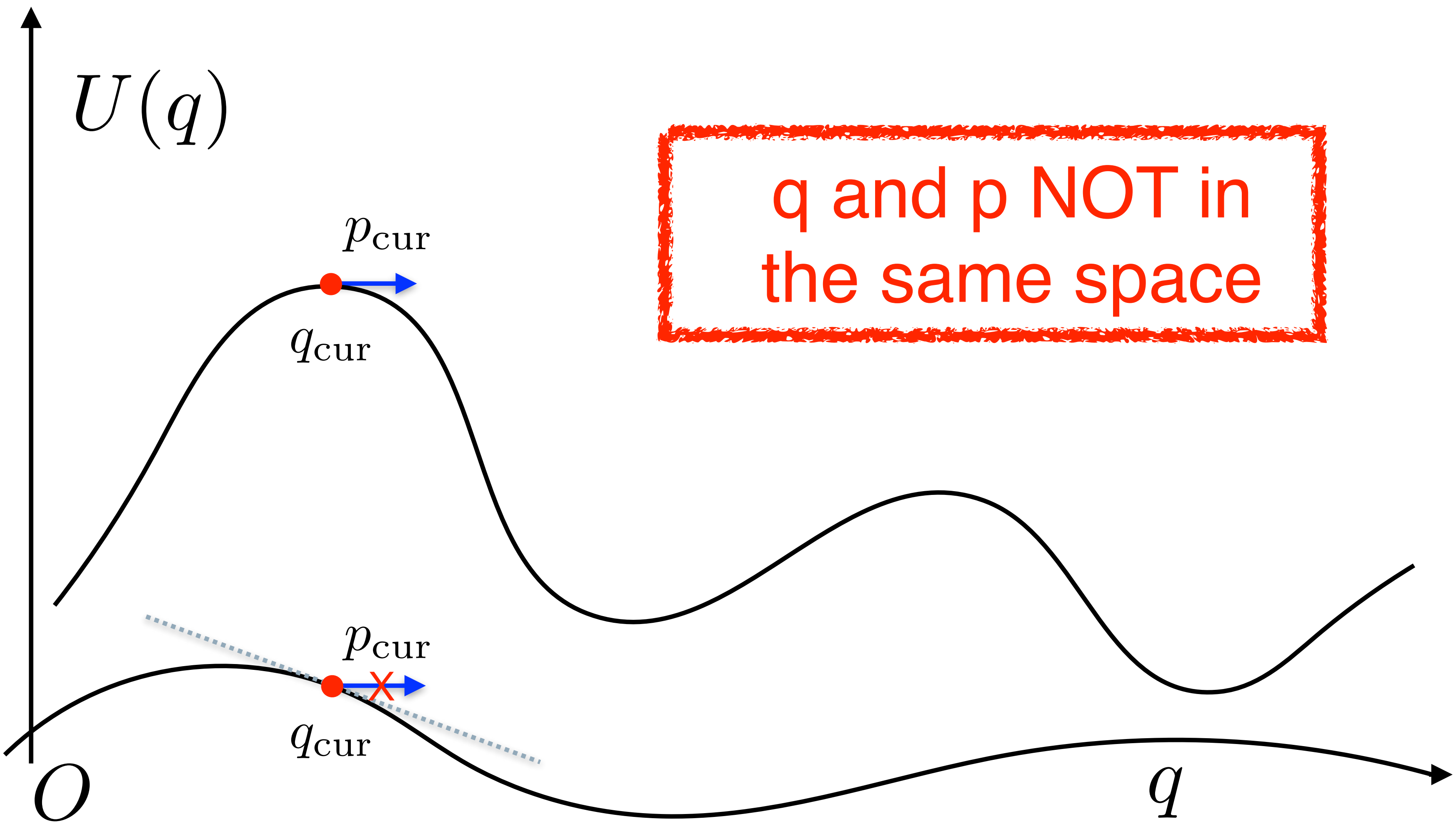
HMC for DP-MGLM

$$q = \theta \in \mathcal{M}$$



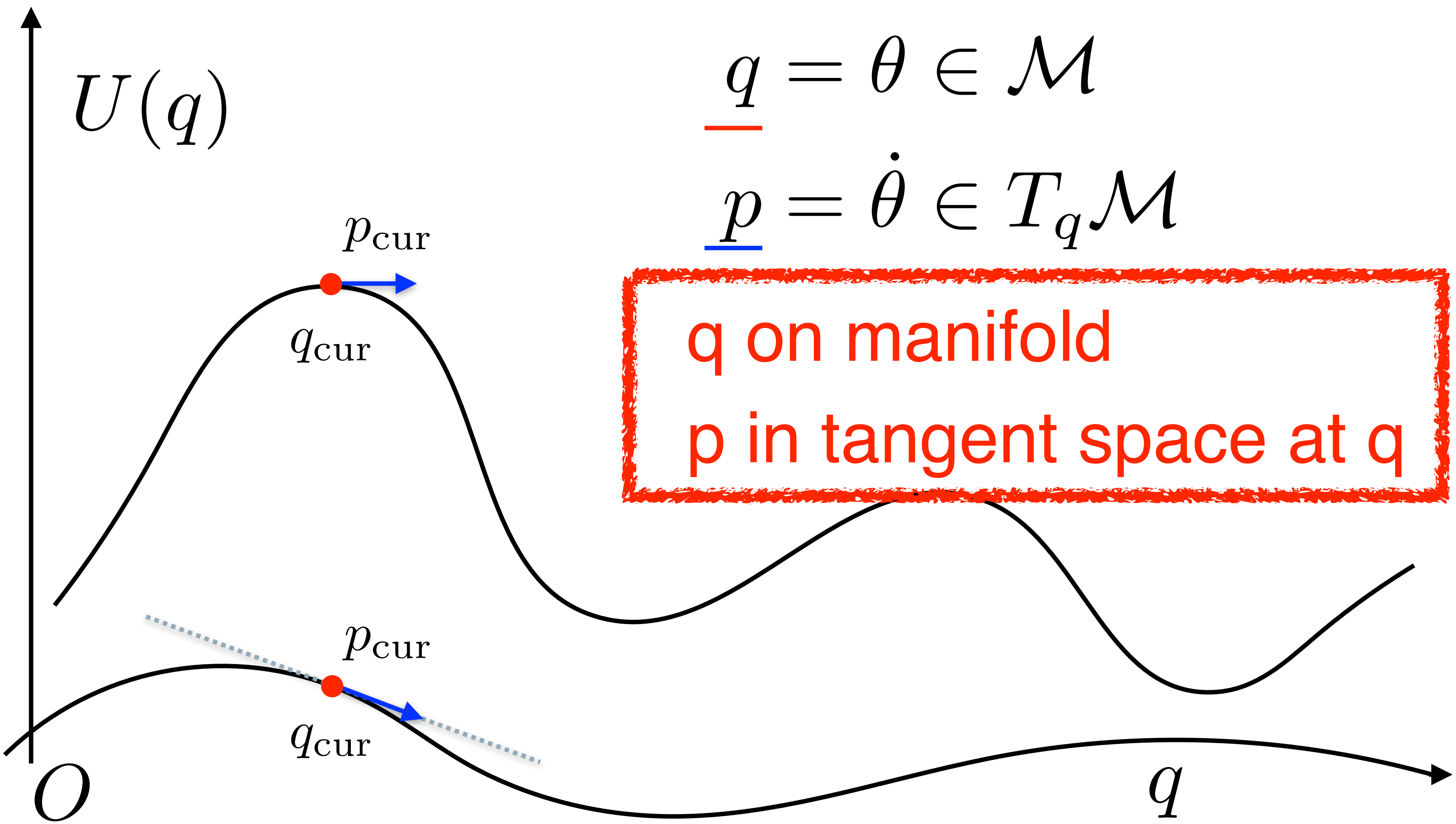
q on manifold

HMC for DP-MGLM



q and p NOT in the same space

HMC for DP-MGLM



$$q = \theta \in \mathcal{M}$$

$$p = \dot{\theta} \in T_q\mathcal{M}$$

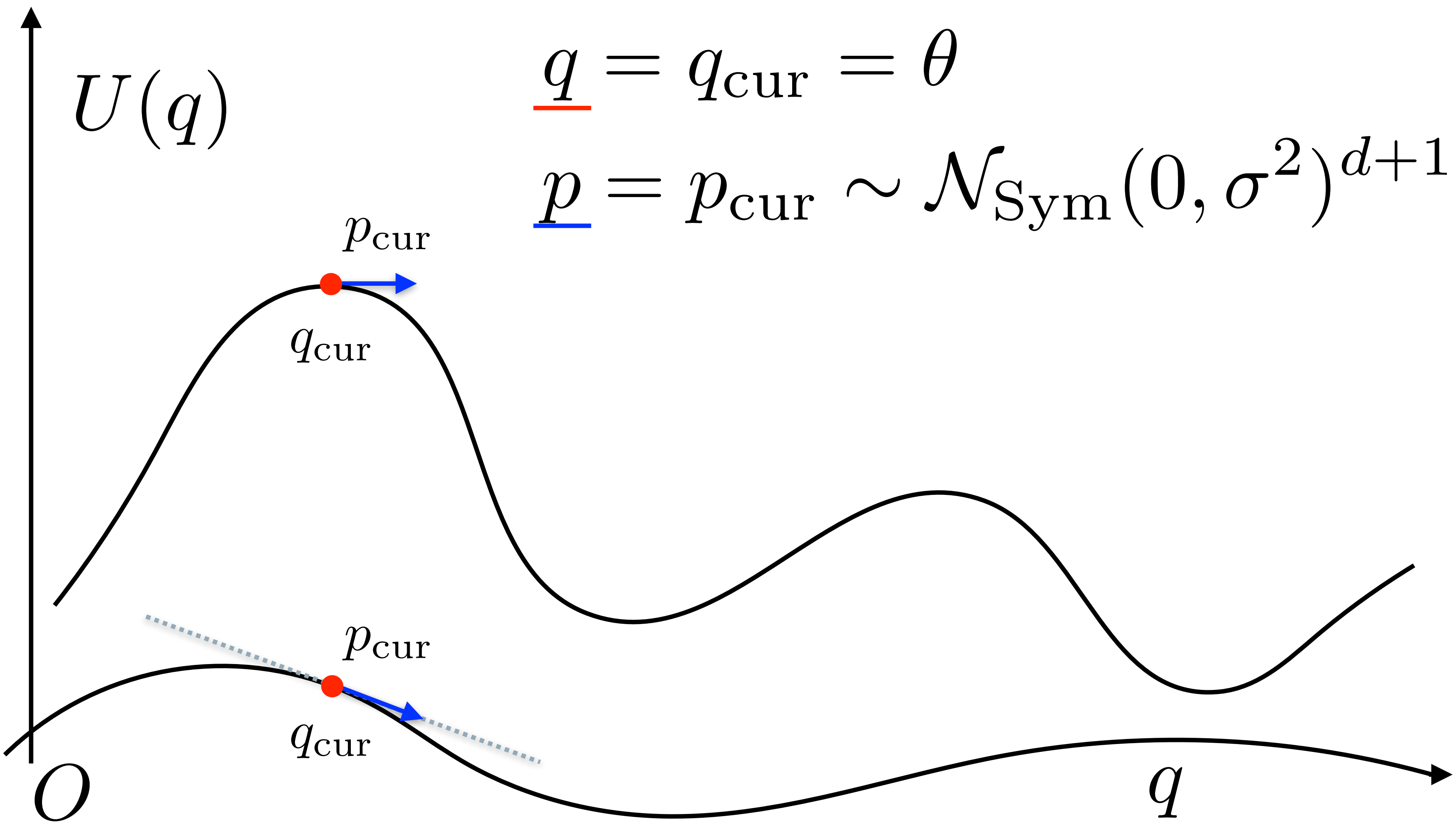
q on manifold

p in tangent space at q

HMC for DP-MGLM

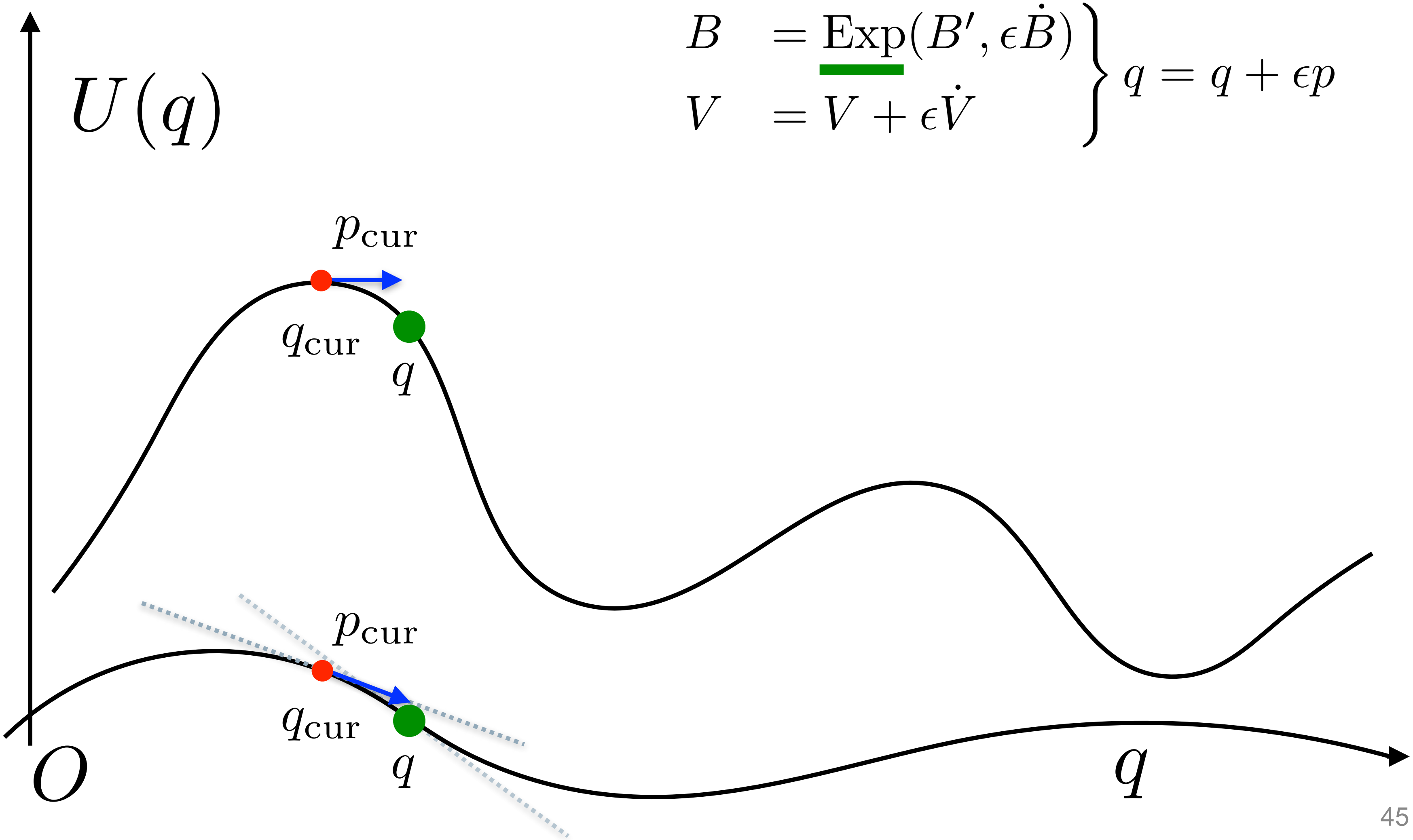
$$q = q_{\mathrm{cur}} = \theta$$

$$p = p_{\mathrm{cur}} \sim \mathcal{N}_{\mathrm{Sym}}(0, \sigma^2)^{d+1}$$



HMC for DP-MGLM

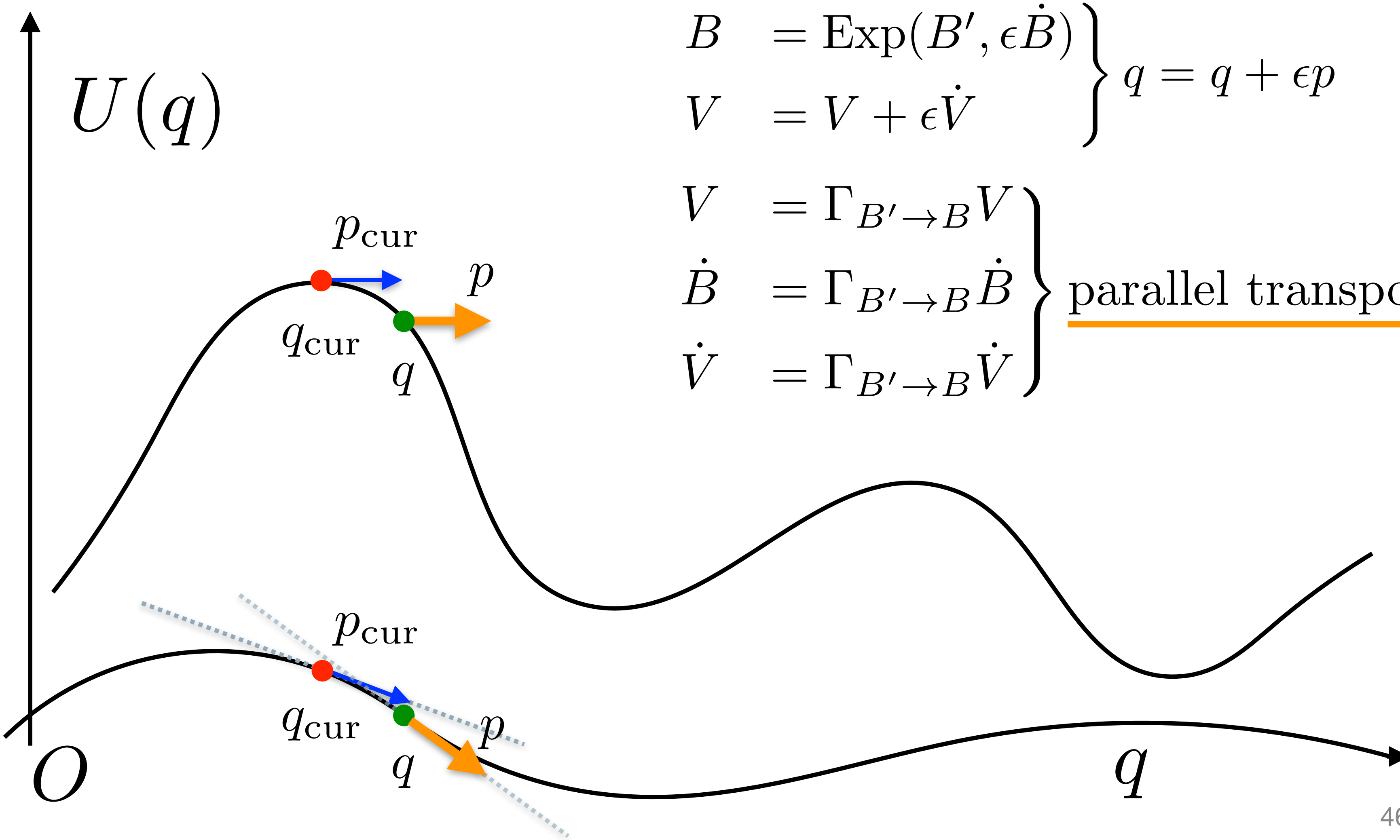
$$\left.\begin{aligned} B &= \underline{\mathrm{Exp}}(B', \epsilon \dot{B}) \\ V &= V + \epsilon \dot{V} \end{aligned}\right\} q = q + \epsilon p$$



HMC for DP-MGLM

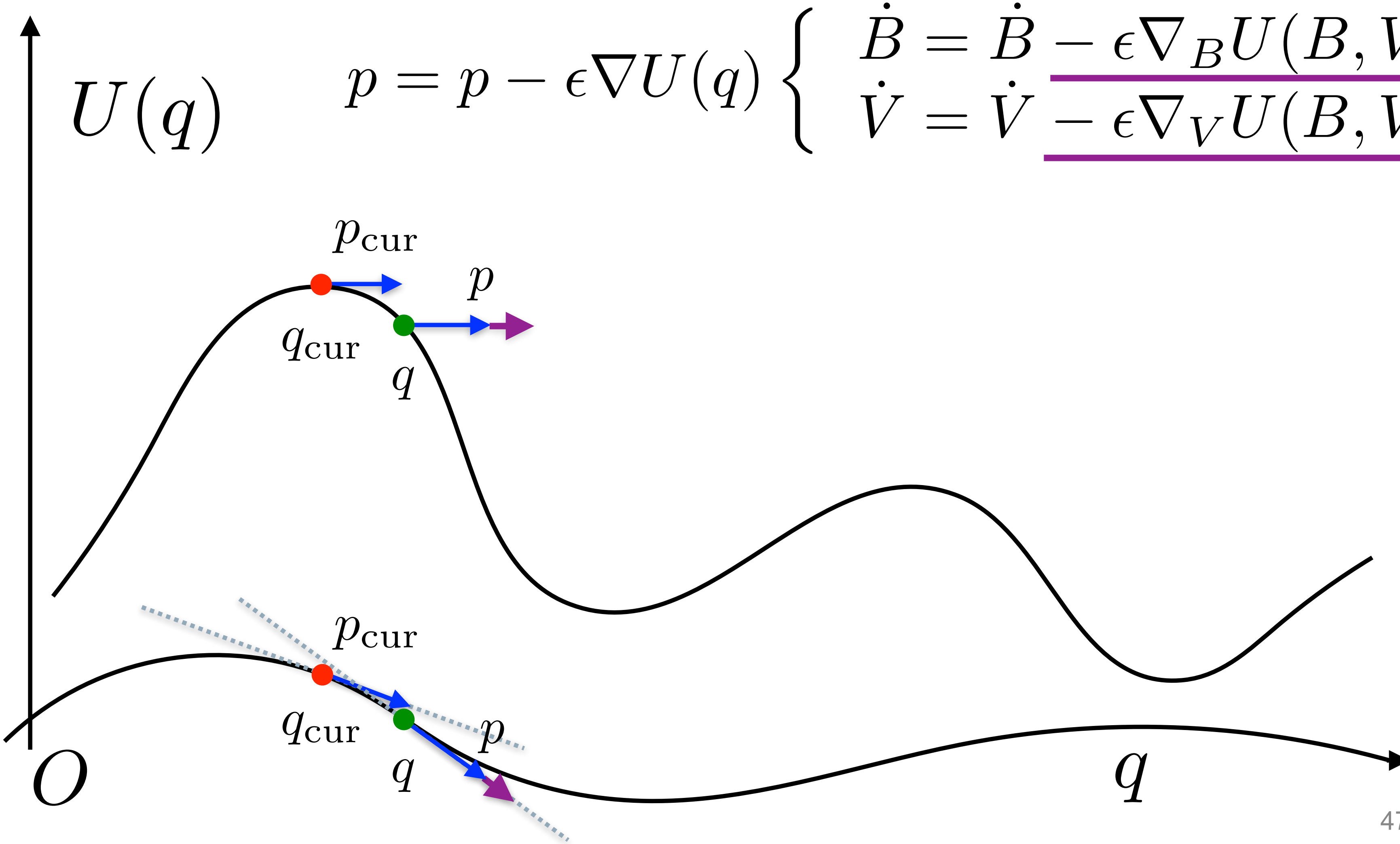
$$\left.\begin{aligned} B &= \mathrm{Exp}(B', \epsilon \dot{B}) \\ V &= V + \epsilon \dot{V} \end{aligned}\right\} q = q + \epsilon p$$

$$\left.\begin{aligned} V &= \Gamma_{B' \to B} V \\ \dot{B} &= \Gamma_{B' \to B} \dot{B} \\ \dot{V} &= \Gamma_{B' \to B} \dot{V} \end{aligned}\right\} \text{parallel transport}$$



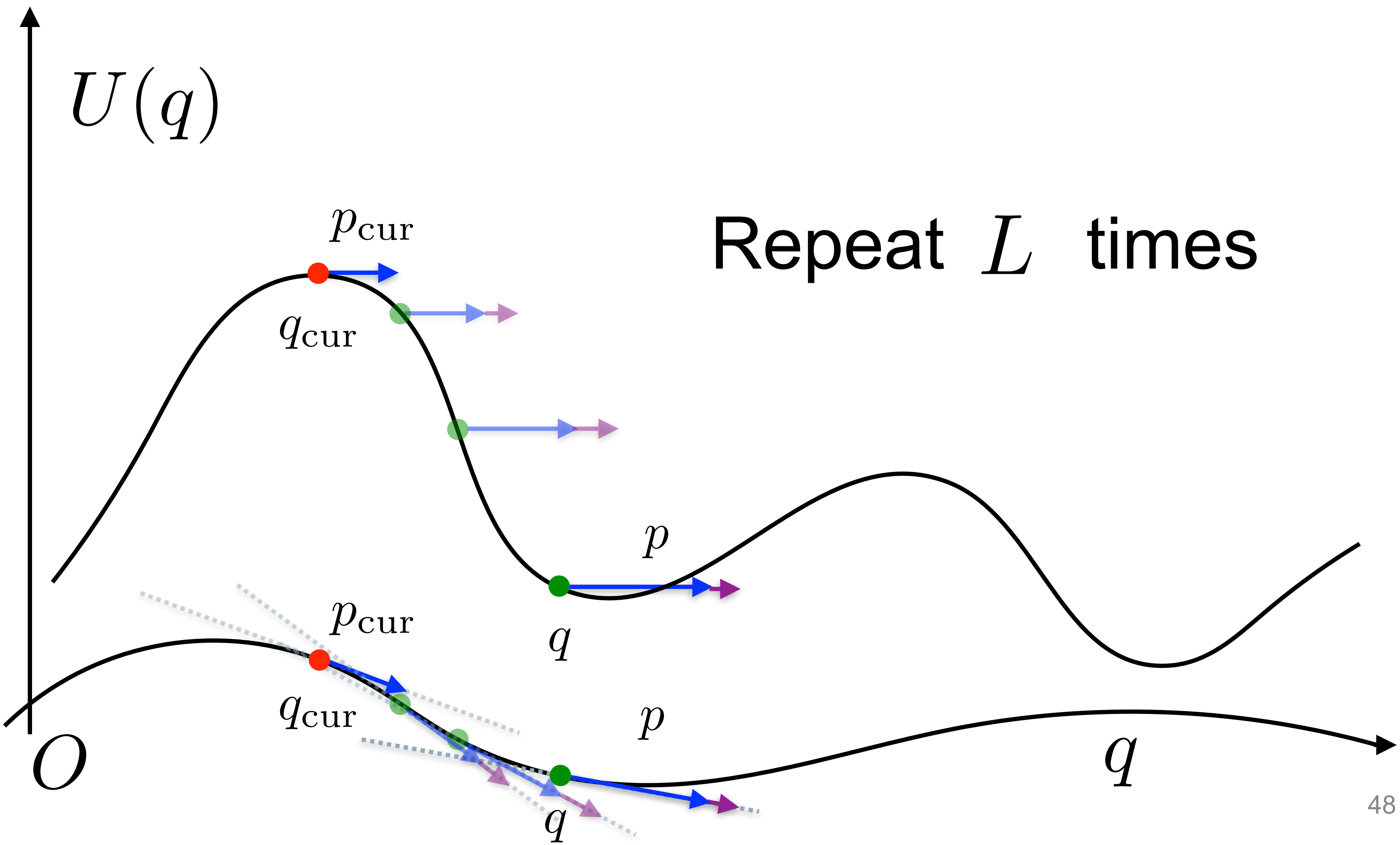
HMC for DP-MGLM

$$p = p - \epsilon \nabla U(q) \begin{cases} \dot{B} = \dot{B} - \epsilon \nabla_B U(B,V) \\ \dot{V} = \dot{V} - \epsilon \nabla_V U(B,V) \end{cases}$$



HMC for DP-MGLM

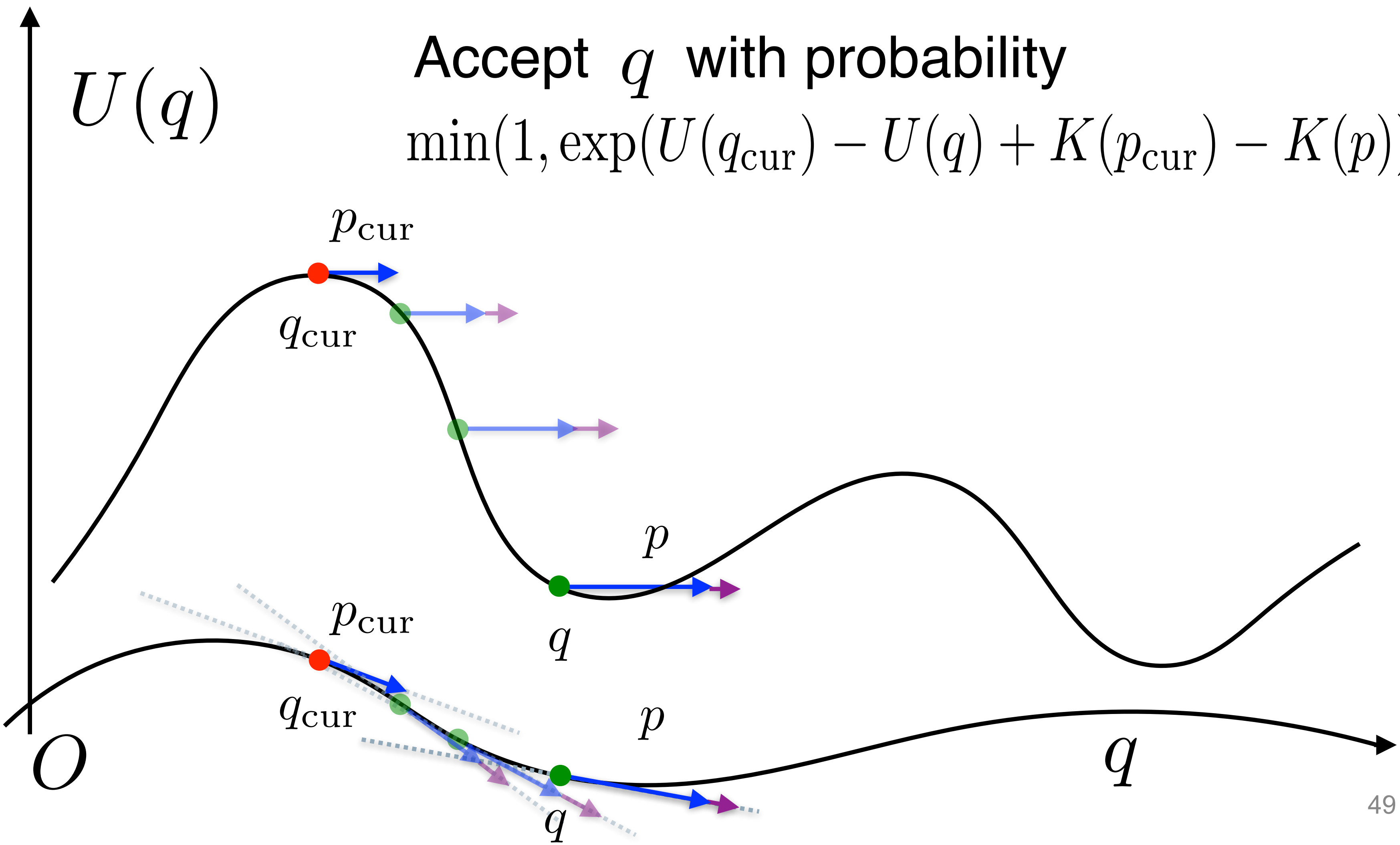
Repeat L times



HMC for DP-MGLM

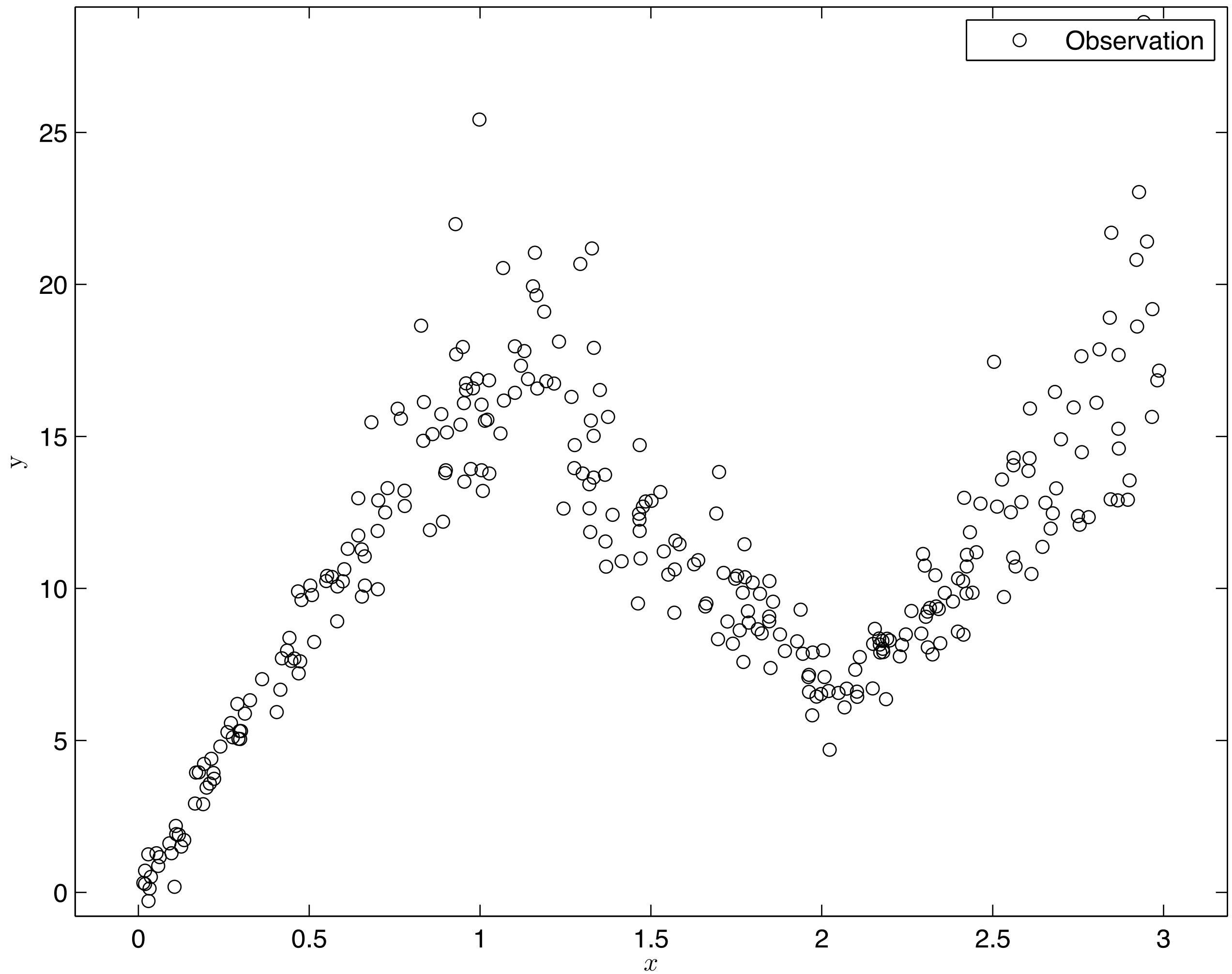
Accept q with probability

$$\min(1, \exp(U(q_{\text{cur}}) - U(q) + K(p_{\text{cur}}) - K(p)))$$

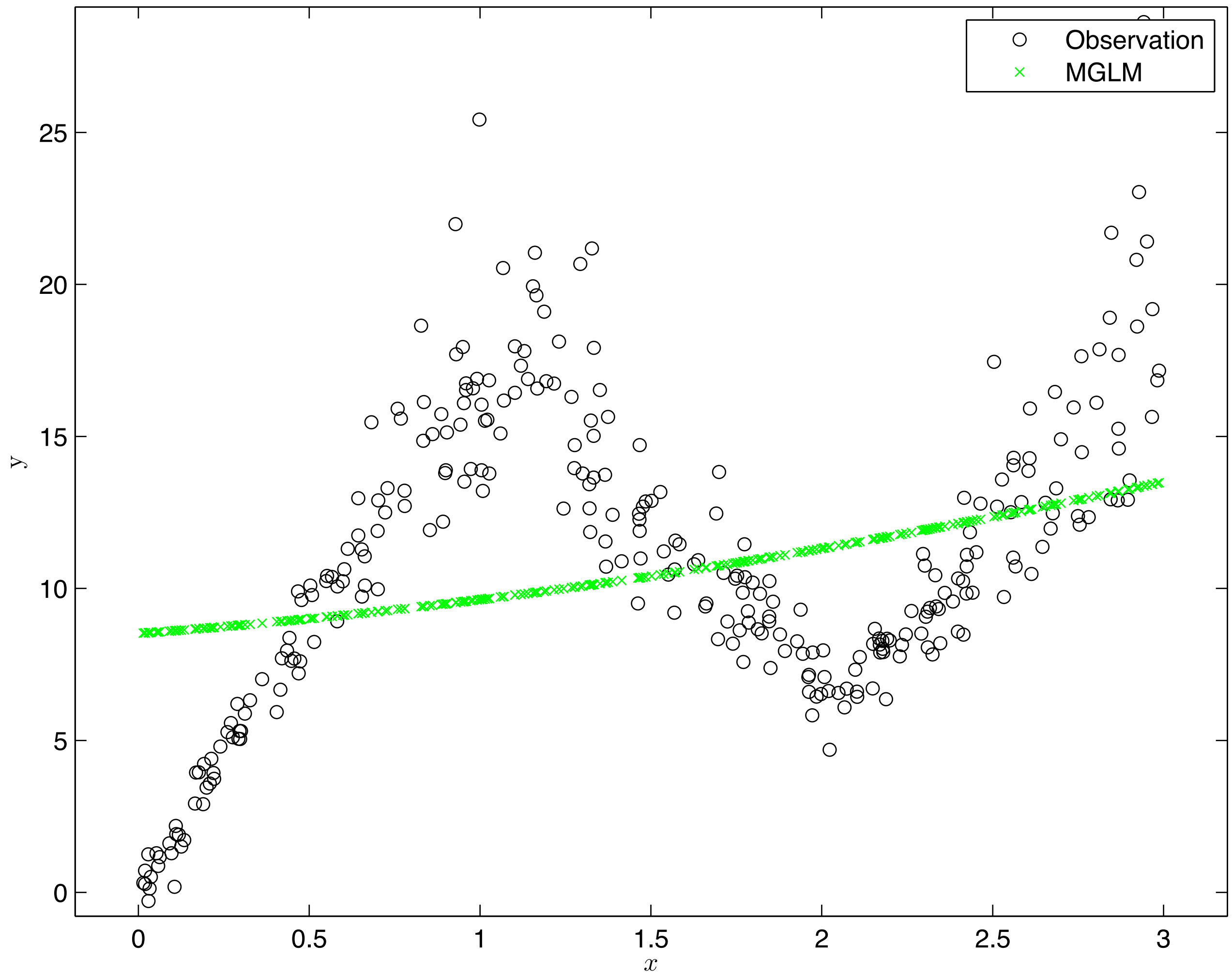


Experiments

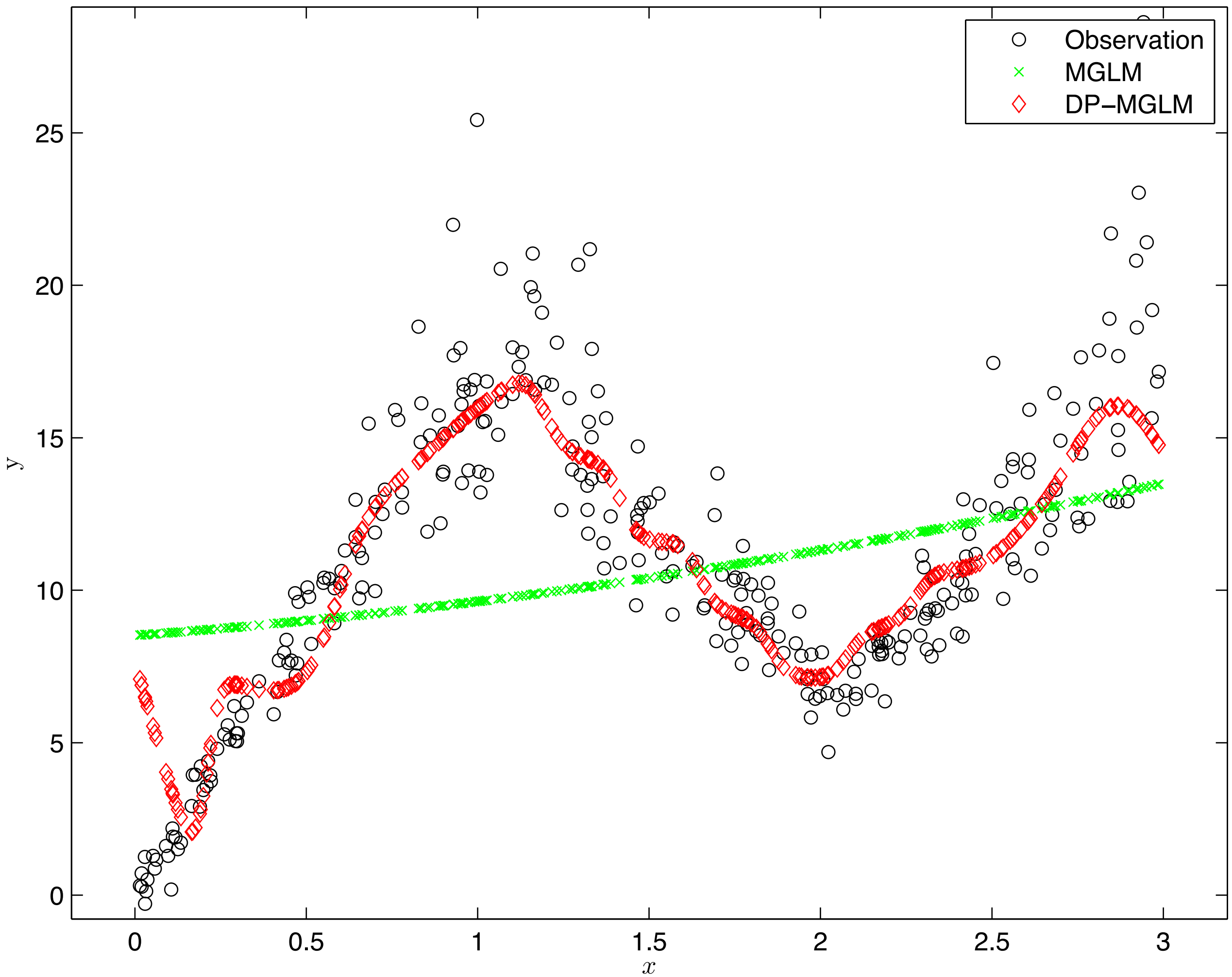
Experiment 1
MGLM vs DP-MGLM



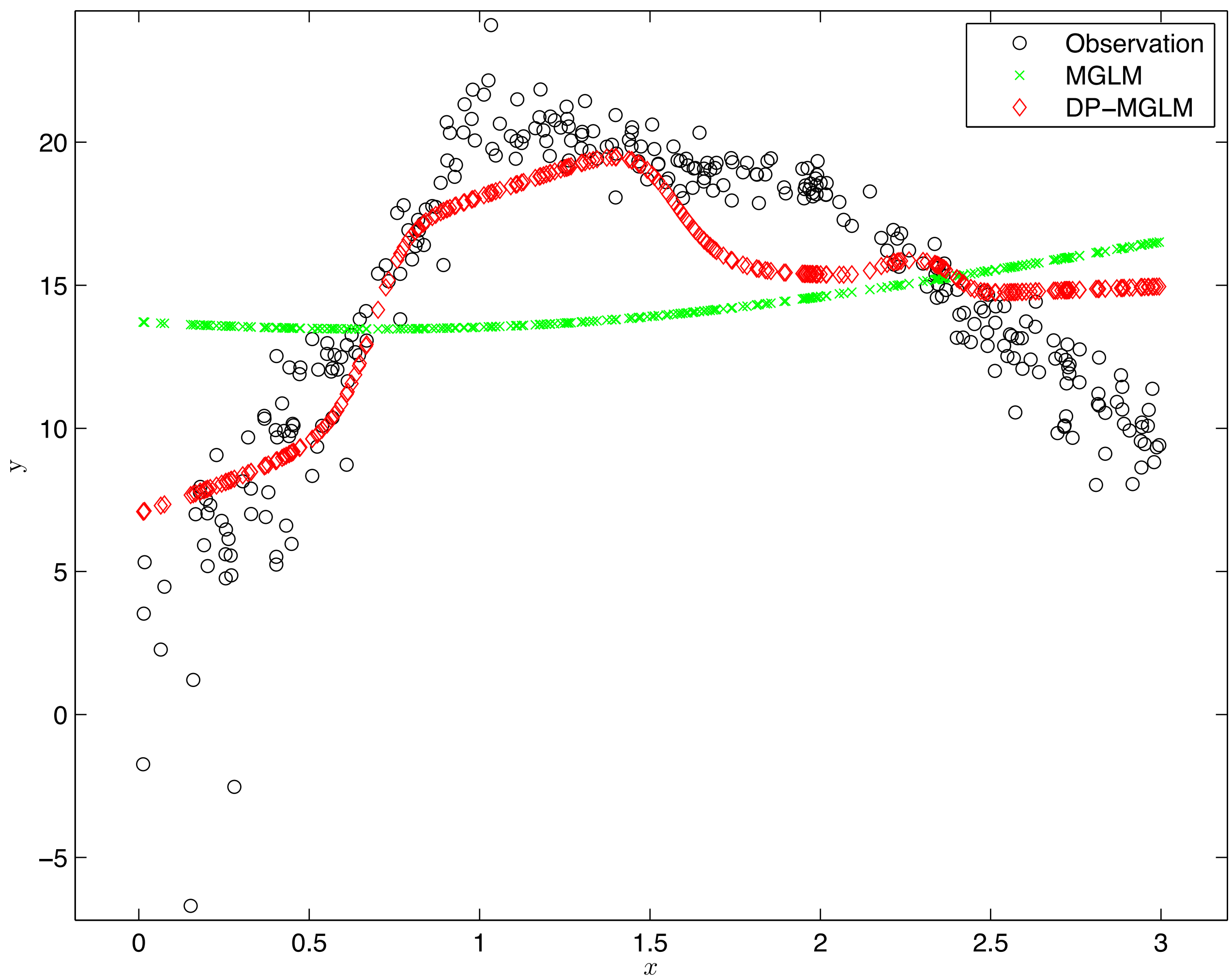
Experiment 1
MGLM vs DP-MGLM



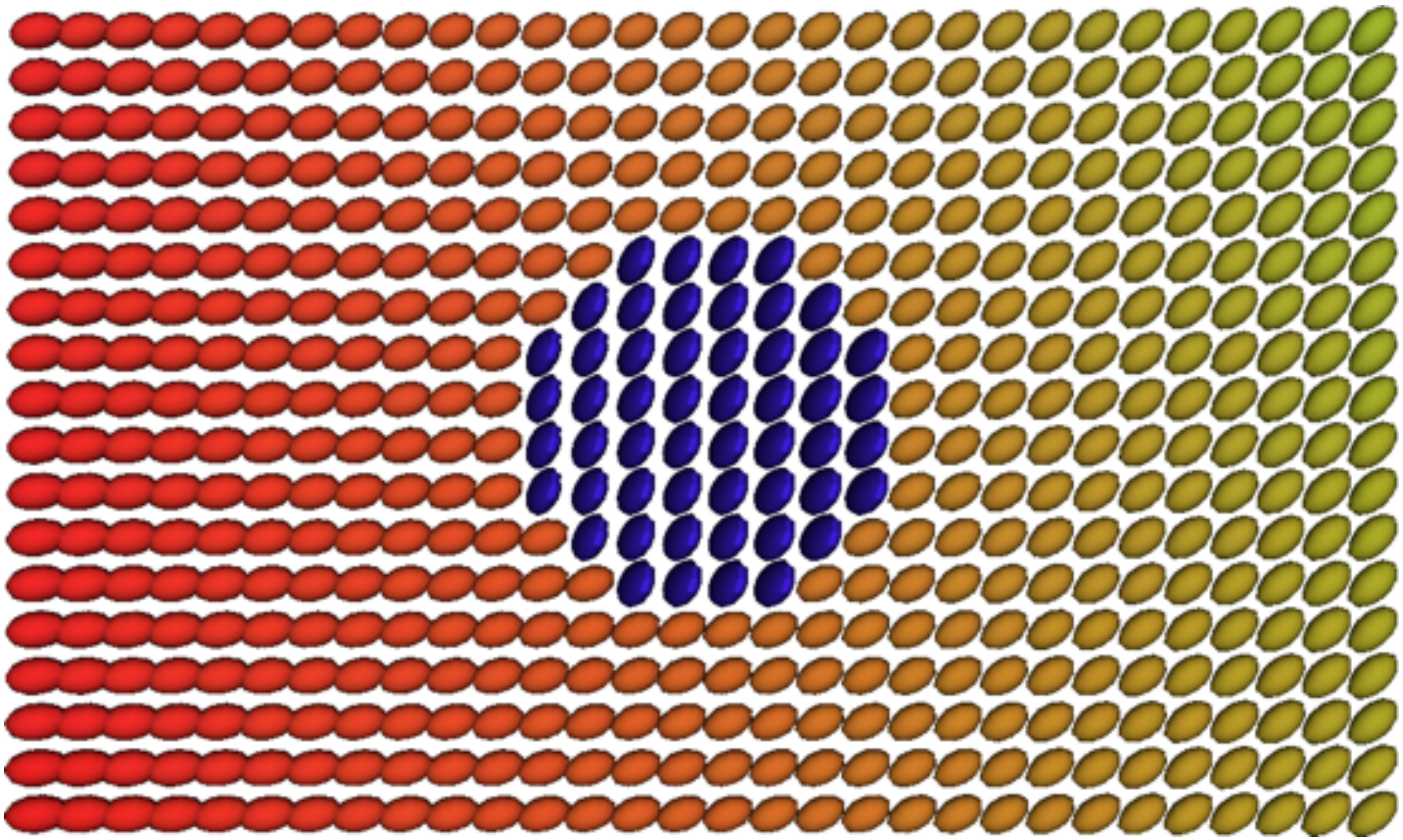
Experiment 1
MGLM vs DP-MGLM



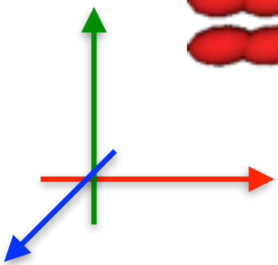
Experiment 1
MGLM vs DP-MGLM



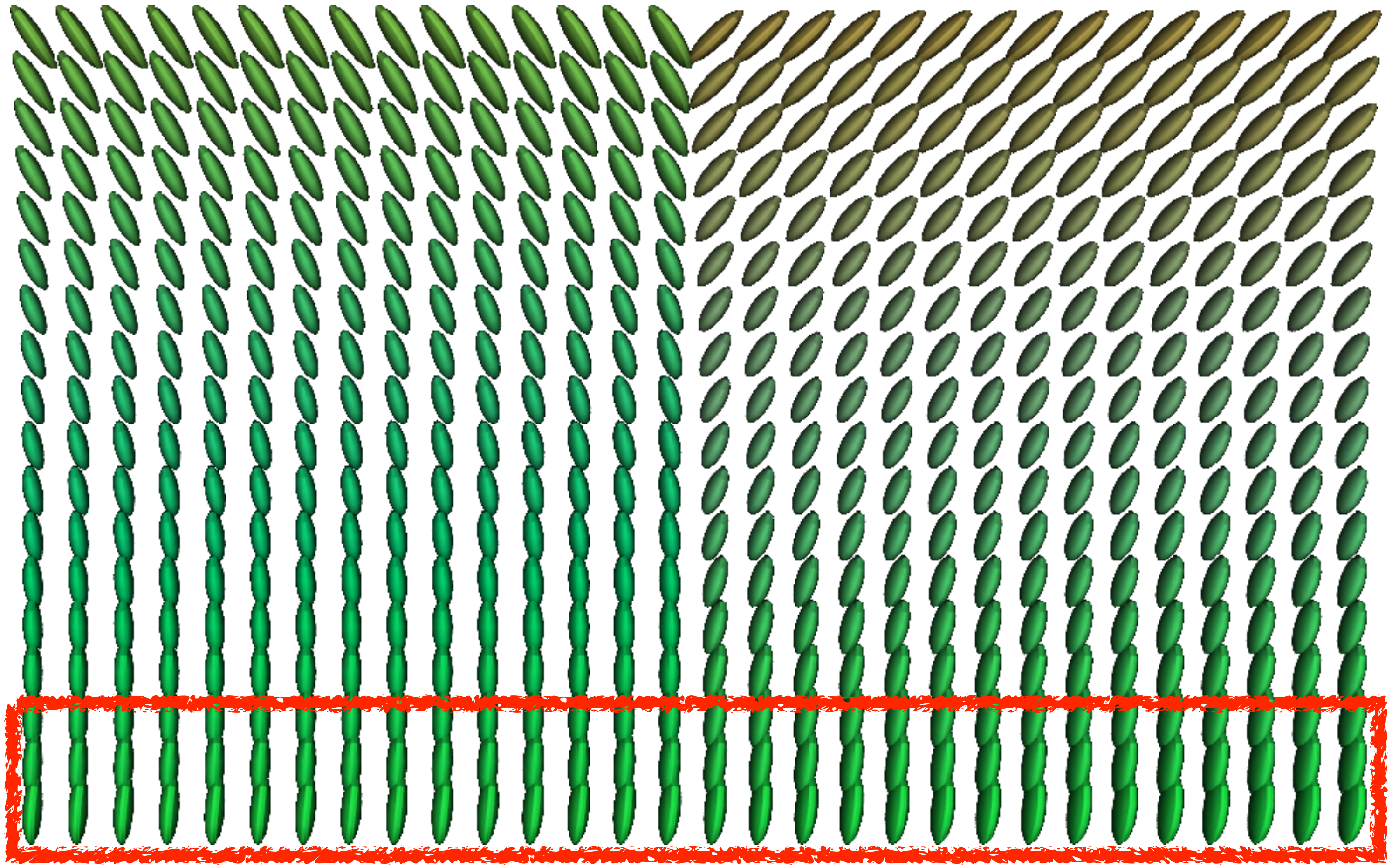
Experiment 2
Clustering on SPD(3)



3D ellipsoid patch 1

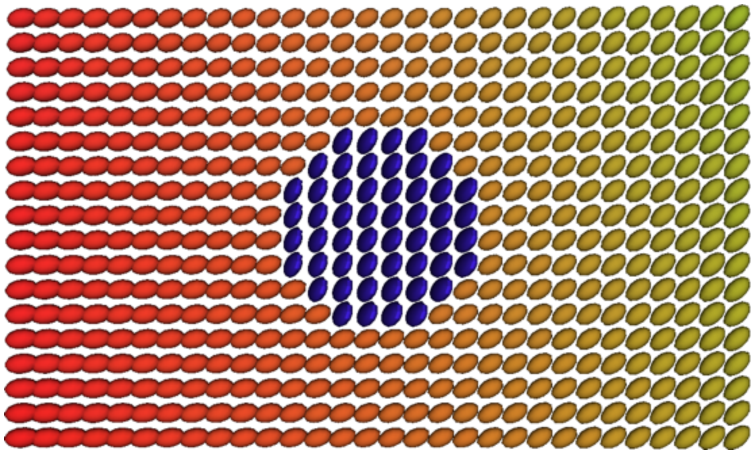


Experiment 2
Clustering on SPD(3)

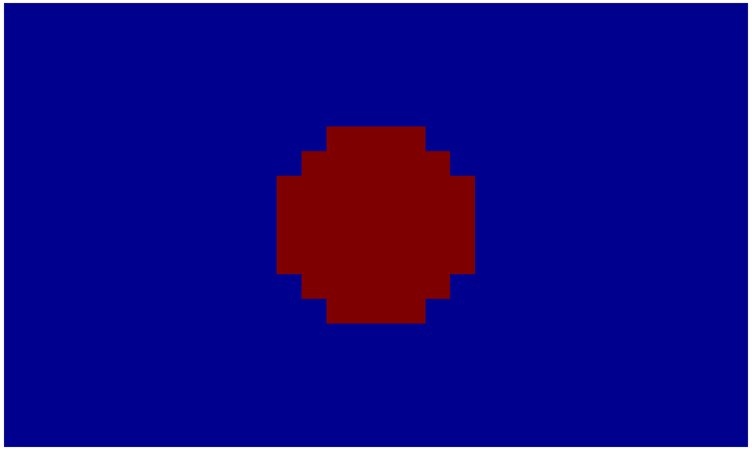


3D ellipsoid patch 2

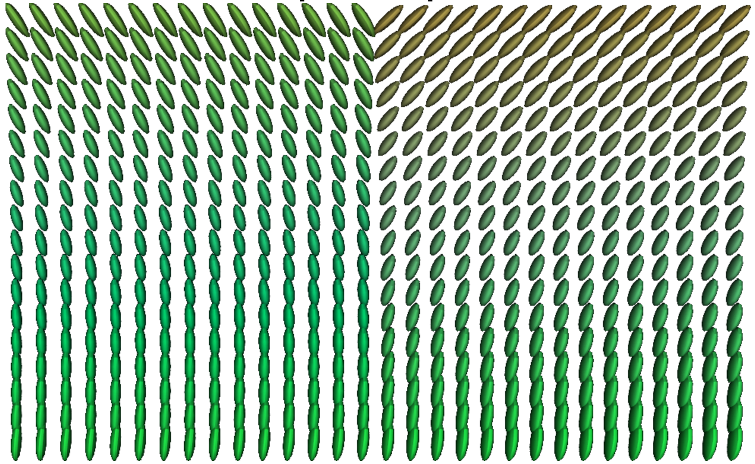
Experiment 2
Clustering on SPD(3)



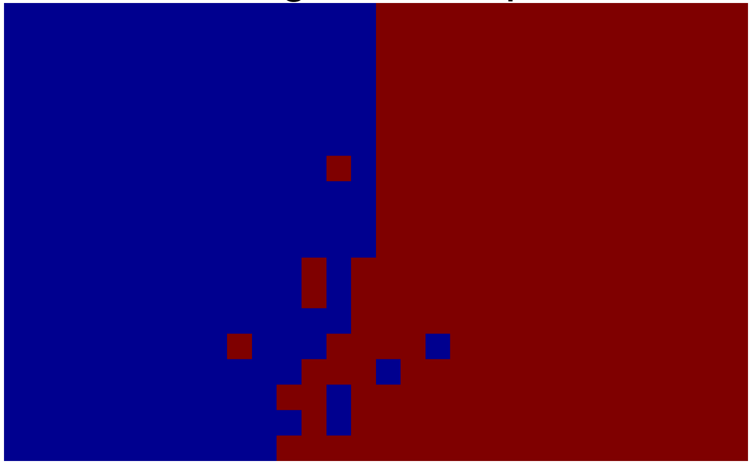
3D ellipsoid patch 1



Clustering result of patch 1

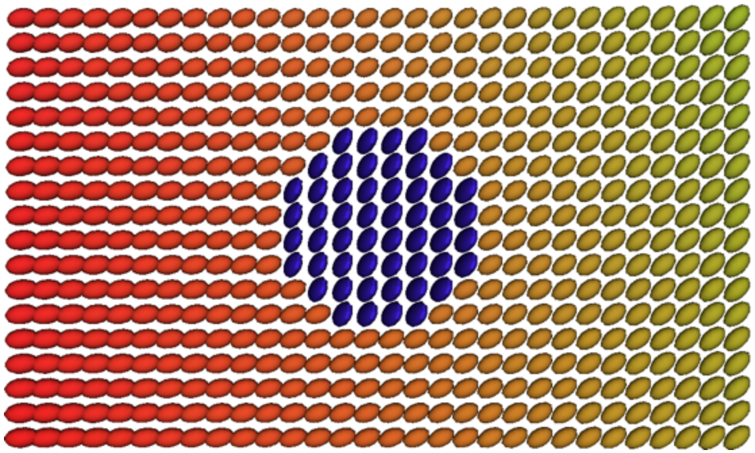


3D ellipsoid patch 2

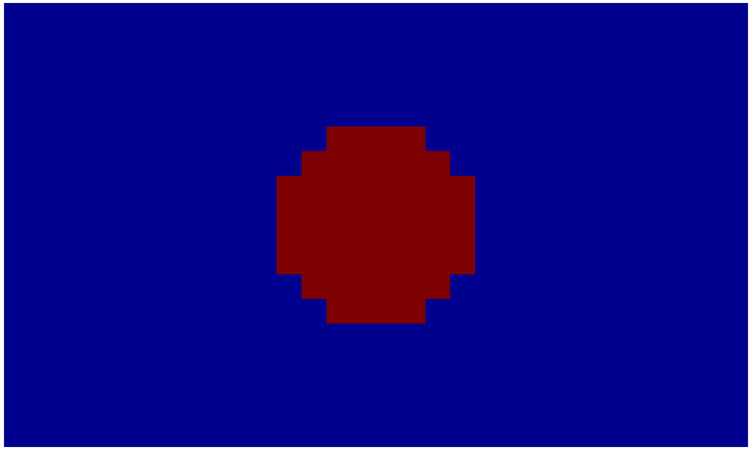


Clustering result of patch 2

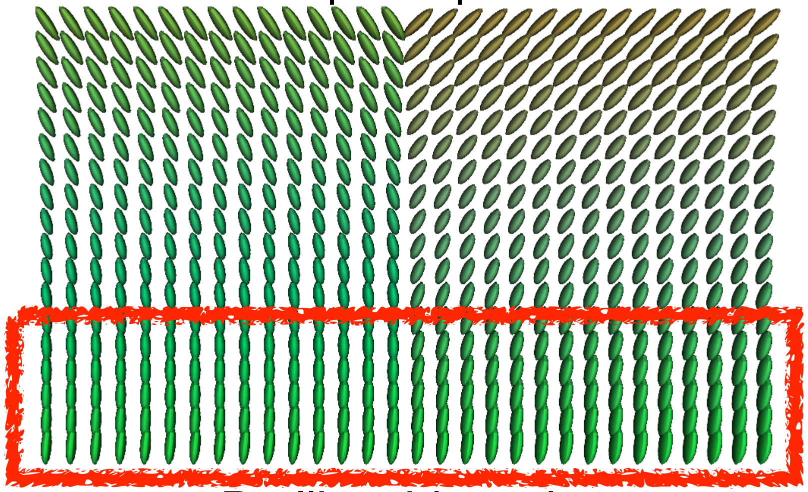
Experiment 2
Clustering on SPD(3)



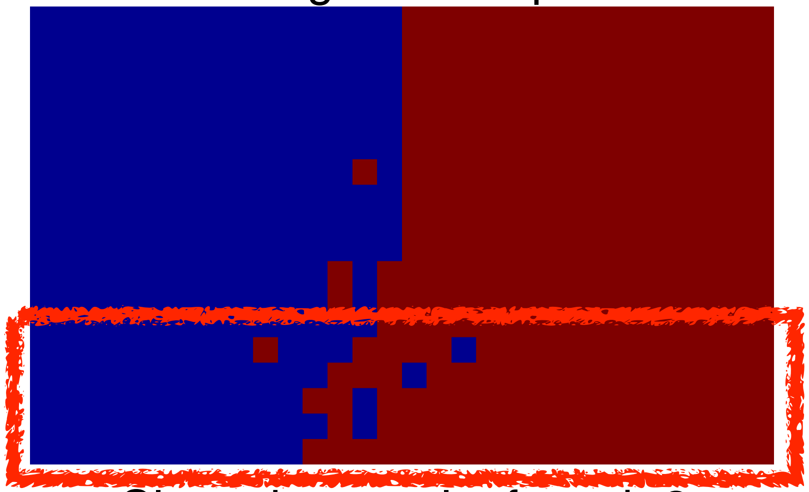
3D ellipsoid patch 1



Clustering result of patch 1

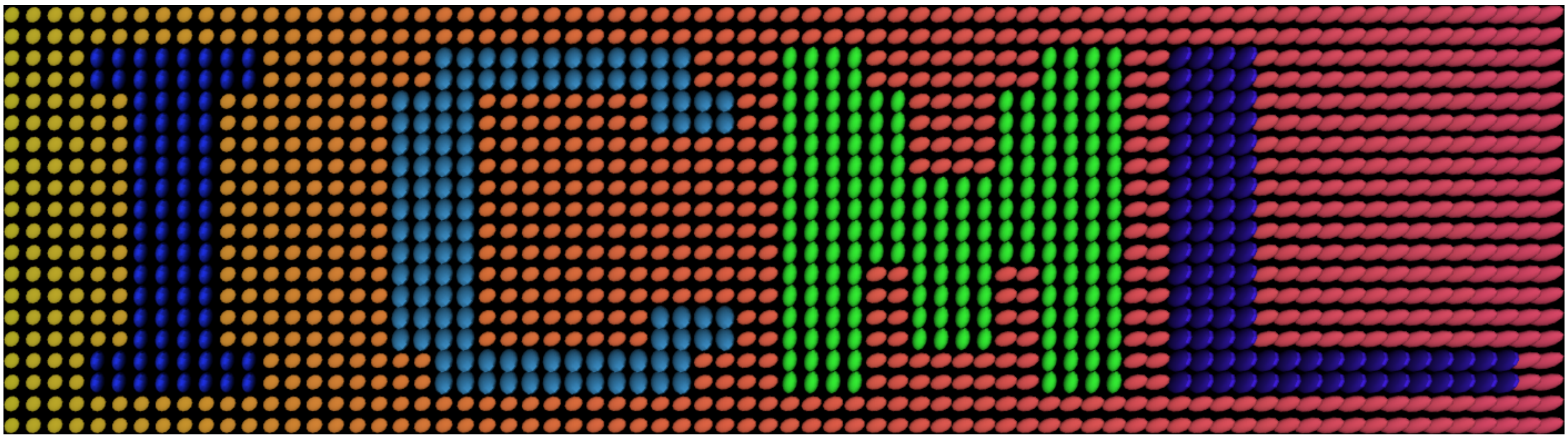


3D ellipsoid patch 2



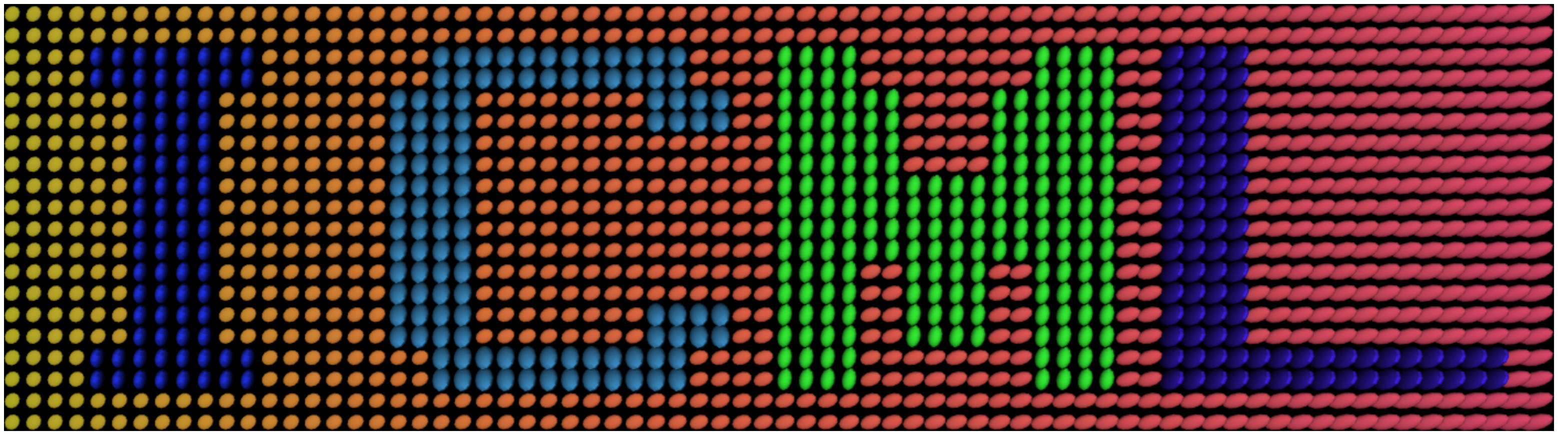
Clustering result of patch 2

Experiment 2
Clustering on SPD(3)

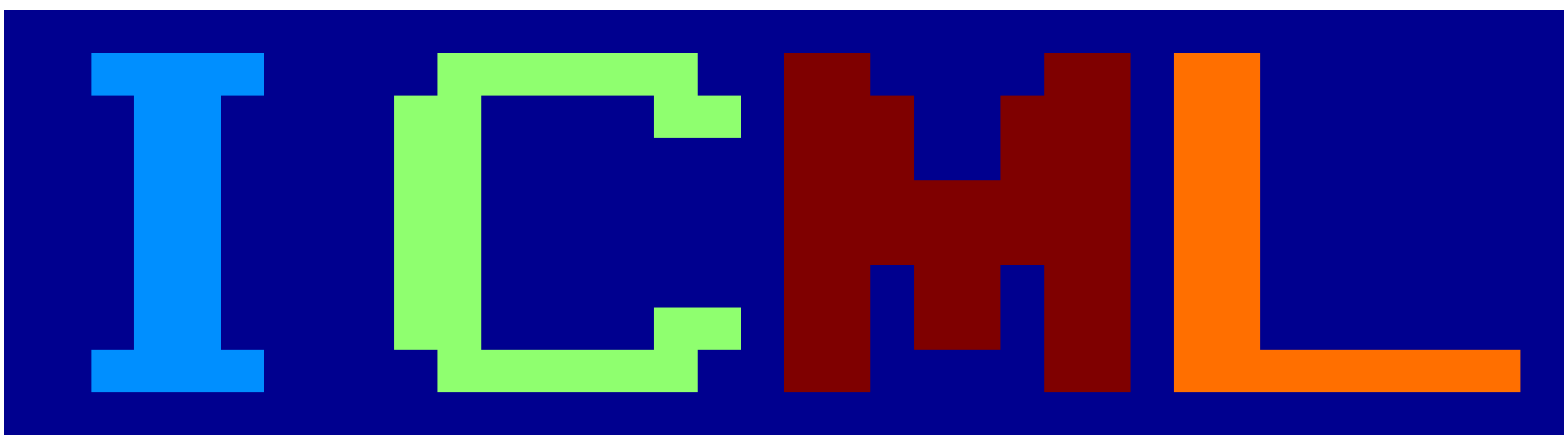


3D ellipsoid patch of ICML

Experiment 2
Clustering on SPD(3)



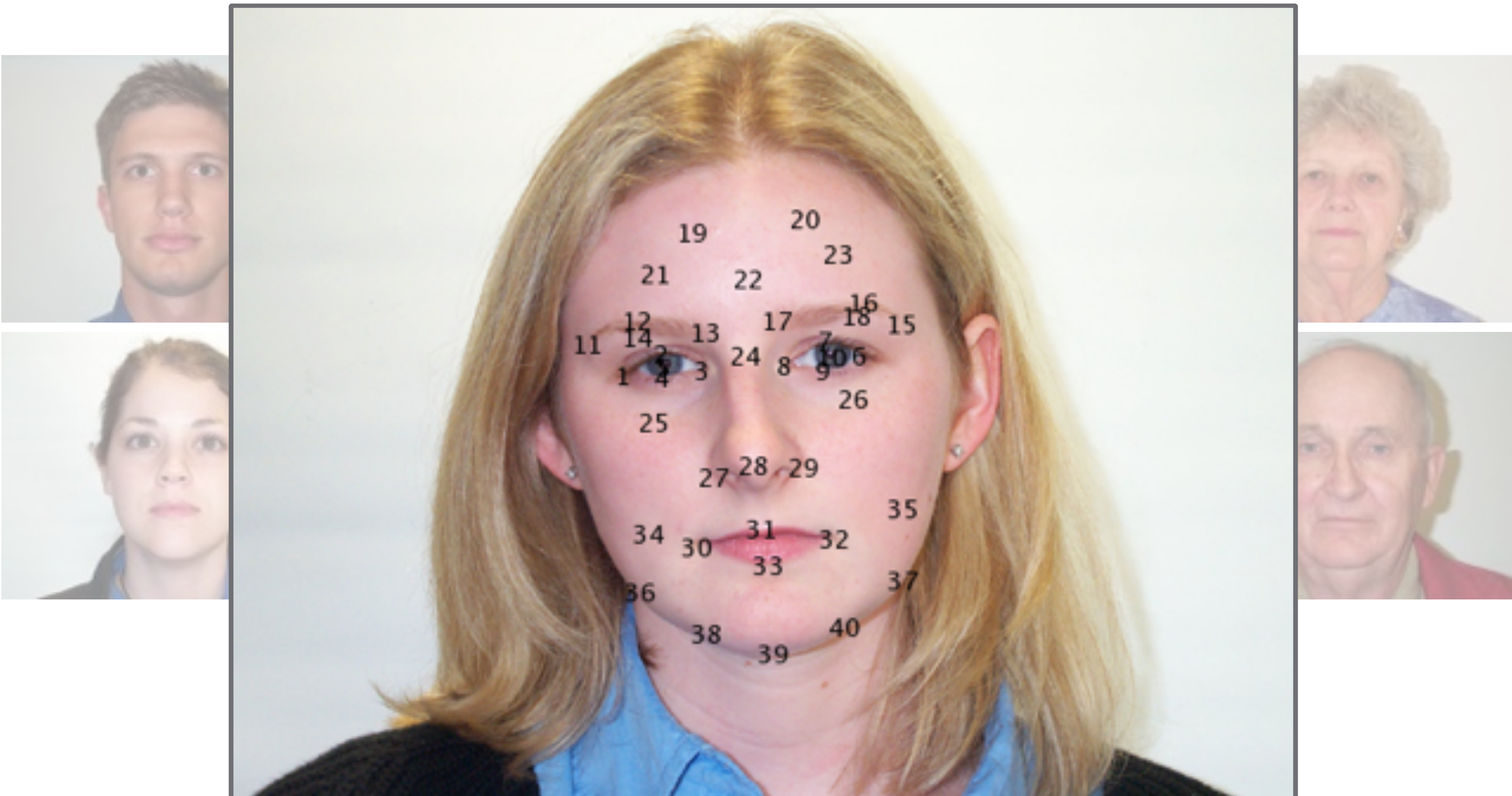
3D ellipsoid patch of ICML



Clustering results

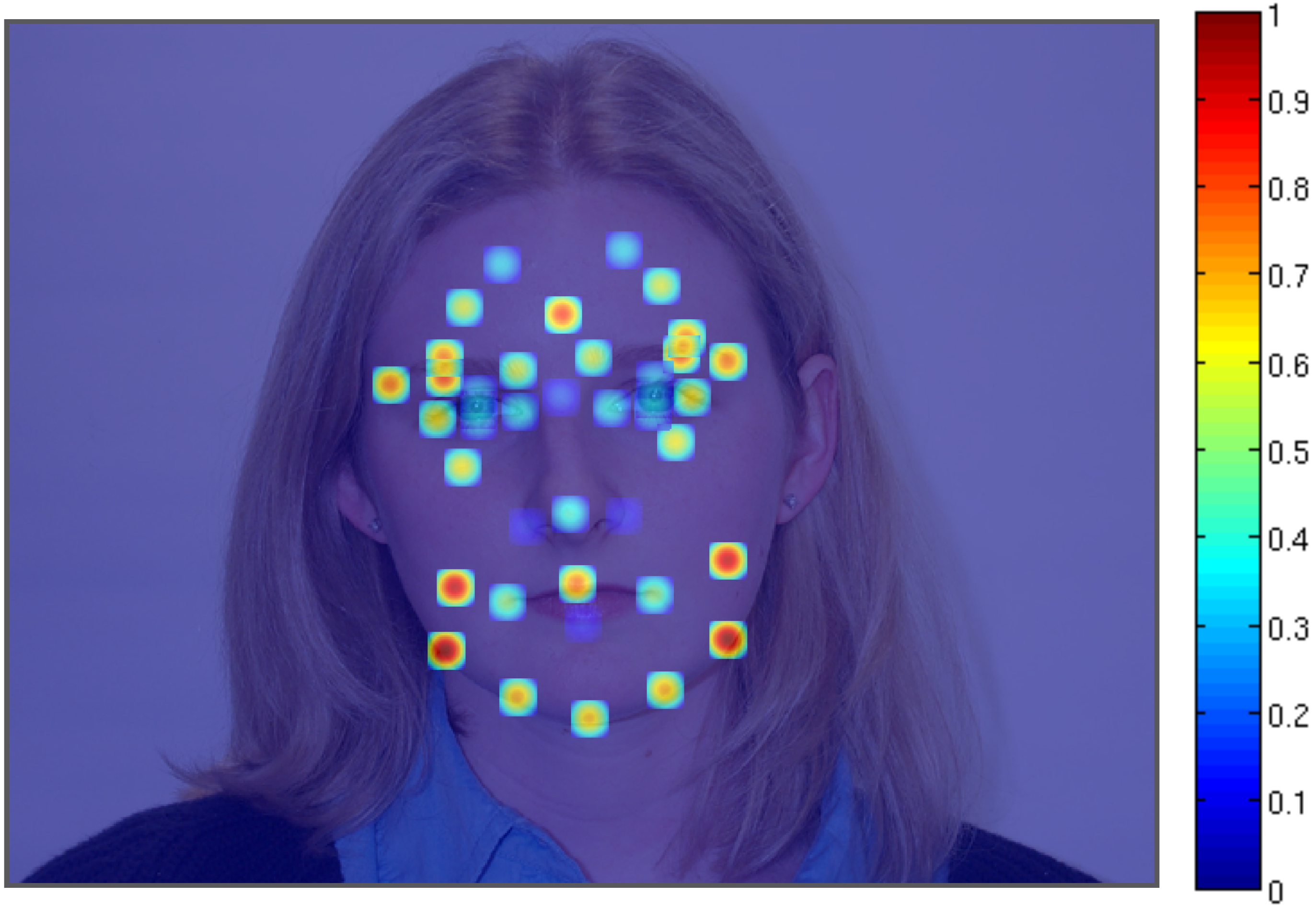
Experiment 3
Age versus landmark appearance

Experiment 3
Age versus landmark appearance



Landmarks

Experiment 3
Age versus landmark appearance



Correlation Magnitude

Conclusion

- DP-MGLM (Dirichlet process multivariate general linear model) for Riemannian manifolds (SPD(n)) learns more complicated models than MGLM
- Clustering based on nonlinear correlation between Euclidian covariates and manifold-valued response
- New distribution (over base point and tangent vectors) and HMC algorithm for DP-MGLM on SPD(n).

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Thank you!