where

$$c \doteq \begin{pmatrix} 1 & 0 \end{pmatrix} e^{t \begin{pmatrix} 0 & 1 \\ -\tau^2 \alpha\beta & -\tau\alpha \end{pmatrix}} \begin{pmatrix} y_0 \\ \dot{y}_0 \end{pmatrix}. \quad (24)$$

The sampling distribution $p(\boldsymbol{\theta}|t^*, y^*) = \mathcal{N}(\boldsymbol{\theta}|A_H\boldsymbol{\Phi}, \Sigma_H)$ is then computed with the modified likelihood distribution

$$p(y^*|t^*, \boldsymbol{\theta}) = \mathcal{N}(y^*|\boldsymbol{M}_{t^*}\boldsymbol{\theta} + c, \sigma_y \boldsymbol{I}). \quad (25)$$

The mean of this distribution differs slightly from the distribution derived in section 3 so that

$$A_H\boldsymbol{\Phi} = \Sigma_H \left(\boldsymbol{M}_{t^*}^T \sigma_y^{-1} \boldsymbol{I} \left(B\boldsymbol{\Phi} - c \right) + \Sigma_{\pi(k)}^{-1} A_{\pi(k)} \boldsymbol{\Phi} \right). \quad (26)$$

Note that we can assume w.l.o.g. that $\boldsymbol{\Phi}$ is of the form $\boldsymbol{\Phi} = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & 1 \end{pmatrix}^T$. The sampling distribution is then defined by

$$\Sigma_H \doteq \Sigma_{\pi(k)} - \Sigma_{\pi(k)} \boldsymbol{M}_{t^*}^T \left(\sigma_y \boldsymbol{I} + \boldsymbol{M}_{t^*} \Sigma_{\pi(k)} \boldsymbol{M}_{t^*}^T \right)^{-1} \boldsymbol{M}_{t^*} \Sigma_{\pi(k)}, \quad (27)$$

$$A_H \doteq \Sigma_H \left(\boldsymbol{M}_{t^*}^T \sigma_y^{-1} \boldsymbol{I} \left(B - \begin{pmatrix} 0 & \dots & 0 & c \end{pmatrix} \right) + \Sigma_{\pi(k)}^{-1} A_{\pi(k)} \right). \quad (28)$$

REFERENCES

[1] B. Akgun, M. Cakmak, J. W. Yoo, and A. L. Thomaz. Trajectories and keyframes for kinesthetic teaching: a human-robot interaction perspective. In *International Conference on Human-Robot Interaction*, pages 391–398, 2012.

[2] B. Argall, E. Sauser, and A. Billard. Policy adaptation through tactile correction. In *Annual Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)*, 2010.

[3] S. Bitzer, M. Howard, and S. Vijayakumar. Using dimensionality reduction to exploit constraints in reinforcement learning. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 3219–3225, 2010.

[4] S. Chernova and A. L. Thomaz. Robot learning from human teachers. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 8(3):1–121, 2014.

[5] C. Daniel, G. Neumann, and J. Peters. Hierarchical relative entropy policy search. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2012.

[6] M. P. Deisenroth, G. Neumann, and J. Peters. A survey on policy search for robotics. *Foundations and Trends in Robotics*, 2(1-2):1–142, 2013.

[7] S. Griffith, K. Subramanian, J. Scholz, C. Isbell, and A. L. Thomaz. Policy shaping: integrating human feedback with reinforcement learning. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2625–2633, 2013.

[8] N. Hansen and A. Ostermeier. Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. In *International Conference on Evolutionary Computation (ICEC)*, pages 312–317, 1996.

[9] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal. Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Computation*, 25(2):328–373, 2013.

[10] A. J. A. Ijspeert, J. Nakanishi, and S. Schaal. Learning attractor landscapes for learning motor primitives. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1547–1554, 2002.

[11] K. Judah, S. Roy, A. Fern, and T. G. Dietterich. Reinforcement Learning Via Practice and Critique Advice. *AAAI*, 2010.

[12] W. B. Knox and P. Stone. Reinforcement learning from simultaneous human and MDP reward categories and subject descriptors. In *Autonomous Agents and Multiagent Systems (AAMAS)*, pages 475–482, 2012.

[13] J. Kober, E. Oztop, and J. Peters. Reinforcement learning to adjust robot movements to new situations. In *Robotics: Science and Systems (RSS)*, 2010.

[14] J. Kober and J. Peters. Policy search for motor primitives in robotics. In *Advances in Neural Information Processing Systems (NIPS).*, pages 849–856, 2008.

[15] P. Kormushev, S. Calinon, and D. G. Caldwell. Robot motor skill coordination with EM-based reinforcement learning. In *Intelligent Robots and Systems (IROS)*, pages 3232–3237, 2010.

[16] A. G. Kupcsik, M. P. Deisenroth, J. Peters, and G. Neumann. Data-Efficient Generalization of Robot Skills with Contextual Policy Search. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2013.

[17] H. Miyamoto, S. Schaal, F. Gandolfo, H. Gomi, Y. Koike, R. Osu, E. Nakano, Y. Wada, and M. Kawato. A Kendama learning robot based on bi-directional theory. *Neural Networks*, 9(8):1281–1302, 1996.

[18] K. Mülling, J. Kober, O. Kroemer, and J. Peters. Learning to select and generalize striking movements in robot table tennis. *International Journal of Robotics Research*, 32(3):263–279, 2013.

[19] A. Paraschos, C. Daniel, J. Peters, and G. Neumann. Probabilistic movement primitives. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2616–2624, 2013.

[20] J. Peters, K. Mülling, and Y. Altun. Relative Entropy Policy Search. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 1607–1612, 2010.

[21] S. Schaal. Dynamic movement primitives - a framework for motor control in humans and humanoid robotics. In *Adaptive Motion of Animals and Machines*, pages 261–280. Springer Tokyo, 2003.

[22] F. Stulp and O. Sigaud. Path integral policy improvement with covariance matrix adaptation. In *International Conference on Machine Learning (ICML)*, pages 281–288, 2012.

[23] G. Teschl. *Ordinary differential equations and dynamical systems*, volume 140. American Mathematical Soc., 2012.

[24] E. Theodorou, J. Buchli, and S. Schaal. Learning policy improvements with path integrals. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 828–835, 2010.

[25] Y. Wada, Y. Koike, E. Vatikiotis-Bateson, and M. Kawato. A computational model for cursive handwriting based on the minimization principle. In *Advances in Neural Information Processing Systems (NIPS)*, pages 727–734, 1993.