Motor synergies are naturally observed during goal babbling

Matthias Rolf and Minoru Asada

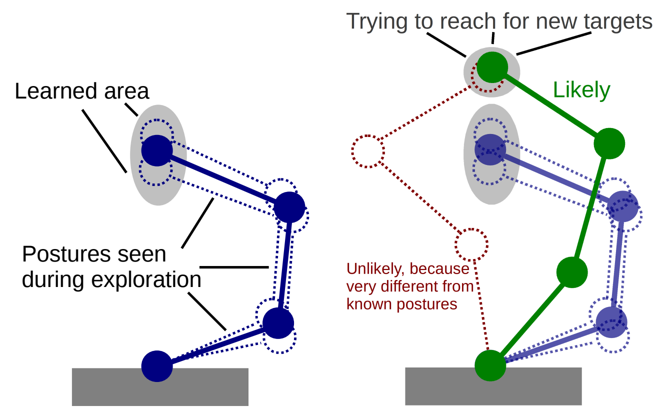
***Abstract*—Motor synergies describe joint activations of muscles or motors during sensorimotor tasks. Hard-wired synergies have been discusses to facilitate and ease motor control at the levels of mechanical actuation of biological systems and spinal motor control. Long-lasting debates additionally issue adaptive synergies as *mechanism for learning* in sensorimotor domains, such as a "freezing and freeing" of degrees of freedom in order to reduce the dimension of sensorimotor tasks. This paper discusses the terms of synergies and DOF freezing in the light of recent work on *Goal Babbling*. We discuss conceptual, experimental, and theoretic evidence showing that synergies and DOF freezing must be distally observable phenomena during goal babbling. We therefore argue that synergies might be viewed as an *effect* rather than the mechanism of efficient sensorimotor learning.**

I. Introduction

Motor synergies describe phenomena in which "multiple muscles are bound together such that a central control signal jointly and proportionally activates all muscles in the synergy." [1] "According to the strong version of this concept, these synergies are invariant, hard-wired patterns of activation across muscles." [2] Motor synergies are widely believed to simplify motor control on several levels. Even at physical actuation levels degrees of freedom are often bound together, for instance in the human hand which is substantially underactuated. Synergies are also found in vertebrates' spinal cords, such as reported for frogs [3] where fixed wirings have been reported that couple subsequent muscle activations. These hard-wired patterns of underactuation and spinal organization are matters of physical constraints that are hypothesized to simplify cortical and cerebellar motor control.

Synergies are not only reported on such physical levels, but also as observations in behavioral experiments. Synergies are often distally observed during (not only [2]) humans' *skilled* execution of sensorimotor tasks. For instance when humans manipulate or just hold objects, strong statistical regularities are found both with respect to muscular forces as well as skeletal posture [4]. Likewise stereotypical muscle activations are found in the arm when performing fast reaching movements [5]. In these scenarios synergies are measured in terms of low-dimensional linear dependencies among muscle activations. These can be found in terms of linear regression methods, or more generally in terms of matrix factorization techniques [6] such as principal component analysis. If, e.g. few principle components with high variance are found (compared to many other dimensions with low variance), then these

M. Rolf and M. Asada are with the Department of Adaptive Machine Systems, Graduate School of Engineering, Osaka University, Japan {matthias,asada}@ams.eng.osaka-u.ac.jp

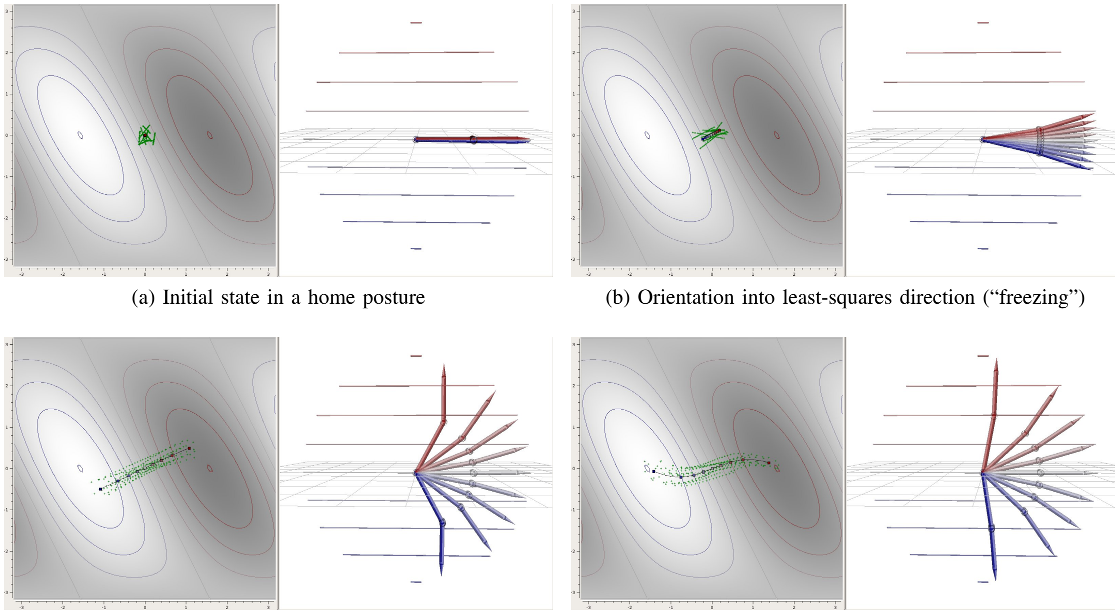


(a) Experience already gathered. (b) Estimate for a new target.

Fig. 1. The concept of goal babbling defines exploratory learning by means of ongoing, goal-directed estimates of how to achieve goals. Since new actions are estimated based on the learned structure of previously explored actions, they must be statistically similar to these previous actions. From a distal perspective this structure can be diagnosed as existence of motor synergies.

components are considered as motor synergies. In contrast to the level of physical constraints it is not generally clear whether such observed synergies are an internal mechanism to structure sensorimotor control, or rather an effect of other mechanisms that reflects (optimal) mastery [7] of a skill.

The question of synergies being mechanism or effect gets even more delicate when investigating the initial *learning* of sensorimotor skills, instead of their highly skilled execution. Learning sensorimotor skills from scratch often concerns the coordination of many degrees of freedom. Such high-dimensional spaces do not permit an exhaustive exploration such that mechanisms to reduce the demand for exploration are necessary. Already Bernstein [8] noticed this problem, and suggested a staged learning process along which infants bootstrap their initial repertoire of motor skills: they would firstly "freeze" certain combinations of degrees of freedom, i.e. using motor synergies, in order to simplify the learning problem and then later on release (or "free") them to achieve versatile performance. In fact, such "freezing" phenomena have been empirically observed in some learning tasks [9], [10], [11]. The idea to alleviate the exploration problem in high dimensions by means of constraining synergies has been picked up in robotics studies [12]. Yet, the choice of these synergies is delicate. Finding the right motor synergy to solve a difficult coordination problem does not seem to be simpler than solving the task right away. Rather, unsuited choices of constraints can



(a) Initial state in a home posture

(b) Orientation into least-squares direction ("freezing")

(c) Expansion along least squares direction

(d) Non-linear decoupling ("freeing") and tuning

Fig. 2. An inverse model of a simple planar arm is learned with goal babbling. The arm as two degrees of freedom. The 2D "action space" of possible postures is shown on the left-hand side of each picture. The task is to control only the effector's height, whose values are color-coded (red: high; blue: low). The images show successive epochs of the learning process. Examples (green points) are suggested by trying to reach with the inverse model. This sampling around the inverse model creates a quasi-one-dimensional data distribution, which can be interpreted from a distal perspective as motor synergy.

easily impair the movement ability. Consequently, this choice demanded substantial manual design in the present robotics studies [12], [13]. So far no study could demonstrate how to choose useful synergies fully autonomously for a previously unknown sensorimotor problem that is subject to learning.

II. Motor Synergies during Goal Babbling

The central argument to use motor synergies as constraint on learning was always that too many degrees of freedom can not be effectively explored. Hence, a motor learning problem should be boiled down to low dimensions so that exhaustive random exploration is feasible. This argument needs to be thoroughly rethought with the advent of highly efficient goal-directed exploration schemes like goal babbling [14]. Such schemes have shown to permit robot learning with human-competitive speed [15] even in very high dimensions. Plus, they give a direct account for how humans' learning might deal with many degrees of freedom by showing the usefulness of early goal-directed movement attempts that have been observed already in neonates [16].

At the most general level the concept of goal babbling is defined as follows:

> **Definition [17]: Goal babbling** is the bootstrapping of a coordination skill by repetitively trying to accomplish multiple goals related to that skill.

This definition does not refer to a particular algorithm or implementation, but to a general approach of "learning by doing". When an agent explores actions or motor commands by trying to accomplish goals, it has to make estimates of *how* to reach them. This very process of making an educated guess based on previously seen and learned experience facilitates the occurrence of distally observable motor synergies: suppose the agent has just explored very few actions (see Fig. 1(a)). It is gradually learning a motor controller in order to reach for them. If the agent tries to reach for a goal it has not experienced yet, the learning controller has to extrapolate. Since any learning is based on statistical regularities, the extrapolated motor command will display many similarities with previous motor commands (see Fig. 1(b)). Motor commands that differ drastically from previous experience will barely be chosen. For instance, if all gathered experience shows that two joints move "together" (like the two joints on the right side of Fig. 1(a) and (b)) it is unlikely that a learner generalizes this data by moving both joints into opposite directions. Hence, all motor commands share statistical properties like a general co-activation of certain joints or muscles that, when observed from a distal perspective, are easily interpreted as motor synergies.

The precise nature of these statistical regularities does of course depend on the particular algorithmic implementation of goal babbling. Several schemes have been proposed so far [18], [19], [20], [21]. Here we focus on original work in [14], which implements goal babbling for the learning of *inverse models*. An inverse model itself can be seen as a nonlinear motor synergy: it suggests actions in some potentially high-dimensional action space for any possible goal which has most often a lower dimension. Since low-dimensional goals are projected into a high-dimensional space of motor commands, these actions must lie on a low-dimensional manifold within that space.

Very specific synergy-like phenomena can be observed during the learning of such models with goal babbling. Fig. 2 shows an example [14] of a robot arm with only two degrees of freedom. Only the height of the end-effector shall be controlled, such that the task comprises redundancy. The left side of each image shows the two-dimensional space of motor commands, in which actions are explored by querying the inverse model (one-dimensional manifold with colored markers) plus adding some exploratory noise. At any point in time, the distribution of data-points (green dots) follows the inverse model and constitutes an essentially one-dimensional structure. This statistical structure could be interpreted as a single motor synergy with significant variance of motor commands, plus an orthogonal direction with only little variance. Both joints seem to be coupled. In fact they are. Yet, the way they are coupled is not a constraint imposed before learning starts (in order to ease it) but a *result* of combining (i) the learning of an inverse model with (ii) goal-directed exploration that already exploits the statistical regularities obtained by learning.

A more detailed inspection of Fig. 2 shows that learning expands into the direction of the steepest gradient from the very beginning (b). A following phase of expansion (c) remains within the same linear subspace inside the space of motor commands, i.e. both joints seem to be linearly coupled. Only in the end (d) this linear coupling is resolved and the inverse model gets well-tuned to the coordination problem in a non-linear manner. This behavior is very similar to reports of freezing (strict coupling) and later on freeing of degrees of freedom, but again is an effect rather than a mechanism inherent to learning.

These observations are highly repeatable based on the algorithm in [14], which collects several examples before a learning step is made. This process rules out many stochastic fluctuations that appear if online learning is applied [15]. In online scenarios the inverse model does not necessarily expand into the unique least-squares direction, but also into less optimal directions. However, the distribution of explored actions is still spanned by means of an inverse model such that a low-dimensional structure is explored inside the high-dimensional space of motor commands. These motor commands thus share statistical properties in a very similar way as sketched previously at a conceptual level (compare Fig. 3 and 1). Even if these solutions are not optimal in the very beginning, they have been shown to finally return into a regime of well-tuned solutions [15], so that also cross-trial comparisons display highly comparable statistics.

III. Fixed Motor Synergies in Linear Domains

Goal babbling is most useful for robotics applications when non-linear problems are considered. Linear domains, on the other hand, are well suited for deeper theoretical investigations. In a linear domain we can consider the behavior of the motor apparatus (or more generally: world) as a linear *forward function* $f(q) = M \cdot q = x$. This function turns actions or

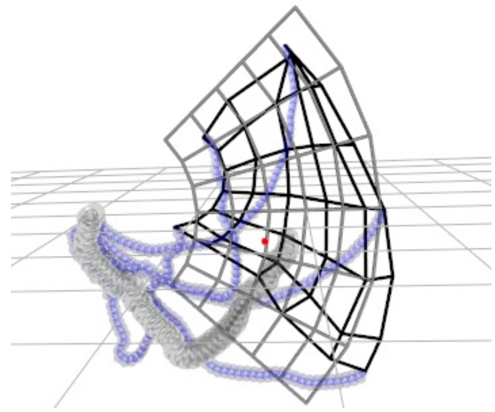


Fig. 3. Online goal babbling on a planar robot arm with 50 degrees of freedom. Motor commands chosen during exploration share common, though initially suboptimal, properties. In this trial one joint in the middle of the arm is heavily used in particular for goals close to the robot arm's basis (left side).

motor commands $q \in \mathbb{R}^m$ into their lower-dimensional causal outcomes $x \in \mathbb{R}^n$ (e.g. the resulting hand position). $M \in \mathbb{R}^{n \times m}$ is a matrix of coefficients that expresses the degree to which each dimension of motor commands influences each dimension of outcomes. Solvable sensorimotor tasks generally have dimensions $n \leq m$. Often, m is substantially larger than n, i.e. there are many more degrees of freedom than actual task variables.

Learning is then formulated by means of a linear *inverse model* $g(x^*) = W \cdot x^* = \hat{q}$ that suggests actions $\hat{q}$ to achieve goals $x^* \in \mathbb{R}^n$. $W \in \mathbb{R}^{m \times n}$ is a matrix of parameters adaptable by learning. Goal babbling can be performed by generating explorative actions with this inverse estimate, plus a perturbation term ε. For the analysis it is assumed that actions $q_k^{(t)}$ are chosen for each goal (from some fixed set) and for several perturbations in each timestep t, which mimics the algorithm in [14]:

$$q_k^{(t)} = W_{gen}^{(t)} x_k^* \quad \text{with} \quad W_{gen}^{(t)} \sim W_t + \varepsilon \ .$$

The components of the perturbation $\varepsilon \in \mathbb{R}^{m \times n}$ are chosen i.i.d. with zero mean and variance σ^2. Each of these actions is performed, and its outcome $x_k^{(t)} = f(q_k^{(t)})$ is observed. The combined data set D of actions q and corresponding outcomes x is then used for a supervised learning step by means of gradient descent:

$$W_{t+1} = W_t - \eta \frac{\partial E^Q(W, D)}{\partial W} \ , \qquad (1)$$

where η is some learning rate and E^Q is the supervised error resulting from deviations $g(x, W) - q$ of the inverse model with respect to the set of data.

This exploration strategy can now be investigated with respect to the generated data distribution and its change over time. Motor synergies are empirically investigated by means of matrix factorization techniques such as PCA [6], which we can directly apply to the linear case analysis of goal babbling. For a PCA we need to consider the multidimensional correlations of the observed motor commands. Assuming a fixed value of the learning parameters W for the exploration strategy sketched above, this correlation can be derived as [17]:

$$\mathbb{Q} = E\left[q \cdot q^T\right] = W\mathbb{X}^* W^T + \sigma^2 \mathbb{1}_m \in \mathbb{R}^{m\times m} \ ,$$

where $\mathbb{X}^* \in \mathbb{R}^{n\times n}$ is the autocorrelation of goals x^*. By construction, the left-hand part of this equation is positive semi-definite with a maximum number of n positive Eigenvalues $\lambda_i > 0$. In terms of a spectral composition (as used for PCA) it can consequently be factorized into matrices $V, D \in \mathbb{R}^{m\times m}$ such that:

$$W\mathbb{X}^* W^T = V\cdot D\cdot V^{-1}, \quad D = \mathrm{diag}(\lambda_1, ..., \lambda_n, \underbrace{0, ..., 0}_{m-n \text{ times}}) \ .$$

This allows to factorize the entire correlation matrix $\mathbb{Q}$ into

$$\mathbb{Q} = V\cdot D\cdot V^{-1} + \sigma^2 \mathbb{1}_m = V \cdot \left(D + \sigma^2 \mathbb{1}_m\right) \cdot V^{-1} \ ,$$

with the same Eigenvectors in V, and Eigenvalues

$$\left(D + \sigma^2 \mathbb{1}_m\right) = \mathrm{diag}(\lambda_1 + \sigma^2, ..., \lambda_n + \sigma^2, \underbrace{\sigma^2, ..., \sigma^2}_{m-n \text{ times}}) \ .$$

Hence, the data distribution has $m-n$ equally small principle components and n larger principle components at any point in time (see Fig. 4(a)). This describes exactly the statistical structure that is used to empirically search for motor synergies in biological motion.

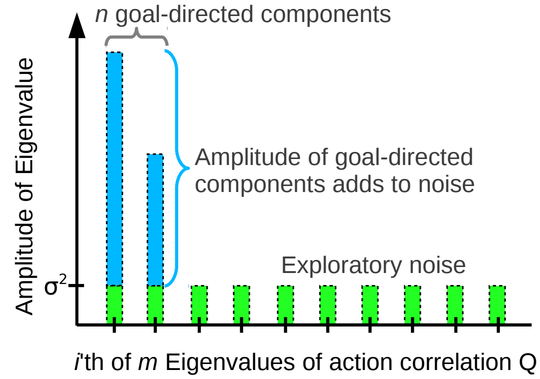
This result holds instantaneously only for a fixed value of the learning parameters W. If W changes as result of learning, actions can be projected into different subspaces so that more different significant principle components are observed over a cumulative learning experiment. However, novel results in [22] show that, although W changes, it remains *within the same subspace* in relevant scenarios. The central tool to study these subspaces during goal-directed exploration is to analyze the *column space* of the parameter matrix W, which mainly drives the exploration and constitutes the principle components belonging to the largest Eigenvalues as discussed above. The columns $\vec{w}_{(i)}$ of W contain the action space directions in which the inverse model projects different dimensions of a goal x^* in observation space:

$$g(x^*) = W \cdot x^* = \sum_{i=1}^{n} \vec{w}_{(i)} \cdot x^*_{(i)} \ .$$

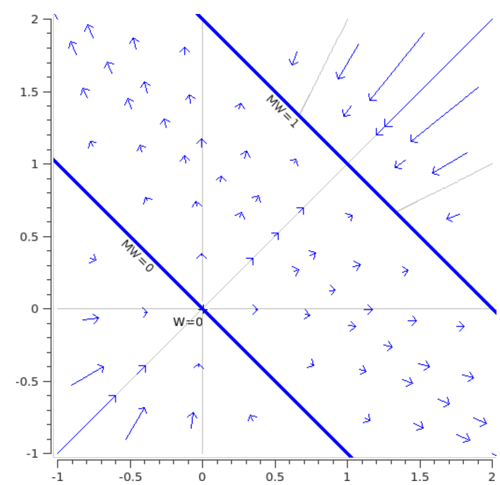
As a first scenario we can consider exploration without exploratory noise ($\sigma^2 = 0$). In that case it can be shown that the column space does never change:

Theorem 1: For $\sigma^2 = 0$, goal-directed exploration remains in the initial column space of W_0, i.e. any W_t can be factorized into $W_0 \cdot P_t$ with $P_t \in \mathbb{R}^{n\times n}$. [22]

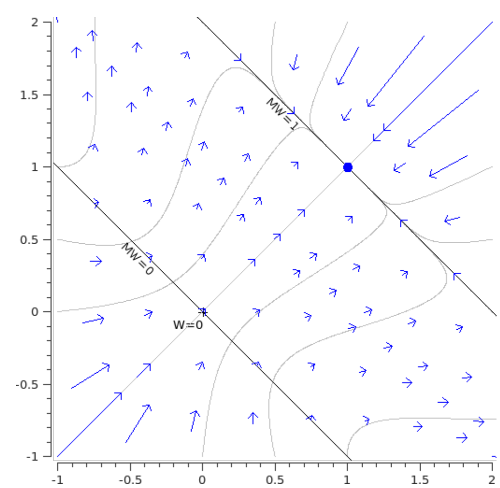
Such development of parameters W is illustrated in Fig. 4(b) for a simple problem with $n = 1$, $m = 2$, $M = (0.5, 0.5)$, and two learning parameters $W = (w_1, w_2)^T$. The plot shows paths of these parameters through the space w_1/w_2 during learning, which are entirely concentric. Learning only re-scales existing components. Learning can get stuck in the Nullspace ($MW =$



(a) Amplitude of the actions' principle components (referred to as synergies) during goal babbling in a linear domain.



(b) Learning parameters for goal-directed exploration without noise stay in the ever-same linear subspace.



(c) When noise is added, parameters still develop linearly between zero and the fixpoint.

Fig. 4. Goal babbling for the learning of inverse models in linear domains: Relevant cases describe a learning process that remains in a fixed linear subspace. This subspace reflects the directions of motor commands in which goals are projected, and results in distinct principal components.

0), or can converge to any possible solution of the coordination problem, which is formally characterized by $MW = 1$.

The situation is more complex for learning *with* exploratory noise. It can be shown that this process always converges to the unique least squares solution instead of any possible solution:

Theorem 2: For $\sigma^2 > 0$, the unique fixpoint of the learning equation 1 is the Moore-Penrose pseudoinverse: $W = M^{\#} = M^T(MM^T)^{-1}$. [17]

Fig. 4(c) shows that this can lead to severely non-linear paths through the parameter space, and thus time-varying synergies. Yet, when starting from an neutral parameter value $W = 0$ that does not impose any direction, it can be shown that the column space remains fixed also for exploration with noise:

Theorem 3: For $\sigma^2 > 0$ and $W_0 = 0$, goal-directed exploration remains in the column space of M^T, i.e. any W_t can be factorized into $M^T \cdot P_t$ with $P_t \in \mathbb{R}^{n \times n}$. [22]

In Fig. 4(c) this is visible by the straight path between $W = 0$ and $W = M^{\#} = (1, 1)^T$. If the column space of the parameters does not change, also the directions of high variance inside the space of motor commands do not change. Even a cumulative statistical analysis of an entire learning process must reveal a spectrum of Eigenvalues as shown in Fig. 4(a). Most variability of motor commands is explained by few variables, which is commonly argued to indicate motor synergies. In the case of goal babbling, this is not a mechanism, but an *effect* of the conceptual organization of learning (see Fig. 1), and the particular dynamics of its implementations.

IV. Discussion

Hard-wired, evolutionary shaped motor synergies to simplify motor control problems very certainly exist. When it comes to learning entirely novel sensorimotor coordination problems, however, the traditional believe that the central nervous system creates novel, "soft" synergies to simplify learning must be challenged. Not only is it hard to autonomously choose synergies that are actually useful for learning. Novel conceptual and algorithmic developments that mimic infants' early goal-directed movements also alleviate the demand to reduce the dimension of the to-be-explored space. This paper has demonstrated on different levels how the goal babbling approach can nevertheless account for the apparent existence of synergies in experiments on biological learning. Instead of assuming synergies as a fixed mechanism, they can appear on a behavioral level as a mere result of learning. During goal babbling this structure is induced by goals, which can be similarly observed during kinesthetic teaching [23], [24] of skills, where an experienced teacher induces typically low-dimensional patterns. Specific transitions like "free(z)ing" of degrees of freedom are also observed in experiments on goal babbling in non-linear domains. Similar effects have also been observed in goal-directed exploration processes for reinforcement learning [25], which indicates that such observations might be a rather general effect of efficient learning schemes.

Acknowledgements

This study has been supported by the JSPS Grant-in-Aid for Specially promoted Research (No. 24000012).

References

[1] M. L. Latash, J. P. Scholz, and G. Schöner, "Toward a new theory of motor synergies," *Human Kinetics*, vol. 11, no. 3, pp. 276–308, 2007.

[2] S. A. Overduin, A. d`Avella, J. Roh, and E. Bizzi, "Modulation of muscle synergy recruitment in primate grasping," *Journal of Neuroscience*, vol. 28, no. 4, pp. 880–892, 2008.

[3] P. Saltiel, K. Wyler-Duda, A. D`Avella, M. C. Tresch, and E. Bizzi, "Muscle synergies encoded within the spinal cord: Evidence from focal intraspinal nmda iontophoresis in the frog," *Journal of Neurophysiology*, vol. 85, no. 2, pp. 605–619, 2001.

[4] E. J. Weiss and M. Flanders, "Muscular and postural synergies of the human hand," *Journal of Neurophysiology*, vol. 92, no. 1, 2004.

[5] A. d`Avella, A. Portone, L. Fernandez, and F. Lacquaniti, "Control of fast-reaching movements by muscle synergy combinations," *Journal of Neuroscience*, vol. 26, no. 30, pp. 7791–7810, 2006.

[6] M. C. Tresch, V. C. K. Cheung, and A. d`Avella, "Matrix factorization algorithms for the identification of muscle synergies: Evaluation on simulated and experimental data sets," *Journal of Neurophysiology*, vol. 95, no. 4, pp. 2199–2212, 2006.

[7] E. Todorov, "Optimality principles in sensorimotor control," *Nature Neuroscience*, vol. 7, no. 9, pp. 907–915, 2004.

[8] N. Bernstein, *The Co-ordination and Regulation of Movements*. Pergamon Press, 1967.

[9] N. E. Berthier, R. K. Clifton, D. D. McCall, and D. J. Robin, "Proximodistal structure of early reaching in human infants," *Experimental Brain Research*, vol. 127, no. 3, pp. 259–269, 1999.

[10] K. Newell and R. van Emmerik, "The acquisition of coordination: Preliminary analysis of learning to write," *Human Movement Science*, vol. 8, no. 1, pp. 17–32, 1989.

[11] B. Vereijken, R. E. A. van Emmerik, H. T. A. Whiting, and K. M. Newell, "Free(z)ing degrees of freedom in skill acquisition," *Journal of Motor Behavior*, vol. 24, no. 1, pp. 133–142, 1992.

[12] L. Berthouze and M. Lungarella, "Motor skill acquisition under environmental perturbations: On the necessity of alternate freezing and freeing of degrees of freedom," *Adaptive Behavior*, vol. 12, no. 1, 2004.

[13] H. Hauser, , G. Neumann, A. Ijspeert, and W. Maass, "Biologically inspired kinematic synergies provide a new paradigm for balance control of humanoid robots," in *IEEE-RAS/RSJ International Conference on Humanoid Robots (Humanoids)*, 2007.

[14] M. Rolf, J. J. Steil, and M. Gienger, "Goal babbling permits direct learning of inverse kinematics," *IEEE Trans. Autonomous Mental Development*, vol. 2, no. 3, 2010.

[15] ——, "Online goal babbling for rapid bootstrapping of inverse models in high dimensions," in *IEEE Int. Joint Conf. Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2011.

[16] C. von Hofsten, "Eye-hand coordination in the newborn," *Developmental Psychology*, vol. 18, no. 3, pp. 450–461, 1982.

[17] M. Rolf, "Goal babbling for an efficient bootstrapping of inverse models in high dimensions," PhD Thesis, Bielefeld University, 2012. [Online]. Available: http://pub.uni-bielefeld.de/publication/2551092

[18] A. Baranes and P.-Y. Oudeyer, "Active learning of inverse models with intrinsically motivated goal exploration in robots," *Robotics and Autonomous Systems*, vol. 61, no. 1, pp. 49–73, 2013.

[19] L. Jamone, L. Natale, K. Hashimoto, G. Sandini, and A. Takanishi, "Learning task space control through goal directed exploration," in *IEEE Int. Conf. Robotics and Biomimetics (ROBIO)*, 2011.

[20] C. Hartmann, J. Boedecker, O. Obst, S. Ikemoto, and M. Asada, "Real-time inverse dynamics learning for musculoskeletal robots based on echo state gaussian process regression," in *RSS*, 2012.

[21] P. O. Stalph and M. V. Butz, "Learning local linear jacobians for flexible and adaptive robot arm control," *Genetic Programming and Evolvable Machines*, vol. 13, no. 2, pp. 137–157, 2012.

[22] M. Rolf and J. J. Steil, "Explorative learning of inverse models: a theoretical perspective," *Neurocomputing*, 2013, in Press.

[23] M. Hersch, F. Guenter, S. Calinon, and A. Billard, "Dynamical system modulation for robot learning via kinesthetic demonstrations," *IEEE Trans. Robotics*, vol. 24, no. 6, pp. 1463–1467, 2008.

[24] A. Lemme, A. Freire, G. Barreto, and J. Steil, "Kinesthetic teaching of visuomotor coordination for pointing by the humanoid robot icub," *Neurocomputing*, vol. 112, pp. 179–188, 2013.

[25] F. Stulp and P.-Y. Oudeyer, "Emergent proximo-distal maturation through adaptive exploration," in *IEEE Int. Joint Conf. Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2011.