FP7-600716

Whole-Body Compliant Dynamical Contacts in Cognitive Humanoids

Year 1
Second year project objectives report

Editor(s)	CoDyCo Consortium
Responsible Partner	IIT
Affiliations	IIT, TUD, UPMC, UB, JSI.
Status-Version:	Draft-1.0
Date:	Apr. 28, 2014
EC Distribution:	Consortium
Project Number:	600716
Project Title:	Whole-Body Compliant Dynamical Contacts in Cognitive Humanoids

Title of Report:	Second year project objectives report
Date of delivery to the EC:	28/04/2014

Workpackage responsible for the Report	All work packages
Editor(s):	Francesco Nori, Vincent Padois, Jan Peters, Jan Babic, Michael Mistry
Contributor(s):	Entire CoDyCo consortium
Reviewer(s):	reviewers
Approved by:	All Partners

Table of Contents

Index of Figures

3.2 Project objectives for the period

3.2.1 Overview

The specificity of CoDyCo relies on the fact that the progress beyond the state of the art is guided by the yearly implementation on the iCub humanoid. Within this context, CoDyCo second year specific objectives were to design and implement the control of whole-body posture while performing goal directed movements. Beyond the activities to achieve this result, other long term activities have been conducted in preparation for the following years objectives. These activities involve human experiments, software infrastructure maintenance and the development of learning/control algorithms.

Task	IIT	TUD	UB	UPMC	JSI	
WP1	3	3	2.29	0.48	-	-
WP2	-	3	7.56	0.48	-	-
WP3	3	10.5	1.85	14.69	-	-
WP4	3.04	21.7	2.04	1.68	-	-
WP5	12	-	-	0.05	-	-
WP6	1.5	-	-	0.31	-	-
WP7	-	-	-	0.13	-	-
	22.54	38.2	13.74	17.94	-	-

3.2.1.1 WP1: toolbox for computing and controlling dynamics of whole-body movements with contacts (UB)

The overall goal of this work package is to develop software libraries and software modules to be used as toolbox by the entire project consortium. The expected outcome for the second year were to develop such (shared) toolbox that can be used by the entire group for implementing balancing and reaching controllers with multiple contacts.

3.2.1.2 WP2: understanding and modelling human whole-body behaviours in physical interaction (JSI)

There were three main objectives within WP2 for the second year of the project: (i) to continue the work on designing of models for human whole body motion in contact where we aimed to form simplified models of high-level understanding of how additional supportive contacts affect human motor control strategies involved in balancing (Task 2.2). (ii) to study strategies that humans use in dealing with uncertanties in contact where we aimed to go beyond the work performed in Task 2.2 by putting human subjects in situations where they had to perform task oriented movements by interacting with either stiff, compliant, or upredictable supports. (iii) to get a deeper insight ito factors involved in human choice of contact utilization and to investigate how interaction through contacts can contribute to learning of whole body motor control (Task 2.4).

3.2.1.3 WP3: control and optimization of whole-body motion in contact (UPMC)

The objectives of WP3 for the second year of the project are threefold. The first one is to demonstrate the applicability of state of the art whole-body motion controllers, such as the one developed in [1] and [2], on the iCub robot in multi-contact, goal oriented senarii (Task

3.4). The second one is to start exploring ways to enrich the retained whole-body controllers with the capability to interact with non-rigid environment (Task 3.3). The third one is to keep exploring potential ways of optimally coupling the local, reactive control level and the global, decision making one (Task 3.4).

3.2.1.4 WP4: adaptation, Generalization and Improvement of Compliant Control and Tasks with Contacts (TUD) The goal of WP4 is to endow the CoDyCo humanoid robot control architecture with the core abilities for the adaptation, generalization and self-improvement of both control laws and tasks that involve physical interaction with humans, and the environment. In this context, we propose learning approaches that work in conjunction with the control architecture devised in WP3 and rather complement analytical robotic approaches with on-policy learning than starting from scratch. A core idea behind this work package is that Learning should complement classical approaches and not supersede them.

The second year objectives of WP4 include:

- Fast regression methods that can deal with well structured input noise, such that physical models can be learned and adapted for tasks that involve many uncertain contacts. A particular focus will be given to prediction-based switching model.
- Novel methods for imitation and reinforcement learning of skills with contact will be devised and tested.
- Learning how to combine elementary tasks by imitation and reinforcement learning. The combinations involved include the learned simultaneous use of elementary tasks, the sequential use as well as the co-articulation of tasks.

3.2.1.5 WP5: systems integration, standardization and evaluation on the iCub robot (IIT) The second year main objective for WP5 was the implementation of a validation scenario consisting of the balancing while performing goal directed actions.

3.2.1.6 WP6: management (IIT) The second year management was primarily dedicated to the project consolidation. Among the main goals the maintenance of the software repository and database.

3.2.1.7 WP7: dissemination and Exploitation (IIT) The main dissemination objectives for the CoDyCo second year were the website maintenance, the dissemination activities and management of the IPR.

3.3 Work progress and achievements during the period

3.3.1 Progress overview and contribution to the research field

All the CoDyCo second year objectives have been attained. Here is a list of the CoDyCo second year achievements.

- Design, implementation and maintenance of the whole-body control software infrastructure. The infrastructure consists of several modules which significantly improved the controller accuracy and robustness thanks to: a module for whole-body torques estimation, a module for force/torque sensors calibration, a module for whole-body dynamics identification and a module for dynamics estimation.
- Design of experimental protocols and data collection of experiments for studying humans in multi-contact interaction with the environment. This includes an experiment on hand-contact assisted balancing, a metric for whole-body stability characterisation, an experiment of human robot physical interaction and a study on collaborative human-robot physical interaction.
- Design and simulation of whole-body control strategies in presence of non-rigid contacts. Experiments on postural control under multiple environmental contacts while controlling the operational space dynamics.
- Development of a theoretical framework for representing movement primitives within probabilistic contexts. Design of a model-free probabilistic representation for simultaneous representation of kinematic and force trajectories. Preliminary studies on the problem of learning strategies to adapt temporal activation of low-level primitives and to deal with interferences in combining multiple whole-body tasks.
- Implementation of the second year validation scenario consisting in whole-body motion control subject to postural, contact and goal-directed (Cartesian) constraints.

3.3.2 Work packages progress

WP1: toolbox for computing and controlling dynamics of whole-body movements with contacts (UB) WP1 objectives were achieved for the second year. In summary, the main accomplishments and impacts for the research community are as follows:

- The `codyco-superbuild` was released as open-source software which contains modules for balancing and reaching control for CoDyCo project.
- Several control libraries for the iCub whole-body motion were developed and tested both in simulations and on the iCub.
- An in-situ force/torque sensor calibration procedure was designed to improve the accuracy of whole-body identification. The results are published in [3].
- Inertial parameters which are estimated from embedded force/torque sensors are used to improve torque estimation of the iCub. The results are outlined in [4].
- An experimental software library was released to perform maximum-a-posteriori dynamic estimation fusing multiple sensors and encoders of the iCub. Computational efficiency and estimation accuracy of the proposed method are studied and published in [5] and [6].

WP2: understanding and modelling human whole-body behaviours in physical interaction (JSI) In T2.2, JSI used the data collected from the biomechanical studies to form a human model for hand-contact assisted balance control in simulation environment. The model serve as platform for devising equivalent robot skills.

In T2.2, UB developed a metric for full-body stability in multiple contact condition. This metric is based on the manipulability of the centre of mass.

In T2.3, JSI developed a novel method for studying human strategies of dealing with contacts with uncertain environment. Instead of performing the contacts with his/her own limbs, the human subject was included into the robot control loop and was asked to perform a task in contact with the environment through the robotic mechanism. To accommodate that, human-robot interfaces were developed. The main advantage of this method, compared to standard biomechanical studies, is that the human observation data can be directly used to build robot skills.

In T2.4, Inria, TUD and UPMC participated in analysing the dataset of the EDHHI experiments where healthy subjects interacted physically with the iCub. The preliminary analysis shows that people, on average, learn quickly how to interact with the robot and move its arms: across three trials, the exchanged forces were smaller and the contacts more precise.

In T2.4, JSI performed a study on multiple healthy subject and analysed the effects of additional supportive contact on full-body balance control. The subjects were continuously perturbed at the waist. In one instance, the subjects did not use any supportive hand contacts while in the other instance, the subject used an additional supportive hand contact. The comparative analysis between the two conditions revealed particular synergies between arm and body muscles which significantly contribute to the improved balance.

In T2.4, TUD and JSI studied whether supporting contacts in human arm reaching tasks are planned or are an effect of a reactive controller. Experiment on multiple subject were performed, where the task was to reach to a target with one hand and use the other hand for additional support. During the experiment, the subject balance was perturbed by a displacement of the ground support.

Finally, in T2.4, UPMC and JSI started an experimental study where the aim is to challenge two well-established but conceptually separated motor control phenomena. We obtained several very promising preliminary results indicating a general mechanism that points out a global trade-off arising from the interactions between movement time, cost and accuracy.

WP3: control and optimization of whole-body motion in contact (UPMC) After two years of project, the level of achievement of the objectives in WP3 meets the expectations. The main achievements are:

T3.3 **Studies** by UPMC and UB on the extension of whole-body control frameworks in order to account for non-rigid contacts. The approach retained by UPMC adapts the desired contact force value and center of mass trajectory in order to establish stable and supporting contacts as fast as possible. This approach assumes a local contact model but does not require the explicit knowledge of its parameters and rather uses the velocity of the contact point to directly adapt the desired contact force. UB computes optimal contact forces based on the momentum of the robot and computes and converts this forces into desired acceleration using an estimated model of the surface in contact. Given these desired accelerations, joint torques are computed in order to best achieve the desired accelerations.

T3.4 **Integration** by IIT and UPMC of the whole-body controllers on the real robots. This includes a large amount of background work by IIT on low-level control aspects (calibration and identification mostly).

T3.4 **Investigation** by UPMC of MPC as an efficient mean to optimally handle the postural balancing problem under varying contact conditions.

T3.4 **Studies** by Inria, TUD and UPMC on the adaptation of task weights in the whole-body controller.

T3.4 **Studies** by TUD on learning inverse dynamics with contacts to predict contact forces.

T3.4 **Studies** by UB on projected operational space dynamics for the control of constrained motion for a manipulator performing a task while in contact with the environment.

WP4: adaptation, generalization and improvement of compliant control and tasks with contacts (TUD) The goal of WP4 is to endow the CoDyCo humanoid robot control architecture with the core abilities for the adaptation, generalization and self-improvement of both control laws and tasks that involve physical interaction with humans, and the environment.

During the second year, IIT developed a theoretical framework for estimating whole-body dynamics from distributed multimodal sensors [5]. TUD continued their research in probabilistic movement primitive representations. A journal article on imitation learning and the co-activation of basic skills is under review. For multi-modal solutions an extension using mixtures of Gaussians with latent variables was published at an international robotics conference [7]. In tasks with contacts, a model-free probabilistic movement representation was developed that models joint distributions over kinematic and force trajectories. This work is under review. Further, TUD investigated noise robust planning methods methods to plan movement skills given task-space constraints. This work was published at an international humanoid robot conference [8] and will be used in year three for generating optimal control laws with learned dynamics models. During year two, TUD investigated the learning of temporal activation profiles of low-level task controller. This work is under review. A similar line of research was conducted by UPMC who studied how to deal with interferences in task combinations in whole-body. The tasks are parameterized with Dynamical Movement Primitives, whose parameters are optimized based on a general compatibility principle. This work was published at an international humanoid robot conference [9]. In a second study that is currently under review, UPMC studied how task variability can be used to modulate task priorities during their execution. UB continued their research on computed torque control leveraging low-gain control.

WP5: systems integration, standardization and evaluation on the iCub robot (IIT) The second year WP5 activities have concentrated on the second year validation scenario. A complete description of the scenario can be found in "D5.2 Scientific report on validation scenario 2: balancing on feet while performing goal directed actions." which discusses the technical implementation of the second year validation scenario (see https://github.com/robotology-playground/codyco-deliverables/tree/master/D5.2/pdf). With respect to the state of the art the work progress represents a step towards whole-body torque control under postural, contacts and goal-directed constraints. The integration of tactile feedback within the whole-body controller is a peculiarity of the implemented CoDyCo validation scenario and therefore represents yet another step forward with respect to the current state of the art.

WP6: management (IIT) The CoDyCo project continued successfully. Management activities included the definition of a second amendment procedure smoothly organized by the consortium and the project officer. The software repository (https://github.com/robotology/codyco) have been significantly improved as clearly documented in the web-based git repository hosting service (https://github.com).

WP7: dissemination and exploitation (IIT) Within WP7, CoDyCo second year achievement include: dissemination at relevant academic and industrial events; population of the CoDyCo database to disseminate robot and humans datasets.

3.3.2.1 Work package 1 progress

3.3.2.1.1 Software architecture design and evaluation of available open-source software pertinent to the scope of the project. (T1.1) The goal of T1.1 was to agree on a specific software architecture with associated software tools whose specifications, dependencies and interconnections meet the requirements and needs for achieving the goals of the project. The software, which is called `codyco-superbuild`, is available via github on https://github.com/robotology/codyco-superbuild. The modules and specifications of the software are as follows:

- `codyco-commons`: A collection of functions and utilities used in the other projects
- `idyntree`: YARP-based Floating Base Robot Dynamics Library
- `paramHelp`: Library for simplifying the management of the parameters of YARP modules
- `wholebodyinterface`: C++ Interfaces to sensor measurements, state estimations, kinematic/dynamic model and actuators for a floating base robot
- `yarp-wholebodyinterface`: Implementation of the wholeBodyInterface for YARP robots
- `WBI-Toolbox`: Simulink Toolbox for rapid prototyping of Whole Body Robot Controllers
- `codyco-modules`: YARP modules and controllers developed within the CoDyCo project

3.3.2.1.2 Simulator for whole-body motion with contacts (T1.2) The CoDyCo project requires a modular, component-based dynamics simulation software providing numerically stable, computationally efficient and physically consistent simulations of whole-body virtual human(oid) systems in contact with rigid or soft environments. To this end, in year one, a new iCub simulator was released and documented as part of deliverable D1.1.

3.3.2.1.3 Control library for flexible specification of task space dynamics of floating base manipulators. (T1.3) During the second year both IIT and UPMC contributed to the development of several software components for controlling the iCub whole-body behavior. UPMC has continued the integration of its whole-body controller within the overall software architecture by properly connecting to the WholeBodyInterface developped by IIT and described in Deliverable 1.2. Tests of this integration has been performed both in the Gazebo simulator as well as on the iCub robots present at UPMC.

3.3.2.1.4 System dynamics estimation software. Extension to environmental compliance estimation (T1.4) The goal of this task is to develop a software tool for on-line identification of whole-body dynamics, as well as the compliance of contacts established between the robot and the environment.

- Within T1.4, IIT continued the activities started during the previous year, with specific focus on whole-body identification [10, 11]. In order to enhance the identification accuracy, an in-situ force/torque sensor calibration procedure was designed [3] (see Fig. 1) and implemented in a software component[2] which has been released with an open source license. Similarly, in order to enhance the torque estimation accuracy, IIT conducted a theoretical analysis that exploits embedded force/torque sensors. It has been proven [4] that the inertial parameters estimated from embedded force/torque sensors can be successfully used for torque estimation.
- During year two, TUD worked with the WBI toolbox in Matlab to study whole-body controllers capable of integrating learned inverse dynamics models. The idea, more specifically, is to combine the low-level torque control at joint level with the learned dynamics models of WP4 (T4.2), which can thus provide the correct torques for rigid and compliant contacts. TUD implemented a whole-body impedance controller and two balance controllers with the WBI toolbox in Matlab (see Figure 2). The controllers were tested on Gazebo and on the iCub mex model simulated in Matlab. The integration with the outcome of WP4 is ongoing.

3.3.2.1.5 Extension and enhancement of the iDyn library. (T1.5) The goal of this task is to provide a reliable software tool for on-line estimation of whole-body dynamics. Before CoDyCo, dynamic estimations on the iCub were relying on the iDyn software library[3], designed for fixed-base robots. Within the first year of CoDyCo, iDynTree[4] was released in

[2] https://github.com/robotology-playground/insitu-ft-calibration.
[3] https://github.com/robotology/icub-main/tree/master/src/libraries/iDyn.
[4] https://github.com/robotology/codyco/tree/master/src/libraries/iDynTree.

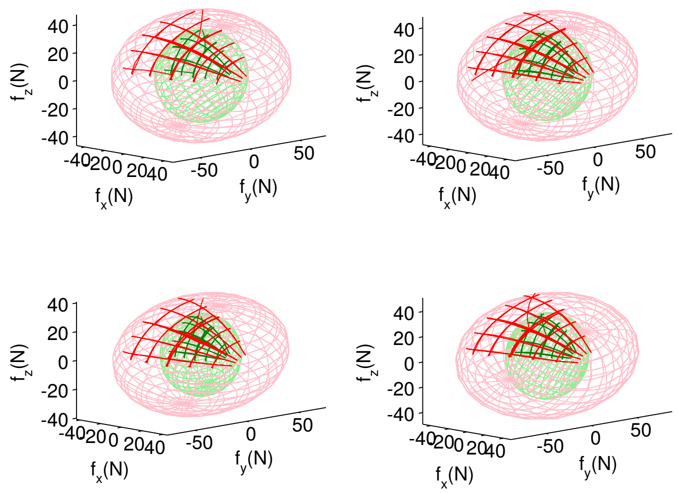


Figure 1: The image shows the accuracy in calibrating the force/torque sensor with the procedure described in [3]. The four plots refer to four different experimental conditions (different values for the calibration weights). Ideally, perfectly calibrated data should lie on the surface of a three dimensional sphere. Dark green: force measurements obtained with the calibration matrix estimated using the proposed technique. Dark red: force measurements obtained with the calibration matrix provided with the sensor. Light red and light green surfaces: ellipsoids fitted to the measured forces. Qualitative calibration accuracy can be obtained by looking at the spherical symmetry of the fitted ellipsoids.

response to the need of representing floating base structures. During the second year of the project, we investigated on the problem of extending iDynTree to the case of multiple redundant sensors. The investigation resulted in an experimental software library[5] currently implemented in MATLAB. The software performs maximum-a-posteriori dynamic estimation fusing multiple sensors such as gyroscopes, linear accelerometers, embedded force-torque sensors and encoders. Computational efficiency is obtained by exploiting the sparsity of the underlying problem [5] (see Fig. 3). Estimation accuracy is obtained by a modified version of the expectation maximisation algorithm [6] (see Fig. 4).

3.3.2.1.6 Resources Overall, the use of resources within WP1 was in accordance to the plans.

WP1 person months	**IIT**	**TUD**	**UPMC**	**UB**	**JSI**
Year 1	8.67	1.00	3.29	0.51	2.00
Year 2	3.00	3.00	0.48	2.29	0
Total	12.00	6.00	14.00	15.00	6.00

[5] `https://github.com/iron76/bnt_time_varying`.

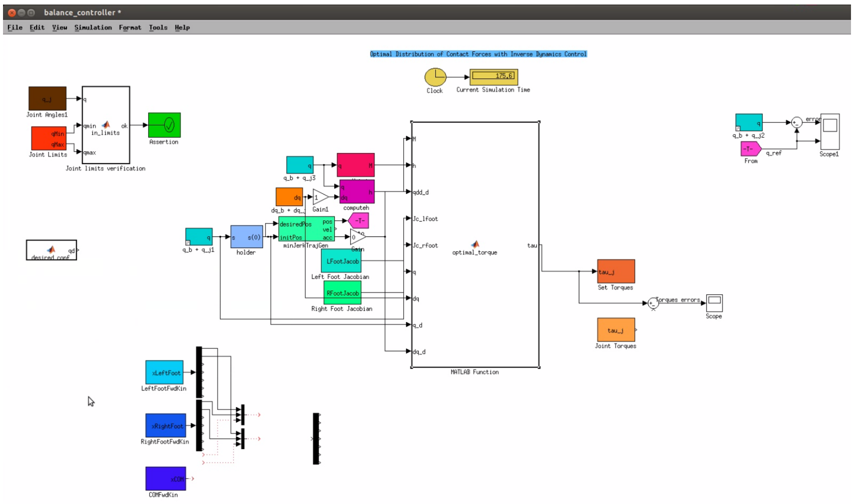


Figure 2: One of the balance controllers implemented by TUD with the WBI toolbox in Matlab.

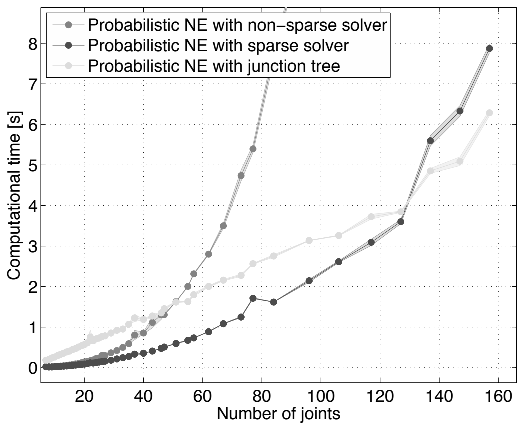


Figure 3: Comparison of non-sparse (S1-gray), sparse (S2-dark-gray) and Bayesian network junction tree (S3-light-gray) solvers in solving maximum-a-posteriori dynamics with redundant measurements (see [5] for details).

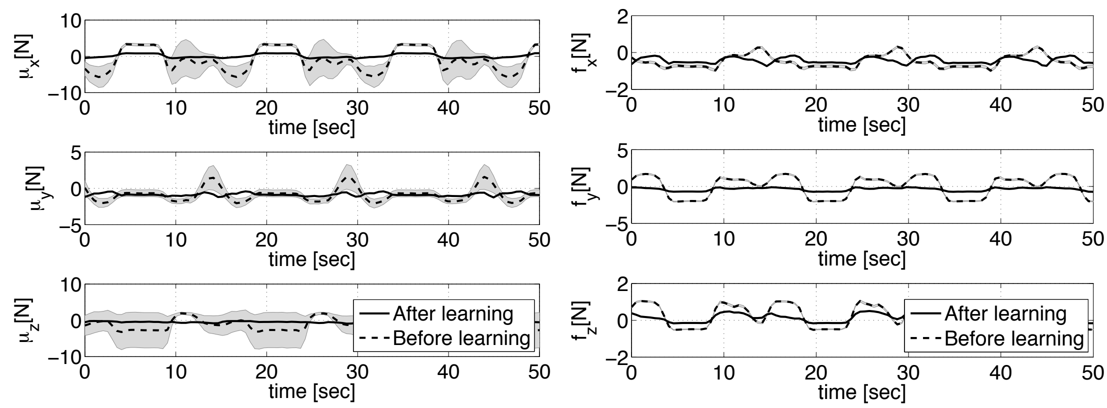


Figure 4: The picture shows the errors in estimating an external wrench. The two curves refer to the estimation obtained before (dashed line) and after (solid line) the estimation of the data covariance with a modified EM algorithm [6].

3.3.2.1.7 Deviations from workplan No significant deviations.

3.3.2.2 Work package 2 progress

3.3.2.2.1 Design of models for human whole body motion in contact (T2.2) In the scope of T2.2, JSI created a 3D dynamic model of a human holding to a stable object during continuous perturbations of stance. The model was devised from measurements on 13 male subjects. Using the kinematical data recorded with frequency of 100Hz, forces that the subjects excerted on the ground, and the forces in the handle, we performed an inverse dynamic procedure and obtained joint torques produced by muscles during the experiment. An illustration of the model is shown in Fig. 5. Using this modelling approach we are now able to efficiently study the biomechanics of humans in contact with the environment.

During the second year of the project, UB studied the effects of hand contact on the stability of a planar humanoid robot (see Fig. 6a) while a momentum based controller is used to control the robot's balancing motion [12]. They compared the simulation results with the results of the experiments on human subjects which are reported in [13]. Both simulations and experiments agreed that different values of hand contact forces in different hand positions cause the same displacements of the CoP of the foot. This implies that regulating the CoP of the foot has the highest priority for both humans and the momentum based controller. This study suggested that the momentum based controller is an adequate controller to replicate human behaviour during balancing motions.

During the second year, UB continued to work on defining a suitable metric to measure the effects of the environmental contacts on the robot's stability. They used the basic concept of end-effector manipulability (for manipulators) in the literature and introduced a new tool to analyze the ability of balance for legged robots which we called it manipulability of the center of mass. This tool defines three different types of ellipsoids which are called 1) velocity ellipsoid, 2) instantaneous velocity ellipsoid and 3) instantaneous velocity ellipsoid due to the unit impulse. The first one shows the velocity of the CoM in different directions due to the unit norm of the joint velocities. The second and third types of the ellipsoids, which are obtained

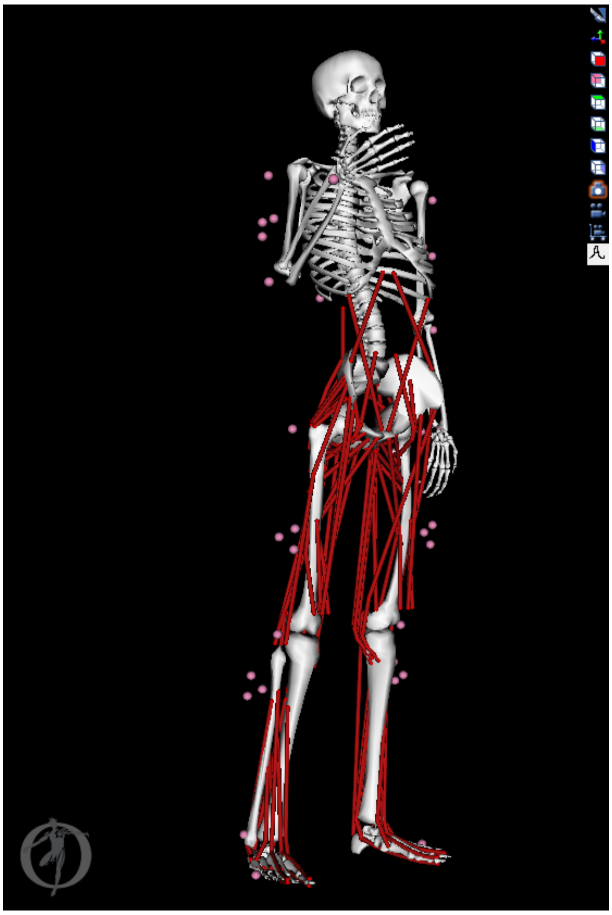
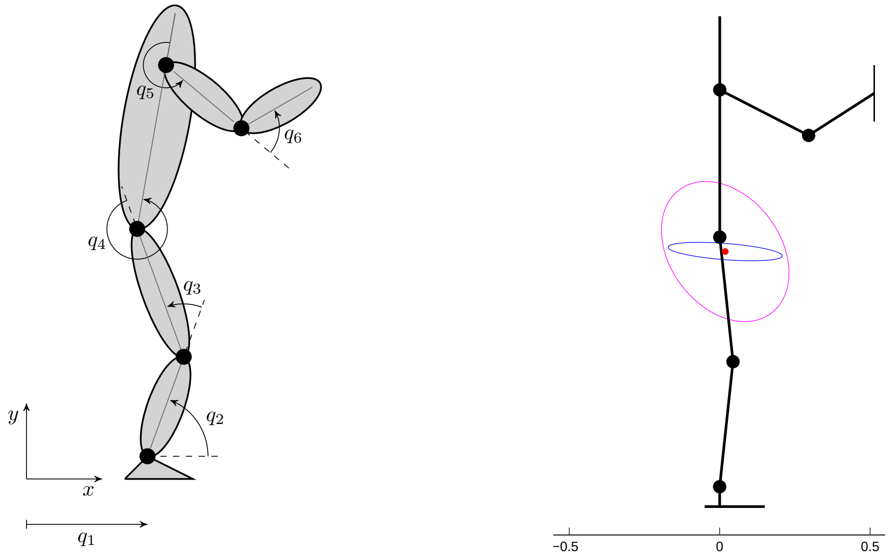


Figure 5: Model of a subject holding a handle.

by using impulsive dynamics, show instantaneous changes of the CoM velocity due to the unit norm of instantaneous changes at the joint velocities and the unit norm of impulse at the actuated joints, respectively. By involving the motion equations into the calculations for the second and third types of ellipsoids (via impulsive dynamics), these ellipsoids allow us to study the effect of under-actuation as well as kinematic constraints on the robot's stability.

As an example, Fig. 6b shows a planar humanoid robot (with its hand is fixed) and instantaneous velocity ellipses for the robot in the specified configuration. Since the robot is fully actuated, the first and second types of ellipses (type 2 and type 3) are the same. This ellipse (type 2) shows how the velocity of the CoM changes when the instantaneous change of the joint velocities due to the impulse has the unit norm. This shows the ability to move the CoM in different directions by a certain amount of movements at the actuated joints. The ellipse type 3 shows how the velocity of the CoM changes due to the unit impulse at the joints. In other words, it shows how a certain amount of impulse at the actuated joints can accelerate the CoM in different directions. All of the ellipses are independent from the controller and they are dependent only on the physical parameters of the robot and its kinematic constraints.

In balancing in a plane, the CoM movement in the horizontal direction is an important measure. By projecting a velocity ellipse on x axis, we obtain a line which its length equals



(a) Schematic diagram of the robot model and its coordinates (b) Velocity ellipses for a planar humanoid robot

Figure 6: Models

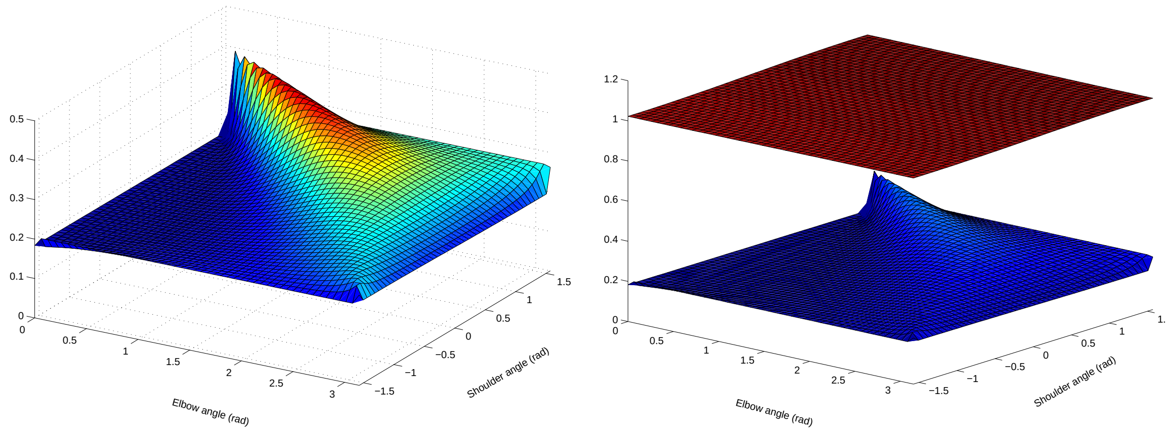


Figure 7: Maximum instantaneous change of the CoM velocity in the x direction due to the unit norm of instantaneous change of the joint velocities for (left) the constrained robot and (right) for both constrained and unconstrained robots.

to the maximum change of velocity of the CoM in the horizontal direction. Figure 7 (left side) shows maximum instantaneous change of the CoM velocity in the horizontal direction for different constrained hand locations (i.e. different elbow and shoulder angles). This is due to the unit norm of instantaneous change of the joint velocities. In the right side of this figure, the graph at the left side is compared with the case that the hand is not constrained. It is obvious that the movement of the CoM is limited due to the kinematic constraint at the hand.

Figure 8 shows maximum instantaneous change of the CoM velocity due to the unit impulse at the joints for different hand locations and for both constrained and unconstrained hands.

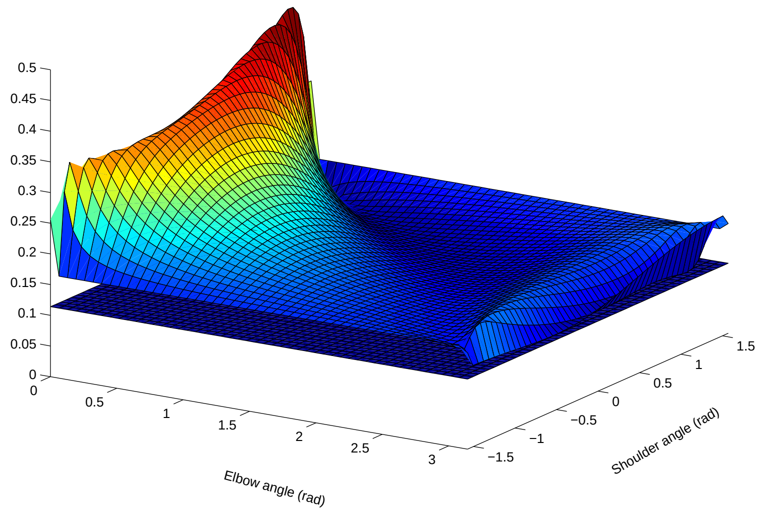


Figure 8: Maximum instantaneous change of the CoM velocity in x direction for the constrained and unconstrained robots due to the unit norm of impulse at the joints.

As it can be seen in this figure, the graph for the constrained robot is always higher than the other one. This implies that the same amount of impulse can cause bigger changes at the CoM velocity in the constrained robot rather than the unconstrained one. The reason is that, in the constrained case, the robot exploits the contact force to accelerate the CoM and therefore less (impulse) torque is needed for the same change at the CoM velocity.

3.3.2.2.2 Strategies of dealing with uncertainties in contact (T2.3) During the second year work on T2.3, JSI developed a novel method to study human strategies of dealing with contacts with uncertain environment. In this method the human subject was made to perform psychical contacts with the environment through the robot. The human was included into the robot control loop through human-robot interfaces. The idea is that the human sensorimotor system and cognitive system controls a novel mechanical system, i.e. the robot, in physical interaction with the environment. This implies additional human motor control learning and adaptation that can potentially provide us with a deeper insight into how humans deal with a novel environment.

Another advantage of this approach is that the human sensorimotor system does not use its own limbs to directly make the contacts with the environment, but uses the robotic limb to do so. Compared to pure biomechanical studies, where the measured human behaviour must be further interpreted, adjusted or transformed before it can be used on the robots, in this approach the measured human behaviour can be directly captured and used in the robot control. This study therefore provides a good complement to our conventional biomechanical studies as performed in T2.2.

The block scheme of the proposed approach is shown in Fig. 9. The human controlled the motion of the robotic limb with the motion of his/her own limb. In addition to controlling the motion, the human also controlled the impedance of the robot. Primary information about the robot state was relayed to the human through a visual feedback. A haptic device was

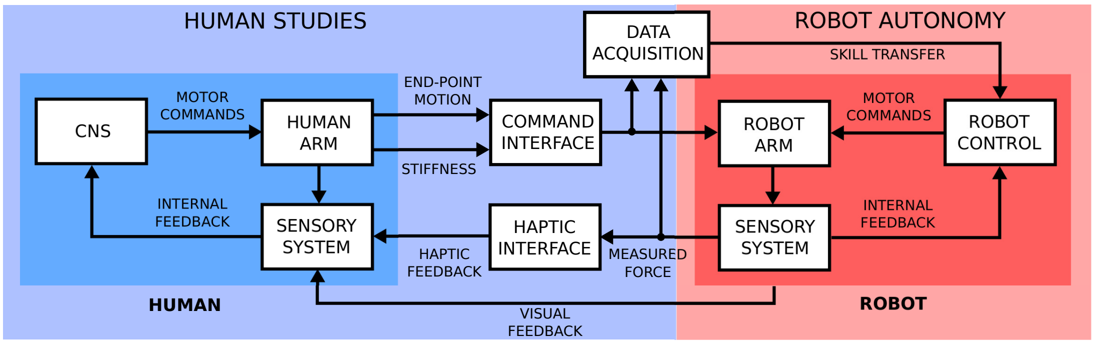


Figure 9: Block diagram of proposed human-in-the-loop robot control framework for study of human behaviour in contacts with environment. During the learning and adaptation stage, the human performs the contacts with the environment through the robot (blue section). The acquired data was used to observe and study the human behaviour. When the human learning process and observation is complete, the learnt skill can be directly captured and used in the autonomous robot control (red section). This is the main advantage compared to the conventional biomechanical studies.

used to provide the human with an additional feedback about the forces sensed by the robot. While controlling the robot in the proposed human-in-the-loop approach, the human central nervous system had to adapt to a new mechanism through sensorimotor learning to perform the desired contact with the environment.

The main goal of studies of human behaviour in contacts with environment in WP2 is to offer a basis from which we can devise equivalent humanoid robot behaviour. The most appealing prospect of the proposed approach to study human motion in contacts with the environment is that the data from the study can be used to directly form skills for autonomous robot control. The sensorimotor data was collected while the human was making the desired physical contacts with the environment though robotic mechanism. This data was then used to form the trajectories. The trajectories were encoded with Dynamical Movement Primitives (DMPs) [14]. The parameters of DMPs were learned by locally weighted regression [15]. The learned trajectories represented the robot skill for dealing with the contacts with the environment according to the human strategy. The trajectories can be included into the robot control system and used for autonomous execution of the learnt task.

One of the key features of the proposed approach is the ability of the human to directly control the impedance of the robot limb in an equivalent way that he/she controls his/her own. For this purpose we developed two novel human-robot interfaces [16, 17] that allow the human to modulate the stiffness of the robotic limb in real-time. The first interface (see Fig. 10, left) was based on measuring human muscle activity by surface electromiography (sEMG). The current measured muscle activity was mapped to the robot stiffness. The second interface (see Fig. 10, right) was based around a linear potentiometer inside a handle held in the human hand. The human controlled the position of the potentiometer knob with a finger position. The finger position is then mapped to the robot stiffness via measured potentiometer voltage.

3.3.2.2.3 Human contact choice and learning through physical interaction (T2.4) Within T2.4, Inria, TUD and UPMC participated in analysing the dataset of the

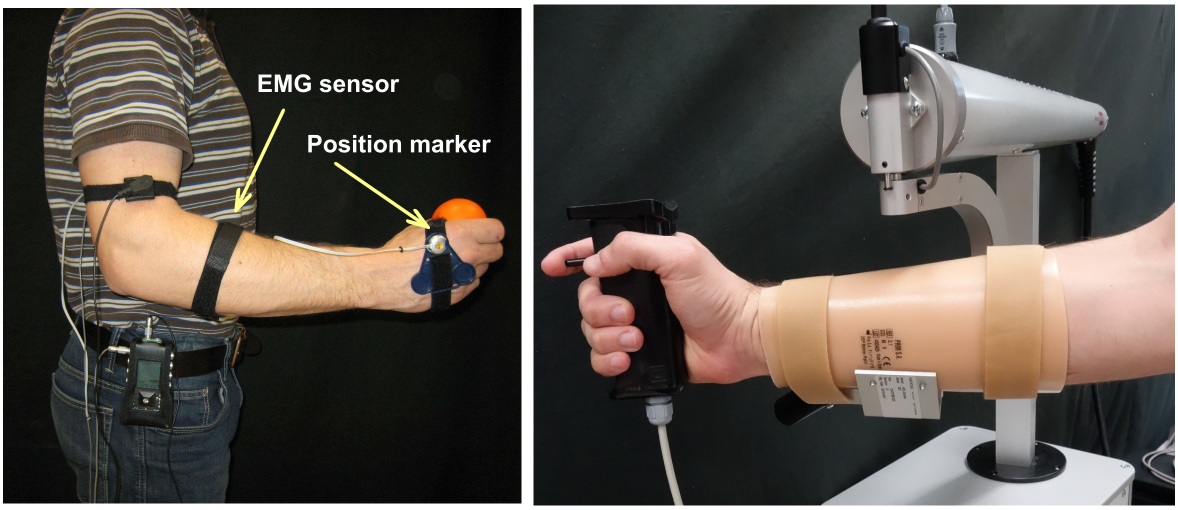


Figure 10: Human-robot interfaces. First developed interface (left) measured human limb motion via optical motion capture system and mapped it to the motion of the robotic limb. The human muscle activity was measured by sEMG and was used as an interface to control the robot impedance. Second developed interface (right) consisted of *HapticMaster* robot and impedance control handle. *HapticMaster* robot measured the human limb position and provided the force feedback. Impedance control handle was based around a spring-return linear potentiometer and was held in the human hand.

EDHHI experiments[6] where healthy adults (18-65 years old) interact physically with the iCub (see Fig. 11). The analysed data include tactile data from the forearm skin and contact forces, estimated by the iDyn modules developed in WP1. The preliminary analysis shows that people, on average, learn quickly how to interact with the robot and move its arms: across three trials, the exchanged forces are smaller and the contacts more precise. Currently, Inria is coupling the analysis of tactile and force signals with individual factors and social signals exchanged by the two peers.

In the scope of T2.4, JSI studied how additional hand contact with the surrounding objects influences whole-body balance conditions. The experiments were performed on multiple subjects where we challenged their balance. The experiments were divided into two main stages. Each stage had 15 sessions in which the subject's balance was perturbed for 5 minutes. In one stage the subjects did not use supportive hand contact. In the other stage they were holding a handle in front of them. We used a motorised wait-pull mechanism [18] to continuously perturb the balance of the standing subjects in either stage by exerting external forces on the approximate position of centre of mass. See Fig. 12 for the experimental setup.

The perturbation waveform of the waist-pull mechanism was constructed in a way that the possible muscle reactions associated with reflexes were eliminated. These reactions could potentially mask the actual role of the hand muscles as the reflex would activate the muscles unrelated to the magnitude of the perturbation. To avoid that, the perturbation waveform was continuous, had relatively low frequency and low pulling forces. During the experiment, we measured muscle activation of the subject's lower leg, trunk and arm muscles, forces in the

[6] http://www.loria.fr/~sivaldi/edhhi.htm

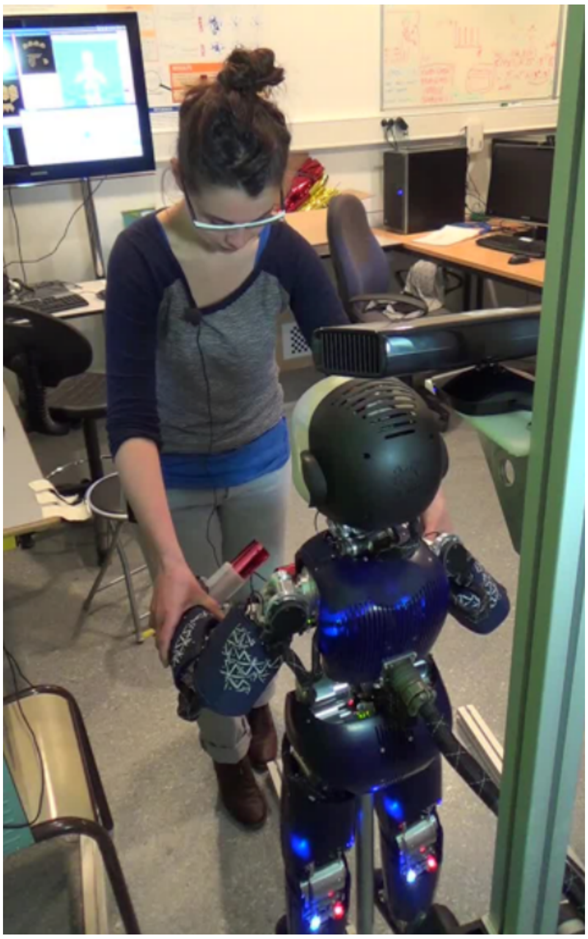


Figure 11: View of a physical interaction between a human and the iCub robot.

handle and the anteroposterior movement of CoP (CoP$_{AP}$).

The results of muscle activation analysis showed that when the subjects were holding to the handle, the activation of the leg muscles was minimal (see Fig. 13). Based on this we can conclude that the subjects mainly used their arm muscles to maintain postural stability. The trunk flexor muscle (Obliques Externus, OE) was more active in the stage when the subjects were holding the handle compared to when they were not. This indicates that a synergy between the arm and trunk muscles was established when additional hand contact was utilised to maintain the equilibrium.

The analysis of the CoP$_{AP}$ movement showed that the displacement of the CoP$_{AP}$ was progressively dropping throughout the repeated sessions of the experiment (see Fig. 14). This was true both in case when supportive hand contact was used and in case when no supportive hand contact was used. These results give a strong hint that a learning and adaptation was present through the sessions of the experiment, as the subject gradually improved the balance control.

We further analysed whether there are any effects of repeated sessions on adaptation of muscle activation and movement of CoP$_{AP}$ (Figure 15), and whether there are any differences between the two stages of the experiment. The results show that the effect of human adaptation in lower leg muscles was statistically significant in the stage when the subjects were

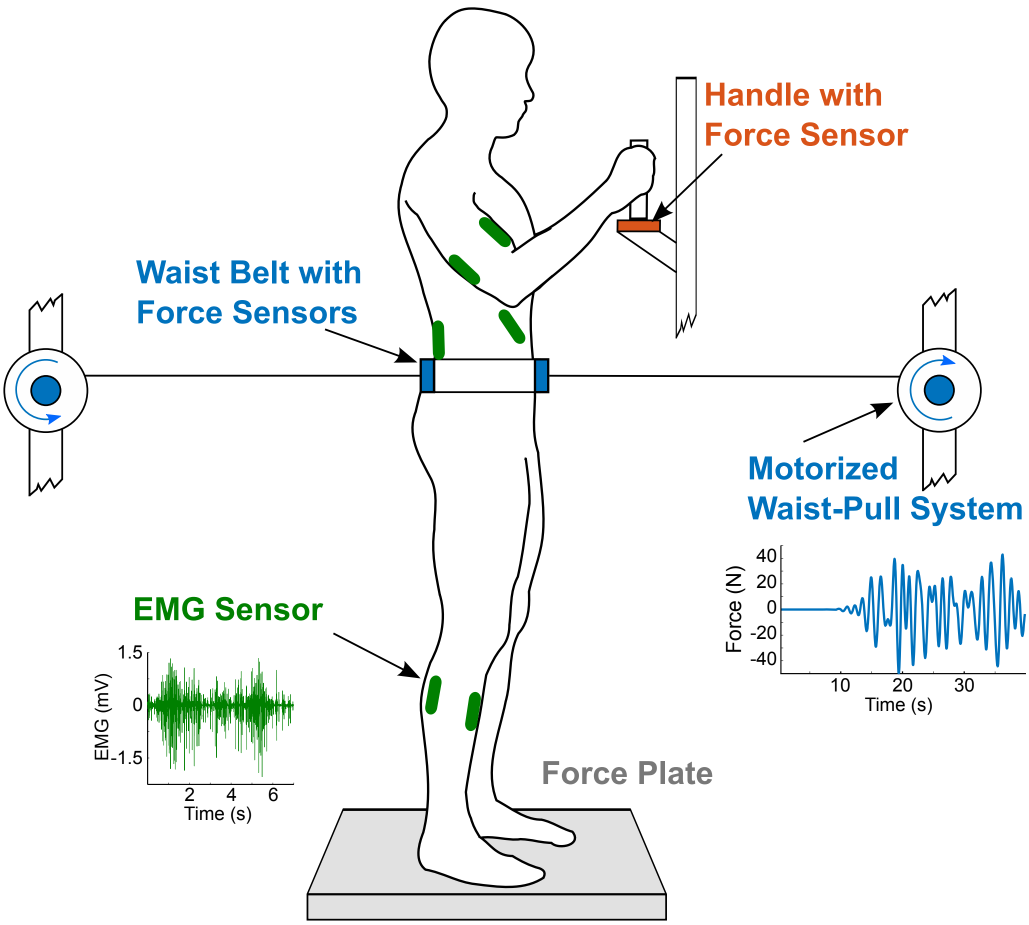


Figure 12: The subject was standing on a force plate, connected to the motorised waist-pull system that generated translational perturbations. The subject was holding the handle with a built-in force sensor mounted on a vertical pole. EMG electrodes were positioned on the major body muscles of the subject's right-hand side.

not using the additional hand support. However, this was not the case for the stage when the subjects were holding to the handle.

The activation of the trunk extensor muscle (MF) was almost the same in both stages and throughout all sessions. On the other hand, the activation of the trunk flexor OE remained unchanged throughout the sessions only in the stage when subjects held the handle. The activation of OE was much higher in this stage compared to stage when subject did not use supportive hand contact.

We performed an analysis of differences in EMG activation levels between the two experimental stages in the frequency spectrum of the perturbation waveform (low = 0.25 - 0.5 Hz, medium = 0.5 - 0.75 Hz, high = 0.75 - 1.0 Hz). A paired samples analysis between the two stage for low, medium and high frequency range revealed that there was an influence of additional hand contact on both lower leg muscles. There were confirmed statistically significant differences between the two stages in all frequency ranges and for all sessions. For the MF muscle these differences were not significant in any of the frequency range nor session.

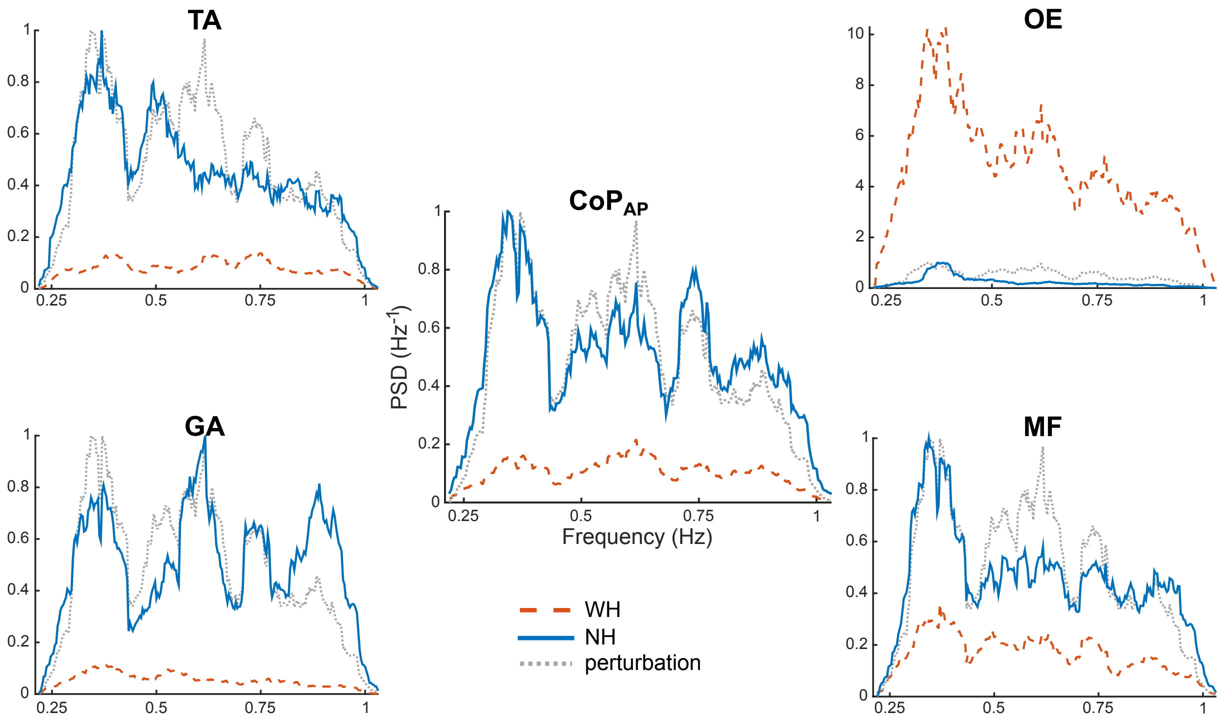


Figure 13: Effect of holding a handle after adaptation stabilised in the last session. The graphs show representative power spectral density (PSD) profiles of CoP$_{AP}$ and muscle activations measured in trunk and lower leg muscles. After the adaptation, effect of additional supportive hand contact stabilised to the perturbation in the last session. All EMG and CoP$_{AP}$ values are presented in a frequency domain, ranging from 0.25 Hz - 1 Hz. The blue (solid) lines represent the power in no-handle and the orange (dashed) lines in handle stage. The grey (dotted) line is the power of the perturbation signal. All signals are normalized to the peak value in the last session. The effect of handle is shown as reduced muscle activation in all muscles in the handle session, except in the trunk flexor muscle (OE), where there is an opposite effect.

However, there were significant differences between the two stages for the OE muscle. These differences occurred in the medium and high frequency range but only in the last session.

When the subjects were holding to the handle, we recorded the forces exerted on the handle during the continuous postural perturbations. Statistical analysis of handle forces revealed that the repetition of sessions had no significant effects (see Fig. 16). Even though the activation of arm extensor muscle changed (decreased) during sessions, there was no significant change in forces applied on the handle.

In collaboration with JSI, TUD studied whether supporting contacts in human arm reaching tasks are planned or an effect of a reactive controller. Investigations on human motor learning has focused on adaptation experiments with fixed contact points leaving research on the computational role of contacts as a free control variable unexplored.

In perturbed target reaching experiments sketched in Figure 17, we studied weather supporting contacts are planned or reactive. Subjects had to reach for distant targets on a screen with their right hand. For reaching the target additional support through contacts with a table

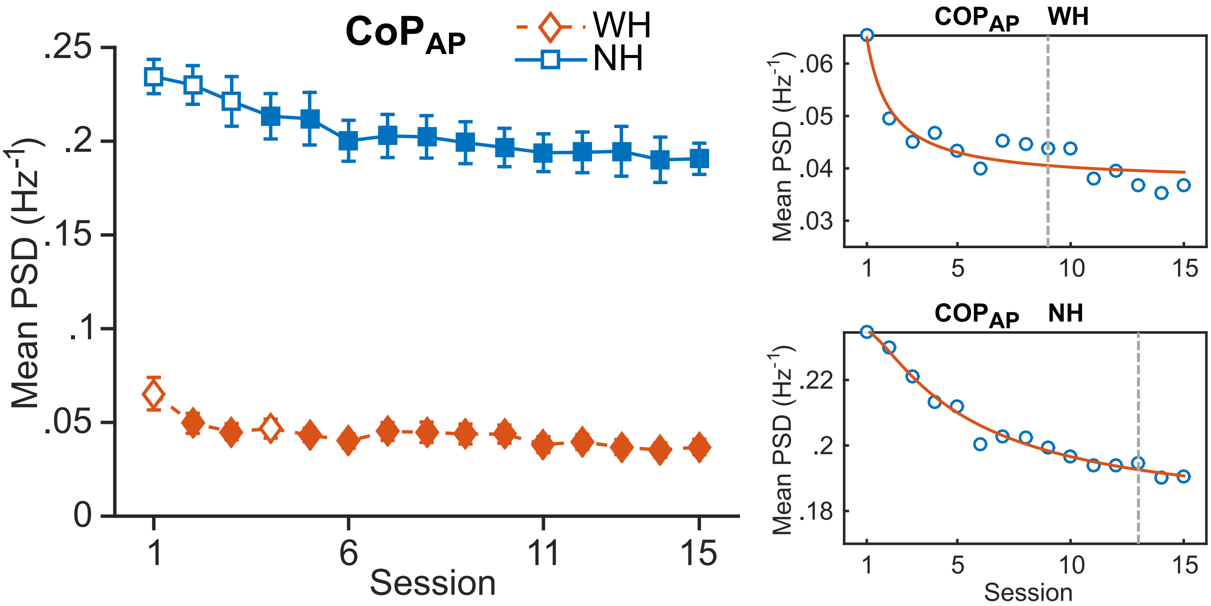


Figure 14: Adaptation of movement of CoP$_{AP}$ is shown on the left graph. Experimental stage with handle (WH) is shown in red, while condition without handle (NH) is shown in blue. Full markers indicate statistically significant differences between the first session and each of the following sessions. In both stages the adaptation is statistically confirmed (p ¡ .001). In the stage where the subjects were holding the handle, the adaptation appeared right after the first session. In no-holding stage it appeared after the third session. The superimposed best-fit curves are shown on the right graphs with orange solid lines. A calculated session number at 3τ of the fitted curve (vertical dotted line) indicates faster stabilisation of adaptation in the handle stage, compared to no-handle stage.

using the left hand was inevitably. If the contacts are planned then the left hand's motion can predict the right hand reaching.

We studied how probabilistic inference in learnt models can be used to answer this question. Evidence for planned contacts could be provided through learning probabilistic models of trajectory distributions and using the models to generate predictions, Figure 17 (a). We found that the target on the screen could be predicted from both, the left hand (mse: 10.4cm $\pm$ 2cm over 20 subjects) and the trunk movement (mse: 6.7cm $\pm$ 1.4cm over 20 subjects), which is illustrated in Figure 17 (b-c). The learnt probabilistic model could also be used to analyse the rate of adaptation of the left hand and the trunk kinematics, where the trunk trajectories converged faster than the left hand motion. This is intuitively explained by the strong need for corrective trunk movements in balancing. A report on the findings is currently in progress of writing.

Driven by the question on how human CNS optimizes arm reaching motions when the supportive hand contact has to be reached in order to maintain postural balance, UPMC and JSI started a combined experimental and computational study where the aim is to challenge two well-established but conceptually separated motor control phenomena: (i) Humans tend to reach faster to a target that looks more rewarding, despite the additional muscular cost of

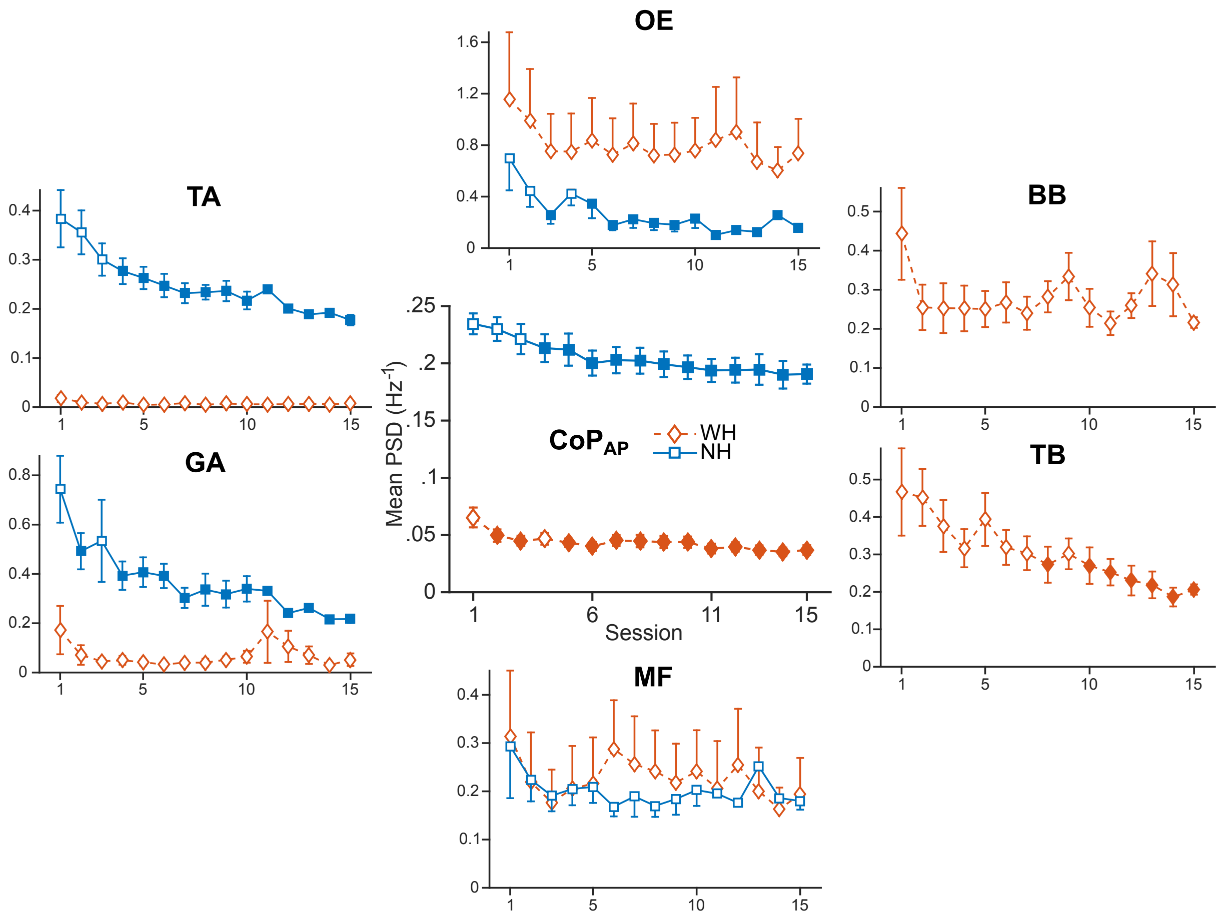


Figure 15: Figure shows the effect of repeated sessions on muscle activation and CoP_{AP} movement. Blue squares with SEM represent mean PSD in no handle and orange diamonds with SEM represent mean PSD in handle condition. Coloured markers indicate statistically significant sessions (mean values of individual session compared to the mean of the 1st session).

a faster movement [19], and (ii) when humans have to be precise, movements take longer to perform [20]. The aim of our study is to experimentally disclose both phenomena and evaluate a novel computational model designed to join them. We obtained several very promising preliminary results indicating a general mechanism that can unify both phenomena and point out a global trade-off arising from the interactions between movement time, cost and accuracy. The experimental setup is shown in Figure reffig:UnifyingTwoPhenomena A report on the findings is currently in progress of writing.

3.3.2.2.4 Resources Overall, the use of resources within WP2 was in accordance to the plans. There was a slight increase in the amount of PM for JSI due to the fact that we could not find a suitable Post-doc but hired a PhD student instead.

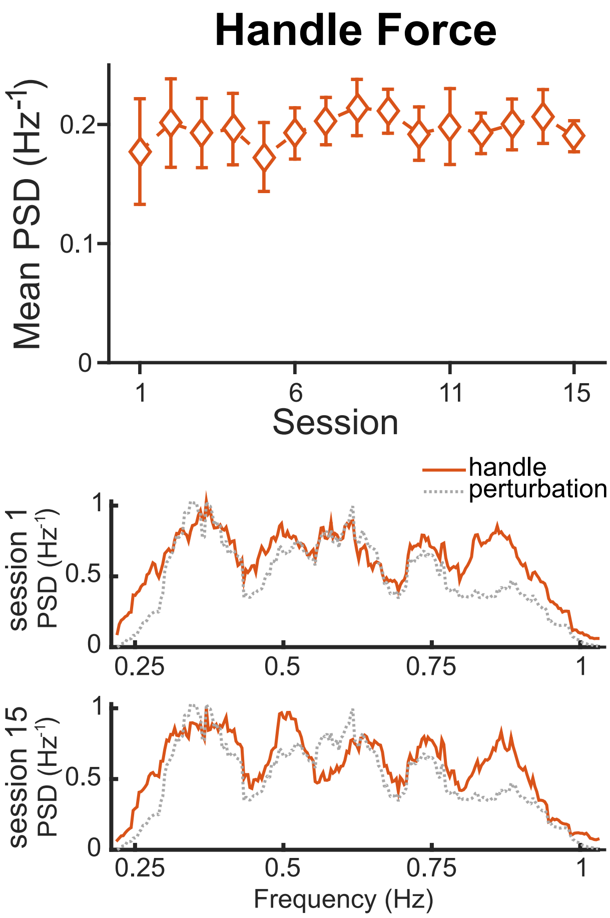


Figure 16

WP2 person months	**IIT**	**TUD**	**UPMC**	**UB**	**JSI**
Year 2	0.00	3.00	0.48	7.56	22.85
Total	0.00	4.00	2.00	45.00	55.00

3.3.2.3 Work package 3 progress The progress for each task is described hereafter.

3.3.2.3.1 Formulating the control problem (T3.2) (UPMC: 3PM) During Year 2, UPMC has explored the different possible usage of the Generalized Hierarchical Controller developed in Year 1 [21] and allowing the specification of both strict and soft hierarchical control problems. In [22], a quasi-static version of this controller is used in simulation for a standing scenario where the iCub robot is using an additional contact which is switched from one hand to the other. This control architecture has also been successfully applied on a KUKA LWR robot [23]. Tasks performances for three conflicting tasks are illustrated on Figure 19.

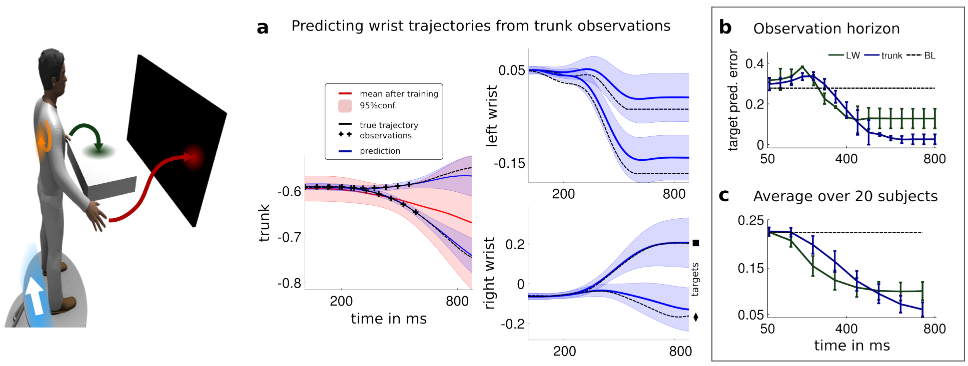


Figure 17: Trunk trajectories predict wrist trajectories. (a) 600ms of trunk trajectories are observed. These observations can predict the wrist trajectories. Shown are predictions for the two exterior targets on the screen. For training 10 trials for each target are used starting from trial 240 backwards in time (before the catch trials). For testing the first perturbed trial after trial number 240 were used. (b) The effect of the observation horizon on the target prediction error is shown for a representative subject. The mean of the training data denotes the base line (BL). (c) Average statistics (mean and 95 percent confidence bound) over 20 subjects.

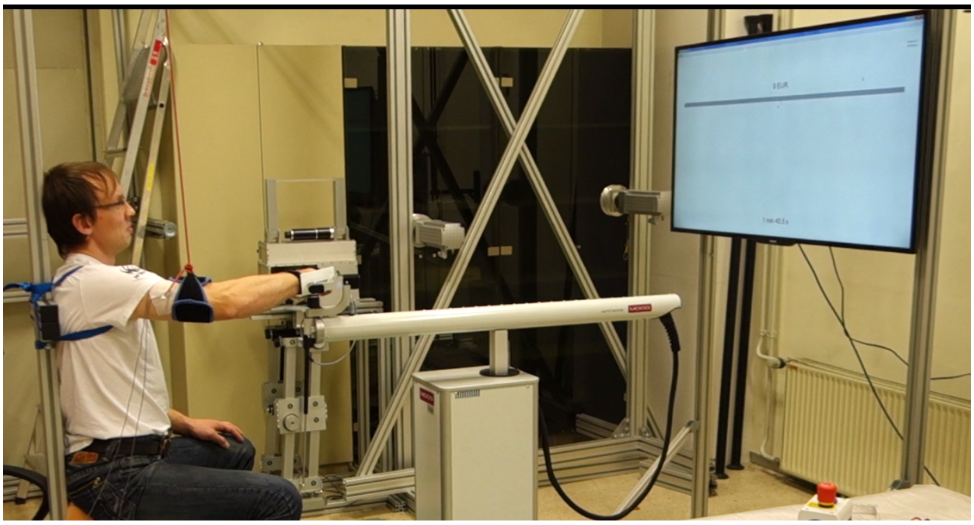


Figure 18: Experimental setup to understand how humans optimize arm reaching motions when the supportive hand contact has to be reached in order to maintain postural balance. The task of the subject was to obtain as high reward as possible in the given time by hitting a target on the virtual wall without knowing its actual size. In effect, the subjects had to find the optimal balance between precision, speed of motion and its cost in order to maximise the reward. To amplify the effect of cost of motion, haptic robot emulated a viscous media through which the subject had to move the hand.

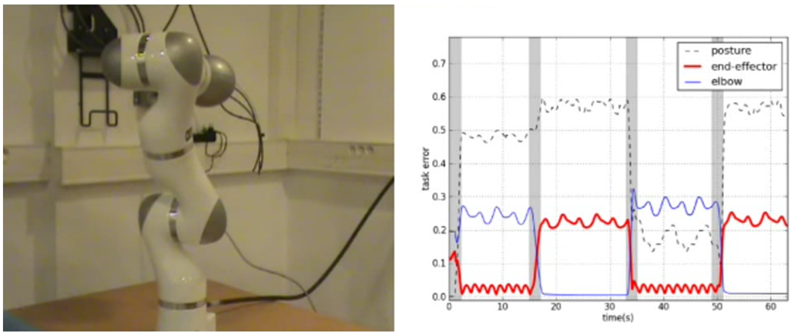


Figure 19: Tasks performance for three conflicting tasks using the GHC control framework.

3.3.2.3.2 Solving the local control problem (T3.3) (UPMC: 4.5PM, UB: 0.9PM, TUD: 2.5PM) During Year 2, UPMC has started to investigate scenarios where the robot is interacting with the environment through rigid and non-rigid contacts. Assuming that no information is a priori available regarding the nature of the contact surface, a first control strategy has been proposed in [24] where the desired contact force is adapted online as a function of the velocity of the contact point. Indeed, the risk with an unknown contact surface is to assume that it will almost instantaneously provide the required contact force to maintain the robot balance. If the surface is non-rigid, the contact point will actually move while being pushed and stable support forces will only be provided to the robot once the contact is properly established. The goal of the adaptation of the desired value for the contact force is to accelerate the obtainment of a stable contact force supporting the robot. The desired trajectory for the center of mass of the robot is also adapted to account for the non-rigidity of the contact surface. One of the advantages of this approach is that it does not actually requires the knowledge of the contact surface impedance. Figure 20 provides a view of the types of considered scenarii and the structure of the considered controller. In this work, the local control problem is solved using the solver described in [1] rather than the one developed in [21]. This choice is related to the fact that the computation cost of the GHC approach remains important and is too high to be actually used in a real-time reactive control architecture for a humanoid robot.

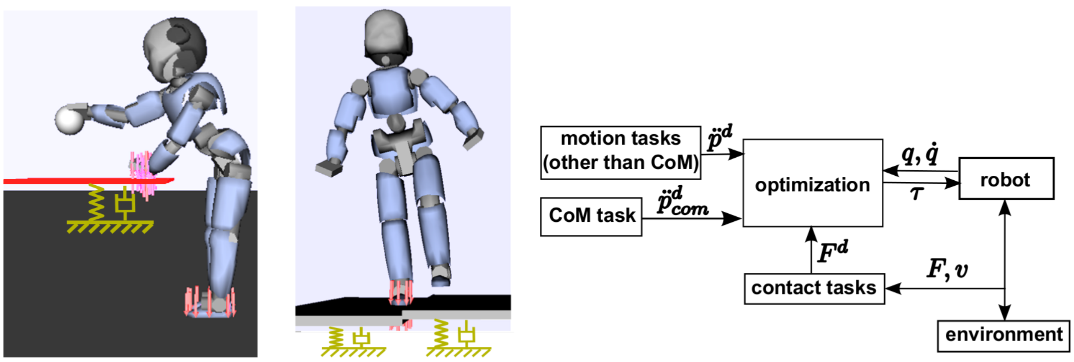


Figure 20: Scenarri of interaction with a non-rigid environment (*left*). Structure of the adaptive control architecture (*right*).

In the meantime, UB proposed an algorithm for implementing a momentum based controller to control balancing motion of legged robots with non-rigid contacts with the environments [25]. The proposed strategy converts the balance problem to a linear constrained optimization problem which its output is the vector of desired joint accelerations. It first calculates the desired supporting forces at the contact points by using the robots momentum. Then, based on the contact model and by using the Jacobian of the contact points, it converts the desired contact forces to the desired joint accelerations. At the end, by using the inverse dynamics of the robot, the joint torques are calculated. To implement the proposed method in practice, stiffness and damping coefficients of the contact model have to be estimated beforehand by using contact model parameter estimation methods.

In order to best use the whole-body controllers developed within the framework of WP3, TUD hired a student for setting up the iCub hardware and simulation environment.

3.3.2.3.3 Bootstrapping and validating the control approach in rigid world and compliant cases (T3.4) (IIT: 3PM, UPMC: 7.19PM, UB: 0.95PM, TUD: 8PM, Inria: 4PM) During Year 2, the whole-body controller developed by UPMC and originally described in [1] has been ported, using the Whole Body Interface described in Deliverable D1.2, on the iCub robot. A non-free-floating first implementation was performed during the *Veni, Vidi, Vici* 2014 summer school. The free-floating version is less straightforward as it requires: a well calibrated torque low-level control loop and a proper on-line estimation of the state of the root-body in order to compute the dynamics and kinematics models of the robot[7]. IIT is working on these two topics and UPMC, in coordination with IIT, is pursuing integration on its iCub robots.

UPMC also kept exploring the contribution of MPC approaches to handle the postural balancing under varying contact conditions. The contributions in this domain are related to the thesis work of A. Ibanez [26] and the postdoctoral work of D. Lau[8]. In order to compute optimal time, duration and position of footsteps along with the center of mass trajectory of a humanoid, a novel mixed-integer model of the system is introduced in the work of A. Ibanez. The introduction of this model in a predictive control problem brings the definition of a Mixed-Integer Quadratic Program, subject to linear constraints. Simulation results demonstrate the simultaneous adaptation of the gait pattern and posture of the humanoid, in a walking activity under large disturbances, to efficiently compromise between task performance and balance. In addition, a push recovery scenario displays how, using a single balance-performance ratio, distinct behaviors of the humanoid can be specified. Results have been obtained in simulation[9] and are being implemented on the TORO robot developed at DLR. This work is also being adapted in order to test control hypothesis to explain the behaviours observed in the experiment developed in T2.4 of WP2[10]. In the work of D. Lau, two simple and novel approaches to solve

[7] The retained solution is to choose one body in contact with the fixed environment as the root-body (one foot for example) and switch the root-body whenever there is a switch in the contact state of the system.

[8] A paper has been submitted to a conference with double blind reviews and cannot be explicitly referred to for obvious reasons.

[9] A video associated to this work can be found here.

[10] This adaptation has been discussed during the visit of Jan Babic as an invited professor during November 2014 at UPMC

for 3D locomotion with multiple non-coplanar contacts are introduced. Both formulations use model predictive control to generate dynamically balanced trajectories with no restrictions on the center of mass height trajectory. The first formulation treats the balance criterion as an objective function, and solves the control problem using a sequence of alternating convex quadratic programs, while the second formulation considers the criterion as constraints to the problem, and solves a succession of convex quadratically constrained quadratic programs. Preliminary results have been obtained in a scenario where a hand contact on a vertical wall is used to improve balance. A staircase climbing scenario has also been studied.

In coordination with T4.4 in WP4, UPMC also explored ways to optimize tasks trajectories [9] and the weights of tasks [27] in the LQP whole-body controller in order to maximize the overall control performance. More details are provided in the description of T4.4.

Continuing the activities which started in the first year of the project, IIT further improved the torque whole-body controller. During the second year of the project, attention was mainly on rigid contacts even if the theoretical and software development was guided and constrained by the requirements for extending the approach to non-rigid contacts. Several improvements to the controller significantly improved its performances. Improvements include in-situ force/torque sensor calibration [3], inertial parameter identification [4] and individual motor transfer function identification [28]. These improvements made possible quite challenging control tasks like the robust single foot balancing represented in Figure 21. Details of the controller will be given in a forthcoming publication.

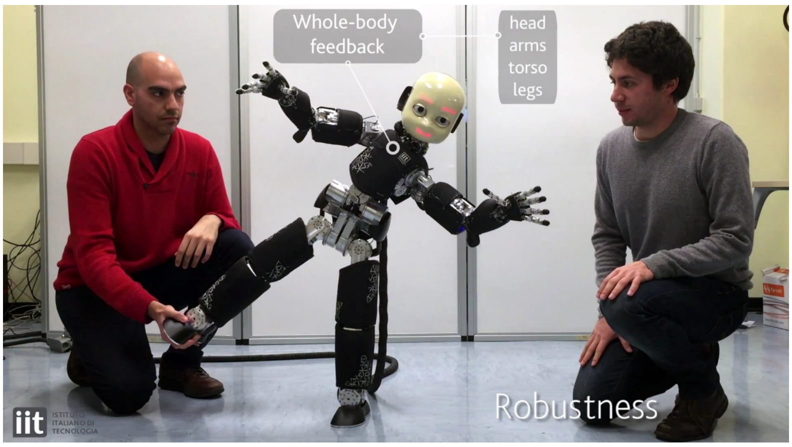


Figure 21: The picture shows the iCub while performing compliant single foot balancing. Details on the controller can be found in [28]. A video of the task is available on youtube[11].

During the second year, TUD continued their research in inverse dynamics model learning in situations with contacts. A mixture of experts approach with Gaussian processes was developed using tactile feedback from the iCub's sensor skin. This approach was evaluated on the iCub robot, where the learned model accurately predicts contact forces, is robust to changes in the environment and outperforms existing analytic dynamic models that make use of force/torque sensor data. An exemplary task is illustrated in Figure 22 when obstacles

[11]https://www.youtube.com/watch?v=SYVCbzGsBF4.

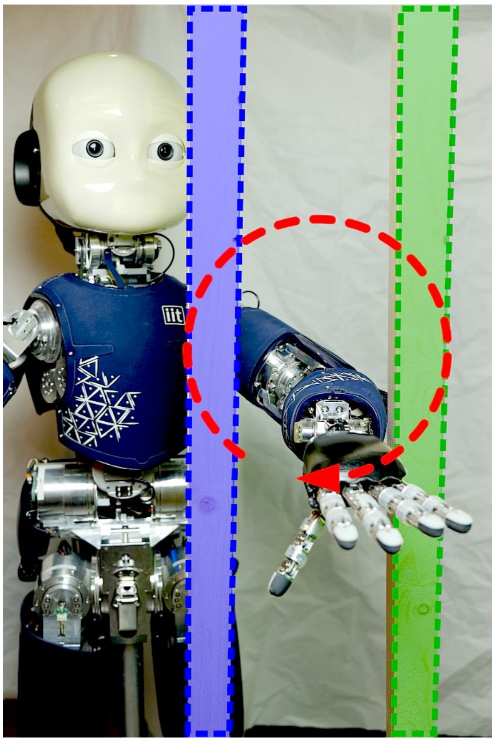


Figure 22: The robot performs a circle with its left arm. The forearm collides alternatively with the left, the right or both contacts.

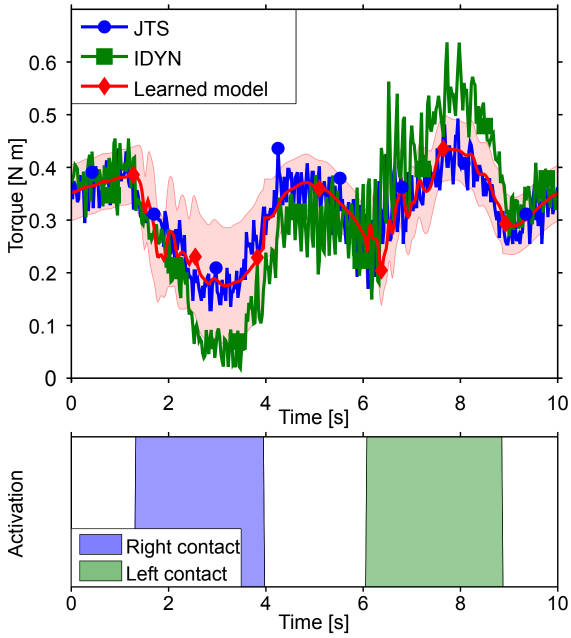


Figure 23: Prediction of torques with multiple contacts and the corresponding activation of the gating network. Our mixture-of-experts model combines the learned single-contact models into a multiple-contact model which outperform the analytic approach.

are introduced on both sides of a planned circular motion. In Figure 23 can be seen that the mixture-of-experts recognize the presence of the two different contacts and opportunely active the corresponding expert to compensate for the contact. As a result, the torques predicted from this approach (red curve) closely follow the ground truth (blue curve) and outperform the analytic model (green curve). A paper was published in an international robotics conference [29]. A study exploiting learned dynamics models for control is in progress of writing. Both studies are detailed in Deliverable D4.2.

INRIA, in collaboration with TUD, started to study the problem of bootstrapping the parameterized controllers proposed in T3.3. A major problem is how the controllers parameters, particularly the task weights, can be initialized or learned through trial-and-error. With the increasing abilities of humanoid robots, the number of tasks increases, together with their weights or priorities: manually specifying them through a sequence of complex manipulations becomes a major bottleneck. As a first step towards the offline optimization of those parameters, Inria evaluated the benefit of using a stochastic optimization derivative-free strategy such as *Covariance Matrix Adaptation Evolution Strategy* (CMA-ES) [30]. Figure 24 provides a view of the robustness of the optimization strategy by comparing several optimizations experiments with constant and random values for the GHC controller. The full study is detailed in the Task 4.4.

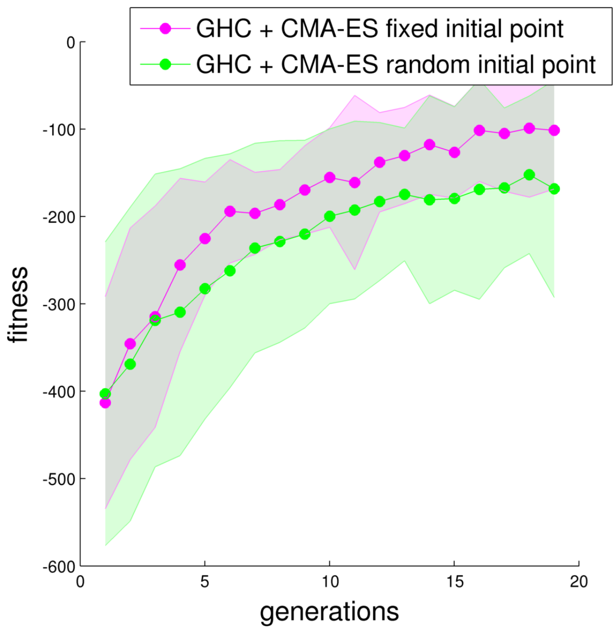


Figure 24: Learning the activation policies for GHC with CMA-ES. The policies were initialized with constant values $\alpha = 0.5$ or random values between $[0, 1]$. The mean and variance of the fitness function used for optimization for $R = 20$ replicates a test experiment (see also Task 4.4).

In the meantime, UB also studied the problem of constrained motion for a manipulator performing a task while in contact with the environment, and proposed a solution based on projected operational space dynamics [31]. The main advantages of this control technique are: 1) it exploits the environment contact constraint itself, so as to minimise the joint torques needed to perform the task; 2) it enables full decoupling of motion and force control; 3) force feedback from a force sensor mounted at the end effector or other contact points is not needed. This work is a step towards a robot control strategy which mimics the human behaviour of exploiting contacts with the environment to help perform tasks. They also presented an experimental implementation of the control method in which a KUKA LWR IV manipulator used an eraser to wipe a whiteboard. They showed that the proposed controller can effectively exploit contact with the whiteboard in order to reduce joint torques while still performing the desired wiping motion.

3.3.2.3.4 Resources

WP3 person months	IIT	TUD	UPMC	UB	JSI	Inria
Year 1	9.90	4.60	15.15	0.00	0.00	-
Year 2	3	10.5	14.69	1.85	0.00	4.00
Total	9.00	24.00	43.5	10.00	4.00	10.5

3.3.2.3.5 Deviations from workplan

No significant deviations.

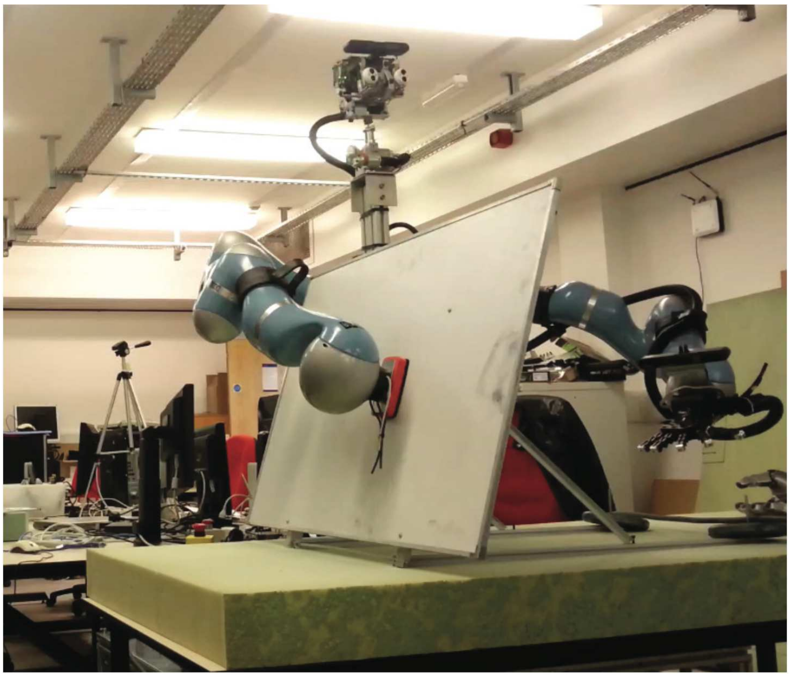


Figure 25: Experimental setting. The right arm of the humanoid platform Boris is used, and placed the board in between its two arms so that the z axis of the robot end effector frame, the z axis of the robot base frame and the axis orthogonal to the board are parallel.

3.3.2.4 Work package 4 progress

3.3.2.4.1 Improved Models from Real-Time Regression with Latent Contact Type Inference (T4.1) Within T4.1 IIT developed a theoretical framework for estimating whole-body dynamics from distributed multimodal sensors [5]. Considered sensors include joint encoders, gyroscopes, accelerometers and force/torque sensors. Estimated quantities are position, velocity, acceleration and (internal and external) wrenches on all the rigid bodies composing the robot articulated chain. The estimation procedure consists of an extended Kalman filter (EKF) which gives the a-posteriori estimation given all the available measurements. Computational efficiency is obtained by formulating the Kalman filter update-step with a sparse Bayesian network. Experiments for validating the proposed theoretical framework have been conducted on a leg of the iCub humanoid robot. The iCub is an ideal platform for the proposed experiment given its distributed force, torque, linear acceleration and angular velocity sensors. Results have shown the accuracy and the computational efficiency of the proposed method. The theoretical framework has been implemented in an open source software (see also Section 3.3.2.1.5).

TUD extended their probabilistic movement primitive (ProMP) approach in two ways. First, a mixture model approach that learns a shared latent structure of related tasks from demonstrations was developed. The shared structure is encoded by a multi-modal vector that modulates the probabilistic primitives. Both, the probabilistic primitives as well as their activations (i.e., the latent variable) are learned from demonstrations. In a table tennis ball prediction tasks this latent variable modulated the slope and the waviness of the ball trajectory. In a

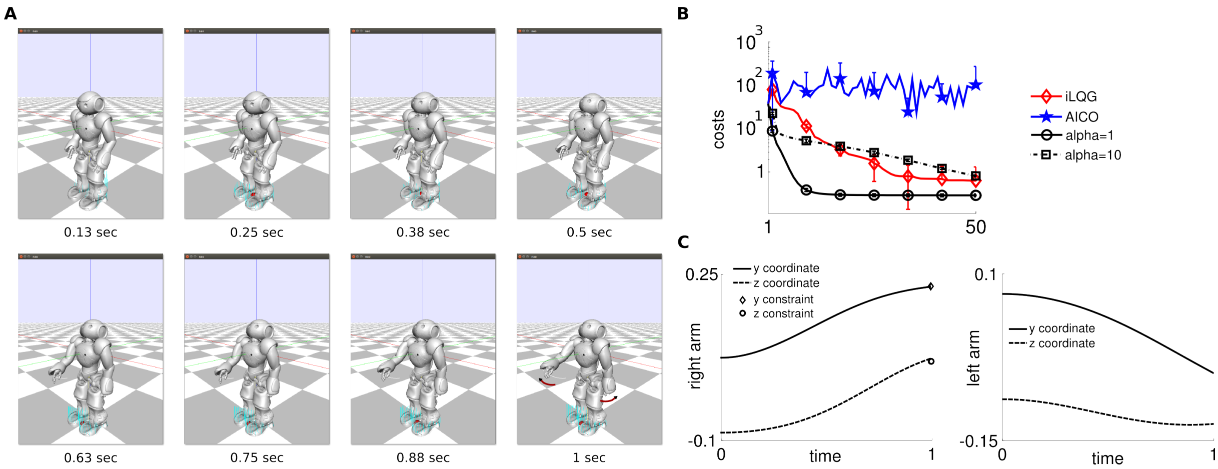


Figure 26: Reaching task with the humanoid robot *Nao*. The robot has to reach a desired end effector position with the right arm while maintaining balance. Eight snapshots of the inferred movement are shown in (A). In (B), the convergence of the costs of the optimization procedure is shown, where we compare *iLQG*, the standard implementation of AICO and the regularized variant. The movement objectives for the right arm are shown in the left panel in (C). To balance the robot lifts its left hand and bends the head back.

Kuka robot arm reaching task, the approach was used to learn bi-modal reaching trajectories that avoid an obstacle placed in front of the robot. This work is detailed in Deliverable D4.2 and will be presented at the IEEE conference on Robotics and Automation in May in Seattle, USA [7].

In a second extension of ProMPs, TUD developed a model-free control method that can be trained from demonstrations and generates time-varying feedback control gains that reproduces the demonstrations. In this approach a joint distribution over states, sensory feedback (e.g., measured joint torques or contact forces) and controls is learned. In conditioning on the current state the next-state control-law can be computed in closed form approximating the true forward dynamics through local linearizations given the demonstrations. TUD evalated this model-free ProMP method on the humanoid robot iCub in lifting objects. A conference paper is currently under review.

3.3.2.4.2 Inferring the Operational Space and Appropriate Controls with Multiple Contacts (T4.2) For controlling high-dimensional robots, most stochastic optimal control algorithms use approximations of the system dynamics and of the cost function (e.g., using linearizations and Taylor expansions). These approximations are typically only locally correct, which might cause instabilities in the greedy policy updates, lead to oscillations or the algorithms diverge. To overcome these drawbacks, TUD added a regularization term to the cost function that punishes large policy update steps in the trajectory optimization procedure. TUD applied this concept to the Approximate Inference Control method (AICO), where the resulting algorithm guarantees convergence for uninformative initial solutions without complex hand-tuning of learning rates.

The new algorithm was evaluated on two simulated robotic platforms. A robot arm with

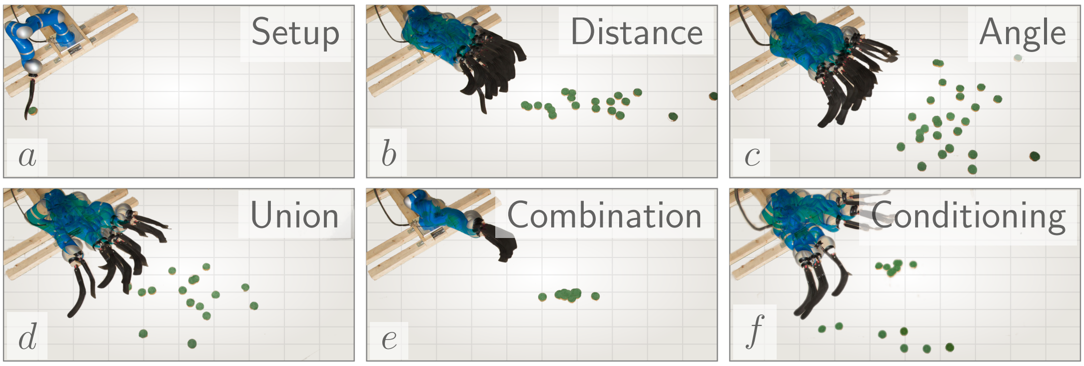


Figure 27: Robot Hockey. The robot shoots a hockey puck. We demonstrate ten straight shots for varying distances and ten shots for varying angles. The pictures show samples from the ProMP model for straight shots (b) and angled shots (c). Learning from the union of the two data sets yields a model that represents variance in both, distance and angle (d). Multiplying the individual models leads to the combined a model that only reproduces shots where both models had probability mass, in the center at medium distance (e). The last picture shows the effect of conditioning on only left and right angles (f).

five joints was used for reaching multiple targets while keeping the roll angle constant. On the humanoid robot Nao, we show how complex skills like reaching (see Figure 26) and balancing can be inferred from desired center of gravity or end effector coordinates. This work was published at the international conference on humanoid robots [8]. Supplemental Matlab demo code is available online at http://www.ausy.tu-darmstadt.de/Team/ElmarRueckert.

3.3.2.4.3 Generalizing and Improving Elementary Tasks with Contacts (T4.3)

The advent of robots in our every day life can only be accomplished with reliable mechanisms for movement generation. Movement Primitives (MP) are a well-established approach for representing modular and re-usable robot movement generators that can be composed into complex movements. An easy-to-learn representation of the primitive is, additionally, the key of recent imitation and reinforcement learning successes. Current MPs approaches offer viable properties such as concise representations of the inherently continuous and high dimensional space of robot movements, generalization capabilities to novel situations, temporal modulation of the primitive, sequencing of primitives, coupling between the degrees of freedom of the robot, and controllers for real time execution. However, no single MP framework exists that offers all these properties. During year two, TUD extended previous results on modeling stochastic movements [32, 33], where a journal version is currently under review.

TUD incorporated all the desirable properties current approaches offer into a single framework and, additionally, TUD introduced new operations on the primitives, such as continuous blending and co-activation of multiple primitives. Most importantly, in this approach, the novel co-activation operator is capable of solving multiple tasks concurrently 27. Furthermore, TUD's approach is capable of reproducing exactly the demonstrated variability of the movement and the coupling between the degrees of freedom of the robot. In this approach, called Probabilistic Movement Primitives (ProMPs) [32, 33], TUD derived all operations in closed form. In order to use the ProMPs for online feedback control, TUD also derived a

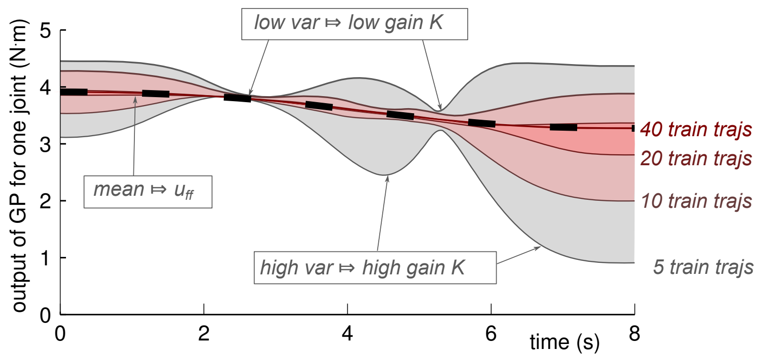


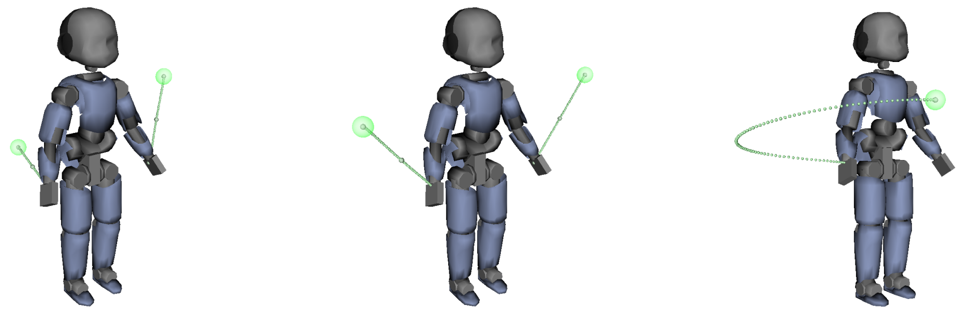
Figure 28: Mean and variance of the Gaussian process ($\mu \mp 2\sigma$) on the same test trajectory of 8 seconds, after having been trained with 5, 10, 20 and 40 training trajectories. With an increasing number of training data, the mean of the GP approaches the true function (black dashed line). The known values for uff are plotted as a black dotted line.

stochastic feedback controller that reproduces exactly the encoded primitive. TUD evaluated and compared this approach on several simulated and real robot scenarios.

Probabilistic movement primitives are a promising approach for learning, modulating, and re-using movements in a modular control architecture. To effectively take advantage of such a control architecture, ProMPs support simultaneous activation, match the quality of the encoded behavior from the demonstrations, are able to adapt to different desired target positions, and efficiently learn by imitation. ProMPs meets all of the aforementioned requirements. The desired trajectory distribution of the primitive is parametrized by a hierarchical Bayesian model with Gaussian distributions. The trajectory distribution can be obtained from demonstrations and simultaneously defines a feedback controller which is used for movement execution. Currently, TUD is investigating extensions of the ProMPs framework to tasks that involve contacts with the environment (see T4.1). In addition, TUD started to investigate the improvement of elementary skills encoded in ProMPs with reinforcement learning, where a conference paper was submitted for review.

3.3.2.4.4 Learning the Prioritization of Tasks (T4.4) During year two, UB continued their research on computed torque control leveraging low-gain control. In computed torque control, robot dynamics are predicted by dynamic models. This enables more compliant control, as the gains of the feedback term can be lowered, because the task of compensating for robot dynamics is delegated from the feedback to the feedforward term. We already showed that Gaussian process regression is an effective method for learning computed torque control, by setting the feedforward torques to the mean of the Gaussian process. During the second year of the project, we extended this work by also exploiting the variance predicted by the Gaussian process, by lowering the gains if the variance is low [34]. This enables an automatic adaptation of the gains to the uncertainty in the computed torque model, and leads to more compliant low-gain control as the robot learns more accurate models over time. On a simulated 7-DOF robot manipulator, we demonstrated how accurate tracking can be achieved, despite the gains being lowered over time, which is illustrated in Figure 28.

Within T4.4, UPMC studied how to deal with interferences between tasks using machine learning tools. Whole-Body Control methods offer the potential to execute several simultane-



(a) Constrained Configuration (b) Workspace Violation (c) Balance Perturbation

Figure 29: Three common multi-task incompatibility scenarios. The desired hand task trajectories are indicated by the green markers. Medium size spheres represent waypoints, and large transparent spheres represent the final waypoints or goals.

ous tasks on highly redundant robots, such as humanoids. Unfortunately, task combinations often result in interferences or incompatibilities which generate undesirable behaviors. Prioritization schemes between tasks, such as strict and soft hierarchies, are typically used to manage these interferences but generally require a deal of time consuming and arbitrary tuning.

To circumvent theses issues, UPMC presented a novel framework for defining and optimizing multiple tasks in order to resolve potential interferences prior to task execution. In a first study [9] the tasks are parameterized with Dynamical Movement Primitives, whose parameters are optimized based on a general compatibility principle, which is independent of the robots topology, tasks or environment. Two test cases on a simulation of a humanoid robot are used to demonstrate the successful optimization of initially interfering tasks. A video summarizing the outcome of this work can be viewed here.

In a second study [27], UPMC studied how task variability can be used to modulate task priorities during their execution, to temporarily deviate certain tasks in the presence of incompatibilities. A method for mapping from task variance to task priority was presented as well as an approach for calculating task variance for generated trajectories. The method successfully resolved three common task conflict scenarios online illustrated on Fig. 29.

TUD addressed the problem of learning the temporal profile of soft task priorities and null-space projectors for the multi-task controllers developed in WP3. Our preliminary results have been submitted to a robotics conference[12].

The first controller of WP3, *Prioritized Task-Space Inverse Dynamics*, is based on *strict task hierarchies*, where a hierarchical ordering of the tasks is set, such that critical tasks (or tasks that are considered as more important) are fulfilled with higher priorities, and the low-priority tasks are solved in the null-space of the higher priority tasks [35]. The strict control approach requires the pre-specification of the task hierarchy. However, in many contexts it is difficult to organize the tasks in a stack and define their relative importance in forms of priorities. When priorities are strict, a higher task can completely block lower tasks, which can result in movements that are not satisfactory for the robot mission (e.g., its "global" task).

The second controller of WP3 is based on *soft task hierarchies*, where the solution is

[12]Modugno, V.; Neumann, G.; Rueckert, E.; Oriolo, G.; Peters, J.; Ivaldi, S. *Learning soft task prioirities and null-space projectors for motion planning of redundant manipulators*. Submitted to IROS 2015.

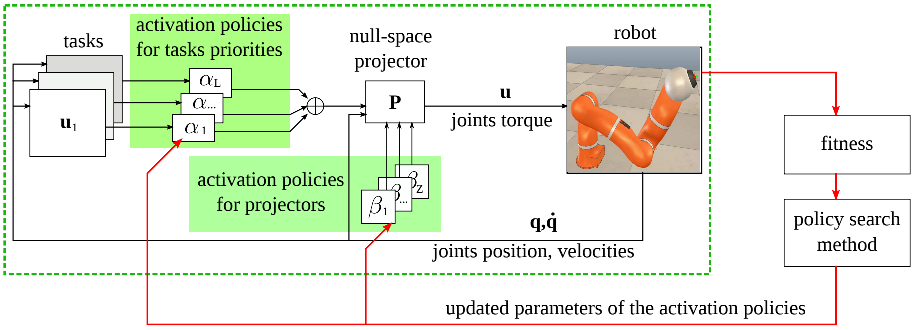


Figure 30: This scheme briefly describes the proposed method. The control torques are computed by a combination of tasks weighted by soft priorities, represented as parameterized activation policies, that are multiplied by a Null-space projector, where some activation functions for different projectors are joined. The global task execution is evaluated and a fitness function is computed: a policy search method is then used to optimize the parameters of the activation policies, both for tasks and projector.

typically given by a combination of weighted tasks [36]. The importance or "soft priority" of each individual task is represented by a scalar weight. By tuning the vector of scalar weights, evolving in time, the global robot behavior can be optimized. Within WP3, Liu et al. [37] propose a generalized projector (GHC) that handles strict and non-strict priorities with smooth transitions when tasks priorities are swapped. They show that adapting these weights may result in a seamless transition between tasks (i.e., reaching for an object, staying close to a resting posture and avoiding an obstacle) and in continuous task sequencing. Despite the elegant framework, their controller needs again a lot of manual tuning: particularly, the evolution of the tasks priorities in time, the timing and the tasks transitions need to be designed by hand. While this approach could still be easy for few tasks and simple robotic arms, it can quickly become unfeasible for complex robots such as humanoids performing whole-body movements that usually require a dozen of tasks and constraints (e.g., control balance, posture, end-effectors, stabilize head gaze, prevent slipping, control interaction forces etc.).

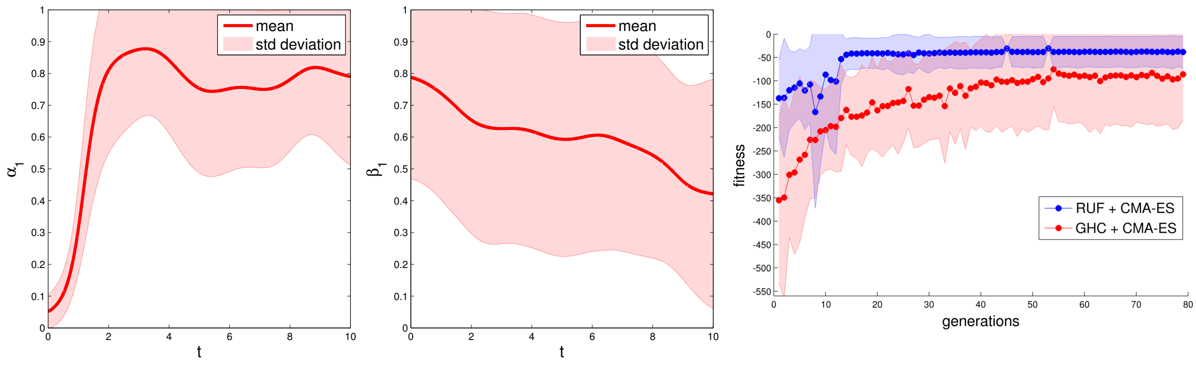
In this task we propose a first solution to the problem of how these weights can be learned through trial-and-error. We study how the *temporal* profiles of the task weights can be learned from a reward function, which is assumed to be given[13].

As a first step towards a controller that is capable of handling multiple tasks and constraints on a complex robot, while allowing us to efficiently learn the task priorities, we propose a regularized version of the Unified Framework (RUF) proposed by Peters et al [39], where the tasks weights and Null space projectors weights are represented by parametrized functional approximators that can be automatically determined through a stochastic optimization procedure. The concept is presented in Figure 30.

As a first results, we show that the optimization process generates weights profiles that cannot be designed manually in advance, see panels (a) and (b) in Figure 31.

We then compare the performance of our controller with the state-of-the-art method GHC

[13]For many robotic task, e.g., tracking desired center-of-mass or end-effector trajectories while avoiding obstacles, such reward functions have been defined in [38].



(a) Attractor activation. (b) Null-projector activation. (c) Learning performance.

Figure 31: The panels in (a) and (b) show the mean and standard deviation of the temporal profile of the activation functions α, β, optimized by RUF+CMA-ES, computed over $R = 50$ replications of the same experiment of the table scenario. (c) Comparison of our method (blue line) to the generalized projector method (GHC).

proposed by Liu et al. [37] (WP3). We consider the following experimental scenario: a 7-DOF KUKA Light Weight Robot arm, starting from a vertical position, must reach a desired point with its end-effector, while avoiding to collide with a table, represented by a surface parallel to the z-axis in between the robot and the goal. The aim of the experiment is to bring the end-effector as close as possible to the desired position, while avoiding collisions with the obstacle. We define $T = 3$ tasks: a regulation task in the joint space, and two reaching tasks for the elbow and the end-effector. For both methods, we find the optimal profiles for the weighting functions with CMA-ES. Our second result is that our controller generally performs better than GHC, even if we optimize the policies in both methods: on an average of 20 replicates, our RUF+CMA-ES finds 90% of the solutions found by our RUF+CMA-ES satisfies the constraints, while only 75% of the solutions of GHC+learning are acceptable. Furthermore, the final best solution found by RUF+CMA-ES outperforms the one of GHC+CMA-ES, as shown in Figure 31 (c).

3.3.2.4.5 Resources

WP4 person months	**IIT**	**TUD**	**UPMC**	**UB**	**JSI**	**INRIA**
Year 1	0.00	8.00	2.22	0.00	0.00	0.0
Year 2	0.00	21.70	1.68	2.04	0.00	2.0
Total	30.00	36.00	20.00-6.00	12.00	10.00	XX

3.3.2.4.6 Deviations from workplan Six person months of UPMC planned for Serena Ivaldi were transfered to TUD. No significant deviations.

3.3.2.5 Work package 5 progress The activities in WP5 are divided into four tasks corresponding to the four years project duration. As a result, during the second year CoDyCo results concentrate on T5.2. The main result consist in the implementation of the validation scenario consisting of the balancing on different type of rigid contacts.

3.3.2.5.1 Scenario 2: iCub posture control while performing goal directed actions (T5.2) The main contributions to T5.2 have been presented in "Validation scenario2: balancing on feet while performing goal directed actions." which discusses the technical implementation of the second year validation scenario (see `https://github.com/robotology-playground/codyco-deliverables/tree/master/D5.2/pdf`). The software developed for the scenario implementation is released with an open-source license and distributed through github (`https://github.com/robotology/codyco-modules`). The main software activities include: a module to identify the whole-body motor transfer functions (`https://github.com/robotology/codyco-modules/tree/master/src/modules/motorFrictionIdentification`), a module for estimating whole-body internal (joint torques) and external (contact) forces (`https://github.com/robotology/codyco-modules/tree/master/src/modules/wholeBodyDynamicsTree`), a module for whole-body joint torque control (`https://github.com/robotology/codyco-modules/tree/master/src/devices/jointTorqueControl`), a C++ module for whole-body control under multiple rigid contacts (`https://github.com/robotology/codyco-modules/tree/master/src/modules/torqueBalancing`).

3.3.2.5.2 Deviations from workplan The original work plan was leaving quite a flexible set of possibilities for both the postural task (e.g. sitting on a chair or balancing on the feet) and the goal oriented action (e.g. opening a drawer while standing or manipulating object while sitting). In the final validation scenario it was chosen to consider a interactive scenario, with the torque controlled iCub standing on his feet while trying to grasp an object moved by the experimenter. As soon as the object exceeds the robot workspace, the iCub takes a forward step in order to increase his workspace.

3.3.2.5.3 Resources Resources were used with small difference with respect to what planned. In particular IIT invested only 2 PM with respect to 12PM planned. The motivation resides in the fact that WP5 took advantage of the significant effort done in WP1 (software) and WP3 (control) and in a sense resources initially planned on T5.1 eventually have been committed to T1.1, T1.2, T1.3, T3.1 and T3.2.

WP5 person months	**IIT**	**TUD**	**UPMC**	**UB**	**JSI**
Year 1	2.00	0.00	0.31	0.00	0.00
Total	48.00	4.00	5.00	0.00	0.00

3.3.2.5.4 Deviations from workplan No significant deviations.

3.3.2.6 Work package 6 progress Activities within work package 6 achieved the expected results both in terms of administrative activities and management activities. As a major achievement, the management successfully concluded a second amendment to include INRIA as a partner. The inclusion was motivated by the the new position of Dr. Serena Ivaldi, currently researcher at INRIA, Nancy.

3.3.2.6.1 Administrative coordination (T6.1) Administration was successfully coordinated by IIT, with significant contribution from Chiara Andreoli (iCub Facility), Francesca Boscolo (project offices) and Maria Carmela Fierro (Robotics, Brain and Cognitive Science Department). The major activity concerned the amendment that the CoDyCo consortium asked the main reason being the fact that Serena Ivaldi, initially hired by UPMC and successively moved to TUD, was recently hired by INRIA as researcher. Part of the administrative coordination activities were also conducted during the mid-year meeting: November 20th-21st, 2014, Ljubljana. Details on the meetings can be found in the CoDyCo website (`http://www.codyco.eu`).

3.3.2.6.2 Software repository implementation (T6.2) The github software repository was several times restructured `https://github.com/robotology/codyco` and the contribution from the different developers can be directly checked in the website.

3.3.2.6.3 Resources Resources were used as follows.

WP6 person months	**IIT**	**TUD**	**UPMC**	**UB**	**JSI**
Year 1	1.46	0.00	0.25	0.00	0.10
Total	5.00	1.00	1.00	0.60	1.00

3.3.2.6.4 Deviations from workplan No significant deviations.

3.3.2.7 Work package 7 progress Dissemination and exploitation activities included the participation to international events addressed to both commercial and academic institutions. A preliminary exploitation plan was delineated and reported in the deliverable D7.1.

3.3.2.7.1 Dissemination activities towards academia, industry, and other users (T7.1) Dissemination activities were conducted in three main events: (1) iCub exposition at ICRA 2013, IEEE International Conference on Robotics and Automation Karlsruhe, May 6 - 10, 2013; (2) iCub exposition at the European Robotics Forum and Innorobo, Lyon 29th March 2013; (3) iCub exposition at the European Robotics Forum, Rovereto 12th-14th of March 2014. The full list of papers published within CoDyCo can be found here: `http://codyco.eu/publications-menu`.

3.3.2.7.2 Exploitation plan (T7.2) The second year activities on T7.1 and T7.2 are all contained in "D7.1 Dissemination and exploitation plan" available here: `https://github.com/robotology-playground/codyco-deliverables/tree/master/D7.1/pdf`.

3.3.2.7.3 Management of IPR (T7.3) No activities to be reported during the second year on this task in consideration of the fact that the task started at the very end of the second year. As a minor starting activity the consortium circulated a list containing each partner responsible contact person for the IPR management. This list is contained in "D7.1 Dissemination and exploitation plan" available here: `https://github.com/robotology-playground/codyco-deliverables/tree/master/D7.1/pdf`.

3.3.2.7.4 Dissemination of a database of human motion with contacts (T7.4) During the second year of CoDyCo, IIT completed the task of setting up a database for storing both human and robot datasets. The details on the database are reported in "D7.2 Standard database with support materials" available here `https://github.com/robotology-playground/codyco-deliverables/tree/master/D7.2/pdf`.

3.3.2.7.5 Resources Resources were used as follows.

WP7 person months	**IIT**	**TUD**	**UPMC**	**UB**	**JSI**
Year 1	1.00	0.00	0.40	0.00	0.00
Total	3.00	1.00	2.00	1.00	1.00

3.3.2.7.6 Deviations from workplan No significant deviations.

3.4 Deliverables and milestones tables

3.4.1 Deliverables (excluding the periodic and final reports)

3.4.2 Milestones

Del. no.	Deliverable name	WP	Type	Date	Respon sible	Person Month
D1.2	Software for controlling of balancing and reaching with multiple contacts.	1	SW	M24	UB	16
D3.1	Local solver in rigid-world cases.	3	R	M24	UPMC	18
D4.2	Learning of tasks with multiple contacts by imitation and reinforcement learning.	4	R	M24	TUD	30
D5.2	Validation scenario2: balancing on feet while performing goal directed actions.	5	R	M24	IIT	13

R = Report, **P** = Prototype, **D** = Demonstrator, **SW** = Software, **O** = Other

Milestone number	Milestone name	Work package(s) involved	Expected date [1]	Leader	Means of verification
MS.2	Validation scenario2: balancing on feet while performing goal directed actions	MS.1 T1.3 T1.5 T4.3 T5.2	M24	IIT	- The iCub successfully reaches an object while exploiting multiple contacts

References

[1] J. Salini, "Dynamic control for the task/posture coordination of humanoids: toward synthesis of complex activities," Ph.D. dissertation, Université Pierre et Marie Curie,

Paris, France, June 2012.

[2] A. Del Prete, "Control of Contact Forces using Whole-Body Force and Tactile Sensors: Theory and Implementation on the iCub Humanoid Robot," Ph.D. dissertation, University of Genova, Genova, Italy, 2013.

[3] S. Traversaro, D. Pucci, and F. Nori, "In situ calibration of six-axis force-torque sensors using accelerometer measurements," in *IEEE International Conference on Robotics and Automation (ICRA 2015)*, Seattle, USA, May 26th - 30th, 2015, 2015.

[4] S. Traversaro, A. Del Prete, S. Ivaldi, and F. Nori, "Inertial parameters identification and joint torques estimation with proximal force/torque sensing," in *IEEE International Conference on Robotics and Automation (ICRA 2015)*, Seattle, USA, May 26th - 30th, 2015, 2015.

[5] F. Nori, "Berdy: Bayesian estimation for robot dynamics," in *submitted*, 2015.

[6] F. Nori, N. Kuppuswamy, and S. Traversaro, "Covariance estimation in newton-euler equations with redundant measurements: application to articulated rigid bodies." in *submitted*, 2015.

[7] E. Rueckert, J. Mundo, A. Paraschos, J. Peters, and G. Neumann, "Extracting low-dimensional control variables for movement primitives," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, Seattle, USA, May 26-30 2015, pp. 1–9.

[8] E. Rueckert, M. Mindt, J. Peters, and G. Neumann, "Robust policy updates for stochastic optimal control," in *Proceedings of the International Conference on Humanoid Robots (HUMANOIDS)*, Madrid, Spain, 2014. [Online]. Available: http://www.ias.informatik.tu-darmstadt.de/uploads/Site/EditPublication/AICOHumanoidsFinal.pdf

[9] R. Lober, V. Padois, and O. Sigaud, "Multiple task optimization using dynamical movement primitives for whole-body reactive control," in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, Madrid, Spain, Nov. 2014, pp. 193–198.

[10] S. Traversaro, A. Del Prete, R. Muradore, L. Natale, and F. Nori, "Inertial Parameter Identification Including Friction and Motor Dynamics," in *Humanoid Robots, 13th IEEE-RAS International Conference on*, Atlanta, USA, October 2013. [Online]. Available: http://www.humanoids2013.com/

[11] S. Traversaro, A. D. Prete, S. Ivaldi, and F. Nori, "Avoiding to rely on inertial parameters in estimating joint torques with proximal f/t sensing," in *Submitted to IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, Chicago, Illinois, 2014.

[12] M. Azad, J. Babic, and M. Mistry, "Effects of hand contact on the stability of a planar humanoid with a momentum based controller," in *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*, Nov 2014, pp. 348–353.

[13] J. Babic, T. Petric, L. Peternel, and N. Sarabon, "Effects of supportive hand contact on reactive postural control during support perturbations," *Gait & Posture*, vol. submitted for publication, 2014.

[14] A. J. Ijspeort, J. Nakanishi, and S. Schaal, "Learning rhythmic movements by demonstration using nonlinear oscillators," in *Intelligent Robots and Systems (IROS), 2002 IEEE/RSJ International Conference on*, vol. 1, Sept 2002, pp. 958–963.

[15] S. Schaal and C. G. Atkeson, "Constructive incremental learning from only local information," *Neural Comput.*, vol. 10, no. 8, pp. 2047–2084, Nov. 1998.

[16] L. Peternel, T. Petrič, E. Oztop, and J. Babič, "Teaching robots to cooperate with humans in dynamic manipulation tasks based on multi-modal human-in-the-loop approach," *Autonomous robots*, vol. 36, no. 1-2, pp. 123–136, jan 2014.

[17] L. Peternel, T. Petrič, and J. Babič, "Human-in-the-loop approach for teaching robot assembly tasks using impedance control interface," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, May 2015, pp. –.

[18] L. Peternel and J. Babič, "Learning of compliant human-robot interaction using full-body haptic interface," *Advanced Robotics*, vol. 27, no. 13, pp. 1003–1012, Jun 2013.

[19] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of experimental psychology*, vol. 47, no. 6, pp. 381–391, 1954.

[20] R. Shadmehr, J. J. Orban de Xivry, and Xu-Wilson, "Temporal discounting of reward and the cost of time in motor control." *The Journal of neuroscience*, vol. 30, no. 31, pp. 10 507–10 516, 2010.

[21] M. Liu, Y. Tan, and V. Padois, "Generalized Hierarchical Control," *Submitted to Autonomous Robots*, 2014.

[22] M. Liu and V. Padois, "Whole-Body Hierarchical Motion and Force Control for Humanoid Robots," *Submitted to Autonomous Robots*, 2014.

[23] M. Liu, S. Hak, and V. Padois, "Generalized projector for task priority transitions during hierarchical control," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Seattle, USA, May 2015.

[24] M. Liu and V. Padois, "Reactive whole-body control for humanoid balancing on non-rigid unilateral contacts," in *Submitted to the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015.

[25] M. Azad and M. Mistry, "Balance control strategy for legged robots with compliant contacts," in *IEEE International Conference on Robotics and Automation*, 2015.

[26] A. Ibanez, P. Bidaud, and V. Padois, "Emergence of humanoid walking behaviors from mixed-integer model predictive control," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, USA, Sept. 2014.

[27] R. Lober, V. Padois, and O. Sigaud, "Variance modulated task prioritization in whole-body control," in *Submitted to the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Hamburg, Germany, Sept. 2015.

[28] F. Nori, S. Traversaro, J. Eljaik, F. Romano, A. Del Prete, and D. Pucci, "icub whole-body control through force regulation on rigid non-coplanar contacts," *Fontiers in Robotics and AI*, 2015.

[29] R. Calandra, S. Ivaldi, M. Deisenroth, E. Rueckert, and J. Peters, "Learning inverse dynamics models with contacts," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2015. [Online]. Available: http://www.ausy.tu-darmstadt.de/uploads/Site/EditPublication/Calandra_ICRA15.pdf

[30] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evolutionary computation*, vol. 9, no. 2, pp. 159–195, 2001.

[31] V. Ortenzi, M. Adjigble, J. Kuo, R. Stolkin, M. Mistry, *et al.*, "An experimental study of robot control during environmental contacts based on projected operational space dynamics," in *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*. IEEE, 2014, pp. 407–412.

[32] A. Paraschos, C. Daniel, J. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in Neural Information Processing Systems (NIPS), Cambridge, MA: MIT Press.*, 2013. [Online]. Available: http://www.ias.tu-darmstadt.de/uploads/Publications/Paraschos_NIPS_2013.pdf

[33] A. Paraschos, G. Neumann, and J. Peters, "A probabilistic approach to robot trajectory generation," in *Proceedings of the International Conference on Humanoid Robots (HUMANOIDS)*, 2013. [Online]. Available: http://www.ias.tu-darmstadt.de/uploads/Publications/Paraschos_Humanoids_2013.pdf

[34] N. T. Alberto, M. Mistry, and F. Stulp, "Computed torque control with variable gains through gaussian process regression," in *IEEE International Conference on Humanoid Robots*.

[35] A. Del Prete, F. Nori, G. Metta, and L. Natale, "Prioritized motion-force control of constrained fully-actuated robots: "task space inverse dynamics"," *Robotics and Autonomous Systems*, vol. 63, pp. 150–157, Jan 2015.

[36] J. Salini, V. Padois, and P. Bidaud, "Synthesis of complex humanoid whole-body behavior: A focus on sequencing and tasks transitions," in *IEEE Int. Conf. Robotics and Automation*, 2011, pp. 1283–1290.

[37] M. Liu, S. Hak, and V. Padois, "Generalized projector for task priority transitions during hierarchical control," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 2015.

[38] J. Kober, D. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," no. 11, pp. 1238–1274, 2013.

[39] J. Peters, M. Mistry, F. E. Udwadia, J. Nakanishi, and S. Schaal, "A unifying framework for robot control with redundant dofs," no. 1, pp. 1–12, 2008.