

Composition of Movement Primitives

May 16, 2025

Contents

1	ProMPs	1
1.1	Recap	1
1.2	Coupling between joints	1
1.3	Hierarchical Bayesian Model	2
1.4	Via-Points Modulation	3
2	Composition of MPs	3

1 ProMPs

1.1 Recap

From (Paraschos et al., 2013, 2018):

- q_t : joint angle over time
- \dot{q}_t : joint velocity over time
- $\tau = \{q_t\}_{t=0\dots T}$: trajectory
- \mathbf{w} : weight vector of a single trajectory
- ϕ_t : basis function
- $\Phi_t = [\phi_t, \dot{\phi}_t]$: $n \times 2$ dimensional time-dependent basis matrix
- $z(t)$: monotonically increasing phase variable
- $\epsilon_y \sim \mathcal{N}(\mathbf{0}, \Sigma_y)$: zero-mean i.i.d. Gaussian noise

$$\mathbf{y}_t = \begin{bmatrix} q_t \\ \dot{q}_t \end{bmatrix} = \Phi_t^\top \mathbf{w} + \epsilon_y \quad (1)$$

$$p(\tau|\mathbf{w}) = \prod_t \mathcal{N}(\mathbf{y}_t | \Phi_t^\top \mathbf{w}, \Sigma_y) \quad (2)$$

$$p(\tau; \theta) = \int p(\tau|\mathbf{w}) \cdot p(\mathbf{w}; \theta) d\mathbf{w} \quad (3)$$

1.2 Coupling between joints

$$p(\mathbf{y}_t|\mathbf{w}) = \mathcal{N}\left(\begin{bmatrix} \mathbf{y}_{1,t} \\ \vdots \\ \mathbf{y}_{d,t} \end{bmatrix} \middle| \begin{bmatrix} \Phi_t^\top & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Phi_t^\top \end{bmatrix} \mathbf{w}, \Sigma_y\right) = \mathcal{N}(\mathbf{y}_t | \Psi_t \mathbf{w}, \Sigma_y) \quad (4)$$

with:

- $\mathbf{w} = [\mathbf{w}_1^\top, \dots, \mathbf{w}_n^\top]^\top$: combined weight vector
- Φ_t : block-diagonal basis matrix containing the basis functions and their derivatives for each dimension
- $\mathbf{y}_{i,t} = [q_{i,t}, \dot{q}_{i,t}]^\top$: joint angle and velocity for the i^{th} joint

1.3 Hierarchical Bayesian Model

The Hierarchical Bayesian Model used in ProMPs is illustrated in Fig. 1.

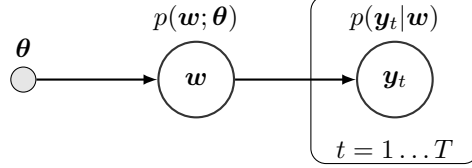


Figure 1: Hierarchical Bayesian Model used in ProMPs.

- $\theta = \{\mu_w, \Sigma_w\}$
- $p(\mathbf{w}; \theta) = \mathcal{N}(\mathbf{w} | \mu_w, \Sigma_w)$: prior over the weight vector \mathbf{w} , with parameters θ , assumed to be Gaussian

$$p(\mathbf{y}_t; \theta) = \int \mathcal{N}(\mathbf{y}_t | \Psi_t^\top \mathbf{w}, \Sigma_y) \cdot p(\mathbf{w}; \theta) d\mathbf{w} \quad (5)$$

$$= \int \mathcal{N}(\mathbf{y}_t | \Psi_t^\top \mathbf{w}, \Sigma_y) \cdot \mathcal{N}(\mathbf{w} | \mu_w, \Sigma_w) d\mathbf{w} \quad (6)$$

$$(7)$$

From (Deisenroth et al., 2020), we have the joint distribution:

$$p(\mathbf{x}_a, \mathbf{x}_b) = \mathcal{N}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}\right) \quad (8)$$

and the marginal distribution $p(\mathbf{x}_a)$ of a joint Gaussian distribution $p(\mathbf{x}_a, \mathbf{x}_b)$:

$$p(\mathbf{x}_a) = \int p(\mathbf{x}_a, \mathbf{x}_b) d\mathbf{x}_b = \mathcal{N}(\mathbf{x}_a | \mu_a, \Sigma_{aa}) \quad (9)$$

Since \mathbf{y}_t and \mathbf{w} are jointly Gaussian, we have:

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{w} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \Psi_t^\top \mu_w \\ \mu_w \end{bmatrix}, \begin{bmatrix} \text{Cov}[\mathbf{y}_t, \mathbf{y}_t] & \text{Cov}[\mathbf{y}_t, \mathbf{w}] \\ \text{Cov}[\mathbf{w}, \mathbf{y}_t] & \text{Cov}[\mathbf{w}, \mathbf{w}] \end{bmatrix}\right) \quad (10)$$

$$\text{Cov}[\mathbf{y}_t, \mathbf{y}_t] = \text{Cov}[\Psi_t^\top \mathbf{w} + \epsilon] \quad (11)$$

$$= \text{Cov}[\Psi_t^\top \mathbf{w}] + \text{Cov}[\epsilon] \quad (12)$$

$$= \Psi_t^\top \text{Cov}[\mathbf{w}] \Psi_t + \Sigma_y \quad (13)$$

$$= \Psi_t^\top \Sigma_w \Psi_t + \Sigma_y \quad (14)$$

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{w} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \Psi_t^\top \mu_w \\ \mu_w \end{bmatrix}, \begin{bmatrix} \Psi_t^\top \Sigma_w \Psi_t + \Sigma_y & \Psi_t^\top \Sigma_w \\ \Sigma_w \Psi_t & \Sigma_w \end{bmatrix}\right) \quad (15)$$

$$p(\mathbf{y}_t; \theta) = \mathcal{N}(\mathbf{y}_t | \Psi_t^\top \mu_w, \Psi_t^\top \Sigma_w \Psi_t + \Sigma_y) \quad (16)$$

1.4 Via-Points Modulation

- $\mathbf{x}_t^* = [\mathbf{y}_t^*, \boldsymbol{\Sigma}_t^*]$: desired observation
- \mathbf{y}_t^* : desired position and velocity vector at time t
- $\boldsymbol{\Sigma}_t^*$: accuracy of the desired observation

Using Bayes rule:

$$p(\mathbf{w}|\mathbf{x}_t^*) = \frac{p(\mathbf{x}_t^*|\mathbf{w}) \cdot p(\mathbf{w})}{p(\mathbf{x}_t^*)} \quad (17)$$

$$p(\mathbf{w}|\mathbf{x}_t^*) \propto \mathcal{N}(\mathbf{y}_t^*|\boldsymbol{\Psi}_t^\top \mathbf{w}, \boldsymbol{\Sigma}_t^*) \cdot p(\mathbf{w}) \quad (18)$$

From (Bishop and Bishop, 2024), the parameters of a conditional multivariate Gaussian $p(\mathbf{x}_a|\mathbf{x}_b) = \mathcal{N}(\boldsymbol{\mu}_{a|b}, \boldsymbol{\Sigma}_{a|b})$ are the following:

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a + \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}(\mathbf{x}_b - \boldsymbol{\mu}_b) \quad (19)$$

$$\boldsymbol{\Sigma}_{a|b} = \boldsymbol{\Sigma}_{aa} - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}\boldsymbol{\Sigma}_{ba} \quad (20)$$

$$\dots \text{Missing development} \dots \quad (21)$$

$$\boldsymbol{\mu}_w^{[new]} = \boldsymbol{\mu}_w + \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t \left(\boldsymbol{\Sigma}_y^* \boldsymbol{\Psi}_t^\top \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t \right)^{-1} (\mathbf{y}_t^* - \boldsymbol{\Psi}_t^\top \boldsymbol{\mu}_w) \quad (22)$$

$$\boldsymbol{\Sigma}_w^{[new]} = \boldsymbol{\Sigma}_w - \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t \left(\boldsymbol{\Sigma}_y^* \boldsymbol{\Psi}_t^\top \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t \right)^{-1} \boldsymbol{\Psi}_t^\top \boldsymbol{\Sigma}_w \quad (23)$$

2 Composition of MPs

ToDo

References

- A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, “Probabilistic Movement Primitives,” in *Advances in Neural Information Processing Systems*, vol. 26. Curran Associates, Inc., 2013. [Online]. Available: <https://proceedings.neurips.cc/paper/2013/hash/e53a0a2978c28872a4505bdb51db06dc-Abstract.html>
- A. Paraschos, C. Daniel, J. Peters, and G. Neumann, “Using probabilistic movement primitives in robotics,” *Autonomous Robots*, vol. 42, no. 3, pp. 529–551, Mar. 2018. [Online]. Available: <https://doi.org/10.1007/s10514-017-9648-7>
- M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for machine learning*. Cambridge University Press, 2020.
- C. M. Bishop and H. Bishop, *Deep Learning: Foundations and Concepts*. Springer International Publishing, 2024. [Online]. Available: <https://doi.org/10.1007/978-3-031-45468-4>