$$
\begin{aligned}
P(h_t \mid Y_t) &\propto P(Y_t \mid h_t) P(h_t) \propto P(Y_t \mid h_t) \\
&\propto \exp\left[\sum_k \frac{\left(Y_{t,k} - v_t\right)^\top R(h_t) \Gamma (C x_t + d)_k}{s_{t,k} \sigma_k^2}\right] \\
&= \exp\left[\frac{\operatorname{tr}\left[R(h_t) \Gamma (C x_t + d)_k \left(Y_{t,k} - v_t\right)^\top\right]}{s_{t,k} \sigma_k^2}\right] \\
&\propto \exp \operatorname{tr}[R(h_t) S] \quad \text{where} \quad S = \sum_k \Gamma\, (C x_t + d)_k \left(Y_{t,k} - v_t\right)^\top / \left(s_{t,k} \sigma_k^2\right) \\
&\propto \exp\left[\cos(h_t)\left(S_{1,1} + S_{2,2}\right) + \sin(h_t)\left(S_{1,2} - S_{2,1}\right)\right]
\end{aligned}
$$

Let $[\kappa\cos(\theta), \kappa\sin(\theta)]$ represent $\left[S_{1,1} + S_{2,2}, S_{1,2} - S_{2,1}\right]$ in polar coordinates. Then

$$
\begin{aligned}
P(Y_t \mid h_t) &\propto \exp[\kappa \cos(h_t) \cos(\theta) + \sin(h_t) \sin(\theta)] \\
&= \exp[\kappa \cos(h_t - \theta)] \propto \text{vM}(h_t \mid \theta, \kappa)
\end{aligned}
$$

Supplemental References

1 Markowitz, J. E. *et al.* Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature* **614**, 108-117 (2023). https://doi.org:10.1038/s41586-022-05611-2

2 Lin, S. *et al.* Characterizing the structure of mouse behavior using Motion Sequencing. (2022). https://doi.org:10.48550/ARXIV.2211.08497

3 Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N. & Liang, J. in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.* (eds Danail Stoyanov *et al.*) 3-11 (Springer International Publishing).

4 Sun, K., Xiao, B., Liu, D. & Wang, J. in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 5686-5696.

5 Zhang, L., Dunn, T., Marshall, J., Olveczky, B. & Linderman, S. in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics* Vol. 130 (eds Banerjee Arindam & Fukumizu Kenji) 2800--2808 (PMLR, Proceedings of Machine Learning Research, 2021).

6 Nath, T. *et al.* Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols* **14**, 2152-2176 (2019). https://doi.org:10.1038/s41596-019-0176-0

7 Wiltschko, A. B. *et al.* Mapping Sub-Second Structure in Mouse Behavior. *Neuron* **88**, 1121-1135 (2015). https://doi.org:10.1016/j.neuron.2015.11.031

8 Hsu, A. I. & Yttri, E. A. B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nature Communications* **12**, 5188 (2021). https://doi.org:10.1038/s41467-021-25420-x PMID - 34465784

9 Luxem, K. *et al.* Identifying behavioral structure from deep variational embeddings of animal motion. *Commun Biol* **5**, 1267 (2022). https://doi.org:10.1038/s42003-022-04080-7

10 Berman, G. J., Choi, D. M., Bialek, W. & Shaevitz, J. W. Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of the Royal Society, Interface / the Royal Society* **11** (2014). https://doi.org:papers3://publication/doi/10.1098/rsif.2014.0672

11 Bohnslav, J. P. *et al.* DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. *eLife* **10**, e63377 (2021). https://doi.org:10.7554/eLife.63377

12 Sun, J. J. *et al.* Caltech Mouse Social Interactions (CalMS21) Dataset. (2021). https://doi.org:10.22002/D1.1991

13 Ackerson, G. & Fu, K. On state estimation in switching environments. *IEEE Transactions on Automatic Control* **15**, 10-17 (1970). https://doi.org:10.1109/TAC.1970.1099359

14 Fox, E. B., Sudderth, E. B., Jordan, M. I. & Willsky, A. S. in *Proceedings of the 25th International Conference on Machine Learning* 312–319 (Association for Computing Machinery, 2008).

15 Andreella, A. & Finos, L. Procrustes Analysis for High-Dimensional Data. *Psychometrika* **87**, 1422-1438 (2022). https://doi.org:10.1007/s11336-022-09859-5