Dynamical self-consistency leads to behavioral development and emergent social interactions in robots.

Ralf Der[1] and Georg Martius[2]

Abstract— **We present an approach that enables robots to self-organize their sensorimotor behavior from scratch without providing specific information about neither the robot nor its environment. This is achieved by a simple neural control law that increases the consistency between external sensor dynamics and internal neural dynamics of the utterly simple controller. In this way, the embodiment and the agent-environment coupling are the only source of individual development. We show how an anthropomorphic tendon driven arm-shoulder system develops different behaviors depending on that coupling. For instance: given a bottle half-filled with water, the arm starts to shake it, driven by the physical response of the water. When attaching a brush, the arm can be manipulated into wiping a table, and when connected to a revolvable wheel it finds out how to rotate it. Thus, the robot may be said to discover the affordances of the world. When allowing two (simulated) humanoid robots to interact physically, they engage into a joint behavior development leading to, for instance, spontaneous cooperation. More social effects are observed if the robots can visually perceive each other. Although, as an observer, it is tempting to attribute an apparent intentionality, there is nothing of the kind put in. As a conclusion, we argue that emergent behavior may be much less rooted in explicit intentions, internal motivations, or specific reward systems than is commonly believed.**

I. Introduction

The long term goal of epigenetic and developmental robotic research is to create autonomous, self-motivated, and intelligent animats [1]. At the core of such a development, there must be some principle or generic drive that guides the learning system towards this goal. Ideally the learning system should fulfill some elemental requirements, such as to be task agnostic, to be open to new environments, to operate with raw sensory information, and to perform online learning in a continual self-determined process. One of the lessons learned from robot control is that the exploitation of the particular agent-environment interaction is of great importance, simplifying fulfillment of tasks considerably, for instance by increasing robustness and energy efficiency with reduced computational demands [2]. In this paper, we present an approach that takes embodiment and agent-environment coupling as the only source of accessible information; and we are going to show how a self-determined development of coordinated and complex behaviors can emerge from a single neural control law without any of the commonly applied concepts like specific drives, intrinsic motivation, curiosity, specific reward systems, or the selection pressure in evolution to obtain a self-determined unfolding of behavior, however, see Sect. II-D for a connection.

This paper continuous the line of our previous works on self-organization (SO) of behavior which started with the principle of homeokinesis [3], [4], [5] about two decades ago. There, the controller adapts by gradient descending the so called time-loop error which balances high predictability and high sensitivity of the dynamical system formed by the sensorimotor loop. This gave rise to the playful machine scenario [5] where many simulated and real robots self-organize their behavior through the interaction with their environments in a playful and task-free way. In the search for further principles that lead to the SO of behavior the maximization of predictive information was proposed [6], [7].

Our most recent research aims at finding self-consistency principles to bring the embodiment even more into the foreground, as realized by differential extrinsic plasticity [8], [9] with its more technically oriented variant introduced in [10]. Notably, while the controller got simpler in each step, the range of applicability steadily increased, advancing from rigid body systems in computer simulations to the complex tendon-driven physical machines featuring also in this paper.

The paper is structured as follows: Section II gives a new derivation of the previous result [8], [10] based on a new self-consistency principle of behavior and internal dynamics. This controller is applied to a number of systems, starting in Sect. III with a muscle-tendon driven arm-shoulder system which is known to be resistant to classical control paradigms. Part of these behavior patterns have been shown already in [10] but we reconsider them here in the light of the self-consistency paradigm. In Sect. IV we apply the same controller to simulated humanoids in order to demonstrate the potential of emerging social interaction.

II. Controller dynamics

Before going into details let us give a brief sketch of the self-consistency idea. We consider underactuated systems with substantial embodiment, where future sensor values carry not only the footprints of earlier motor actions but also of the accompanying physical dynamics. By these competing physical effects, perfect control is an illusion if the system cannot be fully observed. Realistically, all one can aim at is reducing illusions to make the controller of an active, embodied system as much as possible the author of the behavior. For developing the self-consistency idea, think of a robot or human in a typical periodic pattern like crawling or

*GM received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. [291734].

[1]University of Leipzig, Institute of Informatics, 04103 Leipzig, Germany ralf.der@t-online.de

[2]IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria gmartius@ist.ac.at

walking, realized by a complex interplay of all the system's degrees of freedom, i. e. the mechanical ones (the bones), as well as the muscles, tissues and so on, and the internal neural dynamics. In particular, all sensor values recording the state of the body and its relation to the environment will reflect this periodicity.

Each individual state variable, although periodic, displays a complicated time structure. However, when looking at correlations across those variables and/or time, we find that stable periodic behaviors are characterized by steady, time invariant correlation patterns. Concretely, when considering a robot, human, or animal in a stable walking pattern, most state variables, like proprioceptive sensor values x_i, are in a fixed phase relation, mostly in phase or anti phase.

A. *Correlation patterns in perceptual space*

The related velocities[1] $\dot{x}_i$ are either of the same or of different sign, respectively, or they may be roughly zero for phase shifts of $\pm\pi/2$. Correspondingly, the matrix elements $V_{ij} = \langle \dot{x}_i \dot{x}_j \rangle$ where $\langle \ldots \rangle$ is the time average over one period, of the sensor correlation matrix V is a fixed structure for any cyclic motion pattern. Correlations across time may reveal further fixed points characterizing the more involved nature of specific motion patterns. In particular, we are interested in the correlations across a time lag $\theta = t' - t$

$$V'_{ij} = \langle \dot{x}'_i \dot{x}_j \rangle \tag{1}$$

where $t' > t$ is the earliest time at which the motor action generated at time t leaves a "footprint" in later sensor values, i. e. in $x' = x(t + \theta)$.

This is a central point of the paper: while in state space a cyclic behavior is a dynamical pattern, specifically a limit cycle, in correlation space it is a fixed point. So, our aim is a controller that learns to drive an embodied system toward fixed points in correlation space and stabilizes it there. This concept may be of more general interest for robotics as convergence toward a fixed point is more easily realizable than toward a limit cycle. Ideally, as there are many such fixed points—one for each cyclic behavior—the learning procedure should be able of finding them when starting in their respective basin of attraction. This paper will develop such an approach in detail. Having said that, the only remaining question is how the controller must be adapted so that convergence toward a fixed V is achieved.

B. *The controller*

In the applications, we use a neurocontroller realized by a one-layer feed-forward network with m neurons, neuron i defining the nominal value of motor i as

$$y_i = g\left(\sum_{j=1}^{n} C_{ij} x_j + h_i\right) \tag{2}$$

where C_{ij} is the synaptic connection strength to input j and h_i is the bias term, which is set to zero in this work ($h_i = 0$). We use $\tanh$-neurons, i. e. the activation function $g(z) = \tanh(z)$ to get motor commands between +1 and -1. The setup is displayed in Fig. 1.

The correlation matrix Eq. (1) can be converted into an expression for the C matrix by firstly postulating the existence of a mapping M that realizes the back projection of the sensor values x' at time t' to their causes, the motor values at time t. As we are interested in the velocities, we postulate $M\dot{x}' \approx \dot{y}$. Taking the time derivative of Eq. (2) we get the simple result[2] $\dot{y} \approx C\dot{x}$ and thus $M\dot{x}' \approx C\dot{x}$. Multiplying by $\dot{x}^\top$ and using the averages as in Eq. (1) we get[3] $C\langle \dot{x}\dot{x}^\top \rangle \approx M\langle \dot{x}'\dot{x}^\top \rangle$ and eventually[4] $C \propto M\langle \dot{x}'\dot{x}^\top \rangle$ at any of the fixed points in correlation space.

Before proceeding, we state that there is no contradiction, as one might think, between the simple structure of the controller and the expectation that the latter is able of generating a complex cyclic motion pattern with a fixed C matrix. Formally, this is achieved if the mapping $x \to x'$ has complex eigenvalues. Note that the correlation matrix in Eq. (1) may well have such complex eigenvalues as it is not symmetric. Common approaches, see [11], [12] are based on a definite internal dynamics of the controller itself. So, if the controller is detached from the robot, there is still the internal dynamics going on. Quite different, when detaching our controller from the sensor values, there is no dynamics left. Everything is generated only by the interplay with the body-environment coupling.

C. *Self-consistent learning rule*

By exploiting the embodiment, the minimalistic controller with a fixed C matrix is indeed able of realizing a stable cyclic behavior. However, in general we do not know anything about the specific C matrices supporting such a behavior. The idea is to find a learning dynamics that drives the C matrix toward such fixed points in correlation space. Now consider the dynamics in C space

$$\tau \Delta C_{ij} = \sum_l M_{il} \dot{x}'_l \dot{x}_j - C_{ij} \tag{3}$$

where τ sets the time scale and $-C_{ij}$ is a damping term. After convergence, i. e. at a fixed point in correlation space, Eq. (3) yields $C = M\langle \dot{x}'\dot{x}^\top \rangle$ where $\langle \ldots \rangle$ denotes now the moving average with time scale τ. This is taken as the desired result because (i) overall factors in the C matrix do not matter, see below, and (ii) the approximations may be ignored, making the learning rule as simple as possible. As a final step, in order to get the system into activity, we have to maintain a certain feedback strength in each motor channel. For that, we introduce a normalization with a factor

[1] We denote the rate of change of a quantity a by $\dot{a}$, which is interpreted as a velocity. Time derivatives are taken simply as time differences, i. e. $\dot{x}(t) = x(t+1) - x(t)$.

[2] By ignoring the nonlinearity of the $\tanh$ which is justified for small amplitudes of the motor signal y.

[3] In matrix notation, for any two column vectors a and b, $a^\top b$ is the scalar product and $S = ab^\top$ is a matrix with elements $S_{ij} = a_i b_j$.

[4] We use that in the average over one cycle and in the subspace of the state vectors, $\langle \dot{x}\dot{x}^\top \rangle$ is proportional to the unit matrix so that we can ignore that contribution.

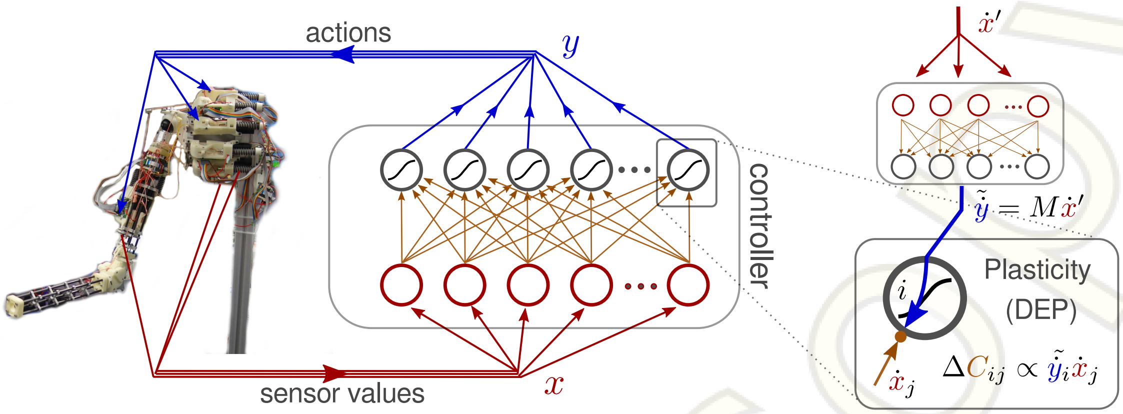

Fig. 1. **Neural controller network connected to the Myo-robotic arm.** The inset on the right illustrates the synaptic plasticity rule, called differential extrinsic plasticity (DEP) [8]. It consists of a modified differential Hebbian law, multiplying the time derivatives of the incoming sensor values $\dot{x}$ with the virtual motor values $\tilde{y}$, which are generated by the inverse model from the next input's derivative $\dot{x}'$. In the case of the arm the inverse model is essentially a one-to-one mapping of sensor to motor values.

κ (meta-parameter) into the controller, replacing Eq. (2) with $y_i = g\left(\sum_{j=1}^{n} \frac{\kappa}{(|C_i|^2+\lambda)} C_{ij} x_j + h_i\right)$, where $\lambda \ll 1$ is a regularization parameter, see [8] for details.

Remember that this result was obtained under the consistency assumption that the controller upholds the system in a periodic pattern with a fixed C matrix. However, when letting the C dynamics run freely using Eq. (3), we have a double dynamics: physical + parameter, which self-consistently generates behavior, with preference to converge toward limit cycles in physical space (where the parameter dynamics stalls). With a convenient feedback factor κ, this is what happens in the experiments. However, in practice, convergence to a fixed matrix is seldom achieved in the strict sense due to the approximations made and because a moving average of a cyclic quantity is still oscillating as long as $\tau < \infty$. Besides the limit cycle attractor scenario, synaptic and physical dynamics can also engage into more complicated interplay leading to fixed point flows in correlation space, see Sect. III-D below.

D. Remarks

Before applying the controller to concrete systems, let us discuss some details.

a) Differential extrinsic plasticity: The learning rule of Eq. (3) has a similarity to differential Hebbian learning, which uses $C_{ij} \propto \langle \dot{y}_i \dot{x}_j \rangle$, a purely internal quantity. Our rule is substantially different as the postsynaptic factor in the learning rule is given by the backpropagated response of the outside world, represented by $M\dot{x}'$. This may be called differential extrinsic plasticity (DEP) as already defined in earlier work [8]. Notably, the fact that the correlation structure of a Hebbian-like law can be directly mapped to the correlations in the outside world may shed new light on the success of Hebbian-like learning rules, see [8] for details.

b) Body inspired control vs. morphological computation: The leading role of the body in the generation of behavior is sometimes referred to as "morphological computation" naming the ulterior motive to shift computational load from the controler to the body. However, as argued by Hoffmann & Müller in [13] this is rather a misleading label, as the physical body is not computing in the classical sense. Of course, one may call the physical dynamics computation but if this is computation, what is not. Instead, the central question is to what extent the body contributes to the overall "orchestration" of intelligent behavior. A potential measure of that contribution could be given by the complexity gap between the controller and the emerging dynamics, provided that a convenient defined complexity measure. Intuitively, we argue that this gap is quite large in the examples of this paper, given that the minimalistic controller generates motion patterns of notable complexity in high dimensional systems. Here, the role of the body is not to just facilitate an orchestration which is essentially controlled by the brain. With our controller, the brain is nothing without the body. In this sense, our approach is more radical in the realization of said orchestration and can be considered as a kind of "radical embodied robotics", see also the remark on radical embodied cognitive science below.

c) Intrinsic mechanism vs. intrinsic motivation: The update rule Eq. (3) can also be obtained from the objective function $\|\delta\|^2$ with $\delta = \tilde{y} - \dot{y}$, where $\tilde{y} = M\dot{x}'$, which quantifies the illusion that the controller is the sole author of the future development. Gradient descending this measure is achieved[5] by just the synaptic dynamics Eq. (3). One may feel free to call this an intrinsic motivation to maintain activity in the sensorimotor loop and to increase velocity correlations between sensor values. Different from common measures like those based on information theory [6], [7] which are given by a single number, this objective function has a dynamical structure as it is a direct function of the current state of the system. Consequently, there is no sampling or long integration of information required. In this sense, if at all, $\|\delta\|$ can be considered as a new, dynamical systems

[5] In matrix notation and ignoring non-linearities as above, $-\frac{1}{2}\frac{\partial}{\partial C}\|\delta\|^2 = \tilde{y}\dot{x}^\top - C\dot{x}\dot{x}^\top$ giving Eq. (3) stipulating that $\langle \dot{x}\dot{x}^\top \rangle \propto \mathbb{I}$ in the subspace of the respective trajectories.

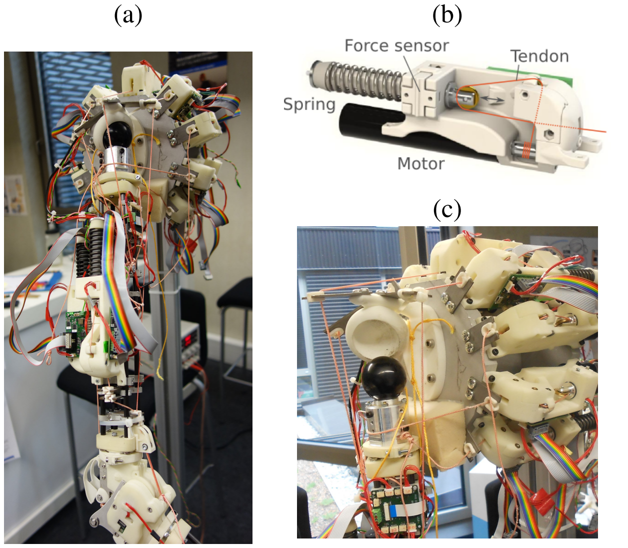

Fig. 2. Myorobotic arm (a) with 9 muscles and a ball shoulder joint, a single muscle element (b), and a dislocated shoulder (c). The dislocation happens wickedly as soon as the tendons are getting slack.

based measure of intrinsic motivation, which complements the developed typology of intrinsic motivations [14].

III. Tendon driven arm-shoulder system

Because of their high complexity and human like structure, anthropomimetic robots are a challenging example for testing the self-organization of behavior using Eq. (3). Different from classical robots, anthropomimetic robots are built following the morphology of the human body. Such robots are more soft than classical systems making them saver to interact with and thus favorable for service robots in human environments. Moreover, because of their human like morphology, they can be used for better understanding human behavior generation and development.

World wide, several of such muscle-tendon driven (MTD) systems have already been built. While mechatronically at an advanced level, the control of both MTD and soft robotic systems in general is still in its infancy. A generic example is given by pertinent EU projects ranging from CRONOS, to ECCEROBOT to MYOROBOTICS. While excellent work has been done in building these robots [15], their control faces fundamental problems [16], [17] and remains restricted so far to primitive behaviors like actuating just an elbow of the MTD system through computed muscle forces, see for instance [18].

In this paper we report on experiments with a tendon driven arm-shoulder system from the Myo-robotics toolkit, see Fig. 2. There are a number of features which make the muscle-tendon driven systems different from classical robots where motor positions directly translate into joint angles and into poses. The most obvious effects stem from the properties of the tendons themselves: they can get slack, wrap or even tangled. Otherwise, different kinds of joints can be used, such as the ball-socket joint, as in the shoulder, see Fig. 2(a), allowing for large reachable space with a single compact joint. However, the potential dislocation of the shoulder is an additional complication. These effects make it hard to predict the joint positions from the geometry and the motor positions. To reduce the difficulty and allow for a defined force transmission a permanent tension on the tendons has to be kept, which in turn poses another problem: The tension can only be achieved by tightening each tendon up against all the others, each individual tension being reported by the spring length. This means that (i) there are infinitely many combinations of tension forces for a single arm pose and (ii) the action of a single motor will be reflected in a change of spring length of all other muscles. In other words, actuating a single muscle is reflected by a pattern of sensory stimulation—a whole-body answer.

Furthermore, the combination of friction and muscle-pose ambiguity leads to hysteresis effects. In general, this makes the translation of a kinematic trajectory for the arm into motor programs extremely difficult, even more so if there are loads and high velocities involved. Although structurally extremely simple, the new controller copes effectively with these problems because it generates motor signals as a whole system answer. In particular, in all our experiments we never had a shoulder dislocation, see Fig. 2(c).

The robotic system is equipped with motor encoders measuring the length of the tendons and with force sensors measuring the spring displacement due to tendon tension. Each muscle i is controlled by a target length of the tendon y_i and provides a sensor value x_i comprised of the actual tendon length l_i combined with the spring length f_i as $x_i = l_i + f_i$. The spring length f_i is normalized to be in the interval $[-0.1, 0.9]$ where 0.1 is the initial pretension (at $f_i = 0$). The length is normalized to be in the interval $[-1, 1]$ and $l_i = 0$ in a manually set central pose. The controller is also supplied with delayed copies of the sensor values, i. e. the new sensor vector is $(x^\top(t), x^\top(t-d))^\top$ with delay time $d > 0$.

A. Behavior as resonance

All experiments with the arm are performed with the same controller introduced in Sect. II with the following parameter settings: $\kappa = 0.5$, $\tau = 1\,s$, $d = 0.5\,s$, $\theta = 0.08\,s$, $\lambda = 10^{-4}$, and a update frequency of the control loop of 100 Hz. The choice of the parameters are not critical, but influence the resulting behavior: κ changes the amplitude and needs to be large enough to create spontaneous behavior; τ regulates the search for a self-consistent behavior where low values corresponds to fast changes and may also be used to kick the system out of a stable behavior; d sets a preferred frequency, which, however, is not necessarily followed because any frequency can be achieved by a suitable coupling given at least 2 dimensions.

Depending on the initial pose and the physical embedding different behaviors develop. In order to follow the argumentation we recommend to watch the referred video clips summarized in Tab. I. By way of example, consider Video 2 where a pendulum was attached to the arm, as displayed in Fig. 3(a). In the beginning, minimal motor

TABLE I

VIDEOS FOR THE INDIVIDUAL EXPERIMENTS AVAILABLE AT `playfulmachines.com/ICDL2016`.

Title	Description	Sect.	Vid./Link
Overview 2–9	Compiled clip of all arm-shoulder experiments		Video 1
Bottle swing	Excitation of a circular pendulum mode	III-A	Video 2
Shaking vertically	A half filled bottle is vertically attached to the tip of the arm: shaking of the bottle mainly along its axis	III-A	Video 3
Shaking horizontally	Same as above but with horizontal attachment	III-A	Video 4
Rotating wheel	Arm attached to a revolvable bar/wheel	III-B	Video 5
Rotating wheel II	Parallel wheel-arm arrangement	III-B	Video 6
Rotating wheel III	Different rotation frequencies	III-B	Video 7
Wiping table	Arm with brush starts to wipe a table	III-C	Video 8
Wiping table modes	Different wiping patterns from reloaded controllers	III-C	Video 9
Free	No external forces applied: pseudo-random sequences of reaching-type behavior	III-D	Video 10
Crawling humanoid	Humanoid robot on the ground develops a crawling behavior from scratch	IV	Video 11
Humanoids at a wheel	Two humanoid robots hold on to the cranks of a wheel and jointly rotate it	IV	Video 12
Socializing I	Harmony in emerging behavior of two humanoids suspended on elastic ropes	IV-A	Video 13
Socializing II	Emerging patters with inverted vision	IV-A	Video 14
Socializing III	On stools, one robot is weakened and gets perturbed repeatedly	IV-A	Video 15
Socializing IV	Same as above, but with delayed vision	IV-A	Video 16
Alien body effects	Two humanoids percieve only sensors of other robot (inversed sign)	IV-B	Video 17
Fighters	Two humanoids fighting	IV-C	Video 18

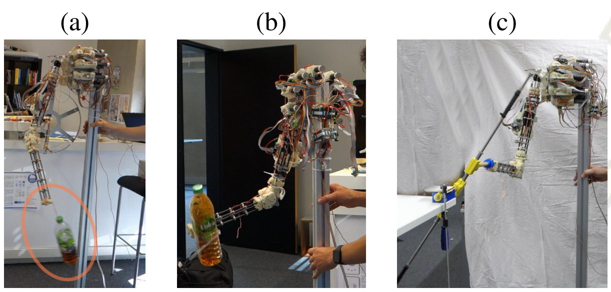

Fig. 3. **Some of the experimental setups with the arm-shoulder system.** Self-excited pendulum mode (a) (Video 2), vertical shaking setup (b) (Video 3), and frontal rotating bar/wheel setup (c) (Video 5).

activities are seen to spontaneously excite minor pendulum motions. These movements directly exert physical forces on the arm which propagate via the springs into the sensor values and eventually into the synaptic dynamics which governs the behavior. This may lead to the amplification of latent pendulum modes until self-consistency, i. e. a stable circular movement of the pendulum, is achieved. These findings elucidate how a physical subsystem, the pendulum, may pilot—by its internal dynamics—the meta-system into a resonant state, i. e. a whole-system mode with defined frequency.

In order to understand why the system is behaving this way, let us take the different steps apart. Initially, the synaptic connections C_{ij} are all 0 so that the motor output is also 0 which corresponds to the initialization length of each tendon. Any initial movement or forces applied to the arm (e. g. gravity) cause an non-zero value for the spring forces and thus in the measured upcoming sensor readings x'. These give rise to non-zero contributions $\dot{x}'\dot{x}$ in Eq. (3). As a remark, this is the decisive difference to differential or other Hebbian learning rules which correlate input and output (y) directly, where the output would remain 0 for all times. The product of $\dot{x}'\dot{x}$ defines the first non-zero entries in the connection matrix C which in turn influences the forthcoming actions y. This explains how the system departs from the unbiased initialization condition. The second major effect is the drive for self-consistency between system dynamics based on positional variables and the parameter dynamics based in velocities, which can be satisfied in limit cycles as explained in Sect. II. The particular limit cycle—the fixed point in correlation space—depends decisively on the responses of the system.

Thus, the emergent behaviors clearly depend on the physical subsystem. For instance, when attaching a bottle half-filled with water, see Fig. 3(b), in either horizontal or vertical orientation, case specific shaking modes arise, as demonstrated by Video 3 and Video 4. In this case, strong signals are circulating through the SM loop whenever the water hits either the side walls or top/bottom of the bottle. These signals may self-amplify and eventually generate motions of the arm in coherence with the physical responses.

B. *Affordances*

By the self-excitation mechanism, the controller may also discover (dynamical) affordances—in the sense of Gibson [19]—of the physical world, see also [20]. This is of interest for developmental robotics, as the discovery of object affordances may form pre-requisites for emerging tool use, higher-level control, prediction, and planning.

We consider the robotic arm connected to a crank of a revolvable wheel/bar, see Fig. 3(c). In terms of the theory of affordances [19], we could say that a wheel affords rotating in the same sense as a chair affords sitting or a knob affords turning. With the new controller, the robot discovers such affordances without any knowledge of the physics of the system and/or any internal motivation for doing just that task. In the experiments, the wheel is modeled by a bar with weights for giving it the necessary moment of inertia. In

Video 5, initially the connection between the arm and the wheel was rather loose so that for small movements there is no reaction from the rotation of the wheel. After improving this connection, an initial push by the experimenter was sufficient for exciting a stable rotation mode. It is as if the controller "understood" how to rotate the wheel, although it is just a result of force exchange and dynamics of the meta-system.

When positioning the wheel in parallel to the arm, the modes were emerging even more readily as seen in Video 6. Moreover, the system can be switched between forward and backward rotation by manual interaction and the frequency of the modes can be manipulated by changing a time-constant of the controller, see Video 7. The spontaneous emergence of the rotation behavior can be argued to be a cognitive act, if we consider—in the sense of (radical) embodied cognitive science [21]—that cognition is to be described in terms of agent-environment dynamics and not in terms of computation and representation.

C. Manual training

In another experimental situation, the robot is equipped with a brush and forced by manual guidance to wipe a table. Video 8 demonstrates how, by the combination of the limiting table plane and the manual force, the robot is driven into a two-dimensional wiping mode. This is due to both the tight closed loop control and the compliance of the synaptic dynamics to external perturbations, see [8]. Most importantly, emerging motion patterns can be identified and stored away by the user simply by taking snapshots of the synaptic weights (C). Video 9 shows the recall of previously acquired wiping modes. The transition between different modes is achieved by switching between those snapshots. Nevertheless, smooth transients are observed which is encouraging for building behavioral architectures from these self-organized behavioral primitives.

D. Behavior as fixed point flow

As mentioned above, behaviors under the controller Eq. (3) are not restricted to limit cycle attractors. As an alternative, we observe also seemingly random sequences of poses which are of interest for building libraries of reaching behaviors. An example is given in Video 10 where the bottle contains only little water so that its responses as a subsystem are too weak to pilot the rest of the system into resonance behavior. Instead of a fixed correlation structure in C, we have now, more or less, a flow of fixed points, but we need further theoretical back up here. In the experiments, this corresponds to a sequence of reaching patterns which might be of interest for such a library.

IV. Socializing Phenomena

Individual development by social interaction is a challenging subject for developmental robotics. In the following we provide a few examples of socializing effects between two robots based on both physical interactions and exchange of visual information. The sensorimotor loop of each robot contains the other robot to some extend, such that the drive for self-consistency of our controller leads to coherent behaviors involving both robots with mutual influences.

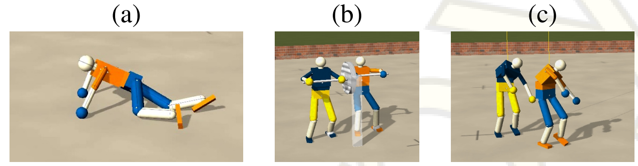

Fig. 4. **Interacting with the environment and another robot.** The humanoid robot crawling at the floor (a) (Video 11), two humanoid robots at a wheel with cranks (b) (Video 12), and two humanoids suspended on elastic robes, perceiving each others' joint configuration (c) (Video 13–17).

We study this in humanoid robots simulated in a 3D physical simulator [22]. Before we start with the interaction of robots let us have a look at a single of these robots controlled by the new paradigm. When left on a flat ground one of the many possibly behaviors that emerge is a crawling behavior, as seen in Video 11 and Fig. 4(a) [8]. Attaching the hands of two of these robots to the cranks of a revolvable wheel, Fig. 4(b), opens a physical communication channel allowing for the exchange of forces. The robots have to negotiate a behavior that is compatible with the constrained system and the motion of the partner. In Video 12 we demonstrate an instance of such a joint behavior, which appears typically within a minute of interaction time.

A. Including vision

Let us consider two robots in "visual" contact meaning that each robot receives the joint anglesof its partner and includes them into its sensor vector. Calling $x^{\text{own}} \in \mathbb{R}^n$ and $x^{\text{oth}} \in \mathbb{R}^n$ the vector of sensor values of its own and of the other robot, respectively, we have $x = \left(x^{\text{own}}, x^{\text{oth}}\right)$. Each robot hypothesizes (correctly) that the other one is of the same construction as itself. Consequently, the vector $x' \in \mathbb{R}^n$ in the learning signal $x'_i x_j$ is replaced with the sum $x'^{\,\text{own}} + x'^{\,\text{oth}}$. Note that due to normalization the increased scale does not matter. In this setting, C is an $n \times 2n$ matrix, mapping the $2n$ sensor values of the extended vector x onto the n motor commands.

With this and only this information, the simulation is started and one of the robots gets perturbed in order to break the perfectly identical conditions. At first, the vision is switched off ($x^{\text{oth}} = 0$) so that we observe how each of the robots explores its behavioral space in a seemingly random manner (remember that the controller and the physics are deterministic, though). After switching on vision, we observe a rapid convergence (in seconds) toward a synchronized motion by which they now explore their behavioral space in unison, see Video 13. Interestingly, when flipping the signs of the received sensor values ($x = (x^{\text{own}}, -x^{\text{oth}})$), the robots start imitating each others' 3D mirror picture, see Video 14. One can also blindfold one of the robots so that it develops independent motion patterns (after a perturbation). After some time the sighted robot learns to imitate the motions of the blinded robots, which acts as a sender. However, if the vision signal is switched off entirely, the memory rapidly decays so that, after seconds, the two robots are developing independently. The mutual influence is also seen when one of the two robots is weakened by reducing the maximal muscle strength. In Video 15 one of the robots is perturbed and/or weakened temporarily. Nevertheless, the weak robot not only is seen to put all its efforts in still following the strong one (as might be expected), but the strong one is also mimicking the weak one to some extend.

Further effects are observed if delaying the vision signals between the robots (same delay in both directions). Interestingly, we observe numerous metastable limit cycle attractors, see Video 16, if the delay time t_d fulfills the requirement $t_d = T(k+1)/2$ with T the period and $k = 0, 1 \ldots$. Depending on whether k is odd or even, in-phase or anti-phase motion patterns are observed.

B. The alien body effect

What happens if we cut or dampen the connections to the robot's own sensors, so that each robot essentially sees only the joint angles of its mate? The perceived sensor signals are now $x = (0.1x^{\text{own}}, x^{\text{oth}})$ and for the learning signal we have $x' = (0.1x'^{\text{own}} + x'^{\text{oth}})$. Each robot has to realize the learning and control of its body mainly on the basis of its mate's movements, reported by the visual system. In a sense, each robot is building its behavior on the illusion that the body of the other one is its own. Although this is an illusion, after a very short time the robots find a way to synchronize and develop behavior similar to the harmony shown in Video 13. This is why we call the emergence of coherent behavior patterns in this setting the *alien body effect*. For demonstrating the stability of the effect, in Video 17 we inverted the sign of the vision channel, like in Video 14. Even though we are making the system more complicated, the two robots converge temporarily into a common motion pattern, imitating each others' 3D mirror picture.

When moving in synchrony, this alien body behavior is no surprise as the sensor values x^{oth} and x^{own} would agree. The point of interest is that these states of synchrony are metastable attractors with a wide basin of attraction, i. e. the robots develop a joint strategy for searching the behavior space. Again, as it turns out, the search behavior becomes more variate if there is a short delay in the visual system.

C. Apparent Intentions

By the observer, the emerging behaviors are often attributed to some underlying concepts such as volition, motivation, intentions, drives. For instance, when two humanoid robots are put in an arena we may observe a fighting scene as shown in Video 18. It can be described in the beginning as if robot 1 is beating its opponent, i. e. robot 2, but later robot 2 is acting as-if in a revenge by trying to hit robot 1 at its head, eventually succeeding. Nothing of those intentions is put into the system nor detectable in the "brain" of either of the two robots. Instead, the apparent intentions emerge from the agent-environment coupling, here the interaction with the respective opponent.

Speculatively, we may try to use this scenario as a testbed for philosophical concepts. In the video scene, we may for instance ask whether agent 2 is responsible for hitting agent 1 at its head. Hopefully, in this context we may contribute to the ongoing [23] debate, starting with Benjamin Libets famous experiments [24], [25] of who is responsible in taking volitional decisions. This is not only of academic interest, as interpretations of those results reach up to the statement that humans cannot be made responsible for their doing as the brain decides long before the conscious I takes over. Of course, there is a vetoing possibility [23], but as the latter is also produced by the brain, it seems questionable if vetoing is the solution of the problem. Anyway, the somewhat disillusioning experience by our experiments is that the irreversible conjunction of brain, body, and environment clearly produces apparently conscious acts, like "discovering" how to rotate the wheel, without any internal functional unit for the conscious representation, internal motivation, reward systems, and the like.

V. Discussion

This paper aims at demonstrating how an extremely simple neurocontroller with the synaptic dynamics Eq. (3) generates highly non-trivial behavior. As the examples demonstrate, the simple principle—making the controller enhance the velocity correlations of the (proprioceptive) sensor values—is a vital tool for the generation of behavior. Metaphorically speaking, it makes the robot "feel" the physical dynamics of its body in interaction with the environment. Piloting the arm movements by either the pendulum or by the dynamics of the water in the attached bottle were exemplifying this effect. Moreover, in a number of further experiments, we demonstrated the "orchestration" of behavior in various sensorimotor machines.

These effects do not necessarily come as surprise. Instead, as discussed in earlier work [5], [9] they are a necessary consequence of spontaneous symmetry breaking mechanisms in the considered systems which reside—roughly speaking—at the edge of chaos, given a convenient choice of the self-amplification factor κ. This is what makes the approach a little difficult to accept in terms of common robotics scenarios. In those scenarios, it is usually possible to clearly say what the controller exactly does to generate the behavior. While this is not possible in our case, we are not lost as (i) there are some rules like the convergence toward cyclic behaviors (under certain physical conditions at least) and (ii) we obtain full fledged behavior of some complexity which—in the Myorobotics case at least—has not been achieved before. In general, we claim that it would be extremely difficult to generate the observed behaviors and effects by the known methods of developmental robotics, in particular considering the rapidness of the development.

Seen as a practical approach to generate complex, force-sensitive interactions with the environment, this controller could also augment the repertoire of classical controllers. Additionally, it may shed light on how biological musculoskeletal systems generate the complex trajectories they use to interact with the environment with an unrivaled flexibility. In this context, we believe our results are relevant for understanding the early sensorimotor development of infants and at the same time give new impulses for development robotics. The essential merit of our controller is that it provides a systematic approach for behavioral self-organization, avoiding the reality gap as demonstrated for instance by the robot at the wheel scenarios. Our experiments with interacting robots gives examples of social interaction in a minimalistic form. In particular it suggests that certain type of interactions may be a simple byproduct of a generic drive to bring internal dynamics in coherence with the external world.

Another point concerns the role of spontaneity and volition in nature. Obviously, acting spontaneously is an evolutionary advantage as it makes prey less predictable to predators. Attempts to explain spontaneity and volition range from ignoring it as an illusion to rooting it deep in thermodynamic and even quantum mechanical randomness [26], [27]. We cannot give a final explanation, but we can give a tentative solution to the dilemma how a system, if governed by a deterministic controller, can be free/act spontaneously in any sense: our deterministic neurocontroller obviously provides a clear example of how a great variety of behaviors can emerge spontaneously in deterministic systems by a deterministic controller due to the said symmetry breaking effects in systems at the edge of chaos. Similarly, there are recent trends in explaining the apparent stochasticity of the nervous system through the complexity of deterministic neural networks [28], [29], [30].

Eventually, the results may not only be of immediate interest for the further development of embodied intelligence. They also offer a new view on the role of self-learning processes in natural evolution and in the brain.

REFERENCES

[1] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey." *Connect. Sci.*, vol. 15, no. 4, pp. 151–190, 2003.

[2] R. Pfeifer and J. C. Bongard, *How the Body Shapes the Way We Think: A New View of Intelligence*. Cambridge, MA: MIT Press, November 2006.

[3] R. Der, U. Steinmetz, and F. Pasemann, "Homeokinesis - a new principle to back up evolution with learning," in *Proc. Intl. Conf. on Computational Intelligence for Modelling, Control and Automation (CIMCA 99)*, ser. Concurrent Systems Engineering Series, vol. 55. Amsterdam: IOS Press, 1999, pp. 43–47.

[4] R. Der and R. Liebscher, "True autonomy from self-organized adaptivity," in *Proc. Workshop Biologically Inspired Robotics*, Bristol, 2002.

[5] R. Der and G. Martius, *The Playful Machine - Theoretical Foundation and Practical Realization of Self-Organizing Robots*. Berlin Heidelberg: Springer, 2012. [Online]. Available: http://playfulmachines.com

[6] N. Ay, N. Bertschinger, R. Der, F. Güttler, and E. Olbrich, "Predictive information and explorative behavior of autonomous robots," *The European Physical Journal B*, vol. 63, no. 3, pp. 329–339, 2008.

[7] G. Martius, R. Der, and N. Ay, "Information driven self-organization of complex robotic behaviors," *PLoS ONE*, vol. 8, no. 5, p. e63400, 2013.

[8] R. Der and G. Martius, "Novel plasticity rule can explain the development of sensorimotor intelligence," *Proceedings of the National Academy of Sciences*, vol. 112, no. 45, pp. E6224–E6232, 2015.

[9] R. Der, "In search for the neural mechanisms of individual development: behavior-driven differential hebbian learning," *Frontiers in Robotics and AI*, vol. 2, no. 37, 2016.

[10] R. Der and G. Martius, "Self-organized control for musculoskeletal robots," 2016, arXiv preprint http://arxiv.org/abs/1602.02990.

[11] N. Van der Noot, A. J. Ijspeert, and R. Ronsse, "Biped gait controller for large speed variations, combining reflexes and a central pattern generator in a neuromuscular model," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 6267–6274.

[12] A. J. Ijspeert, "Central pattern generators for locomotion control in animals and robots: a review," *Neural Networks*, vol. 21, no. 4, pp. 642–653, 2008.

[13] M. Hoffmann and V. C. Müller, *Representation of Reality: Humans, Animals and Machines*. Springer, 2016, ch. Simple or complex bodies? Trade-offs in exploiting body morphology for control, arXiv preprint http://arxiv.org/abs/1411.2276.

[14] P.-Y. Oudeyer and F. Kaplan, "What is intrinsic motivation? a typology of computational approaches," *Frontiers in neurorobotics*, vol. 1, p. 6, 2009.

[15] S. Wittmeier, C. Alessandro, N. Bascarevic, K. Dalamagkidis, D. Devereux, A. Diamond, M. Jäntsch, K. Jovanovic, R. Knight, H. G. Marques, P. Milosavljevic, B. Mitra, B. Svetozarevic, V. Potkonjak, R. Pfeifer, A. Knoll, and O. Holland, "Toward anthropomimetic robotics: development, simulation, and control of a musculoskeletal torso," *Artificial life*, vol. 19, no. 1, pp. 171–193, 2012, PMID: 23186343.

[16] V. Potkonjak, B. Svetozarevic, K. Jovanovic, and O. Holland, "Anthropomimetic robot with passive compliance-contact dynamics and control," in *Control & Automation (MED), 2011 19th Mediterranean Conference on*. IEEE, 2011, pp. 1059–1064.

[17] V. Potkonjak, K. M. Jovanovic, P. Milosavljevic, N. Bascarevic, and O. Holland, "The puller-follower control concept in the multi-jointed robot body with antagonistically coupled compliant drives," in *IASTED international conference on robotics*, 2011, pp. 375–381.

[18] M. Jäntsch, S. Wittmeier, K. Dalamagkidis, and A. Knoll, "Computed muscle control for an anthropomimetic elbow joint," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 2192–2197.

[19] J. J. Gibson, *The theory of affordances, in Perceiving, Acting, and Knowing. Towards an Ecological Psychology.* Hoboken, NJ, 1977.

[20] D. A. Norman, *The Design of Everyday Things*. New York: Basic Books, 2002.

[21] A. Chemero, *Radical embodied cognitive science*. MIT press, 2011.

[22] G. Martius, F. Hesse, F. Güttler, and R. Der, "LpzRobots: A free and powerful robot simulator," http://robot.informatik.uni-leipzig.de/software, 2010.

[23] M. Schultze-Kraft, D. Birman, M. Rusconi, C. Allefeld, K. Görgen, S. Dähne, B. Blankertz, and J.-D. Haynes, "The point of no return in vetoing self-initiated movements," *Proceedings of the National Academy of Sciences*, p. 201513569, 2015.

[24] B. Libet, "Unconscious cerebral initiative and the role of conscious will in voluntary action," in *Neurophysiology of consciousness*. Springer, 1993, pp. 269–306.

[25] ——, *Mind time: The temporal factor in consciousness*. Harvard University Press, 2009.

[26] B. Brembs, "Towards a scientific concept of free will as a biological trait: spontaneous actions and decision-making in invertebrates," *Proc. R. Soc. B*, vol. 278, pp. 930–939, 2011.

[27] C. Koch, "Free Will, Physics, Biology, and the Brain," in *Downward Causation and the Neurobiology of Free Will*, N. Murphy, G. F. R. Ellis, and T. O'Connor, Eds. Springer, 2009, pp. 31–52.

[28] C. Hartmann, A. Lazar, and J. Triesch, "Where's the noise? Key features of neuronal variability and inference emerge from self-organized learning," *bioRxiv*, 2014.

[29] T. Kenet, D. Bibitchkov, M. Tsodyks, A. Grinvald, and A. Arieli, "Spontaneously emerging cortical representations of visual attributes," *Nature*, vol. 425, no. 6961, pp. 954–956, 2003.

[30] R. Bourdoukan, D. Barrett, S. Deneve, and C. K. Machens, "Learning optimal spike-based representations," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 2285–2293. [Online]. Available: http://papers.nips.cc/paper/4750-learning-optimal-spike-based-representations.pdf