EPFL

Enhancing our understanding of human fine manipulation skills and advancing robot dexterity in grasping

Kunpeng YAO

■ École
polytechnique
fédérale
de Lausanne

2022

To my parents
for their love, support, and encouragement

Acknowledgements

As I complete my PhD and begin to reflect on the past four years, I feel very blessed. I have received help from many people during my PhD, without whom I would not have been able to accomplish this dissertation.

First and foremost, I am extremely grateful to my thesis advisor, Prof. Aude Billard, for guiding me along the academic road and supporting me constantly. I very much appreciate her for launching the Skill Acquisition in Humans and Robots (SAHR) project in the interdisciplinary field of humans and robotics, offering me the freedom to explore research topics that interest me in both fields. She held my scientific work to a high standard and trained me to think the way I should possess as a scientific researcher. She always pushed me to explore the essence of the problem and reminded me, "what is the scientific problem?" I have benefited greatly from her foresighted academic perspective and her rigorous thinking in approaching research problems. I would like to thank her for being a scientist that never compromised on the pursuit of her academic goals, and for being an advisor that never reduced her expectations for my research work.

I was very fortunate to have had the opportunity to work with Prof. Dagmar Sternad. Her extensive knowledge, insights, and valuable suggestions helped me acquire a in-depth understanding of the field. I would like to thank Prof. Danica Kragic and Prof. Stalian Coros for serving as jury members of my thesis defense committee. Their feedback and questions encouraged me to reflect more broadly and deeply on my research work. Many thanks to Prof. Jamie Paik for chairing my defense committee.

I have enjoyed working at the Learning Algorithms and Systems Laboratory (LASA) for four years. LASA has been not only a relaxed and flat work environment, but also a warm place full of friendly people. Special thanks to my office mates, Michael Bombile and David Gonon, for the enriching time we have spent together in the office. Michael always patiently answered my questions and carefully derived formulas on the white board; David used to drink beers with me at the Satellite bar from time to time, and we discussed all kinds of interesting topics. Many thanks to Bernardo Fichera, for being a sincere, witty, and humorous friend. We have spent a lot of joyful times together at the cinema and concerts.

I would like to express my sincere thanks to the SAHR project members. I very much appreciate the assistance of Andrew Sutcliffe, who helped us set up the data collection environment and

friendship over all these years.

I feel very lucky to have known Ling Zhou at EPFL and to have spent much joyful time with her around Lausanne. We have many hobbies in common, and often share our happiness with each other.

My deepest gratitude goes to my parents, to whom I would like to dedicate this thesis. Over the years, they have been there for me and supported me in my pursuit of my own life, and have advised and encouraged me when I encountered difficulties. My gratitude is beyond words.

Finally, I am so fortunate to be able to share my life these years with my beloved girlfriend, Yike Tao, for bringing so much joy to my life. She appreciates my merits and is also willing to accept my flaws, and accompanies me to become a better version of myself.

Lausanne, January 17, 2022   Kunpeng Yao

Preface

As I write these lines, my brain is fully occupied with confused thoughts about Science and its bearing to the life of a PhD student. Hardly a thought is given to the rapid, precise and coordinated motion of my fingers, typing these words on the keyboards. Without the luxury of typing lessons, but only the fear of early professional obsolescence and years of practice, my fingers have magically figured out how to translate my thoughts into sets of strokes on a keyboard; yet these movements have little in common with the use of a pen that I learned as a child. This capacity of acquiring and modifying the motion of our fingers and hands to handle new tools, in new ways and for new purposes, is a competence unique to humans. Yet, which brain mechanisms underlay these competences still largely eludes Science. In his thesis, Kunpeng Yao brings some new light to this question.

Through a series of chapters, each reporting on a commensurable and highly commendable research work, Kunpeng takes us through the arduous road of scientific inquiry. He starts with an account on dexterity in crafts. He follows a set of apprentices at watchmaking and documents with high accuracy the difficulties they face to perform the expected fine manipulation task at the highest level of performance. He offers the first to date exposition of how humans search across the set of hand postures and distribute tasks across the fingers. He unfolds a little explored area, as he studies the motion of the fingers of the two hands as a group, when the two hands are tasked to act on the same tool. As a roboticist, Kunpeng Yao takes a mathematical stance to the problem and identifies the key variables and mechanisms at play, offering whenever possible computational implementations to verify his observations and interpretations, thereof.

In the second part of his thesis, Kunpeng Yao sets on the formidable challenge to enable an artificial hand to hold a multitude of objects in-hand, a cornerstone of human dexterity. Steering away from biological resemblance, he draws on engineering techniques for robot planning, machine learning and optimal control, and develops a sequential algorithm to explore the space of possible. In a series of impressive demonstrations, he shows that a robot hand can hold a composite of tiny, large, medium-sized objects in all sorts of configurations.

This thesis is a wonderful example of interdisciplinary research that contributes to both human motor control and robotics. It is a must-read for all scientists who ever wondered whether it was worth the effort to put so many degrees of freedom in a robotic hand.

Lausanne, 10 Jan 2022 Aude Billard

Abstract

From surgery to watchmaking, fine-manipulation skills highly rely on the dexterity afforded by both hands. Coordination is key to human dexterity. Specifically, humans need not only to govern the abundant intrinsic degrees of freedom (DOFs) to allocate controls of task-demanded variables, but also to adapt postures in response to extrinsic task conditions. In spite of the recent advances in robotics research, human dexterity remains unattainable for robots, especially in terms of flexibility and adaptability. Therefore, it is necessary to gain insights into human fine-manipulation skills, to advance the dexterity of robots in similar tasks. This thesis deepens our understanding of human dexterity by investigating the human coordination in fine-manipulation skills taken from watchmaking craftsmanship.

The first part of this thesis investigates both intrinsic and extrinsic coordination of upper-limbs in a non-symmetrical and non-rhythmic bimanual fine-manipulation task. We conduct a comparative study in the assembly of a spring on the watch face. Analysis of motion kinematics reveals that professional subjects mitigate the task challenge by actively modifying task conditions. Moreover, we offer a novel perspective to understand the skill acquisition in humans. We hypothesize that the evolution of coordination during skill improvement is driven by the changes in the structure of the optimal criterion of the central nervous system. We employ a bi-level optimization framework to infer the structure of this optimal criterion.

The second part investigates how roles and control variables are distributed across hands and fingers. We compare task performance of human subjects under two experimental conditions when dismounting a screw from a watch face. When the watch face needed positioning, the role distribution of both hands was strongly influenced by hand dominance; when the watch face was stationary, a variety of hand pose combinations emerged. We propose a taxonomy of bimanual hand pose combinations and develop a graphical matrix-based representation approach to afford analysis of experimental observations. Our analysis suggests that the control of independent task demands is distributed across either hands or functional groups of fingers.

In the third part, we take inspiration from human coordination principle to advance the dexterity of a robotic hand by exploiting its redundant DOFs. We propose a human-like algorithm to enable a robotic hand to grasp objects using arbitrary surface regions, no longer restricted to regular grasp types, such as pinch or power grasp. We present an iterative

process to empower the robotic hand to grasp multiple objects in sequence. Moreover, we formulate a strategy to facilitate the exploitation of redundant DOFs for multitask planning. Our approaches have been validated both in simulation and on a real robotic hand.

In summary, this thesis not only offers a deeper understanding of human dexterity by revealing both the intrinsic and extrinsic coordination of upper-limbs and the role distribution across hands in bimanual fine-manipulation tasks, but also proposes algorithms that enable multi-fingered robotic hands to achieve human-like dexterous grasping of a single or even multiple objects using arbitrary surface regions. This thesis offers the prospect of developing algorithms for robotic dexterous grasping and manipulation, and also inspires the design of novel robotic hands and manipulators.

Keywords: human dexterity, bimanual coordination, human motor control, robotic dexterity, robotic grasping and manipulation

Zusammenfassung

Vom Chirurgen bis zum Uhrmacher - die Geschicklichkeit beider Hände ist bei der Feinmanipulation von großer Bedeutung. Koordination ist der Schlüssel zur menschlichen Geschicklichkeit. Der Mensch muss nicht nur die zahlreichen intrinsischen Freiheitsgrade beherrschen, um die für die Aufgabe erforderlichen Variablen zu steuern, sondern auch die Körperhaltung als Reaktion auf die äußeren Bedingungen der Aufgabe anpassen. Trotz der jüngsten Fortschritte in der Robotikforschung ist die menschliche Geschicklichkeit für Roboter immer noch unerreichbar, insbesondere in Bezug auf Flexibilität und Anpassungsfähigkeit. Daher ist es notwendig, Einblicke in die menschlichen Fähigkeiten zur Feinmanipulation zu gewinnen, um die Geschicklichkeit von Robotern bei ähnlichen Aufgaben zu verbessern. Diese Arbeit vertieft unser Verständnis der menschlichen Geschicklichkeit, indem sie die menschliche Koordination bei der Feinmanipulation in der Uhrmacherkunst untersucht.

Der erste Teil dieser Arbeit untersucht sowohl die intrinsische als auch die extrinsische Koordination der oberen Gliedmaßen bei einer nicht-symmetrischen und nicht-rhythmischen bimanuellen Feinhandhabungsaufgabe. Wir führen eine vergleichende Studie über die Montage einer Feder auf dem Zifferblatt einer Uhr durch. Die Analyse der Bewegungskinematik zeigt, dass professionelle Probanden die Herausforderung der Aufgabe durch aktive Modifikation der Aufgabenbedingungen abmildern. Darüber hinaus bieten wir eine neue Perspektive, um den Erwerb von Fähigkeiten beim Menschen zu verstehen. Wir stellen die Hypothese auf, dass die Entwicklung der Koordination während der Verbesserung der Fähigkeiten durch die Veränderungen in der Struktur der Optimalitätskriterien des zentralen Nervensystems angetrieben wird. Wir verwenden einen zweistufigen Optimierungsrahmen, um die Struktur dieses optimalen Kriteriums zu ermitteln.

Im zweiten Teil wird untersucht, wie die Rollen und Kontrollvariablen auf die Hände und Finger verteilt sind. Wir vergleichen die Leistung menschlicher Probanden unter zwei experimentellen Bedingungen beim Lösen einer Schraube von einem Zifferblatt. Wenn das Zifferblatt positioniert werden musste, wurde die Rollenverteilung beider Hände stark von der Handdominanz beeinflusst; wenn das Zifferblatt unbeweglich war, ergab sich eine Vielzahl von Handposen-Kombinationen. Wir schlagen eine Taxonomie bimanueller Handhaltungskombinationen vor und entwickeln einen grafischen, matrixbasierten Darstellungsansatz, der die Analyse der experimentellen Beobachtungen ermöglicht. Unsere Analyse deutet darauf hin, dass die Kontrolle unabhängiger Aufgabenanforderungen entweder auf die Hände oder

Zusammenfassung

auf funktionale Gruppen von Fingern verteilt ist.

Im dritten Teil lassen wir uns vom menschlichen Koordinationsprinzip inspirieren, um die Geschicklichkeit einer Roboterhand durch Ausnutzung ihrer redundanten Freiheitsgrade zu verbessern. Wir schlagen einen menschenähnlichen Algorithmus vor, der eine Roboterhand in die Lage versetzt, Objekte über beliebige Oberflächenregionen zu greifen und nicht mehr auf reguläre Greifarten wie Zwicken oder kraftvolles Greifen beschränkt ist. Wir stellen einen iterativen Prozess vor, der die Roboterhand befähigt, mehrere Objekte nacheinander zu greifen. Darüber hinaus formulieren wir eine Strategie, um die Nutzung redundanter Freiheitsgrade für die Multitasking-Planung zu erleichtern. Unsere Ansätze wurden sowohl in der Simulation als auch an einer realen Roboterhand validiert.

Zusammenfassend bietet diese Arbeit nicht nur ein tieferes Verständnis der menschlichen Geschicklichkeit, indem sie sowohl die intrinsische als auch die extrinsische Koordination der oberen Gliedmaßen und die Rollenverteilung zwischen den Händen bei bimanuellen Feinmanipulationsaufgaben aufdeckt, sondern auch Algorithmen vorschlägt, die es Roboterhänden ermöglichen, ein einzelnes oder sogar mehrere Objekte mit Hilfe beliebiger Oberflächenregionen der Hand menschenähnlich geschickt zu greifen. Diese Arbeit bietet die Aussicht auf die Entwicklung von Algorithmen für das geschickte Greifen und Manipulieren von Robotern und kann auch das Design neuartiger Roboterhände und Manipulatoren inspirieren.

Schlüsselwörter: menschliche Geschicklichkeit, bimanuelle Koordination, menschliche Motorik, robotische Geschicklichkeit, robotisches Greifen und Manipulieren

Résumé

Qu'il s'agisse de chirurgie ou d'horlogerie, les manipulations fines reposent en grande partie sur la dextérité des deux mains. La coordination est la clé de la dextérité humaine. Plus précisément, les humains doivent non seulement gérer les nombreux degrés de liberté intrinsèques pour allouer les contrôles des variables exigées par la tâche, mais aussi adapter les postures en réponse aux conditions extrinsèques de la tâche. Malgré les progrès récents de la recherche en robotique, la dextérité humaine reste encore inaccessible aux robots, notamment en termes de flexibilité et d'adaptabilité. Il est donc nécessaire de mieux comprendre les capacités de manipulation fine de l'homme, afin de faire progresser la dextérité des robots dans des tâches similaires. Cette thèse approfondit notre compréhension de la dextérité humaine en étudiant la coordination humaine dans les compétences de manipulation fine tirées de l'artisanat horloger.

La première partie de cette thèse étudie la coordination intrinsèque et extrinsèque des membres supérieurs dans une tâche de manipulation fine bimanuelle non symétrique et non rythmique. Nous menons une étude comparative lors de l'assemblage d'un ressort sur le cadran d'une montre. L'analyse de la cinématique du mouvement révèle que les sujets professionnels atténuent le défi de la tâche en modifiant activement les conditions de la tâche. En outre, nous offrons une nouvelle perspective pour comprendre l'acquisition de compétences chez les humains. Nous émettons l'hypothèse que l'évolution de la coordination pendant l'amélioration de l'habileté est déterminée par les changements dans la structure du critère optimal du système nerveux central. Nous utilisons un cadre d'optimisation à deux niveaux pour déduire la structure de ce critère optimal.

La deuxième partie étudie comment les rôles et les variables de contrôle sont distribués entre les mains et les doigts. Nous comparons la performance de sujets humains dans deux conditions expérimentales lors du démontage d'une vis du cadran d'une montre. Lorsque le cadran de la montre devait être positionné, la distribution des rôles des deux mains était fortement influencée par la dominance des mains; lorsque le cadran de la montre était stationnaire, une variété de combinaisons de poses de la main a émergé. Nous proposons une taxonomie des combinaisons de pose de main bimanuelle et développons une approche de représentation graphique basée sur une matrice pour permettre l'analyse des observations expérimentales. Notre analyse suggère que le contrôle des demandes de tâches indépendantes est distribué soit entre les mains, soit entre des groupes fonctionnels de doigts.

Résumé

Dans la troisième partie, nous nous inspirons du principe de coordination humaine pour améliorer la dextérité d'une main robotique en exploitant ses degrés de liberté redondants. Nous proposons un algorithme semblable à celui de l'homme pour permettre à une main robotique de saisir des objets en utilisant des régions de surface arbitraires, sans se limiter aux types de saisie habituels, comme la saisie par pincement ou la saisie par force. Nous présentons un processus itératif pour permettre à la main robotique de saisir plusieurs objets en séquence. De plus, nous formulons une stratégie pour faciliter l'exploitation des degrés de liberté redondants pour la planification multitâche. Nos approches ont été validées à la fois en simulation et sur une main robotique réelle.

En résumé, cette thèse offre non seulement une meilleure compréhension de la dextérité humaine en révélant la coordination intrinsèque et extrinsèque des membres supérieurs et la distribution des rôles entre les mains dans les tâches de manipulation fine bimanuelle, mais elle propose également des algorithmes qui permettent aux mains robotiques multi-doigts d'atteindre une dextérité semblable à celle des humains dans la saisie d'un seul ou même de plusieurs objets en utilisant des régions de surface arbitraires de la main. Cette thèse offre la perspective de développer des algorithmes pour la saisie et la manipulation robotique dextre, et peut également inspirer la conception de nouvelles mains et manipulateurs robotiques.

Mots-clés : dextérité humaine, coordination bimanuelle, contrôle moteur humain, dextérité robotique, préhension et manipulation robotiques.

Contents

Contents

Contents

List of Figures

List of Tables

List of Algorithms

1 Introduction

1.1 Motivation

With the recent advances in robotics research, modern robots can not only replace humans on industrial assembly lines to efficiently accomplish tedious and repetitive manufacturing tasks, but also increasingly participate in human society, such as helping humans with daily chores. While robots exceed human capabilities in both precision and speed of movements, achieving human-like capability to dexterously manipulate objects, and to adapt to unexpected environmental changes in fine-manipulation tasks still remain major challenges (Billard and Kragic (2019)). Therefore, it is essential to gain insights into human fine-manipulation skills, so as to provide inspiration for the development of robotic controllers and learning algorithms.

Human fine-manipulation skills, from everyday tasks such as eating with knife and fork, to professional craftsmanship, such as suturing and watchmaking, require the delicate and harmonious motion of fingers that highly rely on the *dexterity* afforded by both hands (see Fig. 1.1). Having an accurate, yet concise definition of dexterity is not easy, due to the richness of its connotation. The pioneer Soviet neurophysiologist Nikolai Bernstein provided a general definition in his book *Dexterity and Its Development* (Bernstein et al. (1996)): "*dexterity is the ability to solve a motor problem correctly, quickly, rationally, and resourcefully*". More recently, (Wiesendanger (1999)) defined dexterity as "*the essence of purposeful motor behavior and entails all the ingredients of goal- and object-oriented actions, particularly manipulations that have been practiced and acquired*". The essence feature of dexterity that both definitions share in common lies in the ability of "*finding a motor solution for any situation and in any condition*". This fundamental characteristic distinguishes dexterity from "*simple harmony in movements*", as pointed out by (Bernstein et al. (1996)). This feature highlights that in tasks such as manipulation, dexterous motions not only achieve the task objective, but also adapt to environmental conditions.

(a) Polishing[1] (b) Drawing[2] (c) Cutting[3]

(d) Pipetting[4] (e) Peeling[5] (f) Watchmaking[6]
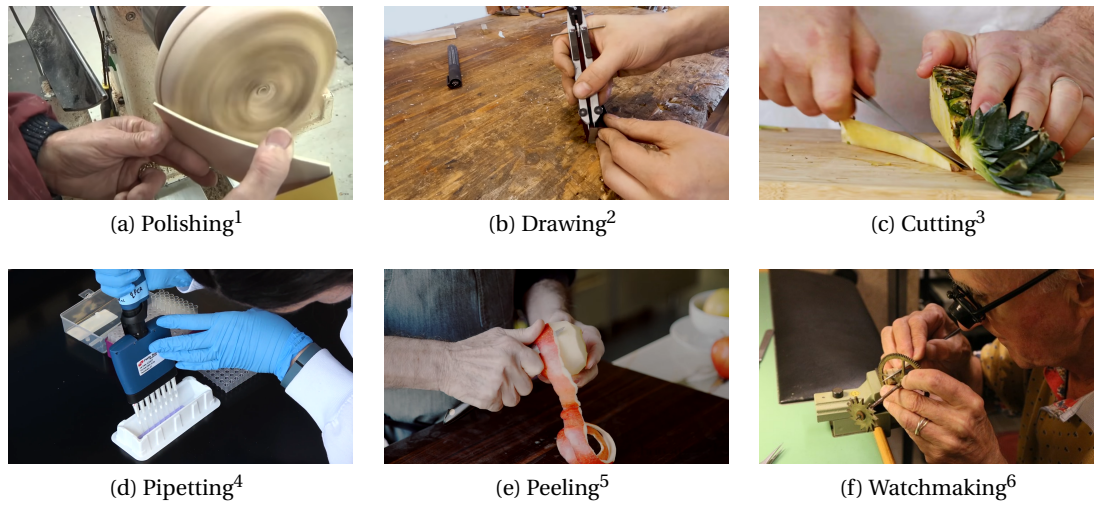
Figure 1.1 – Humans performing dexterous bimanual fine-manipulation tasks.

To generate dexterous motions, the human *central nervous system* (CNS) needs to control the numerous inherent *degrees of freedom* (DOFs) on multiple levels of the human musculoskeletal system. Examples are DOFs that connect linkages to muscles via joints, and the ones that arise due to multiple motor neuron synapses on muscles. Such extreme abundance of DOFs make the human musculoskeletal system highly redundant. For example, each human hand has 27 joint DOFs (excluding the wrist) (Agur and Dalley (2009)). Together with the DOFs of the upper limbs, over sixty DOFs need to be controlled in a bimanual fine-manipulation task. This number exceeds by far the number of DOFs needed to move or to manipulate an object in three-dimensional space in general. Moreover, even for a spatial reaching motion, different velocity and acceleration profiles can lead to almost infinite number of feasible solutions while satisfying the task objectives. This redundancy results in the *degrees of freedom problem* in human motor control (Bernstein (1967)), namely how to reduce this redundancy by constraining the motor system's many degrees of freedom in a way that they work jointly to produce the desired movement (Magill and Anderson (2010)).

Seeking an explanation to the degrees of freedom problem demands the understanding of motor *coordination*, defined by (Bernstein (1967)) as "*the process of mastering redundant degrees of freedom of the moving organ, in other words, its conversion to a controllable system*". Coordination is considered as a means of the CNS to efficiently organize and control the redundant DOFs, resulting in intended movement. Thus, understanding motor coordination is a prerequisite for unraveling the human mechanisms for achieving flexible and adaptive control of movement. (Turvey (1990)) defined coordination as "*the patterning of body and*

---

[1] https://www.youtube.com/watch?v=_SGWNmJYL7M&t=366s&ab_channel=TheWoodWhisperer

[2] https://www.youtube.com/watch?v=u2mWvozSdvE&t=837s&ab_channel=UriTuchman

[3] https://www.youtube.com/watch?v=VjINuQX4hbM&ab_channel=Epicurious

[4] https://www.youtube.com/watch?v=uFzlUT8b_XQ&ab_channel=Addgene

[5] https://www.youtube.com/watch?v=dELMEZXyMuQ&t=58s&ab_channel=BonApp%C3%A9tit

[6] https://www.epfl.ch/labs/lasa/sahr/. Copyright belongs to LASA, EPFL.

limb motions relative to the patterning of environmental objects and events". This commonly acceptable definition points out two important aspects of coordination. First, it states that coordinated motions involve patterns of multiple body parts, achieved by controlling the inherent DOFs of the motor system. We refer to this type of coordination as *intrinsic coordination* in this thesis. Second, it highlights that it is essential to consider the task conditions during motion planning and control. We refer to the relationship between human motor system and environment (i.e. task condition) as *extrinsic coordination.*

Intrinsic coordination patterns are present in the movement of human daily activities and are essential for movement flexibility. For example, the *inter-limb coordination* existing among movements of multiple limbs (Kelso et al. (1979a)), the *intra-limb coordination* organizing DOFs within a limb to move in Cartesian space (Sparrow et al. (1987)), and the *eye-hand coordination* that controls the movements of hands and eyes coordinately through tactile and visual feedback, involved in a typical reach-and-grasp motion (Johansson et al. (2001)).

Extrinsic coordination emphasizes the adaptability of dexterous movements with respect to the environment. In fact, task conditions shape the desire for dexterity. (Bernstein (1967)) stated that "*demand for dexterity is not in the movements themselves, but in the surrounding conditions*". As an example, he compared the skills of "walking on the ground" versus "walking on a rope" to justify that it is the change in task conditions that places higher demands on dexterity.

1.2 Challenges

1.2.1 Coordination of upper limbs in bimanual fine-manipulation tasks

One of the most typical inter-limb coordination, *bimanual coordination,* is essential in motor skills that require the simultaneous motions of both upper limbs (see Fig. 1.1). Manipulating objects using both hands is so common for humans in everyday activities that it has even been argued that "*bimanual coordination is the default mode of the control system, such that unimanual actions require suppression of the contralateral limb*" (Swinnen and Gooijers (2015)).

In bimanual tasks, the role distribution across both upper limbs are either balanced or unbalanced. When balanced, both upper limbs perform under the identical task conditions, usually to achieve the common task objective. For example, lifting a heavy box or catching a flying basketball. Many bimanual tasks in everyday life – in particular, the majority of bimanual fine-manipulation tasks, require the unbalanced role distribution across upper limbs, which is often referred to as the *role-differentiated bimanual manipulation* (RDBM) (Kimmerle et al. (1995)). Such manipulation tasks demand two hands performing different but complementary movements to achieve the common task objective (Babik and Michel (2016)). Examples are suturing, and eating with knife and fork. The inter-limb coordination in RDBM tasks is affected by both the objective and the conditions of performing the task. For instance, to cut

an apple in half, one usually uses the dominant hand to hold the knife, while the other hand assists by holding the apple in place. However, when peeling an apple, the assistant hand keeps adjusting its poses to regularly rotate the apple to free up space for the knife held by the dominant hand.

Maintaining coordinated postures of both upper limbs is crucial to generate complementary bimanual motions. The coordination between two arms has been extensively studied in a variety of scenarios. Related studies generally ignore hand poses and finger movements, and only focus on the spatial movements of arms, especially repetitive motions that have symmetrical, cyclic, or rhythmic natures, such as bilateral symmetrical arm motion (Kelso (1994)), circle drawing (Cattaert et al. (1999)), and rhythmic finger-tapping (Klapp (1979)). Such bimanual coordination patterns are influenced by the relationship of both arms' movements arising from multiple aspects, including timing (Peters (1981)), phase (Haken et al. (1985)), amplitude (Spijkers and Heuer (1995)), and direction (Swinnen et al. (1998)). For example, the bilateral motions during everyday activities manifest trend towards in-phase and anti-phase coordination patterns (Howard et al. (2009)), and it is suggested that towards "same amplitudes" and "isotropic directions" may be the default mode of bilateral motions, so that assimilation effects are less likely to occur (Swinnen and Gooijers (2015)). Although the majority of bimanual tasks seen in everyday human life are neither symmetrical nor cyclical, extant investigations on this motion type are only limited to specific scenarios, such as drawing lines and circles with the left and right arms, respectively (Franz and Ramachandran (1998)). A handful of studies has confirmed the existence of coupling between arms in asymmetrical bilateral movement. For instance, when two hands reach asymmetrically for different target objects, the movement of the hand reaching for the farthest target affects the movement of the other hand, slowing it down to maintain temporal coordination (Bingham et al. (2008)). **What remains unclear, however, is how the two arms coordinate in a general asymmetric and non-rhythmic bimanual task, namely, how to master the intrinsic DOFs of both arms to satisfy task demands while adapting to extrinsic task conditions.** *Task demands* refer to the task variables that must be controlled to achieve the task objective. For instance, the force needed to be applied to press a button, or the torque to be generated to rotate a wrench. Moreover, we use *task conditions* to indicate the remaining parameters that are used to describe the task, such as the position and orientation of the target object.

To address this challenge, in Chapter 2, we conducted experiments to analyze human subjects' movements in a bimanual fine-manipulation task that required asymmetric and non-rhythmic movements of both upper limbs. We formulated kinematic metrics as functions of joint angles to quantitatively analyze the coordination patterns of both arms and hands.

In addition, numerous related studies have revealed that the bimanual coordination patterns progressively tend to stabilize as subjects gain more task-related experience, indicating the improvement of task skills. This process can be revealed by assessing the coordination patterns, typically by analyzing the coordination dynamics of individual joint using phase planes (joint angle-joint angular velocity relationship), or by analyzing the coordination among sev-

eral joints using the angle-angle diagram (plot multiple joint angles against each other) or cross-correlations (Lamb and Bartlett (2017)). These analytical approaches can be very useful in revealing the underlying intrinsic coordination patterns exhibited in repetitive biman-ual motions, especially cyclic or rhythmic motions. **However, such analysis approaches do not reveal extrinsic coordination patterns. In other words, whether task conditions affect coordination patterns, and, whether the effects of task conditions are related to subjects' proficiency, cannot be revealed by applying existing approaches.** This is mainly because these approaches only consider the variables in human motor system in the analysis, but exclude the variables in task conditions. As a result, the interaction between the human motor system and task conditions is ignored.

Chapter 2 provides a solution to this problem by considering task condition variables in coordination metrics. Thus, these metrics enable the investigation of both intrinsic and extrinsic coordination patterns. To reveal the effects of task proficiency on coordination, we compared the analysis results of human subjects at different skill levels. In addition, we formulated a computational framework to infer the human subject's intention of generating such coordination patterns. This helps to disclose the considerations of subjects and reveals, for example, whether the joints are coordinated to satisfy the task demands or simply to maintain a comfortable posture to avoid fatigue.

1.2.2   Coordination of human hands and bilateral control

In fine-manipulation tasks, control of hands and, in particular, control of individual fingers are more challenging compared to the control of arms, due to the abundant DOFs in both hands. A variety of studies have shown that the selection of hand poses and control of fingers are highly task-specific. For example, the hand poses humans use to grasp an object for moving it may be different from the hand poses they employ for using it (Valyear et al. (2011)). In a sequential task consisting of multiple steps, humans also consider the subsequent actions when determining the grasping hand poses (Armbrüster and Spijkers (2006)). Moreover, physical properties of the target object, such as size (Tucker and Ellis (2004)), shape (Gentilucci (2002)), and orientation (Tucker and Ellis (1998)) also affect the shaping of hand pose.

In bimanual fine-manipulation tasks, the inter-limb coordination mechanism across hands is determined by multiple factors. For example, (Rosenbaum et al. (2006)) studied the influence of haptic feedback in an asymmetrical bimanual task, and concluded that haptic tracking enables more independence of limbs. (Hu and Newell (2011)) investigated the force coordina-tion of both index fingers in a symmetrical bimanual task, and hypothesized that the force coordination patterns are generated to minimize muscle force output. (Bhullar et al. (2015); Kang and Cauraugh (2018)) discussed the influence of visual information, and reported that the improvement in bimanual coordination is positively correlated to the improvement of visual information gain. In particular, handedness (Guiard (1987)) strongly affect the role distributions across hands in bimanual fine-manipulation tasks. The dominant hand usually

(a) UTAH-MIT hand[7] (b) DLR hand[8] (c) Shadow hand[9] (d) DLR-II-HIT hand[10] (e) Biomimetic hand[11]
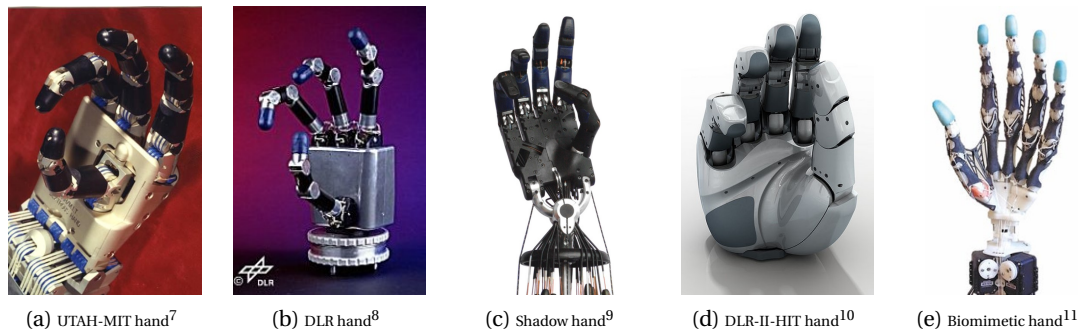
Figure 1.2 – Examples of dexterous robotic hands.

performs the part of the task that demands the highest dexterity and efficiency, while the assistant hand plays an auxiliary role, such as stabilizing the manipulation target (Kimmerle et al. (2010); Gonzalez and Nelson (2015)). Hypotheses have been proposed to explain the effects of handedness on role distribution across hands. The *dynamic dominance hypothesis* (Sainburg (2002)) states that the dominant limb is controlled based on feed-forward mechanism to anticipate the dynamics of the task, while the non-dominant limb is reactive and relies on sensory feedback to achieve higher positional accuracy. (Johansson et al. (2006)) provides a different view of role distribution across hands by stating that the role assignment is neither solely determined by handedness nor is it fixed. Instead, the brain assigns roles to each hand based on the spatial relationships between the forces that each hand needs to generate to achieve the task goal. This mechanism helps to avoid re-grasping or reorienting the grasped object when the extrinsic task conditions change. **Nevertheless, the effects of task demands and task conditions on the shaping of both hand poses remain poorly understood. Do humans select different hand poses to achieve the same task objective under different task conditions (extrinsic coordination)? How do humans distribute roles across hands and fingers (intrinsic inter-limb coordination) to control task demands - especially in the presence of multiple task demands ?**

This thesis provides answers to these question in Chapter 3. We designed experiments and recruited human subjects to perform the same bimanual fine-manipulation task under two different task conditions. We summarized the poses of both hands used by subjects in each condition, and categorized them according to their functionality and finger activity. Our analysis not only illustrates how humans distribute roles across hands to satisfy task demands, but also reveals how task conditions affect the selection of hand poses.

1.2.3 Advancing robot dexterity in grasping

---

[7] https://archive.computerhistory.org/resources/still-image/Robots/102693567.03.01.lg.jpg

[8] https://www.dlr.de/rm/Portaldata/52/Resources/Roboter_und_Systeme/Hand/Hand_I/hand_I_200.jpg

[9] https://robots.ieee.org/robots/shadow/Photos/SD/shadow-photo1-full.jpg

[10] https://www.dlr.de/rm/Portaldata/52/Resources/Roboter_und_Systeme/Hand/HIT_II/HIT_Hand_II_444x600.jpg

[11] (Xu and Todorov (2016))

(a) DLR Hand II, power grasp[12]   (b) DLR Hand II, pinch grasp[13]   (c) Shadow dexterous hand, pinch grasps[14]   (d) Allegro hand, pinch grasps[15]
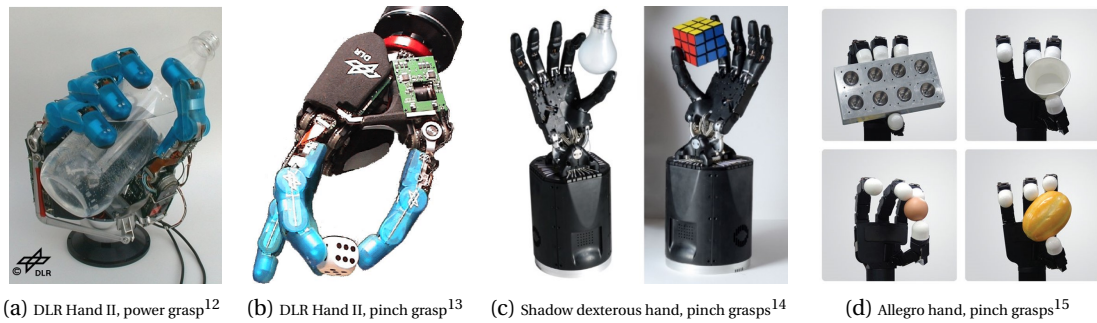
Figure 1.3 – Examples of grasps performed by advanced dexterous robotic hands.

Studies on human hand often inspire the advancement of robotic hands. On the one hand, anatomical and biomechanics studies of the human hand have facilitated the design of anthropomorphic robotic hands. From the early UTAH-MIT hand (Fig. 1.2a, (Jacobsen et al. (1986))), the Shadow hand (Fig. 1.2c, (Kochan (2005))), to the recent DLR-II-HIT hand (Fig. 1.2d, (Liu et al. (2008))), and anthropomorphic hand (Fig. 1.2e, (Xu and Todorov (2016))) not only do they increasingly resemble the human hand in appearance, they are also getting closer to the human hand in functionality. Their design took inspiration from the physiological structure of the human hand, and their performance benefits from the rapid advances in materials, actuators, sensors, and control electronics (Ritter and Haschke (2015)). For example, improvements in electric motors have enabled the miniaturized actuators to accommodate smaller fingers, while providing larger grip forces. The breakthroughs in sensing technology have opened up the possibility for robotic hands to obtain kinesthetic and cutaneous feedback, which can be used to compensate for model inaccuracy and hence to achieve precise control of finger movements.

On the other hand, studies in human movement science and neuroscience have inspired the development of novel control algorithms for robotic hands. For example, the theoretical framework of *synergy* has not only provided an explanation to the human CNS's mechanism of controlling finger movements, but also let to the proposition of synergistic control of multi-DOF robotic hands (Prattichizzo et al. (2010)) and grasping in particular (Ciocarlie and Allen (2009)).

In spite of the advancements in the mechanical structure and control algorithms, the level of dexterity manifested by even the most advanced robotic hand is still incomparable to that of a real human hand. In addition, most of these multi-fingered robotic hands are tasked to grasp a single object at a time, mainly using either *power grasp* to enwrap the object, or *pinch grasp* that employs only fingertips for contact (see Fig. 1.3).

---

[12] https://www.dlr.de/rm/Portaldata/52/Resources/Roboter_und_Systeme/Hand/Hand_II/Hand-II-01.jpg
[13] (Borst et al. (2003))
[14] https://www.morfey.co.uk/shadow-dextrus-hand
[15] http://wiki.wonikrobotics.com/AllegroHandWiki/index.php/File:AH_4pics.jpg
[16] Copyright belongs to LASA, EPFL. Photos made by Sahar El-Khoury, Miao Li, Ravin de Souza.
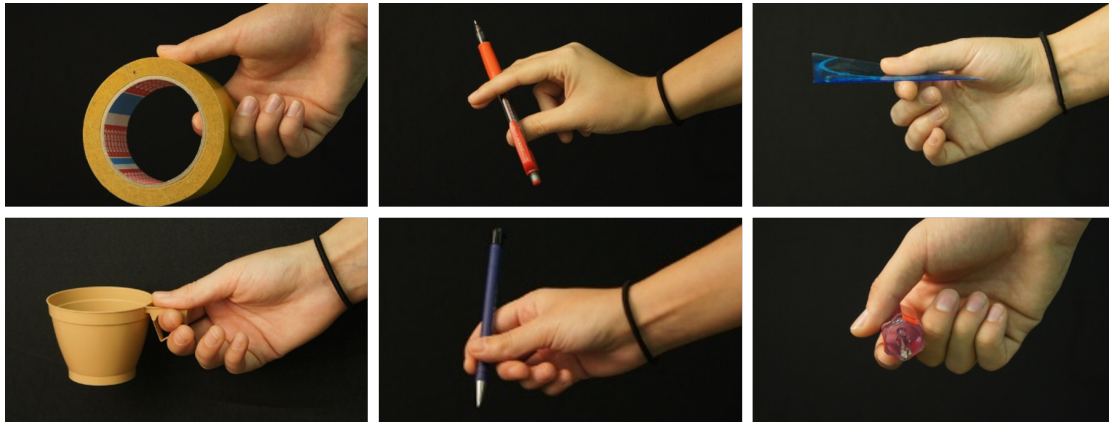
Figure 1.4 – Examples of human dexterous grasping hand poses[16].

In contrast, the human hand, as "*a tool of tools*" (Aristotle, *De Anima*), is able to dexterously grasp the vast majority of objects seen in everyday life nearly effortlessly. Humans modulate hand poses to adapt to extrinsic task condition, such as the shape of objects and task objective. To generate a suitable hand pose, all surface regions of the hand are involved – not only the commonly used fingertips and inner surfaces of the hand, but also the lateral surfaces of the fingers (see Fig. 1.4). For example, holding a pencil to write employs the side of the middle fingertip; manipulating a pair of chopsticks demands the coordination among multiple side surfaces on several fingers. In addition, humans are capable of coordinating the abundant intrinsic DOFs to accomplish multitasking (see Fig. 1.5), such as holding different cutlery when clearing up dishes.

Yet, it remains a challenging problem to control a multi-fingered robotic hand to achieve such human-like dexterity. In particular, (1) how to extend the use of fingertips to the use of entire and arbitrary regions of the hand, and (2) how to exploit the abundant DOFs in hand structure for multitasking, such as grasping multiple objects.

To answer the first question, we analyzed the kinematic model of the hand to determine the reachable space of various hand parts (e.g., finger phalanges and the palm). Groups of hand surface regions that geometrically afford stable contacts with the target object can serve as candidates for synthesizing grasps.

To investigate the latter problem, we developed an iterative process to exploit the kinematic redundancy in the hand model to plan grasps of more objects. This work is described in Chapter 4.

---

[17] https://www.koreaboo.com/stories/red-velvet-irene-silenced-fans-shh/
[18] https://pbs.twimg.com/media/EmQxOByXcAAyJAs.jpg
[19] https://pbs.twimg.com/media/EmQxWuYWEAMQXnW.jpg
[20] https://p.ssl.qhimg.com/t014c42a30dcaf37ce3.jpg?size=1555x972
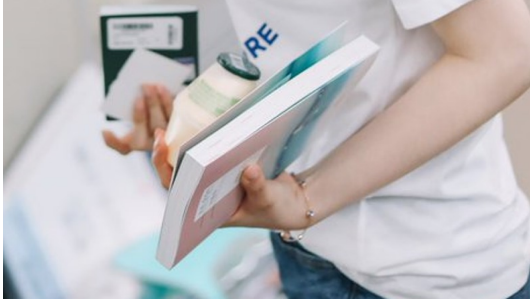
(a) Grasp a book and a small bottle by one hand[17]
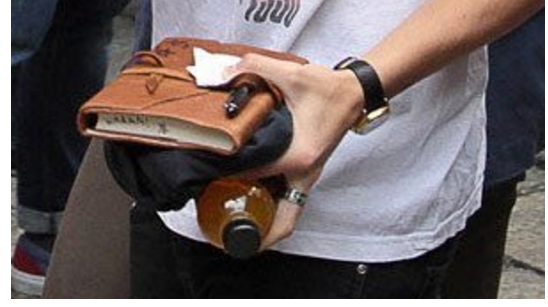
(b) Grasp a book, a bag, and a bottle by one hand[18]

(c) Grasp a bottle and an apple by one hand[19]

(d) Grasp multiple groups of poker cards by one hand[20]

Figure 1.5 – Examples of human hands grasping multiple objects.

1.3 Approaches

This section details the studies and approaches taken in this thesis. To investigate human coordination in fine-manipulation tasks, we design experiments and conduct human studies in watchmaking craftsmanship (Sec. 1.3.1). We analyze the intrinsic and extrinsic coordination in upper limbs from recorded human posture. To understand human coordination at different skill levels, we infer the optimal criterion of the CNS by following the optimal control theoretical framework (Sec. 1.3.2). Then, we focus on the role distribution and task assignment across hands. We analyze the functional grouping of fingers and construct hand pose taxonomies (Sec. 1.3.3). Inspired by the human finger control principles, we develop algorithms to advance the dexterity of robotics hand in grasping tasks. We formulate grasp synthesis as a constrained optimization problem (Sec. 1.3.4).

1.3.1 Study of watchmaking craftsmanship

Understanding human coordination in fine-manipulation tasks helps to answer the challenges regarding human motor control as introduced in the previous section.

We study human bimanual fine-manipulation skills in watchmaking craftsmanship. Manipulation tasks in watchmaking are challenging. It takes years of training for a watchmaking apprentice to master the skill and become a professional watchmaker. In particular, we focus

on two typical fine-manipulation tasks in watchmaking: the assembly of a U-shaped watch spring, and the disassembly of a watch screw.

In the first task, subjects needed to manipulate different tools with each hand individually, in order to assemble a watch spring (see Chapter 2). This allows us to address challenges explained in Sec. 1.2.1, and to shed light on both the intrinsic and extrinsic coordination patterns of upper limbs in general asymmetrical and non-rhythmic motions that remain poorly investigated. Moreover, this task is redundant, in the sense that multiple free variables exist in task conditions. Such variables do not need to be controlled to achieve the task goals; however, they may affect the difficulty of the task. This allowed us to analyze whether human subjects would control these redundant variables and, if so, how they would coordinate the motor systems to achieve this control, i.e., the extrinsic coordination. Furthermore, in this study, we recruited both novice and expert subjects. The between-group comparisons allowed us to observe the effects of task proficiency on coordination patterns.

The second task was to remove a screw of a watch, which also required subjects to perform role-differentiated bimanual fine-manipulation (see Chapter 3). In this task, our main objective was to observe the hand pose selection strategies in different task conditions, so as to tackle the challenges introduced in Sec. 1.2.2, hence, to deepen our understanding of how task demands and task conditions affect the shaping of both hand poses. Compared to the first task, disassembling a watch screw required manipulating of only one tool, thus imposing fewer constraints on the task and potentially allowing for more feasible ways to perform the task. We artificially modified the task conditions to reduce the number of degrees of freedom in the task that needed to be controlled, and contrasted the task performance to the unmodified task condition, in which all variables needed to be controlled. In this way, we were able to understand how task conditions affect the role distribution of hands, as well as the assignment of control variables across the hands and fingers. We systematically analyzed the various hand poses adopted by subjects in each experimental condition to determine whether subjects would use different hand poses if the degrees of freedom in task conditions change. We observed that the role distribution across hands was strongly influenced by hand dominance when there was a high level of redundancy in the task. In contrast, when fewer degrees of freedom in the task needed to be controlled, various combinations of hand poses emerged.

1.3.2 Motor coordination — an optimal control perspective

Although the motor abundance is considered a "*bliss*" (Latash (2012); Flash and Hochner (2005)), how the CNS selects a subset of DOFs as control variables for a specific task, and hence generates a particular solution out of numerous feasible solutions remains an open question.

Many theoretical frameworks and hypotheses have been proposed to explain the motor abundance. For example, the Fitt's law (Fitts (1954)) states that the human motor system will choose to decrease the speed of movement in order to improve its accuracy, and vice versa. The concept of muscle synergy, initially proposed by (Bernstein (1967)), hypothesizes that the

abundant DOFs in human motor system are controlled in a lower-dimensional space through *synergies*. Studies on human and animal movements provide evidence for the existence of synergy in muscle activation (Tresch et al. (1999); d'Avella and Bizzi (2005)) and joint angles (Prevete et al. (2018)). Although synergy explains well the organization of muscles and joints in movements, it is difficult to infer which synergies are common across multiple tasks and which are task-specific. An alternative explanation is the uncontrolled manifold hypothesis (UCM, (Scholz and Schöner (1999))), which states that the brain controls a subset (manifold) of controllable DOFs to stabilize important performance variables, while allowing for variance in task-irrelevant dimensions. In addition, the TNC-Cost (Tolerance, Noise, and Covariation) analysis (Cohen and Sternad (2009)) provides the feasibility for quantitatively analyzing how humans exploit motor abundance to minimize the influence of intrinsic motor noise on the final movement. Nevertheless, these concepts and hypotheses can only be used to analyze the observed motion data and hence, to interpret the motion principles (i.e., the analysis problem), but can not be applied to generate motions (i.e., the synthesis problem).

This thesis interprets the generation of human coordination patterns from an optimal control perspective. Optimal control theory has been applied as a paradigm to understand motor control (Scott (2004); Todorov (2004)) and also to synthesis human motion. It hypothesizes that the CNS generates motor control for a specific *optimal criterion*. This optimal criterion is considered to be task-dependent, and consists of *costs* from multiple aspects. For example, energy expenditure and joint jerk are often considered in studies related to human arm reaching motion (Nelson (1983)). It is noteworthy that there is no standard formulation of this optimal criterion. It may be a complex cost function, or a combination of multiple costs, or may even be too complex to be expressed in closed form. Nevertheless, to investigate the effects of factors of interest, the cost function is usually assumed to be a weighted linear combination of multiple elementary costs (Berret et al. (2008); Mombaur et al. (2010); Albrecht et al. (2012)). These elementary costs are generally selected from a set of artificially predefined metrics, which may be related to task performance or physical constraints. In such a formulation, the proportion (i.e. weight coefficient) of each elementary cost represents its importance in the cost function for motion generation. Therefore, inferring the composition structure of this cost function (i.e., optimal criteria) becomes the pivotal question to understanding human motion principles.

This problem is referred to as the *inverse optimal control* problem. It has been formulated to study human motor control, such as locomotion (Mombaur et al. (2010)) and arm reaching motions (Berret et al. (2008, 2011); Mainprice et al. (2015)). A typical computational model of this problem has a bi-level optimization formulation, consisting of an *upper-level problem* and a *lower-level problem*. The upper-level problem aims at inferring the composition structure of the optimal criterion by estimating the weighting coefficients of each elementary cost. The lower-level problem is a direct optimization problem (e.g., optimal control) that generates motion trajectories under the optimal criterion.

In this thesis, we tailor the formulation of this bi-level computational model to better fit

our study. First, we focus on the optimization process underlying the generation of upper-limb postures (i.e., status) rather than optimizing the motion trajectories (i.e., time-series). Experimental observations revealed that once subjects adjusted their upper limbs in proper postures, they only manipulate the tool at millimeter level, while maintaining their upper limbs almost stationary. Thus, it is the posture rather than the trajectory of the upper limbs that has the greatest impact on task performance. Therefore, we formulated our problem as an inverse optimization problem that relies on solving a bi-level optimization of upper-limb postures. Second, unlike most extant studies that consider only the variables of the biological system (joint angles, joint velocities, etc.), we consider variables in task conditions and introduce task-related metric into the formulation of the optimal criteria. This enables us to consider both the intrinsic and extrinsic coordination of the investigated motion.

1.3.3 The virtual finger concept and hand pose taxonomy

To understand the role distribution across hands and the control of fingers, it is necessary to analyze the hand poses in terms of the placement and functionality of fingers, and relate them to the task demands and conditions.

The concept of synergy has been used to explain the human control of fingers and shaping of hand (Santello et al. (1998)). It assumes that the abundant DOFs in hands (e.g. joints or muscle activation) are controlled in a lower-dimensional synergy space. The intended hand poses for grasping or manipulating objects are generated by combining and regulating synergies that may exist at multiple levels, e.g. postural synergies (Santello et al. (1998)) or force synergies (Santello and Soechting (2000)). The theoretical underpinning behind synergy is the biomechanical coupling in human hands. The movement of the finger joint is controlled by tendons and bounded by soft tissues that cross multiple joints within each finger. Individual movements of fingers are impeded, as they share the same group of muscles and tendons (Li et al. (1998); Zatsiorsky et al. (2000)). Although analyzing synergies can discover the control principles commonly shared by multiple hand poses, it does not reveal the effects of task conditions (Ekvall and Kragic (2005)). Hence, synergy-based approaches may fail to discriminate different grasping intentions for the same hand pose (De Souza et al. (2015)).

To account for the effects of task-related factors, hand pose taxonomies have been constructed to categorize hand poses employed in different scenarios. Criteria for categorizing are multifaceted, such as the functionality (precision grasp or power grasp), and the relationship among fingers (relative motion of the thumb with respect to other fingers). Moreover, taxonomy is advantageous in the sense that it is robust against the variances across task trials and differences among individuals, hence clearly reveals the functionality of each hand pose and also the combination of hand poses at a glance. Several hand pose taxonomies have been offered to summarize human hand poses seen in everyday life. For example, grasp taxonomies (Kamakura et al. (1980); Cutkosky et al. (1989); Feix et al. (2015)) and manipulation taxonomies (Bloomfield et al. (2003); Bullock and Dollar (2011)). However, existing taxonomies

are restricted to the poses of a single hand and can hardly be applied for bimanual tasks. The main reason is that when the task requires both hands to coordinate to achieve a common objective, the shaping of each hand pose cannot be considered in isolation.

In this thesis, we focus on understanding the functionality of fingers and the role distribution of hands in bimanual fine-manipulation tasks. The concept of *virtual finger*, proposed by (Arbib et al. (1985); Iberall (1986)), interprets the human's control of fingers from a functional point of view. One or more real fingers (or part of the hand) that share the same motion and force application are grouped as one virtual finger. It is suggested that the CNS sends commands to control virtual fingers instead of real fingers; and each virtual finger aims at achieving one independent function, such as applying force in a desired direction. For example, to grasp a bottle of water, whether one or multiple fingers are placed on the opposite side of the thumb, they are considered as one virtual finger, because they all share the same functionality in the task – to apply forces against the thumb.

We use the concept of *virtual finger* to analyze the functional roles of fingers in observed hand poses, and then categorize hand poses according to the number of active real fingers and the number of virtual fingers. We also summarize the combinations of both hand poses in a functional bimanual hand pose taxonomy. It reveals the task demands controlled by each hand individually, and hence the role distribution across both hands in the task.

1.3.4 Grasping synthesis — a constrained optimization approach

Grasping is a primitive but central skill for both humans and robots to get in touch with the surrounding environment (Ritter and Haschke (2015)). The problem of determining a configuration for the hand-object system that satisfies certain criteria is referred to as *grasp synthesis* (Shimoga (1996)). In all cases, the hand must be considered in relation to the task demands (e.g., grasping stability) and task conditions (object shapes, contact models, etc.), making grasping a challenging problem.

A feasible grasp can be generated using either *analytic approaches*, relying on solving a constrained optimization problem over grasp quality metrics, or *empirical approaches* (i.e. data-driven approaches) by sampling numerous feasible configurations and selecting the optimal ones (Bohg et al. (2013)). Although empirical approaches do not require accurate model representations and afford directly sampling on real robotic hand, they demand large amount of samples in general. Moreover, they are essentially *object-centered* approaches (Goldfeder and Allen (2011)) and can hardly be applied to generate grasps of multiple objects.

In this thesis, we use the analytic approach to enable the robotic hand to grasp single and multiple objects. Specifically, we formulate the grasping synthesis as a constrained optimization problem (El-Khoury et al. (2013, 2015)), aiming at determining a configuration of the hand-object system, such that the grasp is both collision-free and stable. Collision-free requirements have been formulated as nonlinear inequality constraints. The stability require-

ment guarantees the existence of force-closure, and has been integrated into the problem as nonlinear equality constraints. Grasp quality metrics (Roa and Suárez (2015)) are commonly used as objective functions to enhance desired grasping property of the configuration.

1.4   Contributions

The main goal of this thesis is to deepen our understanding of human dexterity in bimanual fine-manipulation tasks, and also to take inspiration from human motion principles to advance robot dexterity. This thesis addresses the challenges presented in Sec. 1.2 by contributing to both the research field of human movement science and robotics.

- **Intrinsic and extrinsic coordination of upper limbs**
  The coordination mechanism of the arms and hands in asymmetrical and non-rhythmic motion remains poorly understood in the state-of-the-art research, and most extant methods cannot not quantify the effects of task conditions and subject's proficiency on motor coordination. We provide a deeper insight into both intrinsic and extrinsic coordination patterns of upper limbs in this scenario by studying a typical bimanual fine-manipulation task, taken from watchmaking craftsmanship. We formulate metrics to assess the kinematic patterns of hands and arms during manipulation. These metrics are functions not only of joint angles but also of task condition variables, enabling the study of the interaction between the human motor system and external task conditions. In addition, we compare the coordination patterns of novice and expert subjects, allowing us to understand the effects of task proficiency on motor coordination.

- **Understanding human motor learning**
  Extant theories can hardly reveal how multiple factors influencing the acquisition of motor skills. We provide a novel perspective to understand motor learning process from the optimal control perspective. We consider the generation of coordination patterns as an optimization process in the human central nervous system that minimizes costs in certain aspects. To understand the cognitive process behind coordination, we exploit an inverse optimization computational framework to infer the optimal criteria that governs the motion generation process. We formulate task-related metrics in the computational framework, enabling us to quantify the influence of external task conditions. In addition, we compare the inferred optimal criteria between novice and expert groups to investigate the effects of task proficiency. On this basis, we provide a novel interpretation of motor learning process, which considers learning as a modification of the compositional structure of the optimal criteria in the central nervous system.

- **Human hand pose selection strategies**
  We investigate human hand pose selection strategies and disclose the principles of role-distribution across hands and fingers in a bimanual fine-manipulation task. In particular, we examine the effects of task conditions on the selection of hand poses.

(a) Human hand pinches a ball using the thumb and the index finger

(b) Human hand wraps the object with the middle finger and the palm

(c) Human hand grasps multiple objects (example 1)

(d) Human hand grasps multiple objects (example 2)

(e) Robotic hand imitates human pinch grasp

(f) Robotic hand imitates human wrap grasp

(g) Robotic hand imitates human grasping of multiple objects (example 1)

(h) Robotic hand imitates human grasping of multiple objects (example 2)

Figure 1.6 – Multi-fingered robotic hand performing human-like dexterous grasps of single and multiple objects by applying our proposed algorithms.

We exploit the concept of virtual finger to analyze the finger positioning, and then relate the grouping of fingers to the control of task demands under manually designed task conditions. We propose a bimanual hand pose taxonomy to categorize observed hand pose combinations, and also offer a matrix-based graphical representation to summarize and analyze observations from a batch of experimental trials.

- **Robotic dexterous grasping**
  To advance the dexterity of robotic hands, we provide a framework for robotic hands to achieve human-like dexterous grasping of single and multiple objects using arbitrary surface regions, inspired by the human control of fingers for controlling multiple task demands. We first propose a grasp synthesis algorithm that generates stable grasps using pairwise contacts on arbitrary opposing surface regions, no longer limited to the fingertips or the inner surface of the hand. On this basis, we formulate an iterative process that enables the robotic hand to grasp multiple objects in sequence. In addition, we propose a kinematic efficiency metric and an associated strategy to facilitate the exploitation of the kinematic redundancy in the hand model (see Fig. 1.6).

1.5   Thesis outline

Here, we overview the structure of this thesis and list the published peer-reviewed journal articles from the research output of this thesis. The main contributions of each chapter and the relationship among chapters are illustrated in Fig. 1.7.

Chapter 2 and Chapter 3 contribute to the research field of human movement science. In Chapter 2, we investigate the coordination of upper-limbs in a bimanual fine-manipulation task — the assembly of a watch spring. We record kinematic information of human subjects and tactile data. Then, we formulate metrics and evaluate the coordination of both arms and hands in subjects of different skill levels. In addition, we infer the structure of the optimal criteria of human central nervous system by solving an inverse optimization problem. On this basis, we propose our hypothesis and interpretation of motor learning. This study has been published in (Yao and Billard (2020)).

We then shift our focus from the general upper limbs to the hands and fingers in Chapter 3. We conduct a comparative study of an unscrewing tasks under two artificially designed experimental conditions. We categorize all observed hand poses in a bimanual functional hand pose taxonomy by analyzing the assignment of virtual fingers and active fingers. We then analyze the role distribution across hands under different task conditions. This study has been published in (Yao et al. (2021)).

With inspirations taken from our human studies, Chapter 4 contributes to the robotics field. In this chapter, we first present an algorithm for a robotic hand to achieve human-like dexterous grasping using arbitrary opposing surface regions of the hand. Then, we propose an iterative process to exploit redundancy in the kinematic structure of the hand for planning multiple tasks in sequence. In addition, we formulate a metric to assess the kinematic redundancy of the model, and hence an associated algorithm to facilitate the exploitation of kinematic redundancy. At the time of writing, the work presented in this chapter has been submitted to a robotics journal.

Finally, in Chapter 5, we summarize the main contributions of this thesis, discuss current limitations, and point out potential directions for the future work.
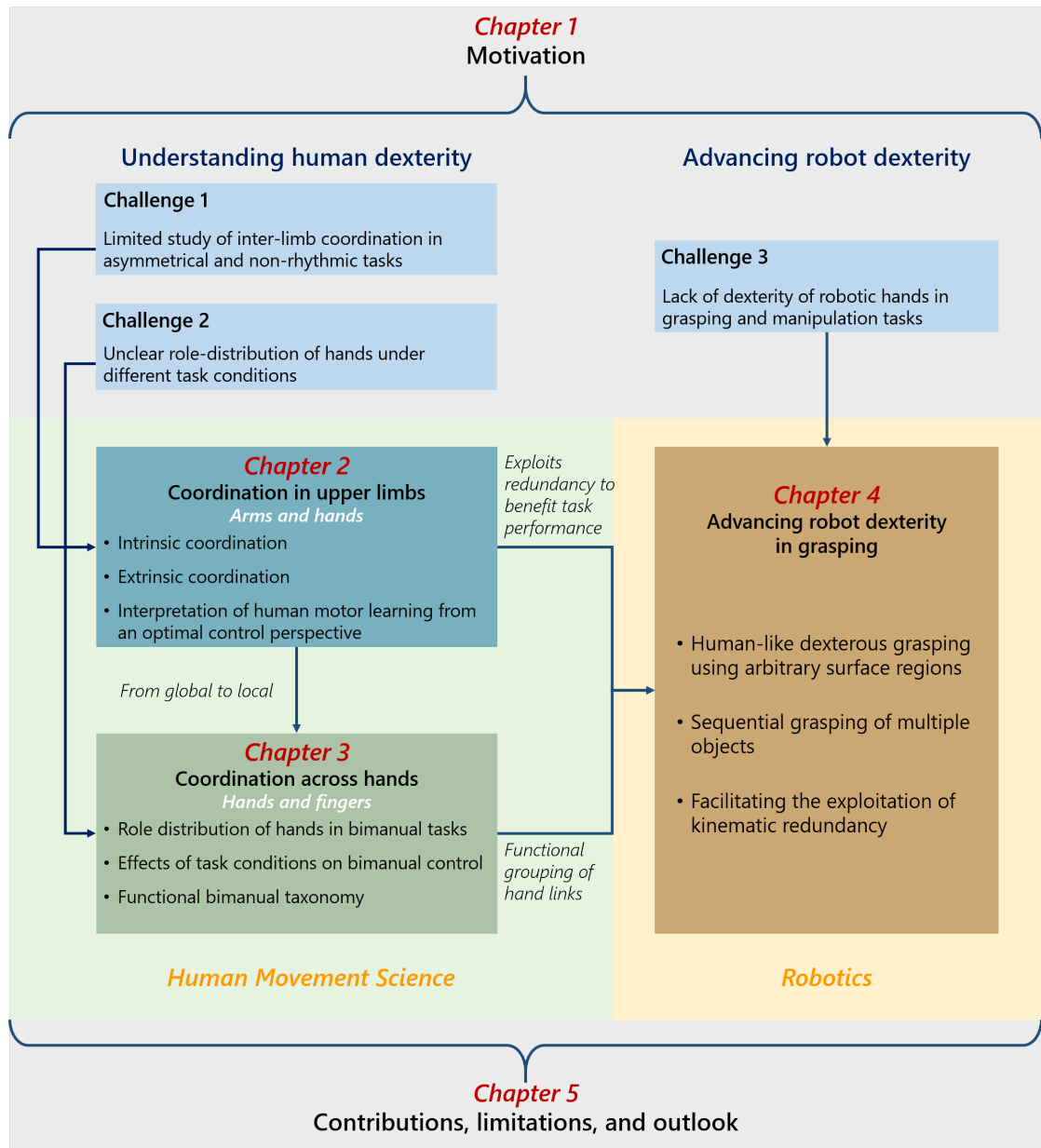
Figure 1.7 – Thesis outline. Chapter 1 states the motivation for the research work and introduces the main challenges to be addressed, along with a brief overview of the approaches taken in this thesis. Chapter 2 and Chapter 3 address the challenges in understanding human dexterity, contributing to the research field of human movement science. Chapter 2 focuses on the coordination patterns in the upper limbs, including both arms and hands, while Chapter 3 concentrates on the role distribution across both hands and among fingers in bimanual fine-manipulation tasks. Inspired by the conclusions of our human studies, Chapter 4 contributes to robotics research by proposing robotic grasping synthesis algorithms to address the challenge of advancing robot dexterity. Chapter 5 summarizes the main contributions of this thesis, discusses current limitations, and points out potential directions for future research.

2 Understanding human acquisition of kinematic coordination patterns

In bimanual fine-manipulation tasks, inter-limb coordination facilitates synchronized spatial and temporal movements of the upper limbs and also allows humans to achieve higher levels of precision that are difficult to achieve with one hand. In this chapter, we start our investigation on human coordination in a bimanual fine-manipulation task taken from watchmaking – the assembly of a watch spring.

Ten novices and five experts participated in the study. We record kinematics of upper limbs and forces applied on the watch face. We formulate kinematic metrics to evaluate the configurations of both arms and hands for each subject. Statistical analysis of these metrics enable a quantitative evaluation of the coordination across subjects in multiple aspects, such as kinematic singularity, task compatibility, and biological discomfort. We also contrast postures of novice apprentices to those of professionals to identify the main distinction in their coordination patterns. Our results indicate that expert subjects wisely place their fingers on the tools to achieve higher manipulation dexterity. Compared to novices, experts also actively align the direction of task-demanded force application with the optimal force transmission direction of their dominant arm.

Second, to understand the cognitive processes underpinning the coordination across subjects at different skill levels, we follow the optimal control theoretical framework and hypothesize that the difference in coordination patterns is caused by changes in the central nervous system's optimal criteria. We formulate kinematic metrics to assess the characteristics of coordination patterns and exploit an inverse optimization approach to infer the structure of the optimal criteria. We interpret human acquisition of novel coordination patterns as an alteration in the composition structure of the central nervous system's optimal criteria, accompanied by the learning process.

The work presented in this chapter has been published in *Yao K, Billard A. An inverse optimization approach to understand human acquisition of kinematic coordination in bimanual fine-manipulation tasks. Biological Cybernetics. 2020 Feb;114(1):63-82.*

2.1 Introduction

Bimanual coordination is central to humans' daily activities and to most craftsmanship. Tasks such as lacing shoes and knitting fabric can hardly be accomplished using one single hand. Human hands and arms are endowed with more than thirty degrees of freedom (DOFs). This is far more than is required when controlling end-point motion in a 6-DOF space. Our upper limbs are hence highly redundant motor systems. Yet, humans display very consistent kinematic patterns when performing the same task, seemingly making little use of this redundancy (Morasso (1983)). A wealth of evidence speaks in favor of the hypothesis that the central nervous system (CNS) overcomes the inherent motor redundancy and masters excessive DOFs through a synergistic coordination of muscles and joints (Bernstein (1967)), resulting in stereotyped movements (Flash (1990)). Each task, however, may require a different synergistic coordination. The process by which these coordination patterns are acquired remains poorly known, especially in bimanual tasks.

2.1.1 Coordination patterns in upper limb motions

Coordination patterns in upper limb motions have been intensively investigated. Studies on arm-hand coordination are primarily within the context of reach and grasp movement (Hoff and Arbib (1993)). These studies revealed phase relationships between hand aperture/closure and arm movements, indicative of temporal synergies between arm and hand. When the reach and grasp movement is perturbed, e.g. when the target is moved during the movement, it was shown that these phases were modulated by the change in target position, and are hence anchored on the target motion (Jeannerod et al. (1995)). (Vainio and Tiainen (2018)) discovered a systematic interaction in proximal and distal prehensile components between two hands, which may be related to the motor preparation of arm extension and grasping motion.

A large amount of studies focused on the inter-limb coordination patterns in symmetric (Kelso (1994)) and rhythmic arm motions, such as drawing circles in symmetric and asymmetric manners (Semjen et al. (1995); Cattaert et al. (1999); Franz and Ramachandran (1998)). Experimental results suggest that the bilateral motions of limbs during everyday activities manifest trend towards in-phase (homologous muscle groups activate simultaneously) and anti-phase (homologous muscle groups activate in alternation) coordination patterns (Howard et al. (2009)), and upper limb motions are more stable if their frequency ratio is an integer instead of non-integer (Summers et al. (1993)). It is also suggested that same amplitudes and isotropic directions may be the default mode of bilateral motions, and assimilation effects are likely to occur otherwise (Swinnen and Gooijers (2015)).

A few works have focused on the coordination patterns underlying asymmetric and non-cyclic motions, such as asymmetric reaching motion towards distinct targets (Kelso et al. (1979b); Bingham et al. (2008)). It is in a way surprising that so little attention has been brought to this topic, since such bimanual movement are quite common in human everyday activities

(e.g. picking-up large objects). These studies confirmed the existence of bimanual coupling. For instance, reaching tasks are performed more easily when both hands reach the targets simultaneously. When the two hands are involved in asymmetric reaches, the movements of the hand reaching for the target the farthest away influences the movement of the other hand, slowing it down to preserve temporal coordination (Bingham et al. (2008)).

2.1.2 Human selection of hand poses for tool use

In comparison to arms, human hands have even more degrees of freedoms that result in much higher redundancy. It has been suggested that the CNS reduces the complexity of controlling fingers by selecting proper hand poses to grasp or to manipulate objects purposefully, through task-oriented grasps (Rizzolatti et al. (1988)). Several taxonomies (Kamakura et al. (1980); Cutkosky et al. (1989); Feix et al. (2015)) have been offered to categorize the large numbers of grasps used in everyday life, and multiple measures for determining the appropriateness of the grasp have been proposed (Ferrari and Canny (1992); León et al. (2012)). These measures enable evaluation of how well the grasp may stabilize the object or enable the subsequent manipulation (Endo et al. (2007)). A proper hand pose guarantees transmission of force or generation of torque that demanded by the task; and the placement of fingers enables the ability of holding tools and performing manipulation.

A wealth of evidence show that humans have already selected the appropriate hand pose prior to closing the fingers on the object. While the arm moves towards the object, the hand pre-shapes the fingers towards the final finger configuration. The formation of hand poses demands coordination cross many DOFs of fingers, and depends highly on the grasping purpose (Dollar (2014)) and target objects' properties (Bullock et al. (2013); Feix et al. (2014)).

A popular approach to explain the coordination mechanism underlying finger control is the concept of synergies (Santello et al. (1998)), which describes the coupling of multiple finger joints in a low-dimensional space. Hand poses may hence be generated from a single or a combination of multiple synergies. The analysis of postural hand synergy is commonly conducted using principal component analysis for static hand poses (Santello et al. (1998)) or singular value decomposition for temporal sequences (Mason et al. (2001)). However, the selection of hand pose is not only based on the functionality of the grasp objective (e.g. generating force or torque for task execution), but also highly related to the subjects' perception of task demands (Friedman and Flash (2007)).

Traditional synergy-based analysis that exploits only kinematic information does not explain how grasps respond to task demands (Ekvall and Kragic (2005)), as the hand pose does not contain information on the tool, nor on the force applied. This refers to grasp on tools that need to exert a specific pattern of forces during subsequent manipulation. One approach to relate grasps to manipulation request was done by (Friedman and Flash (2007)). Task compatibility of selected human hand poses are evaluated based on velocity, force transmission ellipsoids and stiffness ellipsoids. In our previous work, we also offered a mean to categorize grasps

(De Souza et al. (2015)) by combining tactile and postural information of human grasp, and related this to the opposition space theory (Iberall (1987b)). These approaches, however, do not explain how humans can modulate their grasps to better account for task demands. In the present study, we study how hand poses reflect the coordination patterns and control strategies that humans exploit for tool manipulation. We hypothesize that hand poses vary not only for manipulating distinct tools, but also as a result of proficiency. We analyze subjects' finger placements on the tools and compare hand poses, aiming to infer the differences in the understanding of task demands for subjects of different skill levels.

This study investigates coordination of arms and human selection of hand poses in an asymmetric and non-cyclic bimanual task. We focus on a typical task in watchmaking manipulation - the insertion of a watch spring (see Fig. 2.1). To perform the task, the two hands manipulate two different tools, tweezers and a wooden stick (peg-wood). They are constrained spatially to work on the same piece and follow also temporal constraints to move in synchronization to insert the spring. Studying the posture of the arms and fingers on the tools as well as the relative movements of the tools allow determining the asymmetrical and non-cyclic nature of the coordination patterns in such bilateral motions.

2.1.3   Acquisition of coordination patterns

Whether the coordination patterns reviewed above are *innate* or *learned* remains open. One of the earliest coordinated movements, kicking action (Thelen (1985)), can be traced back to the intrauterine stage. Fetuses in the womb display goal-directed kinematic patterns in hand motions by 22 weeks of gestation (Zoia et al. (2007)). Coordination patterns between separate motor systems such as hand-to-mouth coordination (Butterworth and Hopkins (1988)) and hand-to-eye coordination (Von Hofsten (1982)) are also observed at birth. Studies of infants reaching motion, however, suggest that the complex visuo-motor coordination patterns required to control precisely for arm and hand to grasp objects are acquired through a learning process (Georgopoulos et al. (1981); Konczak et al. (1995)). Non-random changes in hand and digit movements are observed in human fetus (Sparling et al. (1999)). Shortly after birth, infants manifest early grasping behaviors (Twitchell (1970)). Around $4-5$ months, infants start to demonstrate typically successful grasping motion (Sgandurra et al. (2012)), and the coordination between arm extension and finger flexion is also observed (Schneiberg et al. (2002)). Yet, little is known on how humans learn coordinated control of both force and kinematic of hand movements to perform fine-manipulation tasks, such as playing musical instruments (Furuya et al. (2011); Jerde et al. (2012)).

Approaches to explain human acquisition of coordination patterns take either an information-processing point of view or a dynamic-pattern perspective (Swinnen and Gooijers (2015)). The information processing theoretical framework suggests the main barrier in acquiring bimanual coordination patterns lies in the interference between limbs, i.e., the so-called neural cross talk (Marteniuk et al. (1984); Heuer et al. (2001)). Through practice or the integration of

sub-tasks, humans are able to overcome the constraints produced by above-mentioned factors, so as to decouple the task in asymmetric motions with distinct amplitude and directions of both arms and learn a new bimanual skill (Swinnen et al. (1993)). However, it has been argued that bimanual coordination is instead the "*default mode*" of the control system (Swinnen and Gooijers (2015)) with uni-manual action being rather only a suppression of the bimanual coordination nature of human motion. The dynamic-pattern framework (Schöner and Kelso (1988); Kelso (1997)) considers the human acquisition of skills as an alteration of the intrinsic patterns (e.g. physical constraints, prior knowledge, and experience) under behavioral information (e.g. constraints imposed by task and environmental) (Schöner et al. (1992)). It is suggested that the human motor system generates bimanual coordination patterns as a result of the interaction of constraints at various levels (Swinnen (2002); Swinnen and Gooijers (2015)). Studies also confirmed that the acquisition of novel coordination patterns are affected by a variety of factors, such as phase (Haken et al. (1985)), direction (Baldissera et al. (1982); Swinnen et al. (1997, 1998)), timing (Yamanishi et al. (1979); Deutsch (1983); Peters (1981)), and amplitude (Franz (1997); Spijkers and Heuer (1995)). It shows that the accuracy requirement of the task reduces relative phase bias in the bimanual coordination patterns (Wang et al. (2017)). Visual guidance plays also an important role, as it reduces the complexity of bimanual coordination (Vaz et al. (2017)). Coordination patterns can be acquired through training; and practice could enhance the stability of the acquired coordination patterns (Smethurst and Carson (2001)). Notwithstanding the foregoing, it remains unclear what task-specified coordinated patterns do humans generate, also whether the coordination patterns change as human subjects continue practice and improve task proficiency, i.e., whether these factors have consistent impact on the generation of coordination patterns at different learning stages.

2.1.4 Optimal control theory for understanding human motion

A central problem in motor control is to understand how humans coordinate redundant degrees of freedoms to generate movements. Optimal control theory (Engelbrecht (2001); Todorov (2004)) considers the human motion generation as an optimization process and suggests that the CNS may optimize underlying optimal criteria (or cost functions) in order to reduce the control complexity by means of coordination. It is presumed that the underlying optimal criterion of human CNS is composed of multiple weighted costs, such as energy expenditure of both inertial and gravitational forces (Berret et al. (2008)). Therefore, inferring the optimal criteria used by the CNS is key to understanding the mechanism of human motion. This becomes an inverse problem to human motion planning and generation.

Inverse optimal control (IOC) can be used to infer the weights of corresponding costs that compose the optimal criteria of CNS. For instance, (Mombaur et al. (2010)) used IOC to determine the optimal criteria for human to generate locomotion. To analyze arm motion, (Berret et al. (2011)) provides experimental evidence for the existence of composite cost for CNS and suggests that human CNS is trying to minimize mechanical energy expenditure (40% on average) and joint level smoothness (35% on average) during arm reaching motion.

(Mainprice et al. (2015)) applied IOC to analyze human arm reaching motion in an interaction task and suggests that humans try to increase the smoothness of arm trajectories rather than to avoid collision with other human subject. (Oguz et al. (2018)) used IOC to investigate the trade-off between kinematics-based and dynamics-based controllers depending on the reaching task. These works focus on modeling human motions as displayed by adults, and not on modeling the process by which these motion have been acquired. They have not taken subjects' proficiency into consideration. Moreover, constraints imposed by task demands are not considered.

The present study follows an optimal control framework to interpret human kinematic coordination, paying particular attention to proficiency of the subject. Unlike the above-mentioned studies that exploit IOC to analyze the process of generating optimal motion trajectories, we focus on understanding the planning of the generated kinematic postures, and thus formulate it as an inverse optimization problem (IOP) without considering the control input (e.g. joint torque) of the human arm. Furthermore, we take into account modeling and acquisition of task-demand explicitly within the optimization process of the CNS. Indeed, as subjects acquire a skill, they acquire simultaneously a model of the task demands and of the pattern of hand-arm coordination that can best meet these task demands. Aside from the inherent constraints of human body, such as muscle/bone capability and musculoskeletal structure, the ability to generate forces meeting task demands is also highly affected by the kinematic properties (Chiu (1987)). We, hence, consider among potential task-level costs that CNS may try to optimize, the singularity of posture, ability to satisfy task demands, and also comfort - from a biomechanical point of view. Our optimization framework is hence a composite of optimal criteria, which includes both inherent (e.g. biomechanical constraints) and extrinsic (e.g. constraints imposed by task demands) costs.

The composition structure of CNS's optimal criteria is indicated by the weights attached to the different cost components. Taking advantage of inverse optimization approach, we infer the weights of these cost components to investigate the composition structure of this optimal criteria. We hypothesize that the different coordination patterns across novice and experts may be revealed by different weights associated to each cost. These weights change as skill improves. We, hence, compare the composition structures of CNS's optimal criteria of human subjects at different skill levels, so as to decipher how such bimanual coordination patterns are generated with improvement in task proficiency.

2.2 Methods

To study bimanual coordination, we requested subjects to perform a typical watchmaking task that requires precise inserting and loading the watch's spring (a key element to the functioning of the watch). Steps of performing the spring assembly task are explained in details in Sec. 2.2.3. Statistics are presented in the form of mean±std values.

We collected both tactile information (applied force) and kinematic information on hands and

(a) Placing spring and localizing tools

(b) Task execution

(c) Watch spring inserted
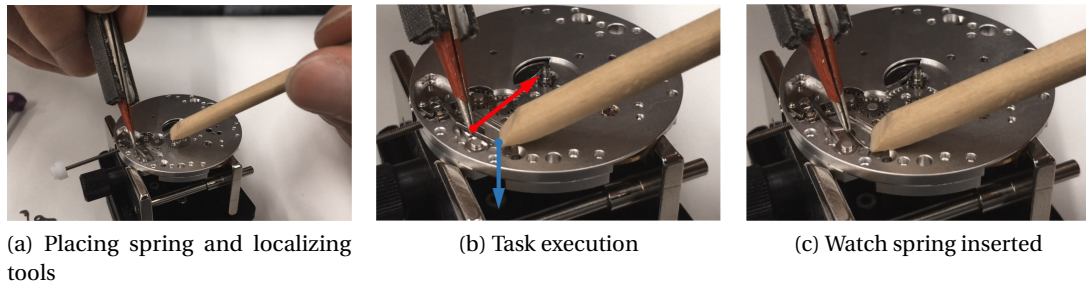
Figure 2.1 – Illustration of the procedure of inserting a watch spring (pictures were taken in front of the subject). The left hand holds a peg wood and the right hand manipulates a pair of tweezers. (a) The spring is unmounted. Subject picks up the watch spring using tweezers and places it onto the watch face. (b) The subject localizes the tools onto the watch face: the peg wood (left) presses the watch spring downwards to maintain it stable; in the meanwhile, the tweezers pinch the foot of the spring. (c) Subject pushes the tweezers to compress the foot of the watch spring for insertion.
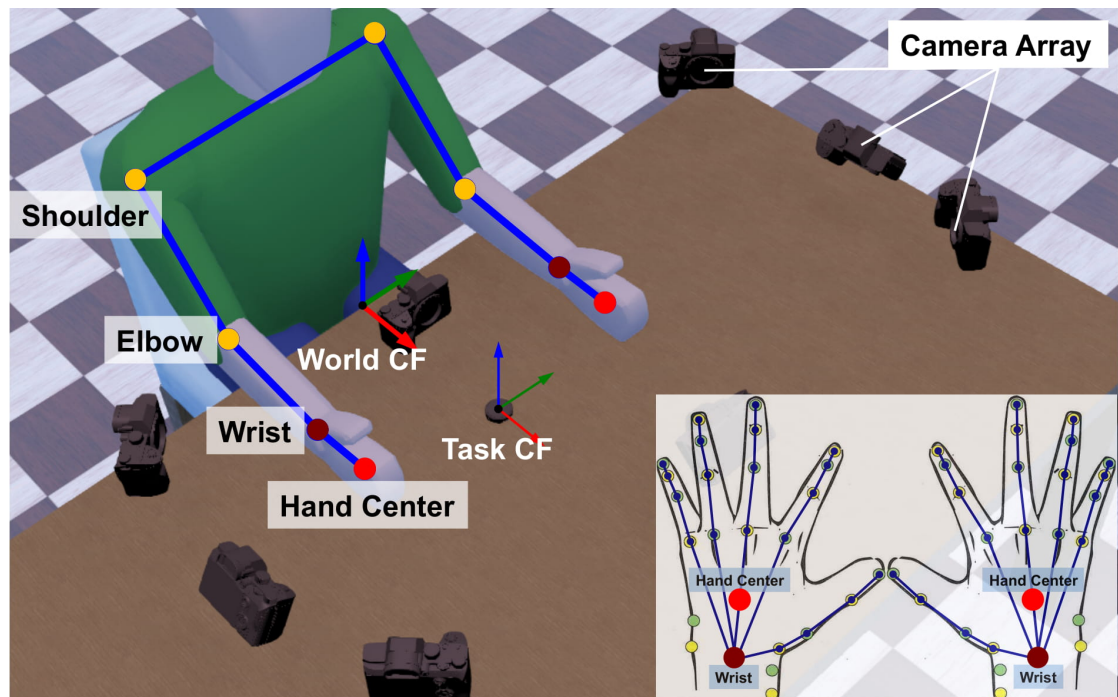
Figure 2.2 – Experimental scenario and the human skeleton model. The end-effector coordinate frame of the arm model (top-left) overlaps the hand coordinate frame (bottom-right).

tools (Sec. 2.2.2). From recorded force information, we extracted task-demands (Sec. 2.2.4); and from the recorded videos, we reconstructed skeleton models of arm and hand (Sec. 2.2.5). We computed hand pose on manipulating tools and express these as coordinated patterns of finger joints. We formulated metrics (Sec. 2.2.6) to evaluate kinematic coordination patterns of human hands and arms. We developed a model based on inverse optimization to explain

differences across novice and experts (Sec. 2.2.8).

2.2.1 Subjects

Fifteen subjects took part in the study. They were recruited from ETVJ (École Technique de la Vallée de Joux) on a voluntary basis following wide announcement. The ability to mount the spring of the watch is acquired during the first two years of the study at ETVJ and is hence considered as mastered for students in the 3rd year and above. Hence, subjects were assigned to novice group (ten) and expert group (five) according to the number of years of practice at watchmaking. Novice subjects (nine men, one woman; age range: $16-30$, mean±std: 18.60 ± 4.33) comprised students in first year of study at ETVJ with no prior expertise in watchmaking. Experts (three men, two women; age range: $17-36$, mean±std: 21.80 ± 7.98) encompassed four students in 3rd and 4th year of study and one professional, all of whom had more than 24 months of practice at watchmaking. All subjects were right-handed.

The experimental protocol was approved by the Human Research Ethics Committee of EPFL. All subjects were provided with an information sheet, and they (or their caretaker when underage) gave informed consent prior to the experiments. Subjects received detailed verbal instructions prior to starting the experiment. They were given the opportunity to ask as many questions as they wished and to train themselves at performing the tasks while wearing the tactile glove (which partially hindered natural sense of touch). No further instructions regarding task objective or performance evaluation were given to the subjects, and they were left free to perform the task.

2.2.2 Apparatus

Workspace is defined as a rectangular-shaped area (50cm × 50cm) on the table surface and in front of the seated subject (see Fig. 2.2). The manipulation platform - watch face, is fixed on a cuboid plastic base with a miniature force/torque sensor (ATI Nano17) mounted underneath. The base is placed in the center of the workspace facing the subjects at the start of each experiment. Subjects are free to change the pose of the base.

Eight GoPro cameras are arranged in a rectangular formation around the subject's workspace. Redundancy across angle of view ensure that one can capture the motion of the subject's arms and hands, the movement of the tools (peg wood and tweezers) and watch base from multiple angles. Twenty-five colored dots (trackers) are pasted on each hand of human subjects to track their hand poses and fingers motions: four trackers on each finger, one on the hand center, four around the wrist (two near the tail of ulna and two near the tail of radius, see the bottom right figure of Fig. 2.2). Wrist joint position is estimated as the center of the ulna-radius trackers. The pose of watch face is tracked by localizing its center and the crown of watch stem manually in the image and then use a smooth tracker across the flow of the image. Tools used by both hands (peg wood and tweezers) are tracked through the same procedure by

localizing their tails. Movements of tracked positions are recorded by the camera array, and their trajectories are extracted from the recorded videos afterwards. Shoulder (glenohumeral joint) and elbow positions are tracked during the post-processing of recorded videos, by manually adding trackers near the real joints on the video frames, based on visual observation.

We define two Cartesian coordinate frames: the world coordinate frame $\mathbb{W}_{\mathrm{CF}}$, located in the middle of the workspace's border; and the task coordinate frame $\mathbb{T}_{\mathrm{CF}}$, located in the center of the watch face[1] (see Fig. 2.2). The ATI sensor records the force and torque applied to watch face at 70Hz, and GoPro cameras record videos at 50Hz.

2.2.3 Study design and data collection

Task description

Subject was asked to sit in front of the table on a stool. The table and stool are part of the regular furniture used by students and teachers at ETVJ for practice session. All sensors were mounted and calibrated before the experiment starts. The subject was allowed to freely move the watch face and tools (see Fig. 2.4). At the beginning of each experiment, the subject was asked to assemble and disassemble several components on the watch face, in order to get familiar with the experimental apparatus. This manipulation trial was discarded in the analysis reported. After this initial familiarization phase, the subject was asked to assemble the watch spring. No time pressure was put on the subject, and each subject performed the task at his own pace.

Insertion of the watch spring is performed as follows. The subject first places the watch spring on its target position on the watch face (see Fig. 2.1a). The subject's assistant hand (left hand in our study) holds a peg wood and presses the watch spring in order to maintain it in place as the other hand loads it. To load the spring, the subject's dominant hand (right hand in our study) uses the tweezers (with the end closed) to push the foot of the spring so that it falls into the reserved slot (see Fig. 2.1b). Failure happens if the spring was not inserted or not correctly loaded. The manipulation is also considered failed if the watch spring drops or if it flies away during loading. Subjects were allowed to try again in case of failures. A trial ends only once the watch spring is stably inserted (see Fig. 2.1c). Each subject was asked to repeat the spring insertion six times. The first two trials were discarded as the subject was still familiarizing him/herself at doing the task with the sensorized gloves.

Motion segmentation and data selection

The recorded manipulation trial was further cut into multiple motion segments. Each motion segment contains one swipe attempt at inserting the spring. It starts when both tools are located at their respective target positions on the watch face and the subject is about to press

---

[1]We use superscripts ("+" and "−") to denote directions of axes, and subscripts ("W" and "T") to indicate the coordinate frame.

(a) Rotation of watch face

(b) Top view of human subject

(c) Top view of force ellipsoid

(d) Definition of angles
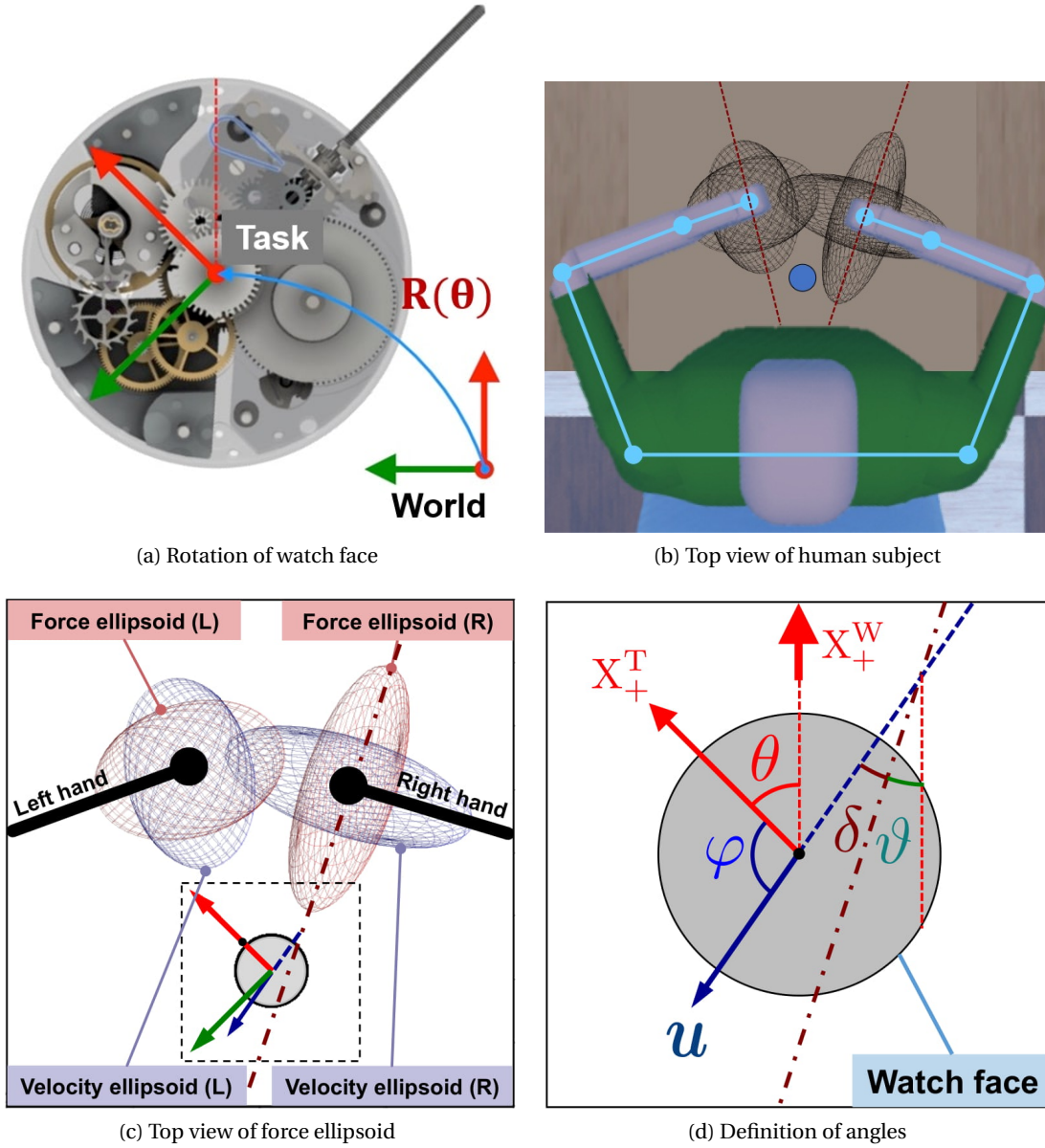
Figure 2.3 – Extraction of coordination angles for analysis. Force/velocity ellipsoids are projected onto the $X_W - Y_W$ plane for a better visualization of their major axes. θ is the angle to rotate $\mathbb{W}_{CF}$ to $\mathbb{T}_{CF}$. The task-demanded force application vector $\boldsymbol{u}$ has an included angle φ with X_T^+. ϑ is the included angle between the major axis of the project force ellipsoid and X_W^+. δ is the included angle between $\boldsymbol{u}$ and the major axis of the projected force ellipsoid.
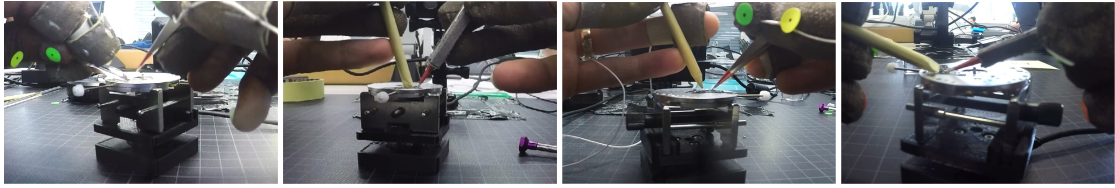
Figure 2.4 – Examples of subjects rotating watch face during task.

the tweezers to close them prior to the assembly execution (see Fig. 2.1b). The segment ends when the spring is released from the tweezers after pressing on it, regardless of whether the attempt was successful or failed (see Fig. 2.1c). Posture adjustments such as changing arm pose, readjusting finger placements, and rotating watch face were excluded from the data analyzed here. Segments that include apparatus malfunctions, irrelevant motions (e.g. unexpected actions), and non-standard operations (e.g. subject uses single hand or odd hand gestures to complete the bimanual task) were discarded. Around 18% of the raw motion segments were discarded from the novice group and 20% from the expert group. After selection, each novice subject has 12.30 segments and each expert subject has 11.20 segments on average.

The average time length of the segmented motion piece is $4.33s \pm 2.17s$ for novice group, and $2.98s \pm 1.72s$ for expert group. Success rate of each subject is calculated as the number of success motion segments divided by the total number of motion segments. Analysis of the motion segments indicates a higher ($F_{1,13} = 15.86$, $p < 0.01$) success rate of subjects from the expert group ($83.49\% \pm 11.97\%$) than the novice group ($57.53\% \pm 11.87\%$) on average.

Extraction of kinematic information

To analyze the coordination patterns across the two arms and hands, we extracted the postures of all limbs from the videos recorded by the GoPro cameras. The camera system is calibrated by computing the intrinsic (Zhang (2000)) and extrinsic camera matrices (Lepetit et al. (2009)) using checkerboard images across the field of view. The location of the marker in the workspace is computed by triangulating their coordinates in cameras for which the view of the marker is not obstructed. To extract the 3D marker position at one frame of recording, a human operator marks the marker position on the images captured by at least two cameras. The marked coordinates are first undistorted using the intrinsic camera matrices and then projected to a 3D line by the extrinsic parameter matrix. The 3D position of the marker is constructed by computing the closest point to the 3D lines from all cameras that has a view of this marker. This process is repeated for all markers on the two hands of each participant, in order to record the fingers' positioning. All implementations are developed using OpenCV. Errors of this system comprise both (1) reconstruction error (3.5 ± 1.4 pixels on average, tested on collected known points), stemming mainly from the calibration; and (2) marking error (0.85mm±0.15mm on average, tested on randomly selected markers), i.e., error between manually marked position and its reconstructed value. We obtain the average error as 1.63mm±0.4mm per marker, by combining both error sources.

Posture of the arm joint was done by hand labelling with an expert tracing the posture in the camera image. Since the arm barely move during the experiment, this is sufficient to obtain accurate estimation. Reconstruction error of the arm joints mainly comes from imprecise visual estimation of the human operator. Approximating the upper limbs as cylinders, the maximum error range of arm joints estimation is 2.44cm, calculated using the mid-upper arm circumference value (Preedy (2012)).

Since the assembly motion requires only a millimeter-scale movement of the tweezers' tip, we consider the postures of the upper limbs as quasi-static during motion execution. Therefore, we extract subjects' postures as the subject is about to start the insertion motion (see Fig. 2.1b), which can be considered as the average position over the course of the motion by approximation. Extraction of the posture is done by manually marking and estimating the 3D coordinates of these points of interest from recorded video frames, as described above. After extraction, we have one complete set of tracker positions (62 trackers in total) that describe the kinematic information of each subject, including 3 points on each arm, 25 points on each hand, 2 points on the watch face, 2 on the peg wood, and 2 on the tweezers. Using these extracted points, we were able to reconstruct the posture of both upper limbs, and the position/orientation of the watch face and tools.

2.2.4 Extraction of task demands

It is difficult to construct an analytic model of the task, due to the nonlinear shape of the spring. Instead, we use the motion data (both sensor signals and videos) recorded from expert subjects as ground truth, to (1) extract task demands, and to (2) evaluate the task performance of novice subjects.

We fit the 3D force signal sequence recorded by ATI F/T sensor during each motion segment as one force ellipsoid. The major axis of this force ellipsoid indicates the task-demanded direction of force application. We constructed force ellipsoids of all successful motion segments recorded from expert subjects. The average direction of their major axes is used as the optimal task-demanded force direction $\boldsymbol{f}^*$. The task requires left hand to hold the peg wood and exert pressure vertically (blue vector in Fig. 2.1b) for maintaining the stability of the spring, while right limb to manipulate tweezers for pinching the spring and pushing it horizontally (red vector in Fig. 2.1b). Hence, we decomposed $\boldsymbol{f}^*$ in $\mathbb{T}_{\mathrm{CF}}$ according to the functionality of each tool: the projection along Z_T is used as the task-demanded vector for peg wood ($\boldsymbol{f}_Z^*$, left hand); and the projection on the $X_T - Y_T$ plane is considered the task-demanded vector for tweezers ($\boldsymbol{f}_{XY}^*$, right hand).

2.2.5 Modeling of human upper limb

The skeleton model of human upper limbs (see Fig. 2.2) is defined in the world coordinate frame $\mathbb{W}_{\mathrm{CF}}$ and overlaps with the upper limb of the recorded subject. It consists of hand

models and arm models of both sides.

Skeleton model of human arm

Each arm model (left and right) comprises three links (upper arm, forearm, and hand) and three joints (shoulder, elbow, and wrist) (Wu et al. (2005)). Only joint rotations are modeled. Upper arm is linked to the human torso at the shoulder joint, which has three DOFs: arm ab/adduction, circumduction, and extension/flexion. Upper arm and forearm are connected by the elbow joint (one DOF: forearm extension/flexion). The wrist joint connects forearm and hand, and has three DOFs: extension/flexion, pronation/supination, and ulnar/radial deviation (Bajaj et al. (2015)). Hand link of the arm model is modeled as a constant link attached to the wrist joint. Fingers are not modeled. An end-effector's coordinate frame is attached at the end of the hand link. It coincides with the base coordinate frame of the hand skeleton model (see Sec. 2.2.5).

The length of each model link was calculated as the distance between the two tracker positions attached at both ends of the link. For example, the length of the upper arm is the distance between shoulder joint and elbow joint; the length of the forearm is the distance between elbow joint and wrist joint. In particular, the length of the hand link is assigned as the distance from the wrist joint to the center of the hand (the tracker attached in the center of the palm, see Fig. 2.2). Since the joints of the human arm and the joints of the constructed arm model do not have a bijective mapping, a point-to-point mapping approach (Peer et al. (2008)) is used to fit the arm model to the recorded human arm posture. Extracted positions of joint trackers (shoulder, elbow, wrist, and palm center) are used to solve an inverse kinematics (IK) problem to obtain joint angles for each DOF of the arm model. The joint value of each DOF is bounded within the corresponding anthropomorphic motion range (Preedy (2012)). Unique IK solution is guaranteed by setting constraints on both joint position ranges and link postures.

Skeleton model of human hand

We used a modified 20-DOF paradigmatic hand model (Malvezzi et al. (2015)) to analyze the kinematic coordination patterns of human hands (see Fig. 2.2, bottom right). The thumb finger has three joints: the distal interphalangeal (DIP) joint (1 DOF: flexion/extension), the MCP joint (1 DOF: flexion and extension), and the carpometacarpal (CMC) joint (2 DOFs: flexion and extension, ab/adduction). Each one of the other four fingers has four DOFs: the ab/adduction and extension/flexion of the MCP joint, the extension/flexion DOF of both proximal interphalangeal (PIP) joint and DIP joint. The length of each finger phalanx is calculated as the distance between the two joints located at both of its ends. Each joint angle of the model is calculated directly using the extracted tracker position. A joint-to-joint mapping approach (Bouzit (1996)) is used to fit the hand kinematic model to the recorded hand data. Contacts are defined as points on the surface of the finger phalanx and manually estimated from the recorded videos. Soft-finger contact model is used by taking the friction

moment with respect to the contact surface normal into consideration (Murray et al. (2017); Prattichizzo and Trinkle (2016)).

2.2.6 Metrics

To measure the goodness of the control, we considered three factors: (1) a singularity measure ω of how far the joints are from singular configurations, (2) a task compatibility measure α, that assesses the capability of the upper limb to satisfy task demands, namely to generate the desired force/motion, and (3) a comfort measure ϵ that estimates potential discomfort that could arise from tension in tendons and muscles caused by movement of certain joints.

Singularity measure ω

We used the manipulability index proposed by (Yoshikawa (1984)) as a metric of the overall kinematic configuration quality. Manipulability index corresponds to the volume of manipulability ellipsoid. The Jacobian $\boldsymbol{J}$ maps the joint space velocity $\dot{\boldsymbol{\theta}}$ to the Cartesian space velocity $\dot{\boldsymbol{x}}$ as $\dot{\boldsymbol{x}} = \boldsymbol{J}\dot{\boldsymbol{\theta}}$. A unit sphere in joint space, described by $||\dot{\boldsymbol{\theta}}||^2 = 1$, can be mapped to an ellipsoid in Cartesian space as:

$$||\dot{\boldsymbol{\theta}}||^2 = \dot{\boldsymbol{\theta}}^\mathsf{T}\dot{\boldsymbol{\theta}} = \dot{\boldsymbol{x}}^\mathsf{T}(\boldsymbol{J}^\dagger)^\mathsf{T}\boldsymbol{J}^\dagger\dot{\boldsymbol{x}}^\mathsf{T} = \dot{\boldsymbol{x}}^\mathsf{T}(\boldsymbol{J}\boldsymbol{J}^\mathsf{T})^\dagger\dot{\boldsymbol{x}} = 1. \tag{2.1}$$

$\boldsymbol{J}^\dagger$ is the pseudo-inverse of $\boldsymbol{J}$. The ellipsoid defined by $\boldsymbol{J}\boldsymbol{J}^\mathsf{T}$ reveals the flexibility of motion in the task space, and its volume is used as the manipulability measure:

$$\omega = \sqrt{\det(\boldsymbol{J}\boldsymbol{J}^\mathsf{T})}. \tag{2.2}$$

Maximizing ω prompts the joints to move towards configurations that are far away from singularities, thus guarantees higher motion flexibility.

Task compatibility measure α

The assembly task requires applying force in the desired direction. Hence, we used force transmission ratio (Chiu (1988)) as a task compatibility measure. The relation between Cartesian force space and joint torque space via the Jacobian $\boldsymbol{J}$, $\boldsymbol{\tau} = \boldsymbol{J}^\mathsf{T}\boldsymbol{f}$, maps a unit sphere in joint torque space to Cartesian force space as a force ellipsoid:

$$||\boldsymbol{\tau}||^2 = \boldsymbol{\tau}^\mathsf{T}\boldsymbol{\tau} = \boldsymbol{f}^\mathsf{T}(\boldsymbol{J}\boldsymbol{J}^\mathsf{T})\boldsymbol{f} = 1. \tag{2.3}$$

Let unit vector $\boldsymbol{u}$ denote the direction of interest, the force transmission ratio α is defined as:

$$\alpha = (\boldsymbol{u}^\mathsf{T}(\boldsymbol{J}\boldsymbol{J}^\mathsf{T})\boldsymbol{u})^{-1/2}. \tag{2.4}$$

A physical interpretation of α is the distance from the center of the force ellipsoid to its surface, along the direction of $\boldsymbol{u}$. A large value of α implies that, with the same joint torque, the arm (or the hand) is capable of applying large force to the environment. In our study, we used the extracted task demands ($\boldsymbol{f}_{XY}^*$ and $\boldsymbol{f}_Z^*$) as the vector $\boldsymbol{u}$ for both arms (or hands), respectively.

Comfort measure ϵ

Large joint movements may stretch tendons and lead to discomfort feelings. The closer a joint approaches its motion boundaries, the less comfort. We adapted the comfort index (León et al. (2012, 2014)) to measure the deviation of joint angles from joint position when the hand is in its rest posture, as an indication of joint stretch and associated joint comfort feeling:

$$\epsilon = 1 - \frac{1}{N_J} \sum_{i=1}^{N_J} \left(\frac{q_i - q_i^r}{Q_i} \right)^2, \; i = 1, 2, \ldots, N_J. \tag{2.5}$$

The angle of ith joint is q_i, and q_i^r denotes the angle when the ith joint is at rest position. N_J denotes the number of joints taken into calculation. This deviation is normalized by the joint motion range Q_i, which measures the distance between q_i^r and the joint limit.

2.2.7 Adaptation of kinematic measures to human model

Adaptation to human hand model

We adapted the above-defined measures to the human hand model, to evaluate the hand poses of recorded subjects. Hand singularity measure ω^h is modified as[2]:

$$\omega^h = \sqrt{\det(\boldsymbol{H}\boldsymbol{H}^\mathsf{T})}. \tag{2.6}$$

The hand-object Jacobian $\boldsymbol{H}$ (Shimoga (1996)) transmits the applied joint torque $\boldsymbol{\tau}$ to the applied force $\boldsymbol{f}$ in an object reference frame as $\boldsymbol{\tau} = \boldsymbol{H}^\mathsf{T} \boldsymbol{f}$. It indicates the volume of manipulability ellipsoid (Murray et al. (2017)) and reveals the ability of the hand to manipulate the tool. The task compatibility measure is reformulated as:

$$\alpha^h = (\boldsymbol{u}^\mathsf{T} (\boldsymbol{H}\boldsymbol{H}^\mathsf{T}) \boldsymbol{u})^{-1/2}. \tag{2.7}$$

The direction of force that correspond to task-demand is projected into hand's coordinate frame and used as vector of interest $\boldsymbol{u}$. The comfort measure of hand is then calculated as

---

[2]We use superscription "h" and "a" to indicate the measures of hand and arm, respectively. The subscriptions "E" and "N" indicate the group, and "L" and "R" indicate the side. For example, $\omega_{E,L}^h$ represents the ω calculated value of left hand in expert group.

follows:

$$\epsilon^h = 1 - \frac{1}{N_f N_J^f} \sum_{f=1}^{N_f} \sum_{j=1}^{N_J^f} \left(\frac{q_{f,j} - q_{f,j}^r}{Q_j^f} \right)^2 ,$$
$$j = 1, 2, \ldots, N_J^f, \ f = 1, 2, \ldots, N^f .$$
(2.8)

The angle of jth joint of the fth finger is $q_{f,j}$, and the distance between its rest position $q_{f,j}^r$ and its joint limit is Q_j^f. N_J^f denotes total number of joints of fth finger ($N_J^f = 4$ in this study), and N_f is the total number of fingers ($N_f = 5$ in this study). We used the statistics in (Lee et al. (2008)) to calculate the joint angle when the hand is at rest pose.

Adaptation to human arm model

The manipulability measure ω^a and task compatibility measure α^a of the arm model follow the measure defined in Eq. (2.2) and Eq. (2.4), respectively, with $\boldsymbol{J}$ being the kinematic Jacobian of the arm. Comfort measure is modified for the arm model as:

$$\epsilon^a = 1 - \frac{1}{N_J} \sum_{i=1}^{N_J} \left(\frac{q_i - q_i^r}{Q_i} \right)^2 , \ i = 1, 2, \ldots, N_J,$$
(2.9)

where N_J is the number of arm DOFs ($N_J = 7$ for each arm in this study), q_i is the angle of ith joint, q_i^r is the rest position angle of ith joint, and Q_i is the corresponding joint limit.

2.2.8 Understanding human kinematic coordination based on inverse optimization approach

We exploited an inverse optimization approach to analyze the human upper limb model, both (1) to infer the composition structure of the central nervous system's optimal criteria, and (2) to understand how this composition structure varies as the human subject improves task proficiency.

Inverse optimization problem formulation

The inverse optimization problem (IOP) aims at identifying the formulation of the optimization problem (OP), i.e., inferring the form of objective function (or a combination of cost components), subject to which the optimization problem can best reproduce the given optimal solutions (Tarantola (2005)). IOP assumes that the optimal objective function can be represented as a combination of multiple elementary costs. The contribution of each elementary cost to the entire composite criterion is represented by its corresponding weight coefficient. Weight coefficients can be identified by formulating the IOP as a bi-level optimization problem (Colson et al. (2007)):

$$
\boxed{
\begin{array}{l}
\textbf{Upper level problem} \\[2mm]
\qquad \min_{\boldsymbol{\omega}} \quad \boldsymbol{\Phi}(\boldsymbol{x},\, \boldsymbol{x}^*) \\[4mm]
\qquad \text{s.t.} \quad \sum_{i=1}^{N} \omega_i = 1,\ \omega_i \in [0,1],\ i = 1,2,\ldots,N
\end{array}
}
\qquad (2.10)
$$

$$
\boxed{
\begin{array}{l}
\textbf{Lower level problem} \\[2mm]
\qquad \min_{\boldsymbol{x}} \quad \boldsymbol{\Psi}(\boldsymbol{x}|\boldsymbol{\omega}) = \min_{\boldsymbol{x}} \sum_{i=1}^{N} \omega_i\, \Psi_i(\boldsymbol{x}) \\[4mm]
\qquad \text{s.t.} \quad \boldsymbol{g}_j(\boldsymbol{x}) \le 0,\ j = 1,2,\ldots,m, \\[2mm]
\qquad\qquad\; \boldsymbol{h}_k(\boldsymbol{x}) = 0,\ k = 1,2,\ldots,n.
\end{array}
}
\qquad (2.11)
$$

Upper level problem

The upper level problem (see Eq. (2.10)) aims at finding the optimal weight vector $\boldsymbol{\omega}^*$[3] that minimizes the distance (measured by $\boldsymbol{\Phi}$) between the current optimal state $\boldsymbol{x}$ (solution of the lower level problem) and the optimal observed state $\boldsymbol{x}^*$ (recorded human data). The optimization variable $\boldsymbol{\omega}$ contains coefficients ω_i, which are assigned to the metrics Ψ_i as defined in the lower level problem. The metric $\boldsymbol{\Phi}$ can be defined either in joint space (error in joint angles) or Cartesian space (error in end-effector pose). We used the sum of mean square errors of joint angles, since the constrained IK guarantees a unique mapping between model joint angles and Cartesian positions, and computing $\boldsymbol{\Phi}$ in joint space also reduces computational complexity. Constraints of the upper level problem include (1) boundary constraints of each weight coefficient $\omega_i \in [0,1]$, and (2) linear constraint $\sum_{i=1}^{N} \omega_i = 1$.

Lower level problem

The lower level problem (see Eq. (2.11)) is a direct optimization problem, and its target is to find the state $\boldsymbol{x}$ that optimizes[4] the objective function $\boldsymbol{\Psi}$, which is assumed to be a linear combination of N elementary costs Ψ_i, weighted by the corresponding coefficients ω_i. We accounted the kinematic metrics described in Sec. 2.2.6, which we used to analyze the coordination patterns of arm models, as high-level cost components Ψ_i, i.e., $\Psi_1 = \omega^a$, $\Psi_2 = \alpha^a$, and $\Psi_3 = \epsilon^a$.

---

[3]We use the superscription "$*$" to indicate the optimal value (solution) of the variable.

[4]The objective function $\boldsymbol{\Psi}$ is supposed to be maximized in the lower level problem. However, we still use "*minimize*" to maintain the consistency of description.

It is noteworthy that the calculated kinematic measures cannot be combined directly, since they have different units and scales. To handle this problem, we introduced a normalization coefficient κ_i to normalize the kinematic measures:

$$\Psi(\boldsymbol{x}|\boldsymbol{\omega}) = \sum_{i=1}^{N} \omega_i \, \kappa_i \, \Psi_i(\boldsymbol{x}), \tag{2.12}$$

$$\kappa_i = \max\{\Psi_i(\boldsymbol{x})\}^{-1}, \; \boldsymbol{x} \in [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}], i = 1, 2, \ldots, N \tag{2.13}$$

with $\underline{\boldsymbol{x}}$ and $\overline{\boldsymbol{x}}$ denoting the element-wise lower and upper boundaries of the state variable $\boldsymbol{x}$, respectively. The maximum value of each kinematic measure Ψ_i is determined by optimizing Ψ_i within the bounds of state variables.

To determine the state variable $\boldsymbol{x}$ of the lower level optimization problem, we performed Sobol analysis to analyze the global sensitivity of cost Ψ_i to each joint of the model. The analysis was conducted using the GSAT (Global Sensitivity Analysis Toolbox, (Cannavó (2012))), and each joint was sampled within its motion range using a quasi-random Monte Carlo method. A joint with a sensitivity ratio above 5% was considered to have significant influence on the kinematic measures and then considered as one state variable.

The feasible movement range of each arm joint is limited both inherently by the biomechanical constraints and extrinsically by the task demands. For instance, in watchmaking, subjects usually place their forearms on the table and close to the watch face; consequently, the extension/flexion of the upper arm is limited. This is revealed by a low variance of the joint angle cross manipulation trials of subjects. Therefore, to better represent the changes in comfort feeling that are caused by the movements of arm joints, we assigned a weighting factor β_j to the comfort value of each model joint:

$$\beta_j = \sqrt{\text{Var}(\{\boldsymbol{q}_j^s\})}, \; \boldsymbol{q}_j^s \in [\underline{\boldsymbol{q}}_j, \overline{\boldsymbol{q}}_j], \; s = 1, 2, \ldots, N_s, \tag{2.14}$$

where N_s is the total number of subjects, $\boldsymbol{q}_j^s$ is the jth joint angle of sth subject, $\underline{\boldsymbol{q}}_j$ and $\overline{\boldsymbol{q}}_j$ are the element-wise lower and upper boundaries of $\boldsymbol{q}_j^s$, respectively. Thus β_j is the standard deviation of jth joint angle cross all subjects. Values of β_j are normalized cross joints in calculation, i.e., $\beta_j' = \beta_j / \sum_{j=1}^{N_J} \beta_j$ with N_J being the total number of joints. Hence, the comfort measure as defined in Eq. (2.9) is modified as:

$$\epsilon^a = 1 - \frac{1}{N_J} \sum_{j=1}^{N_J} \beta_j' \left(\frac{q_j - q_j^r}{Q_j}\right)^2, \; j = 1, 2, \ldots, N_J, \sum_{j}^{N_J} \beta_j = 1. \tag{2.15}$$

Constraints of the lower level problem include boundary constraints of each state variable. Joint angle q_i is bounded within its corresponding anthropomorphic motion range, and rotation angle of the watch face θ is constrained within $[-\pi, \pi]$. Optimization problems of both levels were solved using derivative free optimization method, since common gradient-based optimization approached may fail in this case due to the non-smooth nature of the

(a) Fitted force ellipsoid

(b) Axes of force ellipsoid
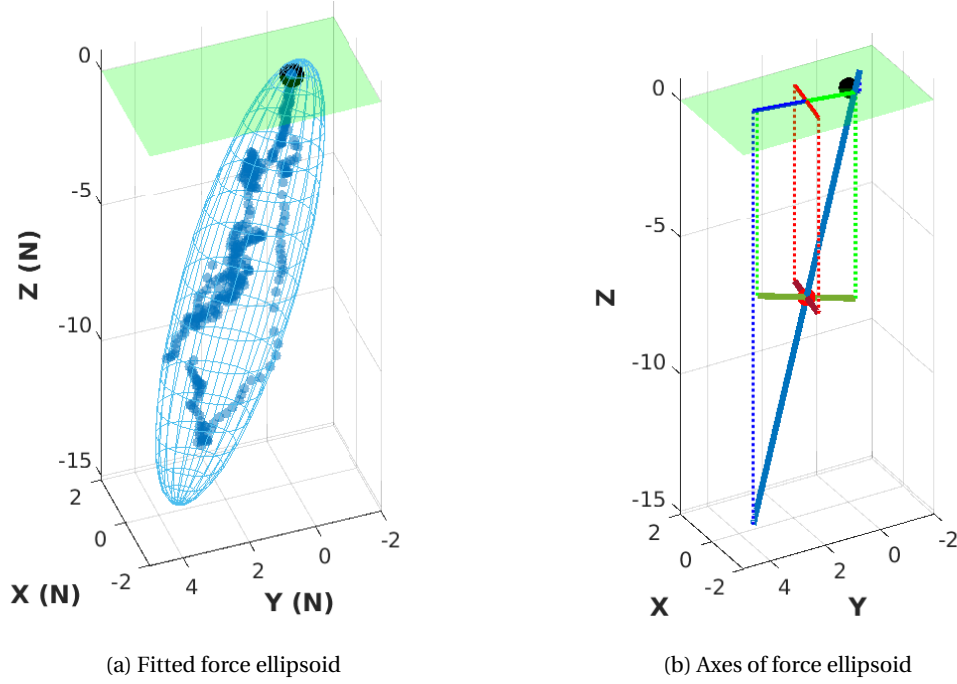
Figure 2.5 – Example of a force ellipsoid and its axes extracted from an expert subject. The major axis of the force ellipsoid indicates the task-demanded direction of force application.

objective function $\Phi(\boldsymbol{x}, \boldsymbol{x}^*)$. We used surrogate optimization approach (Gutmann (2001)) in MATLAB Optimization Toolbox™. Surrogate optimization is a global derivative-free approach for bounded optimization problems. It samples to approximate objective functions and is best suited to time-consuming computation. Global optimal solution is often difficult to guarantee in bi-level optimization problems, as the bi-level optimization framework is intrinsically restricted by its non-convex and non-linear formulation (Bard (2013)), and no efficient method has been proposed to solve this problem so far. We ran the inverse optimization problems five times for each subject and used the solution that results in the best objective function values. The weight $\boldsymbol{\omega}_i$ was normalized before each lower level problem process starts, and hence the total cost of the lower level $\Psi(\boldsymbol{x}|\boldsymbol{\omega})$ was normalized within $[0, 1]$. The finally obtained optimal weight vector $\boldsymbol{\omega}^*$ of each subject was normalized and used for analysis afterwards.

2.3 Results

2.3.1 Extraction of task demands

We applied the Minimum Volume Enclosing Ellipsoid (MVEE) approach (Moshtagh et al. (2005)) to fit the force ellipsoid. The optimal task-demanded force direction extracted from task force ellipsoids is $\boldsymbol{f}^* = (-0.09 \pm 0.09, 0.19 \pm 0.09, -0.97 \pm 0.02)^\mathsf{T}$, as Fig. 2.5 shows. Thus, we used the unit vector $\boldsymbol{u}_\mathrm{L}^* \triangleq \boldsymbol{f}_\mathrm{XY}^* = (0, 0, -1)^\mathsf{T}$ for left hand, and the normalized component on the horizontal plane, $\boldsymbol{u}_\mathrm{R}^* \triangleq \boldsymbol{f}_\mathrm{Z}^* = (-0.42, 0.91, 0)^\mathsf{T}$ for right hand, respectively.

(a) Manipulability metric ω^h (b) Task compatability metric α^h (c) Comfort metric ϵ^h
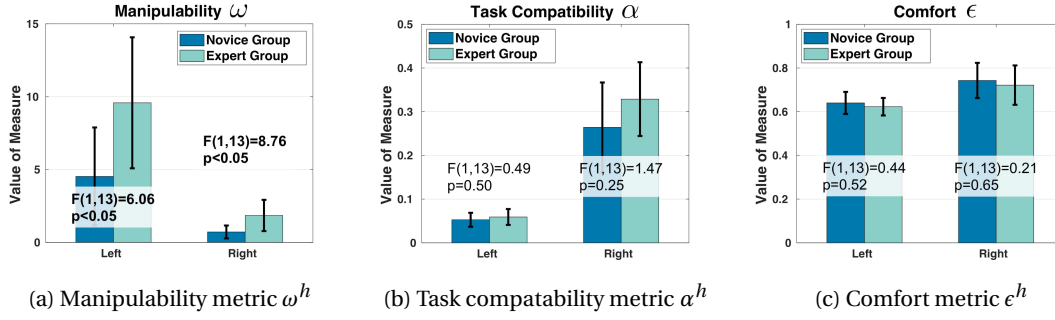
Figure 2.6 – Kinematic measures of hand.

	ω^h		α^h		ϵ^h	
	p	F	p	F	p	F
N	< 0.01	12.64	< 0.01	41.46	< 0.01	11.67
E	< 0.01	13.99	< 0.01	48.78	0.055	5.03

Table 2.1 – Inter-limb comparison of hand kinematic measures for both groups. N: novice group, degrees of freedom of F-distribution $(1, 18)$. E: expert group, degrees of freedom of F-distribution $(1, 8)$.

2.3.2 Analysis of hand poses

The kinematic metrics (ω^h, α^h, and ϵ^h) introduced in Sec. 2.2.7 are calculated for both hands of each subject in hand coordinate frame, which overlaps the end-effector coordination frame (at hand center position) of the arm model. One-way ANOVA was applied to compare performance of subjects in the two groups.

Task compatibility measure α^h and comfort measure ϵ^h are not significantly different between groups for both hands. The two groups differ only in the singularity metric ω^h of hands. Experts have significant larger manipulability metric values than novice subjects for both hands (see Fig. 2.6a).

Inter-limb analysis revealed significance between hands within group. For singularity metric ω^h, both groups have much larger values in left hand ($\omega^h_{N,L} = 4.53 \pm 3.36$, $\omega^h_{E,L} = 9.58 \pm 4.50$) comparing to right hand ($\omega^h_{N,R} = 0.72 \pm 0.44$, $\omega^h_{E,L} = 1.85 \pm 1.07$) at 0.01 significance level (see Table 2.1). While for task compatibility measure α^h, measures of right hands ($\alpha^h_{N,R} = 0.26 \pm 0.10$, $\alpha^h_{E,R} = 0.33 \pm 0.08$) are much larger than left hands ($\alpha^h_{N,L} = 0.05 \pm 0.02$, $\alpha^h_{E,L} = 0.06 \pm 0.02$) for both groups. Subjects in novice group show slightly larger comfort metric ϵ^h in right hand ($\epsilon^h_{N,R} = 0.74 \pm 0.08$) than in left hand ($\epsilon^h_{N,L} = 0.64 \pm 0.05$). This difference is not significant in expert group ($\epsilon^h_{E,L} = 0.62 \pm 0.04$, $\epsilon^h_{E,R} = 0.72 \pm 0.09$).

To gain more insights on this point, we compared hand poses employed by subjects during manipulation, as Fig. 2.7 shows.

(a) Hand pose type I-Left

(b) Hand pose type II-Left

(c) Hand pose type III-Left

(d) Hand pose type I-Right

(e) Hand pose type II-Right

(f) Hand pose type III-Right
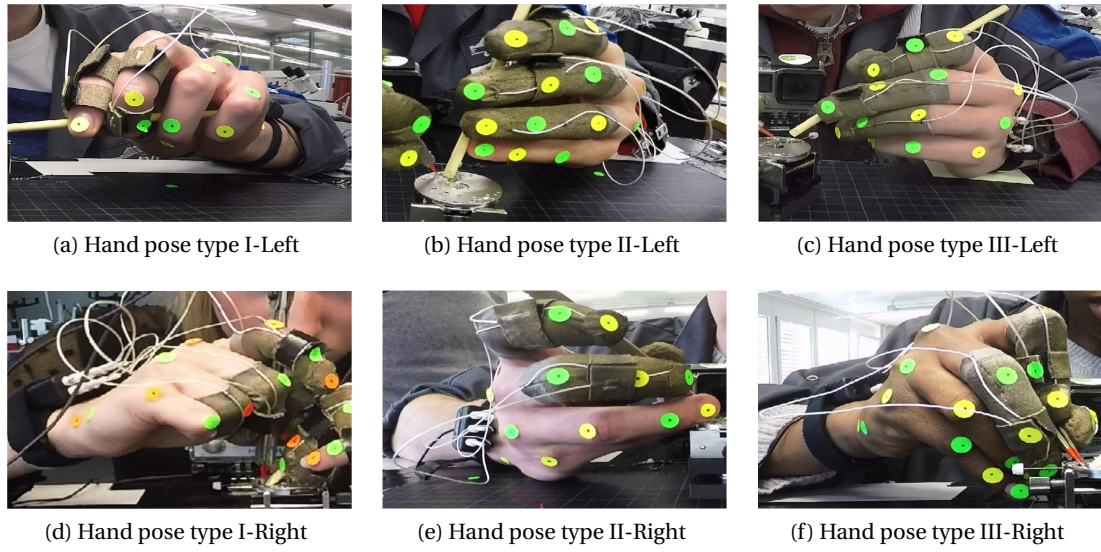
Figure 2.7 – Comparison of typical hand poses observed in experiments.

(a) Manipulability meric ω^a

(b) Task compatibility meric α^a

(c) Comfort meric ϵ^a
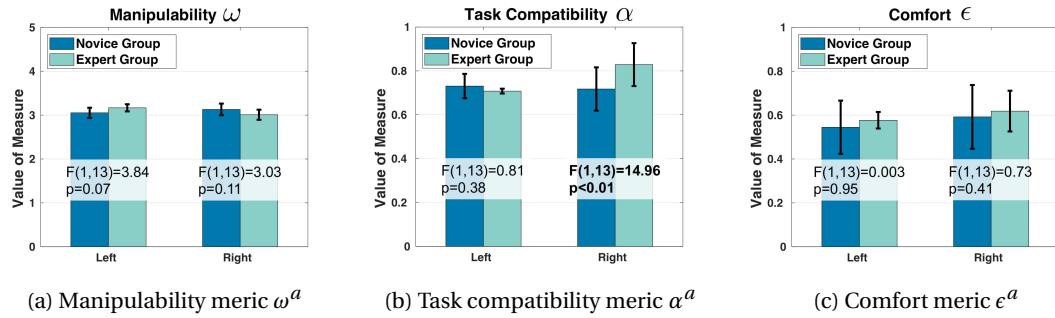
Figure 2.8 – Kinematic measures of arm.

For left hand, most novice subjects prone to use hand poses close to either pose I-Left (Fig. 2.7a, fingers encompass the tool) or pose II-Left (Fig. 2.7b, thumb and other fingers press oppositely). In contrast, expert subjects are more likely to use hand pose III-Left (Fig. 2.7c) to manipulate the peg wood. The differences across hand poses are revealed in the joint postures of each hand. Expert subjects have slightly larger angle ($35.87^{\circ} \pm 9.54^{\circ}$) in the extension/flexion DOF of index fingers' DIP joint than novice subjects ($21.86^{\circ} \pm 8.87^{\circ}$) ($F_{1,13} = 7.94$, $p < 0.05$), while smaller angles of the extension/flexion DOF of the PIP joint for the ring finger (novice: $50.61^{\circ} \pm 19.47^{\circ}$, expert: $19.29^{\circ} \pm 17.55^{\circ}$, $F_{1,13} = 9.16$, $p < 0.01$) and the little finger (novice: $52.52^{\circ} \pm 19.53^{\circ}$, expert: $22.78^{\circ} \pm 27.18^{\circ}$, $F_{1,13} = 6.00$, $p < 0.05$). For the right hand, six out of ten novice subjects have hand poses I-Right (Fig. 2.7d) or II-Right (Fig. 2.7e), and the others use hand pose III-Right (Fig. 2.7f). In comparison, most expert subjects use hand pose III-Right. No significant difference in right-hand joints is observed.

	ω^a		α^a		ϵ^a	
	p	F	p	F	p	F
N	0.18	1.94	0.71	0.14	0.44	0.62
E	**< 0.05**	**6.28**	**< 0.05**	**7.52**	0.38	0.86

Table 2.2 – Inter-limb comparison of arm kinematic measures for both groups. N: novice group, degrees of freedom of F-distribution $(1, 18)$. E: expert group, degrees of freedom of F-distribution $(1, 8)$.

2.3.3 Analysis of arm postures

We compared the joint angle values of the arm model of all subjects. Subjects from both groups do not differ in arm postures (see Fig. 2.9) and have similar kinematic measure values in terms of manipulability (see Fig. 2.8a) and comfort (see Fig. 2.8c). However, the right arm configurations of expert subjects have a larger force transmission ratio ($\alpha_{E,R}^a = 0.83 \pm 0.09$) than the novice subjects ($\alpha_{N,R}^a = 0.69 \pm 0.04$) on average, at 0.01 significance level (see Fig. 2.8).

This difference also holds for the projected 2D component of the force ellipse's major axis on the $X - Y$ plane (see Fig. 2.3b and Fig. 2.3c). For left arm, novice group has an α^a value of 0.72 ± 0.03 and expert group 0.71 ± 0.02 ($p = 0.37$, $F_{1,13} = 0.98$). For right arm, the α^a value is 0.69 ± 0.05 for novice group, and 0.81 ± 0.10 for expert group. This difference is observed at 0.05 significance level ($p = 0.0127$, $F_{1,13} = 8.54$).

Inter-limb comparison indicates that expert subjects have slightly larger manipulability index in left arm ($\omega_{E,L}^a = 3.17 \pm 0.08$, $\omega_{E,R}^a = 3.01 \pm 0.11$), while larger task compatibility in right arm ($\alpha_{E,L}^a = 0.71 \pm 0.01$, $\alpha_{E,R}^a = 0.83 \pm 0.09$), both at 0.05 level (see Table 2.2). The comfort measure does not differ between limbs ($\omega_{E,L}^a = 0.58 \pm 0.01$, $\omega_{E,R}^a = 0.66 \pm 0.04$).

No inter-limb difference is observed in novice group, see Table 2.2 (manipulability: $\omega_{N,L}^a = 3.05 \pm 0.12$, $\omega_{N,R}^a = 3.13 \pm 0.13$; task compatibility: $\alpha_{N,L}^a = 0.73 \pm 0.06$, $\alpha_{N,R}^a = 0.69 \pm 0.04$; comfort: $\epsilon_{N,L}^a = 0.58 \pm 0.07$, $\epsilon_{N,R}^a = 0.59 \pm 0.15$).

2.3.4 Alignment between limb postures and task demands

To identify the factors that affect the task-compatibility, we analyzed the rotation angle of watch face, which affects the task-demanded direction $\boldsymbol{u}$. The angle θ is estimated from the pose of the watch face (see Fig. 2.3a). Slight divergence ($|\Delta\theta| < 10^\circ$) in θ is observed cross motion segments. To better illustrate the distribution of θ, the mean values of θ in all motion segments from both groups were classified into eight main directions, as Fig. 2.10 shows.

We analyzed subjects' postures by calculating three angles: (1) the rotation angle of the watch face θ, (2) the optimal force application direction ϑ (the included angle between the major axis of force ellipsoid and X_W^+), and (3) the angle δ between the task-demanded force

(a) Joint angles of the left arm   (b) Joint angles of the right arm
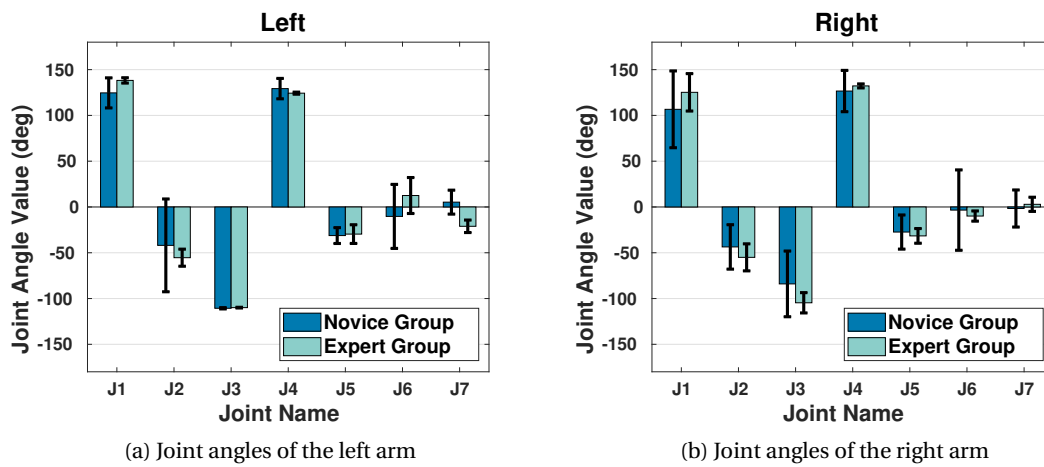
Figure 2.9 – Joint angles of arm model obtained by inverse kinematics. The height of each bar indicates the mean joint angle value, and the error bar indicates the standard deviation.

(a) Watch face rotation angles observed in novice group   (b) Watch face rotation angles observed in expert group
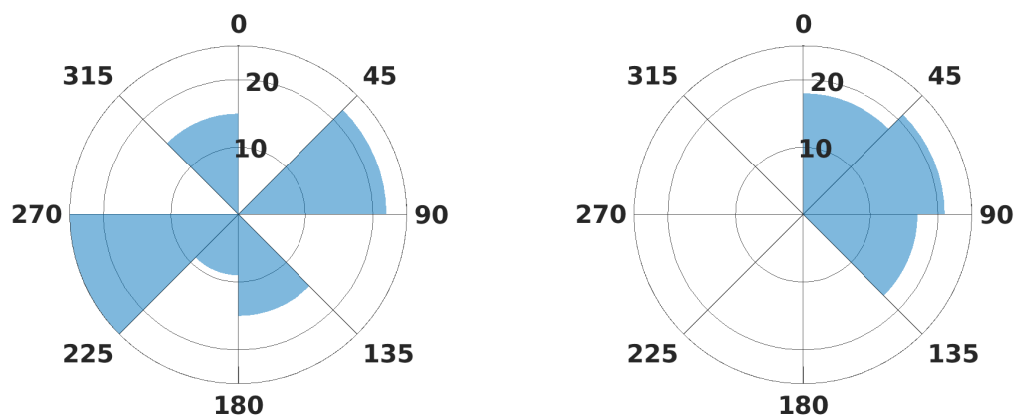
Figure 2.10 – Summary of watch face rotation angles θ observed in experiments. The radius of each bin indicates the number of motion trials that fall into this bin.

	θ	ϑ (L)	ϑ (R)	δ (L)	δ (R)
Novice Group	-43.88 ± 74.93	-60.21 ± 48.99	-137.46 ± 40.41	88.96 ± 14.03	70.48 ± 20.41
Expert Group	39.91 ± 35.36	-44.84 ± 12.02	-151.24 ± 18.64	85.63 ± 17.65	29.74 ± 2.31
p	**< 0.05**	0.50	0.48	0.54	**< 0.01**
$F_{1,13}$	**5.48**	0.46	0.51	0.40	**15.12**

Table 2.3 – Kinematic coordination angles (in degree) of both arms (L: left, R: right). Notice that θ is directed angle ($\theta > 0$ for anti-clockwise rotation and $\theta < 0$ for clockwise rotation). δ is the included angle between major axis of force ellipsoid and $\boldsymbol{u}^*$ and is undirected. We calculate δ in $[0, 90^\circ]$ due to the symmetry of major axes of force ellipsoid. The calculation of φ uses the extracted task-demand vector $\boldsymbol{u}^* \triangleq \boldsymbol{f}^*$ and is obtained as $\varphi = 114.83^\circ$ (in the task coordinate frame).

application direction $\boldsymbol{f}^*$ and the major axis of the force ellipsoid. A geometrical illustration of these quantities is provided in Fig. 2.3d, and the joint angles are compared in Table 2.3. No significant difference is observed in ϑ. This indicates that the optimal force application directions are similar in both groups. However, θ is different between two groups at 0.05 significance level. A significant difference in values for δ is also observed between groups ($F_{1,13} = 15.12$, $p < 0.01$).

2.3.5 Identification of optimal criterion

Global sensitivity analysis suggests that sensitive model joints include arm circumduction (joint 2), forearm extension/flexion (joint 4), extension/flexion (joint 5), and supination/pronation (joint 6) for both arms. These DOFs are considered as state variables in the bi-level optimization problem. Besides, we consider the rotation angle θ as one state variable. We use the human joint angles found through inverse kinematics as the default values for unused joint angles. The lower level of the IOP converges within $150 \sim 200$ iterations (stopping criterion: objective function value stopped decreasing, namely decrease is lower than e^{-5} and the objective function has reach a low enough value, namely within an order of magnitude of -1). The upper level optimization problem is evaluated for 500 trials.

The weights obtained by solving the IOP for the objective function of each subject are reported in Fig. 2.11. Notice that the values are normalized and only indicate relative ratios instead of absolute weights of each cost. An intuitive explanation of a large weight is that the subject makes more effort to reduce/increase the corresponding cost, compared to other costs.

Fig. 2.12 illustrates the average weights for both groups. For the novice group, major contribution to overall cost measures come from task compatibility ($c_{N,L}^\alpha = 35.73\% \pm 6.98\%$) and manipulability ($c_{N,L}^\omega = 33.38\% \pm 11.28\%$) for the left arm, while manipulability ($c_{N,R}^\omega = 39.31\% \pm 12.58\%$) and task compatibility ($c_{N,R}^\alpha = 34.86\% \pm 9.45\%$) for the right arm.

Cost related to increasing comfort is, however, less important for both arms ($c_{N,L}^\epsilon = 30.89\% \pm$

(a) Calculated cost weights for left hand

(b) Calculated cost weights for right hand
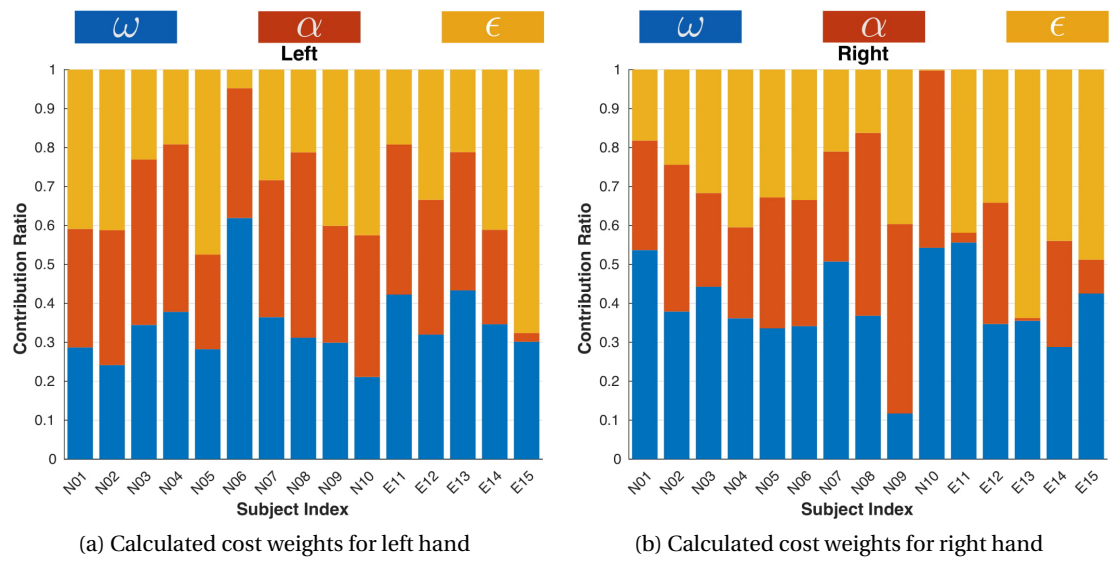
Figure 2.11 – Normalized cost weights calculated for each subject. The prefix of subject number "N" denotes the novice subject (N01-N10), and "E" denotes the expert subject (E11-E15).
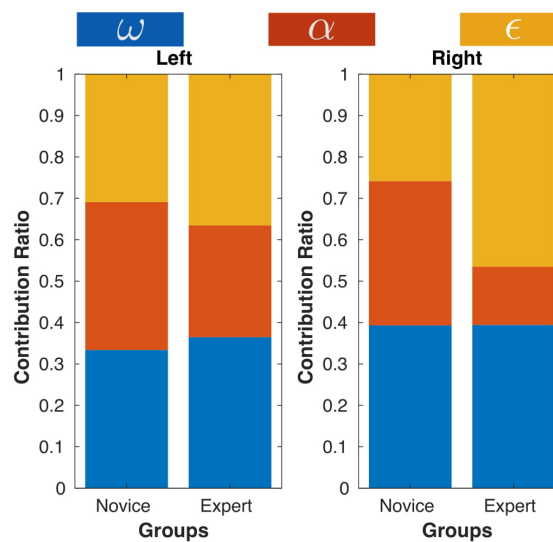
Figure 2.12 – Averaged cost weights calculated for each group.

Measure	ω		α		ϵ	
Arm	L	R	L	R	L	R
p	0.58	0.99	0.14	**< 0.01**	0.52	**< 0.01**
$F_{1,13}$	0.32	0.0003	2.49	**11.65**	0.43	**9.97**

Table 2.4 – Comparison of weight coefficients of kinematics costs between groups (L: left, R: right).

13.68%, $c_{N,R}^{\epsilon} = 25.83\% \pm 12.38\%$). Expert subjects have similar contributions to overall cost for left arm ($c_{E,L}^{\omega} = 36.46\% \pm 6.00\%$, $c_{E,L}^{\alpha} = 27.02\% \pm 14.87\%$, and $c_{E,L}^{\epsilon} = 36.52\% \pm 19.58\%$). However, for right arm, the major cost related to increasing comfort ($c_{E,R}^{\epsilon} = 46.53\% \pm 10.99\%$), which is much higher than novice groups, indicated by a difference at 0.01 significance level ($F_{1,13} = 9.97$, $p < 0.01$, see Table 2.4). The task compatibility cost has a much smaller contribution ($c_{E,R}^{\alpha} = 14.06\% \pm 14.20\%$) than that of the novice group ($F_{1,13} = 11.65$, $p < 0.01$). The contribution of manipulability ($c_{E,R}^{\omega} = 39.41\% \pm 10.28\%$) is rather similar to novice group ($F_{1,13} = 0.0003$, $p = 0.99$).

2.4 Discussion and conclusions

The present study focused on the human kinematic coordination patterns in a bimanual fine-manipulation task. The task requires coordination of both hands and arms in a master-slave relationship, resulting in motions of asymmetric and non-cyclic nature. Kinematics of upper limbs (hand poses, arm postures) plays a crucial role in human bimanual motions, and postures seem to be planned in advance rather than fortuitous (Grea et al. (2000)). As we saw in our study, experts spent less time on task execution than novices on average (expert: 45.88±36.46s, novice: 84.36±54.86s), but dedicated more time on adjusting hand/arm postures and poses of watch faces before manipulation (expert: 9.57 ± 6.76s, novice: 6.38 ± 4.79s). This shows that expert subjects prepare better and end up with efficient execution, indicating an understanding of the critical role of motion kinematics.

The results of our analysis demonstrate the influence of extrinsic conditions (i.e., task conditions) in the generation of coordination patterns within the dynamic pattern theoretical framework (Schöner and Kelso (1988); Kelso (1997)). Experiments show that subjects learn to take advantage of task conditions as they improve their skills.

2.4.1 Dexterity: a marker of skill

The *dexterity* of human hand is a rather abstract concept that encompasses multiple aspects, with which one can associate multiple metrics (Kim and Khosla (1991)). In this study, we used the manipulability index as a metric of dexterity, as it reveals the intrinsic ability of moving fingers to manipulate tools. Our analysis indicates a higher hand manipulability in expert subjects than in novice subjects on average (see Fig. 2.6). We attributed differences

in manipulability across groups to the distinct hand poses. Hand pose I (see Fig. 2.7) is a power grasp that guarantees a strong stability of the held tool. Groups differed also in the distribution of contact along fingers and tools: the tool is en-wrapped by fingers in pose I, while mainly contacted only by fingertips in pose II, which provide higher mobility of the tool. In contrast, hand pose III (mainly used by experts) requires primarily contact on thumb, index, and middle fingers, providing a higher flexibility of fingers (see Fig. 2.6a) and is more suitable for dexterous manipulation. This hand pose is similar to the pose used for grasping a pencil to write, as discussed in (Kelso and Clark (1982); Li and Sastry (1988)). It hence seems that as subjects accumulate more knowledge of the task as a result of practice, they tend to modulate their finger placements to increase the dexterity. They move away from highly stable power grasp poses to poses that ensure a higher flexibility of the tool.

Comparison of left and right-hand poses within the group reveals significant differences of kinematic properties, as a result of distinct hand poses. Subjects in both groups used more fingertip grasping poses for holding peg wood (left hand). In contrast, more contacts on intermediate and proximal phalanges were made to enfold the tweezers (right hand). Fingertip contacts enable higher flexibility of finger motion, correspond to a higher manipulability measure ω^h of left hands on average. This selection of hand pose is essentially determined by the requirements of manipulating the tools. Left-hand poses of higher ω^h enable flexible adjustment of peg wood to maintain the watch spring stable, while right-hand poses of lower ω^h guarantee the stability of manipulating tools by restricting the finger motion.

The significant inter-limb differences of hand poses in both groups indicates that the selection of hand poses depends largely on the manipulation tools, thus highly task-oriented. This finding provides evidence for the task-based grasp planning of human CNS and in consistency with the study of (Friedman and Flash (2007)). Furthermore, the difference observed between groups implies the influence of skill levels (or task proficiency) on hand pose selection. Expert subjects manifest the flexibility of placing fingers on the tools (indicated by significant higher ω^h value than novice group), under the premise of maintaining the ability to satisfy task demands (slightly larger α^h than novice group). We thus conclude that the coordination mechanisms exploited by human to control fingers for tool use is not only affected by the potential functionality provided by the hand pose, but also influenced by skill level, which reflects the subject's understanding of the task demands.

It is noteworthy that although power grasps better stabilize the tools, they do not guarantee higher task compatibility (see Fig. 2.6b). The reason is that the metric α only reveals the ability to transfer joint torque to operational space force. If we consider the left hand pose as an example, increasing finger joint torque in power grasp (hand pose I) empowers the hand to grasp the tools more tightly, but does not generate larger force in the task-relevant direction. To compensate for the insufficient force generated by the hand, we observed that novices subjects would instead adjust arm postures (e.g. rotating wrist, the 6th and 7th joints, see Fig. 2.9) to increase the force generated by the hand. This is confirmed by the slightly higher α^a value and lower ϵ^a in novice group compared to the expert group (see Fig. 2.8b and Fig. 2.8c).

Moreover, the adjustment of wrist joints also causes a lower manipulability of dominant arm (right) than assistant arm (left) as reported in expert group (see Fig. 2.8a). It benefits the task execution in the sense that the dominant arm is less flexible in motion and obtain higher stability of holding tools as consequence. This is also consistent with the analysis of hand poses (see Fig. 2.6a).

The fact that experts' hand pose allows for a better transmission of force in the task-relevant direction results in less demand on the arm. This reduction on arm joint efforts is beneficial particularly in manipulation tasks that can last hours, such as watchmaking. Our analysis indicates, hence, that difference in hand and arm pose across subjects is mainly to be attributed to task demands (direction of $\boldsymbol{u}$) and, to a lesser extent, to the comfort of arm joint angles. Such goal-directed arm and hand movements is a particularly characteristic of our results.

2.4.2 Goal-directed coordination patterns

A few studies discussed generation of goal-directed coordination pattern. For instance, when reaching for grasping a moving target, arm and hand motion are tightly coupled to the target's location and orientation (Jeannerod et al. (1995)). Bingham et al. (2008) suggested that the coordinate timing of bimanual movements to distinct targets is biased towards synchrony and also influenced by visual information. In these studies, however, task conditions (i.e., spatial position of the reaching target) are *extrinsic* to the generated coordination patterns. In this sense, the generation of coordination is essentially an adaptation of the biomechanics of human subjects' musculoskeletal system to the external task conditions. In our study, subjects could actively change task conditions by re-orienting the watch. This is revealed through an analysis of the distribution of θ (see Table 2.3). Results indicate that expert subjects take advantage of the fact that the watch can be re-oriented (see Fig. 2.4) and hence modify the task conditions to achieve a better external coordination. In consistency with this finding is the higher manipulability (ω^a) of the left arm compared to the right arm observed in the expert group (see Table 2.2), indicating that arm postures that enable higher motion flexibility.

We conclude that expert subjects differ from novice subjects in the sense that they exploit adjustable task conditions to "wisely" modify the task-demands, i.e., they rotate the watch face towards a configuration that is more convenient for them to perform the task, instead of adjusting their arm postures to adapt to the original task conditions. From this perspective, "coordination" in our study exceeds the notion of coordination across hand joints and arm joints (i.e., intrinsic coordination), and also entails coordination with the manipulated objects (i.e., extrinsic coordination). This leads to the question of which reference frame is used for control and hence underpin coordination. As pointed out by (Swinnen (2002)), switching across reference frames helps to mediate multiple constraints of coordination. Task conditions may affect CNS's choice of reference frames for movements encoding, such as egocentric (intrinsic) reference, or allocentric (extrinsic) reference (Swinnen et al. (1997)). Our analysis therefore suggests that control of the arms may be performed in a frame of reference attached

to the watch and that, prior to perform the task, the CNS computes an optimal re-orientation of the task frame of reference in global coordinates.

2.4.3 Understanding acquisition of coordination under the optimal control theoretical framework

The present study is unique in the sense that for the first time an inverse optimal control approach is used to investigate the confluence of multiple factors in the generation of kinematic coordination patterns. Unlike previous studies, in which (inverse) optimal control approach is exploited to analyze the human motion principles, we consider the proficiency of human subjects in the IOP to infer the importance of the factors of interest (i.e., elementary costs) that affect the coordination patterns. We make the assumptions that (1) the CNS searches for optimal solutions that balance both intrinsic and extrinsic constraints; and (2) the composition structure of this optimal balance reveals the subject's task proficiency (i.e., skill level). It is likely that the weights associated to satisfying each constraint change as humans learn about the task and improve their ability to perform the task, such as increasing motion adaptability in addition to satisfying task constraints (Seifert et al. (2013)). (Nelson (1983)) also pointed out that skilled movements satisfy more common constraints such as *economy of effort* or *efficiency* in addition to task-oriented objectives.

Comparing the optimal criterion's composition structure between novice and expert groups, we observe that task compatibility cost only contributes a relative small part to the total cost on average in expert groups. However, this does not mean that experts ignore task demands. In fact, the average task compatibility index of expert subjects is higher than the novice subjects on average (see Fig. 2.8). Task demand is hence more important in absolute terms, but less than other costs for experts. We attribute this to experts' ability to modify task conditions, i.e., through the rotation of watch face (Table 2.3), as we discussed in the previous section. Moreover, this decrease in the effort to comply to task conditions is in line with the reduction of *conscious attention* at the last stage of learning, i.e., the amount of conscious attention to movement characteristics reduces. For example, novice drivers pay so much attention to shifting gears that they often miss traffic signs, what rarely happens to expert drivers because the skill of shifting gears is already highly automatized for them (Shinar et al. (1998)).

When comparing costs of right and left arm, we do not see significant difference in the costs in the left (assistant) arm between groups. This is likely due to the fact that the left arm has invariant task demands, despite the change of task conditions (rotation of the watch face). This is because pressing the peg wood downwards does not require directional control on the horizontal plane. In contrast, the motion of the right arm is highly dependent on the task-relevant direction. This is consistent with the calculated kinematic metrics of arms (see Fig. 2.8c). Expert subjects have slightly higher comfort index values of both arms than novice subjects on average, although not significant. These results speak once more in favor of the hypothesis that novice subjects focus on achieving limb postures that can better satisfy task

demands, while expert subjects aim at higher comfort feelings by avoiding stretching arm joints.

On the basis of our analysis, we hypothesize that during the acquisition of novel skills, human subjects first focus on searching for arm postures that satisfy task demands as displayed by novice subjects. Afterwards, as humans acquire task experience and improve their skills, they tend to use more comfortable gestures for task execution, which could be realized by taking advantage of task specifications and adjusting task conditions (e.g. rotating watch face in this study). This is in accordance with the findings in (Seifert et al. (2011)), which shows that experts manifest more adjustment of environmental conditions than novices by exploring properties of frozen water falls during ice climbing. As a result, this leads to increases in both task compatibility and comfort feeling of the musculoskeletal system.

2.4.4 Limitations

Notwithstanding, our current work has limitations that need better clarification.

First, we analyzed kinematic coordination of hands and arms separately. On the one hand, the undirected metrics (ω and ϵ), which are independent of task demands ($\boldsymbol{u}$), can be calculated for hand and arm individually. It could be hard to imply that the singular postures and the comfort feeling of arm may affect the corresponding metrics of hands, and vice versa. On the other hand, no significant difference is observed in the metric α between groups, indicating that the ability to generate task-relevant force (task-compatibility) is not affected by the hand poses. Therefore, we did not consider fingers in the formulation of the inverse optimization problem, this also affords the formulation of a computationally feasible problem

Second, we did not consider individual differences of subjects in analysis. It is, however, known that prior knowledge, personal experience, and individual-specific abilities affect task performance (Ferguson (1956)). Similarly, multiple factors, such as expertise, motivation, and personality, may also affect the learning process (Williams and Hodges (2004)). In the present study, we hypothesized that the costs for novices and experts were the same and that learning would only result in a modulation of the weights.

Furthermore, we formulated our cost components at task level. It is arguable that the formulation of kinematic metrics may not reveal the true low-level costs for the human's musculoskeletal system (Nelson (1983)). The purpose of our study, is to interpret the generation of human upper-limb coordination at multiple high-level aspects (i.e., singularity aspect, task compatibility aspect, or biomechanical aspect). This may also help to reduce the correlation among low-level cost components. For example, costs that comprise joint angle accelerations $\ddot{\theta}$ and the ones that comprise joint angle jerk $\dddot{\theta}$ (Berret et al. (2011)). However, we cannot exclude other aspects that also influence the task performance, which are not taken into consideration in the present work. Further decomposition of our current task-level costs may help to gain deeper insights into the actual structure of the CNS's optimal criteria.

3 | Hand pose selection in a bimanual fine-manipulation task

We have investigated the coordination in upper limbs in the previous chapter. Our analysis reveals that experienced subjects modulate task conditions to benefit manipulation. However, it remains an open question whether subjects would choose distinct strategies under different task conditions. In this chapter, we shift our attention from the posture of upper-limbs to the selection of hand poses and placement of fingers.

We examine the selection of hand poses in a high-precision screwing task, taken from watch-making, with a special focus on the role distribution across hands under different task conditions. We adjust the task difficulty by artificially devising two task conditions: the *free-base* condition and the *fixed-base* condition. Twenty right-handed subjects were divided into two groups to perform manipulation in each condition. Experiments showed that although subjects employed similar hand poses across steps within the same experimental conditions, the hand poses differed significantly in the two conditions. In the *free-base* condition, subjects needed to stabilize the watch face on the table. The role-distribution across hands was strongly influenced by hand dominance: the dominant hand manipulated the tool, whereas the non-dominant hand controlled the additional degrees of freedom that might impair performance. In contrast, in the *fixed-base* condition, the watch face was stationary. Subjects employed both hands even though a single hand would have been sufficient. Importantly, hand poses decoupled the control of task-demanded force and torque across hands, through multiple functional groups of fingers. This preference for bimanual over unimanual control strategy could be an effort to reduce variability caused by mechanical couplings and to alleviate intrinsic sensorimotor processing burdens. To afford analysis of this variety of observations, we develop a novel graphical matrix-based representation of the distribution of hand pose combinations. Atypical hand poses that are not documented in extant hand taxonomies are also included.

The work presented in this chapter has been published in *Yao K, Sternad D, Billard A. Hand pose selection in a bimanual fine-manipulation task. Journal of Neurophysiology. 2021 Jul 1;126(1):195-212.*

3.1 Introduction

Humans are capable of performing a variety of tasks where their two hands cooperate and complement each other; everyday examples include cutting a steak with knife and fork, or opening a bottle cap. Many crafts require exquisitely fine coordination of both hands, from stitching to surgery and watchmaking, where the manipulated tools can be extremely small and, thus, manipulation often requires a microscope. Acquisition and fine-tuning of such skills necessitates many years of practice. In fact, improvements probably never stop, as suggested by a seminal cross-sectional study by (Crossman (1959)) reporting data over many years of experience. When performing bimanual skills, both hands have to adopt intricate poses to cooperatively maneuver target objects with balanced forces and torques applied to achieve the task goals. How the human central nervous system (CNS) controls all the degrees of freedom of the two hands to adopt appropriate hand poses and finger positioning is still poorly understood.

3.1.1 Hand pose selection for tool use

A wealth of evidence shows that the fingers' postures are determined before the hand closes on objects through the pre-shaping of fingers, reflecting synergistic control of the fingers (Mason et al. (2001)). Hand pose selection for tool use is task-dependent, and the same tool may also be held differently depending on the specific sub-task. The selection of hand poses is guided by a variety of factors, including the objective of the manipulation, the task's requirements and the physical properties of the manipulated object (Bohg et al. (2013)), all of which we deem as task demands. Proper choice of hand pose is crucial to task performance. Multiple lines of evidence indicate that our brain recognizes task demands prior to movement execution (Klatzky et al. (1995)) and that hand pose selection is influenced by the objective of the manipulation (De Souza et al. (2015)) and the forces required by the task (Friedman and Flash (2007)). For instance, power grasps that require both fingers and palms, are indicative of tasks requiring high stability or large forces, such as wringing a towel or screwing the cap of a jar. Conversely, pinch grasp poses that use primarily fingertips are typical of tasks requiring fine motor skills, such as writing with a pencil or playing musical instruments.

Human hand pose selection prior to grasping an object is also guided by the purpose of the subsequent manipulation, its end-goal (Ansuini et al. (2006)), the anticipated end-state (Cohen and Rosenbaum (2004)) and the dynamics of grasping (Ansuini et al. (2008)). It is also associated with the target's physical properties, such as size (Tucker and Ellis (2004)), shape (Gentilucci (2002); Daprati and Sirigu (2006)), orientation (Tucker and Ellis (1998)), and spatial location (Johnson-Frey et al. (2004)). Priming of the grasping kinematics has been studied in a variety of task scenarios, including the grasp-to-use and the grasp-to-move (Valyear et al. (2011)). Longer reaction times were reported in the former scenario, indicating a more extensive movement planning process (Armbrüster and Spijkers (2006)). The ability of selecting task-oriented hand poses can be improved with practice (Jeannerod et al. (1994)).

3.1.2 From unimanual to bimanual hand pose selection

While there has been a substantial literature on the role of human hand poses for tool use, the vast majority of studies focused on unimanual shaping of the fingers, in particular, grasping. When the task requires both hands to manipulate the same object, shaping of each hand pose cannot be considered in isolation. Both hands will be recruited in such a way to balance roles and efforts in a symmetric or asymmetric manner. Symmetric and balanced roles are observed when the two hands perform identical motion, e.g. lifting a heavy box. Balanced role distribution has received more attention and has also been studied extensively in rhythmic movements (Kelso (1984); Jirsa and Kelso (2004); Sternad et al. (1992)).

The asymmetric role distribution across the two hands and how these roles change in response to different task demands remains an open challenging question. For example, a knife is held differently when one cuts an apple versus when one peels an apple. Further, the dominant and non-dominant hand adopt different roles. To cut an apple in half, one typically uses the dominant hand to operate the knife, while the non-dominant hand assists by holding the apple in place. Similarly, when peeling an apple, the fingers of the assisting hand rotate the apple to allow the dominant hand to cut the apple. Tasks that require asymmetric and complementary roles between the hands have been referred to as role-differentiated bimanual manipulation tasks (Kimmerle et al. (1995); Babik and Michel (2016)). The dominant hand controls the part of the task that requires higher dexterity and efficiency, while the assisting hand plays an auxiliary role by stabilizing the object (Gonzalez and Nelson (2015); Kimmerle et al. (2010); Guiard (1987)).

Handedness originates from hemispheric differences of the human brain (Toga and Thompson (2003)). Each hand is controlled predominantly by the contralateral hemisphere, yet, both hemispheres are involved when acquiring new bimanual skills, albeit different regions for different functionalities (Theorin and Johansson (2007); Blinch et al. (2019)). According to the dynamic dominance hypothesis, each hemisphere is responsible for different aspects of task performance and optimizes different costs (Sainburg (2002)). In this framework, the control of the dominant limb is based on a feed-forward control mechanism to anticipate the dynamics of the task. In contrast, control of the non-dominant limb is more reactive and relies on sensory feedback to achieve higher positional accuracy (Kagerer (2016)), and is thus better suited for maintaining stability during the task (Sainburg and Kalakanis (2000); Bagesteiro and Sainburg (2003); Goble et al. (2006); Scheidt and Ghez (2007); Sainburg (2020)). Hence, human hand selection seems to be affected by both performance asymmetries and task demands (Przybyla et al. (2013)).

Another view is provided by (Johansson et al. (2006)), who suggest that the brain assigns differentiated roles to the two hands according to spatial relations between the forces that each hand needs to produce to achieve the task goal. Hence, role assignment is no longer solely determined by handedness, nor is it fixed. Instead, the two hands can switch roles as the task evolves. Such an active control of role assignment may be more effective in a large

variety of tasks. For example, switching roles across hands could prevent having to reorient or regrasp the object. This is particularly useful when the external task conditions change (Theorin and Johansson (2010)).

3.1.3 Taxonomies and virtual fingers

Despite the large number of degrees of freedom in hand movements, in daily life, humans tend to use a subset of the possible ways they can shape their hands. To encapsulate these typical hand movements, hand poses have been categorized according to the number of used fingers and their positioning (e.g. precision versus power grasp), and the relation among the used fingers (e.g. relative position of thumb with respect to other fingers). There are three main methods to taxonomize the human grasp (Kamakura et al. (1980); Cutkosky and Wright (1986); Cutkosky et al. (1989); Feix et al. (2015)): (1) Hand-centric and motion-centric manipulation taxonomies (Bullock and Dollar (2011)); (2) the haptic action-focused taxonomy for disassembly tasks (Bloomfield et al. (2003)); and (3) taxonomies to simplify control of robotic hands for prosthetics and rehabilitation (Heinemann et al. (2015)). While all these taxonomies are powerful and representative of a large set of regular hand poses, they only focus on common unimanual tasks and do not account for combinations of hand poses in bimanual tasks. They also have not considered more rare hand poses adopted in special crafts, such as watchmaking, which is discussed in this paper.

One promising explanation to hand and finger shaping is the concept of virtual finger (VF), introduced by Arbib and Iberall and colleagues (Arbib et al. (1985); Iberall (1987a,b)). Instead of focusing on the abundant mechanical degrees of freedom of individual fingers, the virtual finger concept analyzes hand poses at a functional level. Each virtual finger represents one unit of the hand aimed to achieve an independent function, for instance, applying force in a desired direction. One virtual finger may consist of one or more real fingers, or even part of the hand; components of one virtual finger move together to contribute to the same function in the task. For instance, to grasp and lift a bottle of water, the thumb and other fingers are usually placed on the opposite side of the bottle to apply forces in opposition; fingers that press the bottle against the thumb can be considered as one virtual finger. It was suggested that control commands are directly sent from the brain to the virtual finger, and feedback is also received from the virtual finger instead of the real fingers (Arbib and Hoff (1994)). Such an approach explains how the brain may simplify the control of the complex hand by reducing the number of degrees of freedom to match the task's demands. The VF concept was supported through experimental evidence in a series of studies of grasping (Santello et al. (1998); Baud-Bovy and Soechting (2001); Gilster et al. (2012)). In this study, we use the VF concept to describe the positioning of the fingers on the tool in relation to different task demands.

3.1.4 Hypotheses and task considered

This study seeks to understand how different task demands affect the hand pose selection for each hand and the role assignment across the two hands. To this end, we chose a task that requires precise control of position, orientation and force of the tool. Specifically, we selected a precision screwing task that is taken from the behavioral repertoire of Swiss watchmakers, a profession that requires years of training. Screwing is likely the most common task in watchmaking. During both positioning and insertion of the piece, the two hands move in coordination through fast and precise movements. Positioning of the watch face and control of force are crucial to avoid breaking the fragile elements of the watch.

To reduce the problem and make it accessible to experimental study, we selected one watch face and one screw used in the first year of watchmaking training at the École Technique de la Vallée de Joux (ETVJ, a 150-year-old school in watchmaking in Switzerland). The screw was small enough to require the desired precision, but also large enough to be amenable to our subjects, who were knowledgeable in screwing small elements, but were not professional watchmakers.

We hypothesized that hand pose selection and role distribution across the two hands is strongly influenced by *task demands*. Task demands were manipulated by creating restrictions on the manipulated object in two experimental conditions. We assume that the more degrees of freedom required to be controlled by the task, the higher the task demands. In the fixed-base condition, the degrees of freedom of the task were reduced by attaching the watch to the table. In the free-base condition, the watch was free to slide on the table, forcing simultaneous control of both watch and screw, thus corresponding to higher task demands.

Specifically, we hypothesized that the number of degrees of freedom that need to be controlled in the task will affect the choice of hand poses across the two hands. We expected that less constrained tasks, i.e. tasks with fewer degrees of freedom to be controlled, would lead to more variation in the selected combinations of hand poses. Conversely, tasks with more degrees of freedom to control, and for which fewer solutions exist, would lead to less variation in the hand pose combinations and would increase the influence of hand dominance on role distribution.

3.2 Materials and methods

3.2.1 Participants

Twenty participants (age: 24.2 ± 6.0 years, 5 women) were recruited from engineering students at École Polytechnique Fédérale de Lausanne (EPFL) and from the École Technique de la Vallée de Joux (ETVJ). All subjects had experience in using precision screwdrivers, as per their training, but were not proficient at watchmaking. All subjects were right-hand dominant, as assessed by the Edinburgh handedness inventory (Caplan and Mendoza (2011)). They were divided
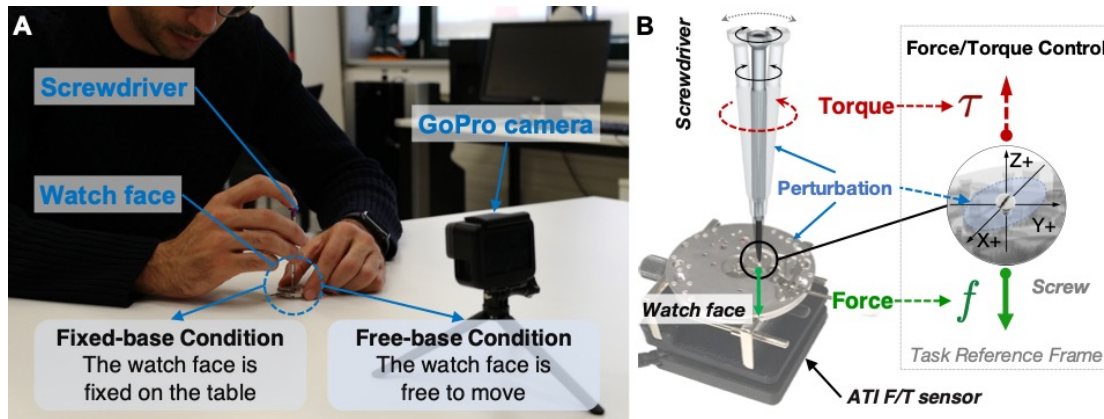
Figure 3.1 – Experimental setup and apparatus. A: One GoPro camera was placed in front of the subjects to record hand movements while they performed the task. B: Original watch face as used for the experiment. In the ex post facto experiments for analyzing hand poses, an ATI F/T sensor was mounted underneath the watch face to measure the applied resultant force and torque.

into two groups with ten subjects each, performing one of the two experimental conditions. Prior to the experiment, subjects were informed about the experimental procedures and gave their written informed consent. The study was approved by the Institutional Ethics Review Board at the École Polytechnique Fédérale de Lausanne.

3.2.2 Apparatus

An original watch face (type ETA 17 JEWELS 5494, diameter 36.6mm) and a flat screwdriver (length: 89.0mm, diameter of tip: 1.6mm) served as experimental apparatus and tool (Fig. 3.1A). The screwdriver is an original jewelers' precision screwdriver. One GoPro camera (GoPro Inc., US) was placed on the table about 20cm in front of the subjects to record hand and finger movements at 60Hz. The camera was mounted so that it could capture a clear front view of the hands and all finger movements.

The experimenter conducted an ex post facto study to analyze the forces and torques generated by different hand poses. To this purpose, an ATI Nano17 miniature force/torque sensor (Fig. 3.1B, resolution: 1/80N for force sensing and 1/16Nmm for torque sensing) was mounted underneath the watch face using a 3D-printed support. The multi-axis sensor measured the forces and torques applied on the watch screw in all three Cartesian coordinates (X, Y, and Z), using as reference the horizontal plane on which the watch face rested (Fig. 3.1B). The sensor captured the torques and forces generated along the Z-axis. An anti-clockwise movement of the screwdriver generated torque in the Z+ direction (the vector in red), and pressing the screwdriver cap led to a force in the Z- direction (the vector in green), opposing the torque's direction. Note that the sensor could only capture the resultant force and torque applied on the watch face, thus the force or torque controlled by each individual hand could not be

Figure 3.2 – Typical hand pose combinations observed in each experimental condition. A: In the fixed-base condition, subjects displayed more diversity in their hand poses. (1) In most cases, one hand (right hand) rotated 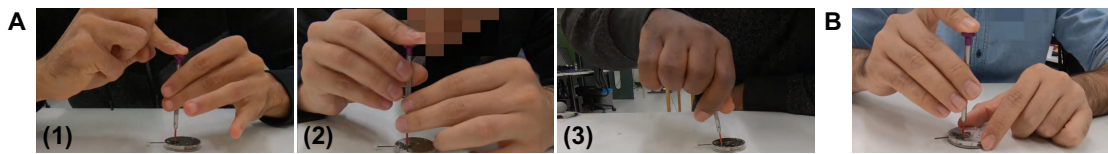the screwdriver, and the other hand (left hand) assisted manipulation by stabilizing the screwdriver. (2) In other cases, the fingers of the two hands moved in coordination, while the right finger provided stabilization. (3) In a few trials, subjects used only one hand. B: In the free-base condition, the dominant (right) hand typically controlled the screwdriver and the non-dominant (left) hand stabilized the watch face.

inferred from the resultant sensor recording. Therefore, subjects performed the task without this sensor in the experiment. The sensor was only used by the experimenter in the ex post facto study for analyzing each one of the observed single hand poses. Analysis of recorded subjects' motion enabled the experimenter to mimic the hand poses adopted by the subjects.

3.2.3 Experimental task

Subjects sat on a chair facing a table with the watch face laid flat on the desk surface (Fig. 3.1A). They were instructed to use a slot-head screwdriver to dismount a tiny screw (Fig. 3.1B, head diameter: 1.6mm, total length: 3.1mm) on the watch face. No specific instructions were given on how they should do this. Regardless of how they completed the task, there were always two steps: (1) the localization step, in which the subject picked up the screwdriver and inserted its tip onto the screw head's slot; (2) the execution step, in which the subject maneuvered the screwdriver to dismount the screw. Subjects had to rotate the screwdriver in an anti-clockwise direction and generate torque in the vertical (+Z) direction (Fig. 3.1B). During this process, a constant downward pressure had to be maintained on the head of the screwdriver to avoid cam-outs. Cam-outs were a common type of failure, where the screwdriver slid out of the groove of the screw head during rotation (Mironov et al. (2018)). It slowed down the task considerably and could even break the tiny watch components. Two experimental conditions were contrasted.

Fixed-base condition

One group of subjects (10 subjects) performed with the watch face immobilized on the table's surface (Fig. 3.2A). In this scenario, subjects only needed to control the six degrees of freedom of the screwdriver and did not need to be concerned about potential movements of the watch face while dismounting the screw.

Free-base condition

Another group of subjects (10 subjects) performed the task where the watch face was free to move on the table surface (Fig. 3.2B). To complete the task, subjects had to prevent potential movements of the watch face while dismounting the screw. This task required the control of not only the six degrees of freedom of the screwdriver, but also three additional degrees of freedom for the translational and the rotational movement of the watch face.

In both conditions, subjects were free to place their hands as they wished, using either one hand or both hands. They had sufficient time to get familiar with the screwdriver before starting the experiment. Before each recording, subjects first mounted the watch screw tightly into its anchor. This part of the action was not monitored, and subjects could take as much time as they wished. The experimenter made sure that the screw was tightly mounted, then the subject started to dismount the screw. Each subject performed five consecutive trials in total. They were encouraged to complete each trial as fast and as accurately as possible. If a cam-out happened, the subject stopped, corrected the failure, and then resumed the task. Task performance was evaluated based on the number of cam-outs, and the average rotation time of the screwdriver. A tightly mounted screw needed to be rotated about five complete rounds until it could be dismounted.

3.2.4 Data analysis

Video analysis and dependent measures

Subjects were video-recorded to determine failure rates, time of finger rotation, and finger positioning on the screwdriver. Note that for the full experiment, only video analysis was available. In the analysis, the continuous video recordings were first segmented into trials, with each segment containing one complete unscrewing manipulation. The start of each trial was defined when the subject picked up the screwdriver. The trial ended when the subject had successfully unmounted the screw and the screw was completely out of its anchor. The segmentation error was maximally ±3 frames, equivalent to 50ms (video was recorded at 60Hz). This accuracy was the experimenter's reported uncertainty in deciding the starting/ending frame of movements.

For each trial, five types of information were extracted. (1) t_s, the start time when the subject had successfully localized and inserted the tip of the screwdriver into the screw head, and was about to rotate the screwdriver. (2) t_e, the end time when the subject stopped rotating the screwdriver, completing the task. (3) N_m, the number of rhythmic finger movements. Each cycle started with finger flexion when the finger contacted the screwdriver and was about to rotate; the movement was completed when the finger extended and stretched back to its original position. (4) N_f, the number of cam-outs or failures that occurred during the trial. (5) The combination of left and right hand poses used for the task, summarized in the hand pose matrix H (see Sec. 2.4.3).
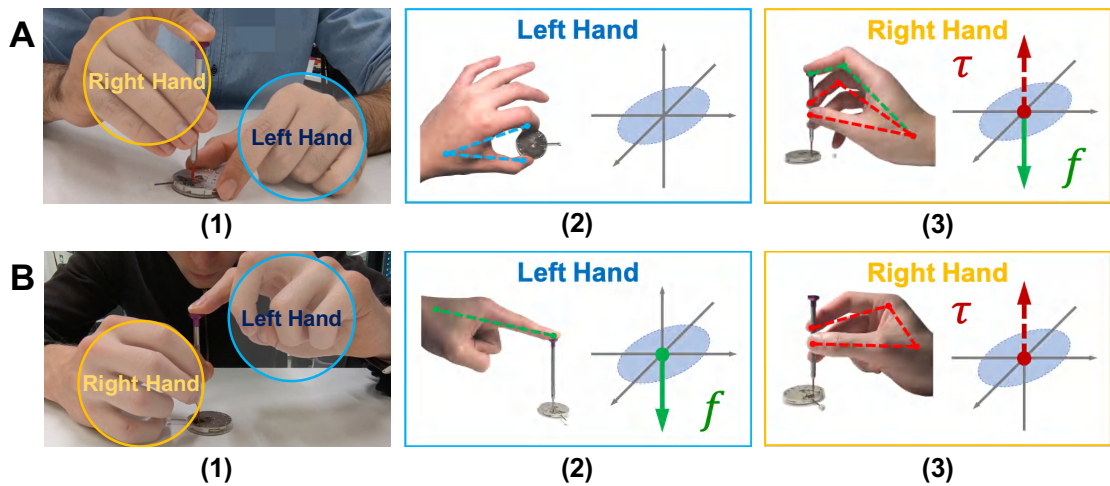
Figure 3.3 – Typical combinations of hand poses observed in A: the free-base condition and B: the fixed-base condition. In each case, (1) shows the observed typical hand pose combination, (2) and (3) illustrate the individual hand pose as well as the task demands controlled by each hand.

Three raters independently annotated the video recordings. Two raters had no experience in the study and were naive to the experimental task, one rater was the experimenter. All raters labeled the hand poses according to the hand pose taxonomy described below. The results demonstrate a 98% match rate among the three raters; only 2% of hand poses were classified into different categories. This occurred when subjects added or removed one active finger during manipulation.

Analysis of task performance

To evaluate the subjects' task performance, two quantitative metrics were defined.

- **Failure rate**
  The average number of cam-outs across trials, δ, served as a descriptive measure for the general difficulty and performance of the task. The number of cam-outs could be larger than the number of trials, as subjects could fail more than once before completing a task trial.

- **Movement time**
  The average finger movement time for each trial, t, calculated as the total trial time divided by the number of rhythmic finger movements N_m per trial. One complete finger rotation included one flexion and one extension of the finger. In case of a cam-out, only the motion segment before the cam-out was used to calculate t of this trial.

Virtual finger analysis

For each hand pose used by subjects, its function was analyzed by mapping virtual fingers (VF) onto the real fingers. Anatomically, one VF may involve one or more real fingers, or even including the palm. Each independent function, e.g., applying force in one direction, was represented by one virtual finger. Therefore, one hand pose could involve several VFs. Following Arbib, Iberall and colleagues, the components within a virtual finger are controlled as one unit (Arbib et al. (1985); Iberall (1987a,b)).

For example, in the typical hand pose combination in the free-base condition (Fig. 3.3A(1)), the function of the subject's left hand (Fig. 3.3A(2)) was achieved by two VFs that constrained the movement of the watch face. Each VF contained one real finger: VF1: thumb, VF2: index finger, both shown in blue dashed lines. The function of the subject's right hand (Fig. 3.3A(3)) was mapped to three VFs: the index finger (VF1, green color) maintained the force in the vertical direction; the middle finger and the ring finger constituted VF2, since they moved in unison to apply force in the same direction on the screwdriver; the thumb was mapped to VF3 (both VF2 and VF3 are in red color), as it moved opposite to VF2 and exerted an opposing force. VF2 and VF3 moved in synergistic coordination and applied opposing forces to rotate the screwdriver. Using this hand pose combination, both task-demanded force and task-demanded torque could be controlled by the three VFs in the right hand.

In the fixed-base condition in Fig. 3.3B(1), the subject used a different manipulation strategy: the subject assigned the left index finger to one VF for maintaining the pressure (Fig. 3.3B(2)). With the right hand (Fig. 3.3B(3)), the subject employed two VFs: the thumb was assigned to VF1, and index and middle fingers were assigned to VF2, since both index and middle fingers were in contact with the screwdriver to apply force in the same direction. These two VFs of the right hand applied forces in opposing directions to rotate the screwdriver, while the task-demanded force could be applied by the VF on the left hand. This analysis of virtual finger assignment was conducted for all observed hand poses, and the results are summarized in Fig. 3.4.

Hand pose taxonomy

Fig. 3.4 summarizes the hand poses adopted by either the dominant or non-dominant hand in both task conditions, both during localization of the tool and execution of the task. The hand poses are ranked according to the number of virtual fingers (#VF) and the number of active real fingers (#AF) involved. Hand pose 1 is task-specific and has not been categorized in the literature. Hand poses 2, 3, 4, 5, and 6 correspond to hand poses reported previously in the GRASP taxonomy by Feix and colleagues (Feix et al. (2015)): inferior pincer (type 33), palmar pinch (type 9), prismatic 2-finger (type 8), prismatic 3-finger (type 7), and prismatic 4-finger (type 6). Pose 7 is adjusted from the stick type (type 29), whereas poses 8, 9, and 10 were also newly observed in this study. These new poses can be considered as integrations of the prismatic grasp poses (type 9, 8, and 7) and the index finger extension pose (type 17).
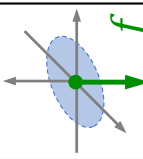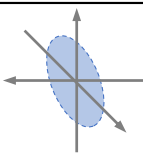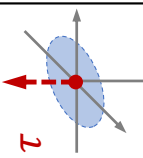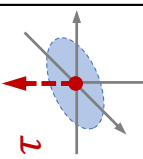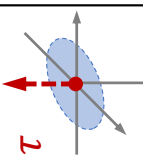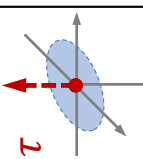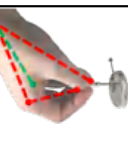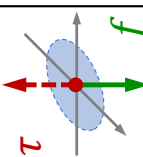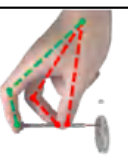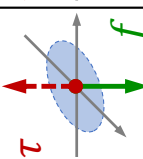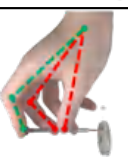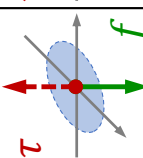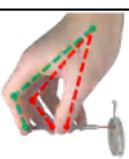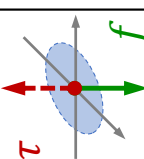
Pose	Description	Illustration	# VF	# AF	F/T Control
1	Index finger (and thumb) pressing		1	1 or 2	f
2	Inferior pincer		2	2 or 3	
3	Palmar pinch		2	2	τ
4	Prismatic 2 finger		2	3	τ
5	Prismatic 3 finger		2	4	τ
6	Prismatic 4 finger		2	5	τ
7	Pinch/tripod with palm pressing		3	2 or 3	τ f
8	Palmar pinch with index pressing		3	3	τ f
9	Prismatic 2 finger with index pressing		3	4	τ f
10	Prismatic 3 finger with index pressing		3	5	τ f

Figure 3.4 – Functional hand pose taxonomy. For each hand pose, the number of virtual fingers (#VF) could be 1, 2, or 3. Fingers are represented by numeric values (1: thumb, 2: index finger; 3: middle finger, 4: ring finger, 5: little finger) and letter P for palm. The number of real fingers used to complete the task was registered as the number of active fingers (#AF). Vectors in the task coordinate frame (F/T Control) represent the force or torque that could be generated by manipulating the tool using the corresponding hand pose. The solid and dashed vectors denote continuous and discontinuous maintenance of force or torque. The light blue ellipsoid indicates the region of potential disturbances. Poses 1, 7, 8, 9, and 10 are specific to tasks in this study and have not been reported previously.

59

Notice that although the right hand is shown for illustration in most cases in Fig. 3.4, all hand poses could be adopted by either hand, regardless of handedness.

Measuring forces and torques of hand poses

The experimenter replicated each of the observed hand poses summarized in Fig. 3.4 with an ATI F/T sensor mounted underneath the watch face for an ex post facto experiment. For each hand pose, the force/torque recorded by the experimenter is summarized in Fig. 3.4 (bottom row, F/T Control). We used dimensionless unit vectors to represent the force (in green color) or torque (in red color) that could be applied in the directions of interest (force in the Z+ direction and torque in the Z- direction). Force and torque components that were likely to occur inside the X-Y plane were perpendicular to the directions of interest, thus they were considered disturbances. We used a dimensionless ellipse (in light blue color) on the X-Y plane to indicate the existence of these potential disturbances.

Hand pose 1 could only apply constant force by exerting axial pressure onto the screwdriver. Hand pose 2 was the only pose used to manipulate the watch face. It stabilized the watch face, but could not control any task demands. Hand poses 3-6 included two VFs that moved in coordination to generate the torque needed to rotate the screwdriver. Hand poses 7-10 controlled both force and torque simultaneously. They generated and maintained continuous force on the screwdriver cap in the Z- direction with either the palm (pose 7) or the index finger (poses 8-10). Torque could only be generated in a discontinuous manner, as fingers that rotated the screwdriver had to move back and forth, and the switch of finger movements interrupted the torque production.

Fig. 3.5 summarizes the hand pose combinations that were observed in both experimental steps and in both experimental conditions. Each individual hand in the combination was categorized according to Fig. 3.4. Hand poses with the same number of virtual fingers were considered to be in the same category in Fig. 3.5, regardless of the number of active fingers. For example, hand pose combinations (0,3), (0,4), and (0,5) were categorized in the same combination in Fig. 3.5 (first column), although they had different numbers of active fingers.

Hand pose matrix

The hand pose matrix H was defined to summarize the hand pose combinations observed in each experimental condition (free-base and fixed-base) and each experimental step (localization and execution). First and second index of each matrix entry indicate the left and right hand pose categories (0-10, according to the taxonomy summarized in Fig. 3.4, 0 for unused hand) of this combination. Each entry's value represents the total number of trials that this hand pose combination was observed, separately for the two steps and the two experimental conditions. Four H matrices distinguished between the localization and execution steps in the free-base and fixed-base conditions: H_{free}^{loc}, H_{free}^{exe}, H_{fixed}^{loc}, and H_{fixed}^{exe}. For example,
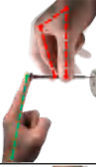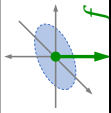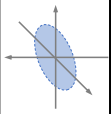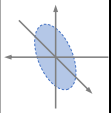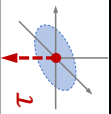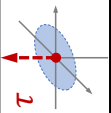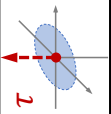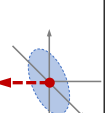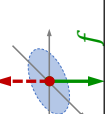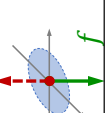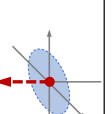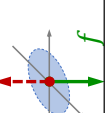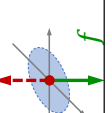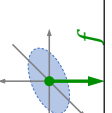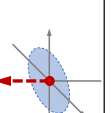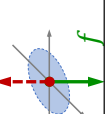
Hand Pose Combination									
Left	Pose	0	0	0	1	2	2	3 or 4 or 5	3 or 4 or 5	3 or 4 or 5
	# VF	0	0	0	1	2	2	2	2	2
	F/T Control				f			τ	τ	τ
Right	Pose	3 or 4 or 5	7	8 or 9 or 10	3 or 4 or 5	7	8 or 9 or 10	1	4	8 or 9 or 10
	# VF	2	3	3	2	3	3	1	2	3
	F/T Control	τ	τ f	τ f	τ	τ f	τ f	f	τ	τ f

Figure 3.5 – Bimanual functional hand pose taxonomy. Combinations of hand poses are categorized according to their virtual fingers analyzed in Fig. 3.4. For each hand pose (left or right), #VF indicates the number of virtual fingers in this pose, and F/T Control illustrates the task condition (force, torque, or the motion of the base) being controlled.

$H_{fixed}^{exe}(2,7) = 3$ indicates that the hand pose combination of left pose 2 and right pose 7 was observed in the execution step in 3 trials of the fixed-base condition.

The hand pose matrix provided a comprehensive overview of all observed hand poses, as well as a visualization of functional distributions across both hands. A symmetric role assignment between hands was reflected in the entries along the diagonal of H, when both hands used identical poses to operate the tool. Role assignment was asymmetric, when both hands used different poses. A right-hand-lead role assignment was reflected by entries in the upper triangular region of the matrix. Conversely, entries located in the lower triangular region corresponded to a left-hand-lead role assignment. Along the diagonal of the matrix, the total number of active fingers in both hands increases from the top left corner (both hands were unused) to the bottom right corner (all fingers from both hands were used). Moreover, unimanual manipulation trials corresponded to the first row (left hand unused) and the first column (right hand unused) of H.

Entropy and structural similarity index

To quantify how sparse the matrix H was, the entropy value, h, was calculated (Cover (1999)):

$$h = -\sum_{i=1}^{n} p_i log(p_i) \tag{3.1}$$

where n is the number of observed different hand pose combinations, which corresponds to the number of H's non-zero entries; p_i denotes the probability of the ith hand pose combination observed in the experiment. A high value of h indicated the task was completed using multiple feasible hand pose combinations. The notation of h follows the same convention as H. For example, the entropy of hand pose matrix H_{free}^{loc} is denoted as h_{free}^{loc}.

The structural similarity index (SSIM) was originally developed to compare the similarity between two images (Wang et al. (2004)). We applied this index to measure structural similarity between two matrices H_X and H_Y, calculated as:

$$s(H_X, H_Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)},$$
$$s \in [-1, 1], \tag{3.2}$$

where μ_X and μ_Y are local means of H_X and H_Y, respectively; σ_X and σ_Y are the standard deviations of H_X and H_Y, respectively; σ_{XY} is the cross-covariance of H_X and H_Y. C_1 and C_2 are regulation constants that depend on the value range of the matrix. They were used to stabilize the division when μ or μ were close to zero. The value of $s(H_X, H_Y)$ ranges from -1 (disparate structures) to $+1$ (identical structures).
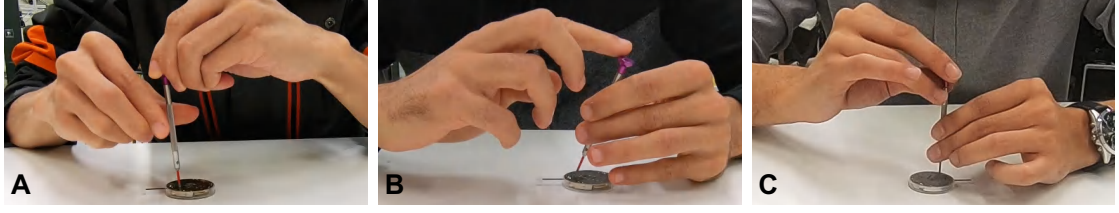
Figure 3.6 – Hand pose combinations observed when failures happen during the execution step under the fixed-base condition. A: $H^{exe}_{fixed}(1,3)$. B: $H^{exe}_{fixed}(4,1)$. C: $H^{exe}_{fixed}(4,8)$.

Statistical analysis

The relation between task performance metrics (failure rate δ and movement time t) were fitted with linear regression. All data were analyzed using MATLAB. A one-way ANOVA was used to evaluate statistical significance in the comparison of the task performance between groups.

3.3 Results

3.3.1 Speed and accuracy of performance

Each subject spent around ten minutes participating in the experiment, including some time for getting familiar with manipulation tools. Subjects were required to repeat the experimental task for only five trials to avoid fatigue of hand and fingers in this intense movement. In each trial, subjects spent around $4.9 - 9.2$s to complete the rotations (6.4s on average).

In the free-base condition, the subjects' average finger movement time was $t_{free} = 0.50 \pm 0.08$s (range: $0.39 - 0.67$s). The subjects' overall failure rate was $\delta_{free} = 0.08 \pm 0.14$ (range: $0 - 0.40$). Fig. 3.6 illustrates three examples of cam-outs. The average movement time was negatively correlated with the average number of cam-outs for all 10 subjects with a correlation coefficient $\rho(t_{free}, \delta_{free}) = -0.68$, suggesting failures were more likely to occur in faster movements.

In the fixed-base condition, the subjects' average finger movement time was $t_{fixed} = 0.36 \pm 0.07$s (range: $0.27 - 0.49$s). The subjects' overall failure rate was $\delta_{fixed} = 0.12 \pm 0.17$ (range: $0 - 0.40$). There was a negative correlation between average movement time and average failure rate for all 10 subjects with $\rho(t_{fixed}, \delta_{fixed}) = -0.66$.

Subjects performed the task significantly faster in the fixed-base condition compared to in the free-base condition. The difference in the average movement times was significantly different from zero ($F_{1,18} = 17.24, p < 0.001$). However, the subjects' overall failure rates were not significantly different in these two conditions ($F_{1,18} = 0.33, p = 0.57$).

Figure 3.7 – Hand pose matrix and the corresponding composition by subjects in the localization step of the free-base condition. A: Hand pose matrix of the free-base condition H_{free}^{loc} visualized as a heat map with the color bar indicating the magnitude of values. Matrix entries are the number of observations. Empty cells indicate no observations. Hand pose combinations in the small figure correspond to one of the categorized combinations in Fig. 3.5. B: Number of observations split by subjects. The X-axis lists the three hand pose combinations observed in this step and condition. The Y-axis represents the number of observations for each hand pose combination. The colors indicate each subject's contribution.

3.3.2 Hand pose analysis

The hand poses in each step of the task (localization and execution) were identified to analyze how hand pose selection was influenced by the number of degrees of freedom that needed to be controlled.

Hand poses during localization

The matrices H_{free}^{loc} and H_{fixed}^{loc} summarized the hand pose combinations observed during the localization step, in both free-base and fixed-base conditions, respectively.

Free-base condition In the free-base condition, subjects consistently utilized the hand pose combinations, represented as matrix entries $H_{free}^{loc}(2,8)$, $H_{free}^{loc}(2,9)$, and $H_{free}^{loc}(2,10)$, respectively (Fig. 3.7A). Hand pose 2 is used by the left hand to stabilize the watch face on the table by using the thumb and index finger to pinch the watch face along its circumference. This pose is combined with the right hand in a prismatic pose, with the index finger pressing in all trials. Among these trials, the right hand's number of active fingers varies from 3 (pose 8, only the thumb and middle finger pinched the screwdriver, accounted for 2%), to 5 (pose 10, the thumb pressed against all the remaining active fingers, accounted for 52%). In the remaining 46% trials, the right hand was in pose 9, and the thumb pressed against the middle and ring fingers.

Figure 3.8 – Hand pose matrix and the corresponding composition by subjects in the localization step of the fixed-base condition. A: Hand pose matrix of the free-base condition H_{fixed}^{loc} visualized as a heat map with the color bar indicating the magnitude of values. Matrix entries are the number of observations. Empty cells indicate no observations. Hand pose combinations in the small figure correspond to one of the categorized combinations in Fig. 3.5. B: Number of observations split by subjects. The X-axis lists all the hand pose combinations observed in this step and experimental condition. The Y-axis represents the number of observations for each hand pose combination. The colors indicate each subject's contribution.

Concurrently, the index finger of the right hand in pose 8, 9, or 10 pressed the screwdriver head to constrain its movement in the axial direction during insertion, while the remaining active fingers of the right hand, i.e. the thumb, middle and ring finger (including the little finger in pose 10) were in contact with the side of the screwdriver to stabilize it. Poses 9 and 10 were used the most across all trials. The palmar pinch pose (pose 8) was used only once. In spite of the differences in the number of active fingers, the right hand poses 8, 9, and 10 shared the same number of virtual fingers: the thumb and the index finger each served as one virtual finger, while the remaining active fingers moved as a group to form the third virtual finger.

The variability of the observed hand pose combinations during localization in the free-base condition was quantified by the entropy metric. The very small number $h_{free}^{loc} = 0.17$ indicated the high concentration of hand pose combinations (Fig. 3.7A). The bar chart in Fig. 3.7B splits the observations by subject and illustrates the hand pose combinations used for localization. The height of each bar represents the number of trials. Four subjects changed their hand poses across trials (S4, S7, S9, and S10), while the remaining subjects used the identical hand pose combination throughout all trials.

Fixed-base condition Fig. 3.8A shows the hand pose matrix for localizing screwdriver in the fixed-base condition. Subjects in up to 32% of trials used the prismatic grasping pose combination $H_{fixed}^{loc}(4,9)$ to operate the screwdriver. The left hand used pose 4 to pinch the tool involving the thumb, index, and middle fingers. Simultaneously, the right hand having more active fingers and virtual fingers adopted pose 9 to pinch and press the screwdriver. In this pose, the index finger pressed on the screwdriver head to restrict its movement along the axial direction when it was inserted into the screw slot, while the thumb, middle, and ring fingers pinched the screwdriver to control the position of the tip. In a few trials, subjects used hand pose combinations similar to $(4,9)$, although with slight adjustments captured in the matrices as $H_{fixed}^{loc}(4,8)$, $H_{fixed}^{loc}(5,9)$, and $H_{fixed}^{loc}(4,10)$. These adjustments included adding or removing active fingers to form variations in the prismatic poses.

In contrast, 28% of localization hand pose combinations used fewer active fingers and also fewer virtual fingers, denoted in the matrix as $H_{fixed}^{loc}(0-1,3-5)$. In such combinations, the left hand was either not used (pose 0, in $H_{fixed}^{loc}(0,4-5)$) or used the index finger (or both thumb and index finger) to press the screwdriver head (pose 1, in $H_{fixed}^{loc}(1,3-5)$). This achieves the same functionality as the index finger of the right hand (pose 9, in $H_{fixed}^{loc}(4,9)$), as discussed above. The right hand in pose 3, 4, or 5 simply pinched the screwdriver without pressing its head. Such combinations resulted in an equivalent virtual finger assignment across hands as hand poses 8, 9, and 10 that were commonly used in the free-base condition (see $H_{free}^{loc}(2,8-10)$). These hand pose combinations had one virtual finger pressing the screwdriver and multiple virtual fingers for holding the screwdriver.

Fig. 3.8B illustrates the observed hand pose combinations split by subjects. The choice of hand pose combinations visibly varied more than in the free-base condition. Seven out of 10 subjects used different hand pose combinations across the trials (S1, S2, S3, S5, S6, S8, and S9). The higher variability of hand pose combination in the fixed-base condition was quantified by the higher entropy value: $h_{fixed}^{loc} = 0.52$.

Hand poses during execution

The matrices H_{free}^{exe} and H_{fixed}^{exe} summarize the hand pose combinations observed during the execution step, in both the free-base and fixed-base conditions, respectively.

Free-base condition For the execution step, there was high consistency in the choice of hand poses both across subjects and trials, as is clear from the matrix in Fig. 3.9A. Subjects used the hand pose combination $H_{free}^{exe}(2,8)$ in 48% trials and $H_{free}^{exe}(2,9)$ in 52% trials. These hand poses were similar to those during localization of the tool (Fig. 3.7A) and revealed a clear predominance of the right hand, concentrated in the upper triangular region of the matrix. In most trials, the right hand formed a prismatic pose, where the index finger pressed the screwdriver head to maintain a downward force when rotating the screwdriver. Simultaneously,
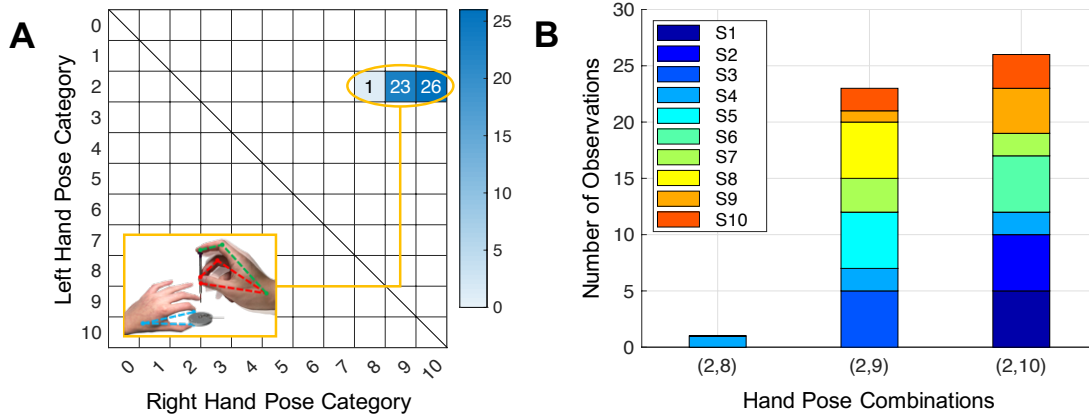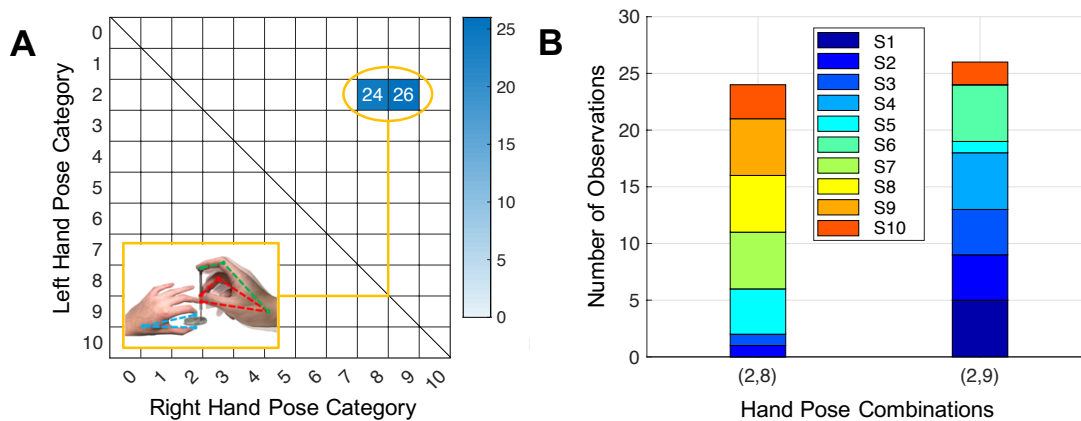
Figure 3.9 – Hand pose matrix and the corresponding composition by subjects in the execution step of the free-base condition. A: Hand pose matrix of the free-base condition H_{fixed}^{exe} visualized as a heat map with the color bar indicating the magnitude of values. Matrix entries are the number of observations. Empty cells indicate no observations. Hand pose combinations in the small figure correspond to one of the categorized combinations in Fig. 3.5. B: Number of observations split by subjects. The X-axis lists all the hand pose combinations observed in this step and experimental condition. The Y-axis represents the number of observations for each hand pose combination. The colors indicate each subject's contribution.

the left hand was in the inferior pincer pose to stabilize the watch face along the circumference. The predominance of these two hand pose combinations is expressed by the small entropy value $h_{free}^{exe} = 0.12$.

This is also visible in the distribution of hand poses across subjects and trials (Fig. 3.9B). Subjects seldom changed hand poses: 6 subjects (S1, S4, S6, S7, S8, and S9) used identical pose combinations in all trials; 3 subjects (S2, S3, and S5) changed only once, and 1 subject (S10) used $H_{free}^{exe}(2,9)$ twice and $H_{free}^{exe}(2,8)$ three times.

Task performances quantified by failure rate (accuracy) and movement time (speed) for each hand pose combination are illustrated respectively in Fig. 3.10A and Fig. 3.10B. The two observed hand pose combinations (Fig. 3.9A) both correspond to an 8% failure rate (Fig. 3.10A) with a similar average movement time of 0.51s and 0.49s (Fig. 3.10B), respectively.

One-way ANOVA analysis of movement time indicated that the subjects who changed hand poses across trials performed the task faster than those who consistently used one hand pose combination ($F_{1,48} = 8.54$, $p = 0.005$). However, the average finger movement time was 0.54 ± 0.15s (mean $\pm$ std) before changing hand poses and 0.54 ± 0.10s after changing hand poses, across all trials. Statistically, there is no significant difference revealed ($F_{1,10} = 0.01$, $p = 0.92$). Moreover, all subjects continued using the identical hand poses after failed trials, and the average failure rate remains on the similar level after changing hand poses ($F_{1,48} = 2.95$, $p = 0.09$). Therefore, the task performance in both aspects are not related to the change of hand
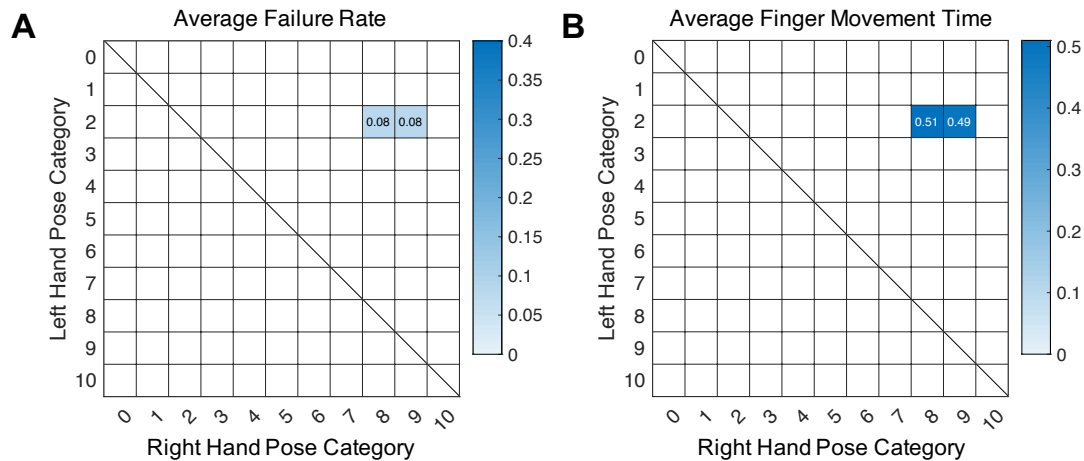
Figure 3.10 – Task performance associated with each execution hand pose combination in the free-base condition. A: Average failure rate. B: Average finger(s) movement time. The shade of the color bar is proportional to the values.
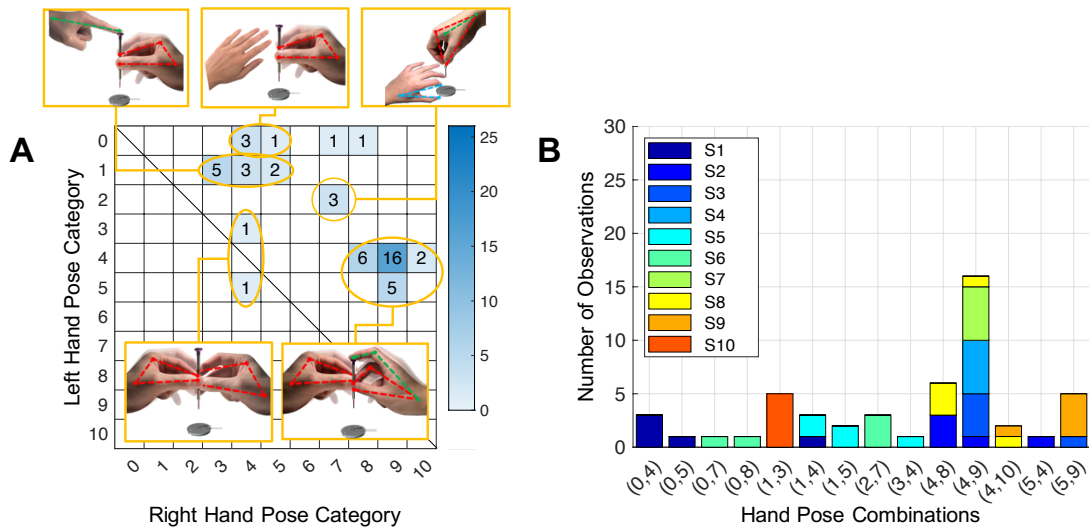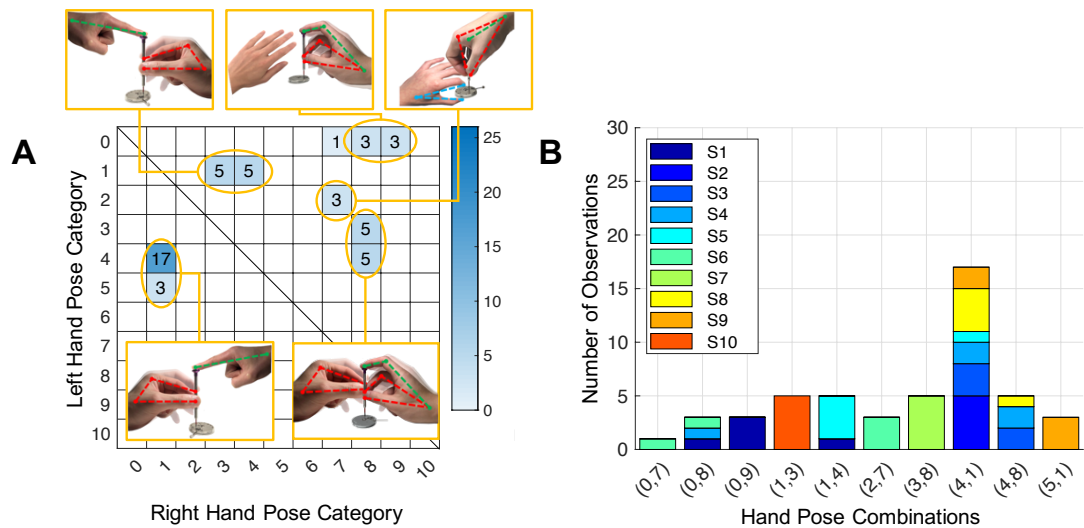

Figure 3.11 – Hand pose matrix and the corresponding composition by subjects in the execution step of the fixed-base condition. A: Hand pose matrix of the free-base condition H_{fixed}^{exe} visualized as a heat map with the color bar indicating the magnitude of values. Matrix entries are the number of observations. Empty cells indicate no observations. Hand pose combinations in the small figure correspond to one of the categorized combinations in Fig. 3.5. B: Number of observations split by subjects. The X-axis lists all the hand pose combinations observed in this step and experimental condition. The Y-axis represents the number of observations for each hand pose combination. The colors indicate each subject's contribution.

poses.
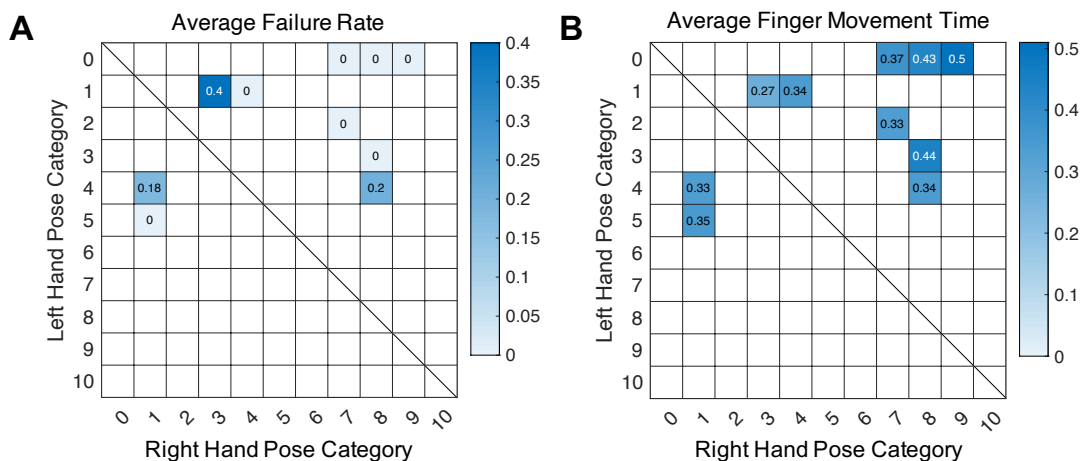
Figure 3.12 – Task performance associated with each execution hand pose combination in the fixed-base condition. A: Average failure rate. B: Average movement time. The shade of the color bar is proportional to the values.

Fixed-base condition In contrast to the free-base condition, hand poses in the fixed-base condition revealed more variation in their role distribution. This is best visualized by the execution hand pose matrix H_{fixed}^{exe}, where the entries are quite dispersed (Fig. 3.11A). This was quantified by the entropy metric $h_{fixed}^{exe} = 0.41$. Inspection of the distribution of hand poses across subjects (Fig. 3.11B) shows that most subjects changed their hand pose combinations between trials. Only 3 subjects (S2, S7, and S10) used the same hand pose combination across all trials. The most frequently observed hand pose combinations were $H_{fixed}^{exe}(4-5,1)$, employed in 20 trials, corresponding to a 40% occurrence rate. The entries are located in the lower-triangular region of the matrix (Fig. 3.11A). In these trials, the right hand was in charge of maintaining the downward force on the screwdriver. Simultaneously, the left hand adopted the prismatic pose to rotate the screwdriver with the thumb, the index finger, and/or the middle finger.

In 30 out of 50 total trials, the right hand rotated the screwdriver, shown in all matrix entries in the upper-triangular region of H_{fixed}^{exe}, while in the remaining 20 trials, the left hand took this role shown by entries in the lower triangular region of H_{fixed}^{exe} (Fig. 3.11A).

Task performances associated with each hand pose combination in the fixed-base condition are summarized in Fig. 3.12A and Fig. 3.12B, respectively. The combination $H_{fixed}^{exe}(1,3)$ had the highest failure rate of 40% (Fig. 3.12A). This occurred when the right hand used only two active fingers (the thumb and the index finger) to rotate the screwdriver. This hand pose is associated with the shortest average movement time of 0.27s (Fig. 3.12B). Notably, the failure rate decreased to 18% if the left hand served as the lead role to execute a similar pose, denoted by $H_{fixed}^{exe}(4,1)$ (Fig. 3.12A). This combination had an average movement time of 0.33s (Fig. 3.12B). The hand pose combination $H_{fixed}^{exe}(4,8)$ was associated with a failure rate of 20%. Using this combination, fingers from both hands alternately rotated the screwdriver,
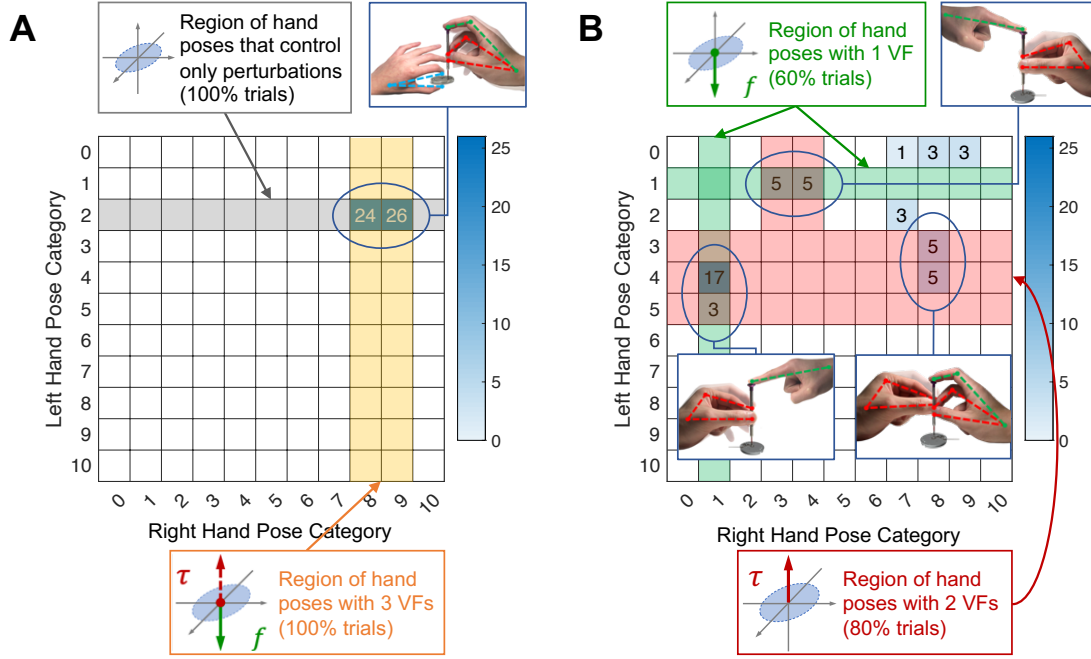
Figure 3.13 – Assignment of virtual fingers in execution hand pose combinations, and their controlled task demands in A: the free-base condition and B: the fixed-base condition. Hand pose combinations in the small figures correspond to the categories in Fig. 3.5. The shade of the color bar is proportional to the values. The hand pose combinations that were only used to control the perturbation are highlighted in gray color; hand pose combinations that are assigned 1, 2, and 3 virtual fingers are highlighted in yellow, green, red, and yellow color, respectively. The number of trials in which such hand pose combinations were used are denoted as percentage.

with the right index finger maintaining downward pressure on the screwdriver head during rotation. Hand pose combinations with a zero failure rate were located in the upper-triangular region of the matrix H^{exe}_{fixed} (Fig. 3.12A). They were associated with the relative longer finger(s) movement time (Fig. 3.12B).

The change of hand poses does not seem to be caused by failure, as only one subject (S4) changed hand pose from $H^{exe}_{fixed}(4,8)$ to $H^{exe}_{fixed}(4,1)$ once after a cam-out, and then used $H^{exe}_{fixed}(4,1)$ in the remaining trials. The average finger movement time was 0.37 ± 0.10s (mean $\pm$ std) across all subjects before changing hand poses and 0.38 ± 0.09s after. Statistics did not reveal any significant differences ($F_{1,20} = 0.01, p = 0.94$). Moreover, no significant difference in task performance has been found between the subjects who consistently used the same hand pose combinations and those who changed hand poses across trials (average finger movement time: $F_{1,48} = 1.49, p = 0.23$, average failure rate: $F_{1,48} = 0.92, p = 0.34$). Thus, task performance does not seem to be affected by the change of hand poses.

Structural similarities and differences of hand pose matrices

In both the free-base and the fixed-base conditions, similar choices of execution and localization hand pose combinations were observed. In the free-base condition, six out of ten subjects maintained the same hand pose combinations throughout all trials (Fig. 3.9B), while in the fixed-base condition, only three subjects used the same hand pose combinations (Fig. 3.11B). The majority of subjects showed diversity in choices and used at least two different hand pose combinations. This observation was quantified by the structural similarity index s, which assessed similarities across distributions of hand pose combinations during localization and execution in the two experimental conditions. We found the highest similarity score between localization and execution steps within the same experimental condition: $s(H_{free}^{loc}, H_{free}^{exe}) = 0.62$ for the free-base condition and $s(H_{fixed}^{loc}, H_{fixed}^{exe}) = 0.18$ for the fixed-base condition. The low similarity scores $s(H_{free}^{loc}, H_{fixed}^{loc}) = 0.09$ and $s(H_{free}^{exe}, H_{fixed}^{exe}) = 0.07$ indicated that hand pose combinations used in both experimental steps in the free-base condition were distinctly different from those in the fixed-base condition.

Virtual fingers assignment

Specific patterns in the hand pose combinations' distribution in hand pose matrices could be observed when paying attention to the functionality of hand poses. This functionality could be assessed by analyzing the virtual finger assignment. In Fig. 3.13, hand pose combinations during execution were regrouped according to their VF assignment. As stated previously, in the free-base condition, subjects displayed consistently similar hand pose combinations with similar VF assignment. Fig. 3.13A illustrates this by showing all hand pose combinations located in the intersection region. The right hand formed three VFs, all of which were dedicated to the control of both task demands (highlighted in yellow). The other hand was in charge of compensating for eventual perturbations, but did not dedicate VFs to controlling explicit task demands (highlighted in gray).

In the fixed-base condition, the pattern of VF assignment was different. VFs in charge of controlling task demands were distributed across the two hands. In 60% of the experimental trials, one of the two hands used one VF to generate the required vertical force (highlighted in green), while the other hand used two VFs to control rotating torques (highlighted in red, Fig. 3.13B). These hand pose combinations correspond to the entries $H_{fixed}^{exe}(4-5,1)$ and $H_{fixed}^{exe}(1,3-4)$. In 20% of trials, corresponding to $H_{fixed}^{exe}(3-4,8)$, the thumb and middle finger from the right hand only played a secondary role to assist the torque control by the lead hand. If only the degrees of freedom of the task are considered and the assisting fingers are ignored, then these results show that subjects divided force and torque across the hands in 80% of the trials. Subjects used one hand to control one independent task demand. This contrasts to the free-base condition where one hand controls both task demands (intersection of highlighted regions, Fig. 3.13A).

3.4 Discussion

To increase our understanding of how different hand poses are selected and combined in a bimanual high-precision manipulation task, this study examined how humans maneuvered a jeweler's screwdriver to dismount a watch screw. Asking subjects to perform the task under two task conditions with different numbers of degrees of freedom to control allowed to determine the effects of the task's demands on their hand pose selection strategies. We hypothesized that tasks with more degrees of freedom for hand allocation would lead to more variability in hand pose combinations. In contrast, tasks with fewer opportunities for variations in hand poses will enhance the influence of hand dominance in role distribution.

As expected, subjects used distinct hand pose combinations to satisfy the control of task demands under both conditions. The fixed-base condition was less constrained and subjects displayed a larger variety of hand pose combinations. In contrast, in the free-base condition, subjects needed to maintain the watch face in a stationary position on the table, which required the control of additional degrees of freedom. These constraints limited hand pose variations, and only a few hand pose combinations could satisfy the control of task demands and successfully complete the task.

We also observed a strong preference for bimanual over unimanual operation when completing the task. In the free-base condition, The left hand provided support by controlling the residual degrees of freedom, while the right hand performed the major manipulation motion. However, this role distribution was modulated when the task's degrees of freedom decreased in the fixed-base condition, allowing for more variations of hand pose combinations. The assisting left hand started to contribute more actively to generate the movement, either by coordinating its finger movements or by balancing the forces with those of the right hand that was in charge of generating the motion. At times, roles across the two hands were switched as subjects explored different hand poses.

To display the diversity of hand pose combinations and ease visual assessment of the role distribution across the two hands, we created a novel visualization approach, the hand pose matrix. Sparseness of the matrix entries denoted low diversity in hand pose combinations, and vice versa. The prevalence of handedness was expressed in the concentration of entries in the upper and lower diagonals in the matrix. To account for hand poses particular to watchmaking, we expanded the extant taxonomies of hand poses. To relate our taxonomy to the task-specified control demands, hand poses were categorized according to their virtual finger assignment in relation to force/torque control abilities. This matrix representation helped to summarize and quantify the unbalanced role distribution across the two hands. This matrix representation is not restricted to this study and can be applied to analyze general bimanual tasks.

3.4.1 Trade-off between speed and accuracy

An analysis of task performance in terms of movement time and failure rate revealed that a decrease in task completion time was associated with an increase in failure rate. This observation is consistent with the widely observed speed-accuracy trade-off, i.e. improvements in accurate performance are achieved at the expense of speed, or vice versa (Fitts (1954); Fitts and Peterson (1964); Plamondon and Alimi (1997); Heitz (2014); Guiard et al. (2011)). This basic finding validated that these realistic and complex data conform to generally accepted performance features.

3.4.2 Task's degrees of freedom and their effects on hand pose selection

Our study contrasted two task conditions to quantify the effect of the task's degrees of freedom on the selection of hand poses. The free-base condition increased the number of degrees of freedom to be controlled in the task, due to the necessity of maintaining the watch face in a stationary position. In this condition, less variability in the hand pose combinations was observed in both the localization and the execution task steps.

Experimental condition: free-base condition

Without exceptions, subjects in the free-base condition adopted the same pose combinations during both task steps with their left hand to stabilize the watch face on the table (Fig. 3.7A and Fig. 3.9A). Given the extra degrees of freedom imposed by the watch face, pose 2 was the only feasible hand pose when the supporting (left) hand controlled these degrees of freedom to maintain the watch face's stability. Using hand pose 2 to control the degrees of freedom from the watch face was, however, not the determining factor for a successful task completion because it did not generate the force and torque required to rotate the screwdriver. Force and torque were then to be generated by the dominant (right) hand. In principle, the right hand could use any hand poses (except pose 1 and 2) to generate the task-demanded force and torque. However, the subjects solely used hand poses 8 and 9. This may be attributed to the fact that hand poses 3-6 led to inconsistent force generation, since these four hand poses did not have a VF specified for the control of force. When using these poses, the only way to generate the necessary force was pressing the screwdriver downwards while rotating them. Hand poses 7-10 employed an extra virtual finger for force control, which corresponded to either the index finger (for poses 8-10) or the palm (for pose 7). These poses delivered constant pressure without affecting rotational movements. Interestingly, subjects did not use hand pose 7 at any stage, despite its feasibility for controlling both task demands. This could be attributed to fatigue caused by lifting the forearm or bending the wrist that was required when using this hand pose. Holding the object in place while inserting a screw is part of a wide range of tasks, including repairing a cell phone, a watch, or glasses. This use of hand pose in the free-base condition may reflect the subjects' habits acquired through life experience (Sainburg (2020)).

Experimental condition: fixed-base condition

In the fixed-base condition, the watch face was mounted on the table. This provided more options for both hands and thus led to a wider selection of hand pose combinations, as quantified by larger entropy values h_{free}^{loc} and h_{free}^{exe}. Primary differences were in the assignment of function: hand poses separated the control of the force and torque across the two hands. In about 60% of trials, one hand (in pose 1) provided downward pressure on the screwdriver, while the other hand rotated the screwdriver.

The observed combinations of hand poses in the free-base condition were rarely seen in the fixed-base condition. Specifically, hand pose 2 was used in the free-base condition to restrain the watch face, and was only used in three out of 50 trials by one particular subject (only in combination with pose 7 from the right hand). In addition, unimanual manipulation manner was employed in 14% trials, with the right hand having the similar functionality as being used in the free-base condition.

Synergies, decoupling of control variables, and hand pose selection

The concept of synergy has been widely used to explain human hand pose, pre-shape, and finger coordination in humans (Santello et al. (1998)) and to control artificial hands in robots (Prattichizzo et al. (2011)). It assumes that the human brain couples hand joints or muscles and controls them in a lower-dimensional space through synergies. Task-specific synergies can be organized at multiple levels, from neural constructs to muscular coordination patterns. There is a wealth of evidence that some synergies develop during childhood, such as the simultaneous control of all fingers for a power grasp, and the coordinated control of thumb and index for a pinch grasp. Such simultaneous control of multiple degrees of freedom, while beneficial to speed up daily object manipulation, limits the independent control of fingers that is necessary for fine-manipulation, as found in crafts and playing musical instruments.

In addition to synergies, there is evidence that fingers are coupled biomechanically. Tendons cross multiple finger joints, and finger range of motion is limited by soft tissues. Individual movements of each finger are impeded by the fact that they share the same group of muscles or are connected by tendons (Li et al. (1998); Zatsiorsky et al. (2000)). In addition to the anatomical linkages (Lang and Schieber (2004)), the specific innervation of adjacent fingers leads to synchronous flexion of adjacent finger joints (Winges et al. (2008)). Such biomechanical coupling affects also the ability to generate force independently with each finger, and force applied by one finger may inadvertently lead other fingers to also passively exert force. This is known as the force-enslaving effect (see (Schieber and Santello (2004)) for a review). In our study, such biomechanical coupling affects performance. Indeed, control of the thumb, the middle, and the ring fingers to generate the rotation of the screwdriver inevitably affects the control of the index finger and makes it more challenging for this finger to maintain a steady and stable downward pressure on the screwdriver to stabilize it. Such interference among fingers in the same hand can be largely reduced, if all fingers work in coordination to

produce the same movement, i.e. mapping all fingers to the same VF. This speaks in favor of control solutions that reduce the number of VFs required for the task, as observed in our data. Therefore, subjects' choice of hand poses in the fixed-base condition is likely the result of the brain trying to reduce undesired variability caused by biomechanical couplings.

Individualized finger control is demanding and requires practice. To achieve individual control of the fingers requires decoupling this neural organization and activating additional muscles to minimize unintended movements of other fingers (Schieber and Santello (2004)). Despite these hurdles, dexterous fingers movements can be attained through practice, as evident in the extraordinary skills of pianists (Furuya et al. (2014); Jäncke (2009)). Intensive practice of these skills can lead to change in the motor cortex' representation of finger digits (Buonomano and Merzenich (1998)). Although producing certain finger coordination patterns becomes easier with training, generating such novel coordination increases intrinsic computational burdens, especially when a large number of degrees of freedom need to be controlled simultaneously, as it is the case in bimanual fine-manipulation tasks. One may hence prefer to not train novel finger coordination patterns or synergies. Instead, one would usually seek alternative approaches, such as decoupling the control variables required for the task into simple and independent components. In this study, the task afforded the decoupling of force and torque production. Such decoupling may simplify control and allow using hand poses from an existing synergy repertoire. Control of each hand is then made easier by the fact that each of these primary hand poses controls one independent function. In our task, however, such decoupling requires the control of both hands. Coordinating both hands to perform the task is not necessarily easier than learning new coordination patterns, especially when this coordination is unusual.

In the free-base condition, subjects used extensive hand pose combinations, with the motion-control hand having three VFs. Controlling these three VFs requires subjects to adopt task-specific postural synergies, some of which may be novel or less trained depending on the subjects' life experience. In contrast, in the fixed-base condition, most hand poses required control of one or at most two VFs. Moreover, most of these hand poses are typical of those adopted in everyday tasks (Feix et al. (2015)). Thus, commonly used synergies may be sufficient for controlling such hand poses. This study did not evaluate cognitive load required to perform the task in either conditions. We can however expect that subjects may have found the free-base condition more challenging to control.

3.4.3 Hand pose selection strategies across localization and execution

In an effort to determine if hand pose selection was influenced by sequential variations in task demands, we also compared hand pose combinations across the two steps of our task: localization and execution. Recall that localization required to control precise positioning of the watch face and the tool, whereas execution denoted the segment where the hands controlled both force and torque on the tool, in addition to movement.

In the free-base condition, we did not observe major changes in hand poses from the localization to execution due to the constraints imposed. Hand dominance seemed to be the primary factor to assign roles to the left and right hands in both steps. Moreover, while most subjects used the same left hand pose (pose 2) in both steps, they modified the right hand pose when transitioning from localization to execution (from pose 9 to pose 8, and from pose 10 to pose 9) by decreasing the number of active fingers. Having more fingers in contact with the screwdriver better guaranteed stability, but imposed more constraints on its movement. This was valuable during localizing the tool, but detrimental during execution, as the tool had to slide along more fingers. In addition, coordinating multiple fingers to achieve the fast rhythmic rotating movement proved more difficult with several fingers.

In the fixed-base condition, subjects used the dominant right hand to localize the tool in 98% trials. Most subjects adopted similar hand pose combinations from localization to execution, with only slight adjustments of finger placements. For example, one subject (S7) removed one active finger from each hand after transiting to the execution step in all trials. Another subject (S1) did not employ the index finger during localization, and then placed the index finger as one independent VF to control force during execution. While most modifications reduced the number of active fingers, it did not change the roles and functions of the two hands, as the VFs and their specific functions remained the same. This was likely due to the fact that both task steps required precision. Only two out of ten subjects, S6 and S10, adopted consistently the identical hand poses across the two experimental steps without any adjustment.

Our analysis revealed that subjects tended to adopt the same type of hand pose for both steps. While we observed differences in hand poses across localization and execution steps, these were minor changes overall. This is in line with the observation that, despite the seemingly infinitely many controlled variables, humans appear to consistently choose the same types of hand poses for similar tasks and to do so effortlessly (Friedman and Flash (2007)). In our study, preserving the same hand pose throughout the task appeared more time-efficient and less cumbersome, as it obviated repositioning the screwdriver or adjusting finger placement on the tool. This choice was likely dictated by an economy of efforts and a desire to increase comfort (Rosenbaum et al. (1990)). This observation is consistent with evidence showing that hand poses are primed differently depending on how to maneuver a tool to accomplish a skillful task (Valyear et al. (2011)). Our study demonstrated that this mental pre-shaping, shown previously in unimanual tasks, also applied to bimanual tasks.

3.4.4 Hand dominance determines role distribution

As expected, hand dominance was a factor to determine functions across the two hands (Guiard (1987)). The dominant hand adopted the role that required finer control of forces and torques. Dominance was particularly visible when the task degrees of freedom were high, as in the free-base condition, but also during localization in both conditions. Interestingly though, in the fixed-base condition, hand dominance seemed to play a lesser role during the

execution phase, as in only 60% of the trials the right hand rotated the screwdriver, while the left hand did so in 40% of the trials. In 20 trials, subjects even swapped hands, using the left instead of the right hand for guiding the screwdriver. In 10 trials, the two hands even alternated when rotating the screwdriver. Six subjects used their left hand instead of their dominant right hand to control motion in tasks. Additional analyses did not reveal any significant difference in the handedness index between the subjects who swapped hands and others ($F_{1,8} = 1.05, p = 0.3364$). Moreover, subjects' failure rate and movement time also do not relate to their handedness indices (average failure rate: $F_{1,8} = 2.56, p = 0.1483$, average movement time: $F_{1,8} = 2.47, p = 0.1547$).

Manipulating the screwdriver and trying to avoid cam-outs requires high accuracy of control. As suggested by the dynamic dominance hypothesis (Sainburg (2002)), the dominant limb is specialized for controlling the task dynamics, while the non-dominant limb achieves higher positional accuracy (Kagerer (2016); Bagesteiro and Sainburg (2003); Goble et al. (2006); Scheidt and Ghez (2007)). Therefore, swapping roles between hands could favor the non-dominant (left, positioning the watch face) hand's control for improving the accuracy by attenuating the dominant (right, controlling the motion of screwdriver) hand's control strength. This swap may also be a strategy to promote task effectiveness by avoiding re-orienting or re-grasping the screwdriver, as suggested by (Theorin and Johansson (2010)). According to our observation, in these switched trials, subjects shaped their left hand in poses 3, 4, or 5 during the localization step. If the left hand played the lead role during localization in some of these trials, the subject would have used the left hand as the prime actor in the following execution step without re-assigning the roles of both hands.

3.4.5   Why bimanual control when unimanual control suffices?

Our results showed that subjects preferred to use both hands, even when a single hand would have been sufficient to complete the task. In the fixed-base condition, the task could have been performed with one hand. However, unimanual manipulation manner was observed in only 14% trials (Fig. 3.11A, $H_{fixed}^{exe}(0, 7-9)$), and in 86% trials, subjects still performed the task bimanually.

Such preference for using two hands can be observed in numerous daily life activities. For example, a bottle cap can be unscrewed with one hand. However, doing this operation unimanually is more complex from a control viewpoint. First, it requires control of the palm, little finger, ring finger and sometimes the middle finger to enwrap the bottle. Then, as the palm and smaller fingers stabilize the bottle's neck, the index and thumb must rotate in synchrony to unscrew the cap. Because of the biomechanical coupling and neural crosstalk, to generate thumb and index movements may lead to a decrease in the force applied by the other fingers and to destabilize the object. The same task, however, can be achieved with more accuracy when using both hands.

Our data indicates that the same principle is at play. Unimanual control of the screwing task

required independent control of the index finger from the other fingers, as well as to control up to three VFs to generate the required force and torque simultaneously. This forced to break the traditional thumb-finger coupling, which may have contributed to the difficulty of performing the task with a single hand (Feix et al. (2015); Dollar (2014)). Instead, when using both hands, torque control can be decoupled from force control, and one needs only to use the common thumb-index coupling to generate the desired motion on the tool.

Efforts required to decouple the control of task demands across the two hands paid off. Indeed, when examining task performance and efficiency, the results showed that bimanual task execution was faster than unimanual task execution. Subjects moved their fingers significantly faster, while maintaining similar average failure rates. Selection of unimanual or bimanual strategy may be dictated by a minimization of cost (Theorin and Johansson (2010)). Subjects may have opted for bimanual over unimanual control in an attempt to maximize success and reduce failure rate, in place of optimizing for economy of efforts.

While decoupling and distributing the task across the two hands may seem to simplify the overall control, one should not forget that this now requires coordinating the two hands. In our study, likely little effort was required to coordinate the two hands, since one of the two hands was static, fixating the watch or holding the screwdriver. The brain could hence rely on existing bimanual strategies for spatio-temporal coordination inherent to bimanual tasks (Andersen and Siebner (2018)). In our study, bimanual coordination was reduced to spatial coordination, as both hands had to align the tool with the required position and orientation relative to the watch. Control of the screwdriver's orientation and position required high precision to avoid cam-outs. This was made particularly challenging as subjects had to rely more on proprioception than vision, given the size of the screw and the fact that the screwdriver could sometimes block the view on the screw. Using both hands was likely advantageous over using a single hand, as contact between the hands could mitigate sensorimotor noise and improve spatial localization (Chinn et al. (2019); Jackson et al. (2002)).

Subjects' preference for bimanual over unimanual control was hence likely an effort to both reduce variability caused by biomechanical couplings and to alleviate intrinsic sensorimotor processing burdens due to the large number of degrees of freedom in both hands that needed to be controlled. Since both hands were tasked to work on the same spatial localization task, efforts may also have been reduced (Theorin and Johansson (2010)).

3.5   Conclusions

This study presented evidence that a task's number of degrees of freedom play an important role in the selection and shaping of bimanual hand poses. The results suggested a strong preference for bimanual over unimanual operation that would require a decoupling of individual fingers' control. The choice of hand poses is informed by handedness; the dominant hand performs the major movement components, and the non-dominant hand tends to assist and control residual degrees of freedom. However, as tasks become less constrained (task's

degrees of freedom reduced), the non-dominant hand contributes more actively and shares forces with the dominant hand. At times, the roles of the two hands were switched as subjects explored different hand poses. Subjects exhibited a diversity of hand pose combinations when the task became less constrained. We speculate that this search for different hand poses may be an effort to improve performance, possibly seeking postures that minimize the influence of sensorimotor noise on precision of control (Sternad (2017)). While our study focused on a bimanual task, our results also offer insights in unimanual performance. Interestingly, subjects rarely exploit the full potential of the hands' and fingers' dexterity.

To facilitate visual assessment of the diversity and role distribution of hand poses, we proposed a hand pose matrix as a tool to summarize experimental results. This visualization is not specific to this study and may be useful for other bimanual tasks. The taxonomy also included analysis of virtual fingers in relation to control of motion and force/torque distribution demanded by the task. It is worth noting that the hand pose combinations used in this study are task-specified, and were created for distinct task conditions. It does not include other hand poses found in extant hand taxonomies that were not relevant in the present context. The hand pose matrix does not provide information regarding the change of hand poses across trials. It would be interesting to extend this matrix to include a temporal tracking of hand poses to provide time-varying task information.

Finally, our analysis did not relate the choice of hand pose to the force and torque required by the task, mainly because our tasks required only small forces that are well within the strength of the human hand. Moreover, it is also suggested that the control of hand poses and the regulation of contact forces are independent (Santello et al. (1998)). However, for tasks that require larger forces, such as opening the cap of a bottle, the selection of hand poses and role distribution is likely modulated by the precision and strength that each hand can contribute. Analyzing applied forces for manipulating the same object with different hand poses may provide more insights on the effects of force demands.

4 Exploiting kinematic redundancy for robotic grasping of multiple objects

In previous chapters, we have investigated both the inter-limb coordination in upper limbs and the role distribution across hands in bimanual fine-manipulation tasks. In particular, our study suggests that humans flexibly group fingers into multiple functional units, i.e., *virtual fingers*, to achieve the simultaneous control of multiple task demands.

The concept of virtual finger was essentially proposed for analysis purpose. It enabled us to disclose the principle of finger role assignment by analyzing hand poses. It also provides inspiration for solving the synthesis problem. A typical scenario is to synthesize grasps of a multi-fingered robot hand.

In this chapter, we take such an inspiration to develop algorithms that enable a multi-fingered robotic hand to achieve high dexterity in the presence of multiple task demands. Specifically, we enable a robotic hand to grasp and hold multiple objects simultaneously by exploiting its *kinematic redundancy*, referring to all its controllable DOFs. We first propose a human-like dexterous grasp synthesis algorithm to generate stable grasps using pairwise contacts on arbitrary opposing hand surface regions, no longer limited to fingertips or the inner surface of the hand. To this end, we construct a reachability map by modeling the available space of the hand. This map guides the formulation of a constrained optimization problem in joint space, solving for feasible, stable, and collision-free grasps. We formulate an iterative process to enable the robotic hand to sequentially grasp multiple objects. Moreover, we propose a kinematic efficiency metric of the hand model, and an associated strategy to facilitate the planning of multi-objects grasping. We validated our approaches both in simulation and on a real robotic hand.

The work presented in this chapter has been submitted as a journal article entitled *Exploiting Kinematic Redundancy for Robotic Grasping of Multiple Objects* to *The International Journal of Robotics Research* at the time of writing this thesis.

4.1 Introduction

Grasping is a primitive but central skill for robotic hands and manipulators to interact with the environment. It is also a prerequisite for the realization of any desired manipulation movement. Grasping requires the hand to coordinate its multiple degrees of freedom to establish multiple contacts on the surface of the target object to form a stable and collision-free grasp. The problem of generating such a grasp configuration for the hand is known as grasp synthesis (Shimoga (1996)).

Despite recent advances in robust grasping and in the design of complex robotic hands with numerous sensing and actuation, the community remains largely focused on grasping a *single* object, albeit with either a power grasp, i.e. using the inner surface of the hand to wrap the object (Zhuang et al. (2019)), or through a pinch grasp, using thumb and index in coordination (Deng et al. (2020)). Although such simple grasping poses enable robotic hands to handle a majority of objects, it lacks the ability to dexterously employ the abundant DOFs of the hand in a grasping task. In comparison, humans grasps objects with a variety of grasp poses, using not only their fingertips but also the palm, and almost any regions on the hand surface, as documented in multiple taxonomies of human hand poses (Gonzalez et al. (2014); Feix et al. (2015); Starke et al. (2020)). In this respect, the dexterity of robots is still far from the dexterity of human hands (Billard and Kragic (2019)).

In addition, although many hands have abundant DOFs in their structure, most advanced robotic hands including industrial robotic manipulators are only capable of grasping one single object at a time; and even the local adjustment of fingers on the grasped object is a challenging problem (Sundaralingam and Hermans (2018)). The problem of robotic grasping of multiple objects has rarely been studied. Some representative work has focused on the use of a simple robotic manipulator to envelope multiple objects simultaneously, and the physical interaction among the objects enveloped as a whole under such conditions (Harada and Kaneko (1998); Yamada et al. (2011); Yamada and Yamamoto (2015)).

Human hands, in contrast, can grasp multiple objects simultaneously in everyday tasks. For instance, holding different cutlery when clearing up dishes, or manipulating a pair of chopsticks to pick up food. These types of grasps hold objects within the lengths of adjacent fingers (Feix et al. (2015)). While such grasps that use all the contact of the fingers are common in humans, they are rarely seen in robotics that tend to favor fingertip grips. Empowering robotic hands and manipulator with human-like dexterity would allow the robotic hand to multitask and thus increase the efficiency.

This chapter aims at improving the dexterity of a robot by exploiting the kinematic redundancy of hand in grasping tasks. We define *kinematic redundancy* as the set of available degrees of freedom in a robot model that can be employed in task planning. In this chapter, we propose a framework to exploit the kinematic redundancy of a robot hand model to achieve human-like dexterity, for example, grasping objects using any region of the hand or grasping multiple objects. This will not only improve the efficiency of robotic grasping tasks, e.g. on

industrial assembly lines, but also inspire the design of algorithms for dexterous manipulation of multiple objects.

4.1.1 Robot grasping synthesis

The robotic grasping problem has been extensively and intensively studied in the past decades (Bicchi and Kumar (2000); Sahbani et al. (2012); Ozawa and Tahara (2017)). A feasible grasp can be generated using either *analytic approaches*, relying on solving a constrained optimization problem over quality metrics, for example, optimization-based approaches (Panagiotopoulos and Al-Fahed (1994)), or *empirical approaches* (i.e. data-driven approaches) by sampling numerous feasible configurations and selecting the optimal ones (Bohg et al. (2013)). Empirical approaches also include learning-based approaches, such as learning from human demonstration (Aleotti and Caselli (2010)), statistical learning-based approaches (Murali et al. (2018)), and deep learning-based approaches (Caldera et al. (2018); OpenAI: Andrychowicz et al. (2020)). Although empirical approaches do not require accurate model representations and affords directly sampling on real robotic hand, it demands large amount of samples in general. Moreover, it is essentially an *object-centered* approach (Goldfeder and Allen (2011)). Thus, it can hardly be applied to generate grasps on multiple objects.

Given the model of the robot hand and the model of the object to be grasped, the grasping synthesis problem aims to find feasible configurations by determining the contact points between the hand surface and the object surface. Human reaching and grasping movements typically consists of two parts (Rosenbaum et al. (2001)): the hand is pre-positioned near the object based on their relative spatial relationship (high-level planning), and then the hand moves its fingers to establish contacts with the object (low-level planning). Such a two-stage grasping process is commonly used in the planning of robotic grasping and manipulation tasks, and we hence divide our review accordingly.

High-level planning

The high-level planning determines the relative spatial pose between the hand and the target object to grasp. (Ciocarlie and Allen (2009)) introduced the concept of eigen grasp to describe hand configurations in a low-dimensional subspace, based on the hand postural synergy (Santello et al. (1998)). This is an object-centered approach, and the optimal grasping position of the hand with respect to the target object is determined by generating large amount of samples of the hand sub-spaces around the target object.

Kinematic properties of the robot hand model can be used to facilitate more efficient grasp positioning. For example, (Zacharias et al. (2007)) sampled a robotic arm model in joint space and constructed a *capability map* to describe the spatial reachability of the arm. Spatial regions with high reachability index values are prioritized for task planning, such as generating robot manipulation sequences (Ruehl et al. (2011)). The capability map can be constructed

offline and easily adapt to new tasks and objects. Object models have also been exploited in grasping synthesis. (Gienger et al. (2008)) introduced an object-specific *task map* that represents a manifold of feasible power grasps on the target object and defined the goal as a hyper volume on the map. This is useful to generate a feasible path for all fingers to close simultaneously to envelop the object. It however does not consider other type of grasps than power grasp nor planning of individual finger positioning. (Zacharias et al. (2009)) proposed an object-specific *grasp map* to represent the characteristics of the target object and showed that the combination of object grasp map and robot capability map in high-level planning greatly speeds up the low-level planning with guaranteed grasping quality. (Roa et al. (2011)) constructed *graspability map* by combining a sampling-based approach with the hand's reachability map. It employs the reachability map to restrict the spatial location of the hand relative to the object, and then samples the positions and orientations of the hand around the object in a discretized spatial mesh grid. Only poses that result in collision-free and force-closure grasps are retained. The optimal grasp is then selected from a set of feasible hand poses. Afterwards, the robot hand moves to this desired optimal position, adjusts its posture, and then closes its fingers to grasp. The graspability map has demonstrated its efficiency in grasping tasks (Siciliano (2012)) and has also been applied to guide the design of grippers (Eizicovits et al. (2016)). However, the accuracy of the map depends highly on the voxelization resolution of the spatial volume around the object. Resolution is often non-uniform and hence, at the same spatial location, the quality of the grasp may vary significantly for different hand orientations. It is also time-consuming to evaluate the collision and stability conditions for all hand pose samples, especially since most of these end up being discarded. Moreover, the graspability map has to be recreated for any novel target objects. Although it is possible to generate an approximation of the graspability map based on shape primitives (Eizicovits and Berman (2018)), quality degradation is inevitable.

To handle these issues, we construct a hand-centered *reachability map* by sampling in joint space, consisting of three-dimensional reachable spaces of each individual link of the hand (i.e. finger phalanx or the palm), represented by three-dimensional alpha shapes (Edelsbrunner and Mücke (1994)). We use the hand reachability map to describe the reachable internal sub-spaces of the hand model, providing a boundary for the spatial location of the target object relative to the hand. The optimal grasp configuration is then computed based on the property of this reachability map.

Low-level planning

Low-level planning in grasping is often formulated as a constrained optimization problem that aims at determining the desired contacts points and forces on the object. It optimizes a desired task quality measure, which is referred to as the grasping quality metric, while satisfying constraints, e.g., joint limits, collision avoidance, closure properties (Bicchi (1995)).

A number of metrics based on grasping configurations have been widely used in grasping synthesis (Mishra (1995); Miller and Allen (1999); Borst et al. (2004); Roa and Suárez (2015)) and

also in the understanding of human grasping (León et al. (2012)). (Li and Sastry (1988)) tackled the optimal grasping problem, optimizing task-oriented quality metrics for multi-fingered robotic hands. (Ponce and Faverjon (1995)) computed stable grasps of 2D polygonal objects using a three-fingered robotic hand. By using geometric conditions for closure properties, (Zhu and Ding (2004)) proposed an algorithm for computing form/force closure grasps of 3D objects with curved surfaces and multiple contact points. (García-Rodríguez et al. (2015)) investigated the problem of generating stable grasps of circular objects through two contacts. More recently, (El-Khoury et al. (2015)) solved the optimal grasping problem for a multi-fingered robotic hand in the form of joint space optimization, which is subject to collision-free and force-closure constraints. A comprehensive review of grasping synthesis algorithms can be found at (Shimoga (1996); Sahbani et al. (2012)). Most work assume, by default, that the grasping points will be done at the fingertip and exclude other surface areas on the finger from the grasp synthesis. When surface areas are used, e.g. in power grasp, this done uniformly for the entire length of the finger and is not the result of a selective choice of which surface to use for contact. In our previous work (Sommer and Billard (2016)), we offered an approach to maximize contacts along both the frontal and back surface of the fingers selectively to increase the number of contacts along the object. However, the location of these contact points was pre-specified and not part of the optimization.

Although fingertips have a higher motion freedom compared to other parts of the hand, as pointed out by (El-Khoury et al. (2013)), using only fingertips cannot fulfill task-specified demands in all grasps. In contrast, we provide a highly human-like solutions that dexterously exploits full hand kinematic structure to compensate the insufficiency of fingertips. We allow contacts to present over the entire finger surface and at any point on the palm of the hand without pre-assignment and further limitations.

4.1.2 Grasping multiple objects

While grasping of a single object has been intensively studied, the problem of grasping multiple objects with a single hand has received much less attention. In an early attempt to enable a robot to grasp two objects, (Harada and Kaneko (1998)) tackled the problem of grasping objects using a robotic manipulator. Given the grasping model, this study analyzed the kinematics and internal forces required to stabilize the two target objects when grasping them simultaneously. In a similar task scenario, (Yamada et al. (2011); Yamada and Yamamoto (2013)) analyzed the stability condition for grasping multiple three-dimensional objects based on the grasp potential energy, using a grasp stiffness matrix (Yamada et al. (2009)). Other issues related to multiple objects grasping, such as feasibility of placing fingertips (Yu and Fukuda (2013)) and the analysis of contact surface geometry model (Yamada and Yamamoto (2015)) have also been discussed in similar scenarios.

The above studies use on a specific grasp type, i.e. the envelope grasp. Multiple target objects are viewed as a whole, and are arranged next to one another in a chain-like manner or piled

up into a pyramid shape. The fingers make contact with the two objects only at the outer of the object chain or pile. The objects lying in the middle are stabilized through contacts with its neighbor objects. There again, points of force application are modeled on the fingertips that are in contact with objects on either side of the chain. On the one hand, analyzing the interaction among multiple objects enveloped inside the hand largely increases the difficulty of the grasp synthesis. On the other hand, it is hard to perform any manipulation of the grasped objects because there are only two contact points between the robot hand and the object.

The problem of grasping and manipulating multiple object has also been discussed in different scenarios. For example, (Donald et al. (2000)) proposed an algorithm for moving multiple objects as a whole in a multi-robot system that is able to perform specific manipulation operations, but precise grasping and manipulation of each single object is not possible. In a recent study, a novel robotic manipulator has been designed to grasp handle multiple objects (Mucchiani and Yim (2020)). However, this highly relies on specific robotic devices and is thus difficult to be generalized to other robotic hands or manipulators.

In this work, we consider a broader range of multiple objects grasping scenarios, whereby each object can be held by arbitrary part of the fingers and in various poses. This potentially enables the manipulation of each one of the grasped objects through contacts, and also avoids tackling the interaction of multiple objects. We use the opposition spaces spanned by opposing hand surface regions to reformulate the problem of grasping multiple objects at a sequence of single object grasping problems. This sub-problem can be solved by our proposed grasping synthesis algorithm.

4.1.3 Human hand pose selection for grasping

Studies on human hand poses in grasping have provided insights to help understand the human dexterity, suggesting that the shaping of the hand plays a significant role in dexterous skills, such as playing piano (Furuya et al. (2011)) or watchmaking craftsmanship (Yao and Billard (2020)). Human dexterity is such that it selects hand poses and adapts the finger placement according to the geometric properties of the target object (e.g. shape (Gentilucci (2002)), size (Daprati and Sirigu (2006)), and spatial location (Johnson-Frey et al. (2004))) and task requirements (Valyear et al. (2011)), leading to better task performance. Such hand pose selection strategy is influenced by multiple factors, such as the intention of subsequent movement (Ansuini et al. (2006)), the predicted final status of the grasp (Cohen and Rosenbaum (2004)), or the goal of the task. For example, human uses distinct hand poses when grasping for using the object than when grasping for moving the object (Valyear et al. (2011)). Moreover, human also adjust the task distribution across fingers to meet the requirements of controlling multiple independent task demands (Yao et al. (2021)).

4.1.4 Virtual finger and opposition space

Human dexterity showcase humans' unique ability at learning how to organize and coordinate efficiently the multiple DOFs of the hand. To interpret how humans organize and control the coordination of the fingers during grasping, (Arbib et al. (1985)) introduced the concept of *virtual finger*. A virtual finger represents a group of one or multiple real human fingers or even parts of the hand (e.g. the palm) that move in unison to achieve an independent function, such as applying force in a desired direction. This concept provides a functional analysis of the hand that consider controlled DOFs per group. It associates this group to a function, distinguishing functions that contribute to the task and secondary functions that do not directly contribute and may hence be ignored. The concept of *opposition space* (Iberall (1986)) expands this concept and explains how virtual fingers come to coordinate placement and forces to stabilize the grasp or generate desired motions on the object. Opposition space can be used to determine the intended actions on the object when observing how human place their fingers on the object (Biegstraaten et al. (2006); De Souza et al. (2015); Smeets et al. (2019)). An opposition space (OS) is defined as the sub-spaces inside the hand model that are structured by opposing patches on virtual fingers, such that opposing forces can be applied inside this space. It has been proposed that there are three basic oppositions that constitute human prehension (Iberall (1997)), i.e. (1) *pad opposition*, formed by two finger pads opposing each other, (2) *palm opposition*, structured by the palm of the hand and the finger patch opposing it, and (3) *side opposition*, constructed by the lateral patches of fingers. Most grasping types discussed in robotics are restricted to pad opposition (e.g. pinch grasp) and palm opposition (e.g. power grasp), while grasps using side opposition are hardly ever discussed.

We propose a grasp synthesis algorithm, inspired by the concepts of virtual finger and opposition space. We enable contacts on arbitrary parts of the hand, and map the kinematic chain from the hand base to the contact point onto a virtual finger. In this way, only the DOFs affecting the contacts on the objects are considered into the current planning. Irrelevant DOFs are ignored and considered as kinematic redundancy that can be exploited in planning future tasks. In our modeling, any pair of opposing regions on the hand surface can generate an opposition space. These spaces can be used to construct grasping on each object separately in such a way to ensure local force closure property.

4.1.5 Our contributions

Our contribution consists of three main aspects:

- We propose a human-like dexterous grasping synthesis algorithm that empowers a multi-DOF robotic hand to grasp objects using arbitrary surface regions of the hand, within the opposition spaces spanned by pairwise hand links.

- We propose an iterative process that allows the robotic hand to grasp multiple objects

(a) Human right hand model (20-DOF)   (b) Robotic left hand model (16-DOF)
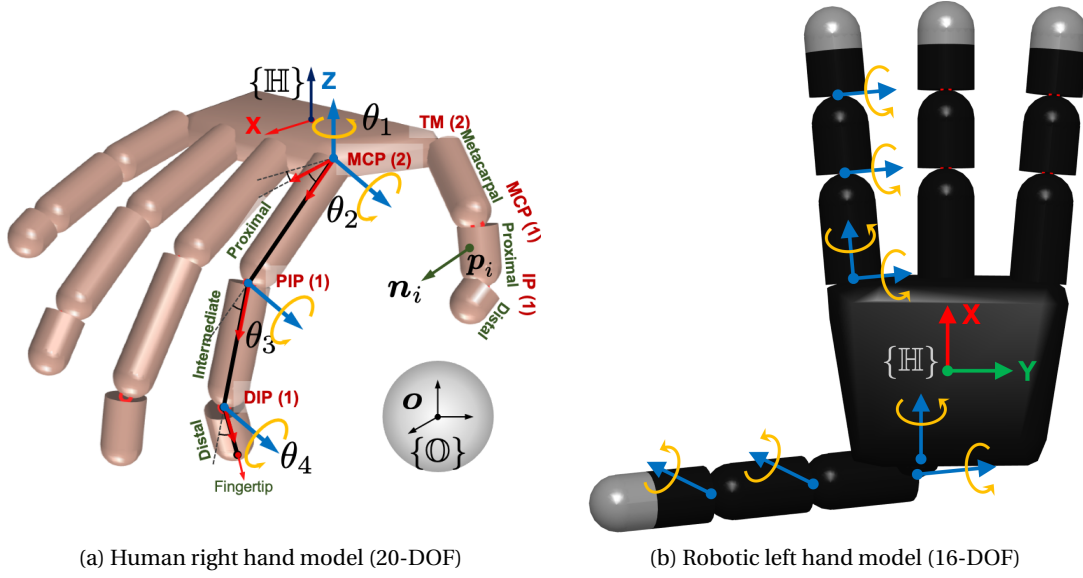
Figure 4.1 – Kinematic models of human hand and robotic hand (adapted from (Malvezzi et al. (2013))). The hand reference frame $\{\mathbb{H}\}$ is defined at the geometric center of the palm. The object reference frame $\{\mathbb{O}\}$ locates at the geometric center of the object. Joint reference frames of other fingers are defined following the same convention as the index finger.

in sequence.

- We propose a kinematic efficiency metric to quantitatively measure the exploited kinematic redundancy in the robotic model, and an associated strategy to facilitate the exploitation of kinematic redundancy in the planning of multiple sequential grasps.

4.2 Notation and models

4.2.1 Notation

We first list important notations in Table 4.1 before we introduce our algorithms. A spatial vector pointing from $\boldsymbol{a} \in \mathbf{R}^3$ to $\boldsymbol{b} \in \mathbf{R}^3$ is denoted as $\overrightarrow{\boldsymbol{ab}}$. The function $\boldsymbol{d}(\cdot, \cdot)$ indicates the Euclidean distance measure between input entries, e.g. a pair of points or geometric models. We use $\{\cdot\}$ to represent the set of a certain type of elements. Moreover, we use $|\cdot|$ to denote the cardinality of the element, and $||\cdot||$ the Euclidean norm of a vector or a matrix.

4.2.2 Modeling of hand

We use a 20-DOF anthropomorphic human hand model (right hand) to explain our proposed method in detail. Moreover, we also validate our algorithms on a 16-DOF Allegro hand model (left hand) by generating a variety of grasps of everyday objects, which have been tested on

Table 4.1 – List of notations.

REFERENCE FRAMES

$\{\mathbb{H}\}$:	hand reference frame
$\{\mathbb{O}\}$:	object reference frame
$\{\mathbb{C}_i\}$:	local reference frame of the ith contact
$\{\mathbb{L}_i\}$:	local reference frame of the ith link

MODELS AND SETS

$\mathscr{H}$:	hand model
$\mathscr{O}$:	object model
$\mathscr{K}$:	kinematic redundancy set that comprises all redundant DOFs
$\mathscr{R}_i$:	the reachable space of the ith link
$\{\mathscr{R}\}$:	the reachability map of the hand
$\mathscr{M}_\mathscr{C}$:	the collision map of the hand
$\mathscr{S}_{i,j}$:	the opposition space spanned by $\mathscr{R}_i$ and $\mathscr{R}_j$
$\mathscr{L}_i$:	the ith general link (either a finger phalanx or the palm)

INDICES

N_l:	total number of general links
N_j:	total number of joints
N_c:	total number of contacts
N_o:	total number of objects

PARAMETERS AND VARIABLES

$\mathscr{Q}$:	grasp quality measure, used as objective function
$C_\mathscr{S}$:	capacity of opposition space $\mathscr{S}$
$\boldsymbol{o}$:	position of object center represented in $\{\mathbb{H}\}$
$\boldsymbol{p}_i$:	position of the ith contact on the hand represented in $\{\mathbb{H}\}$, origin of $\{\mathbb{C}_i\}$
$\boldsymbol{p}_i^*$:	projection of $\boldsymbol{p}_i$ on the object surface along $\overrightarrow{\boldsymbol{p}_i\boldsymbol{o}}$
ρ_i:	radical distance associated with $\boldsymbol{p}_i$
ϕ_i:	angular coordinate associated with $\boldsymbol{p}_i$
α_i:	height ratio associated with $\boldsymbol{p}_i$
$\boldsymbol{n}_i$:	contact normal of the ith contact on the hand
$\boldsymbol{f}^i$:	force applied on the ith contact
$\mathscr{F}^i$:	the approximated friction cone on the ith contact
$\mathbf{f}_n^i$:	the nth edge of $\mathscr{F}^i$
$\boldsymbol{d}^i$:	the vector from $\boldsymbol{o}$ to $\boldsymbol{p}_i$
τ_n^i:	torque generated on the object center by $\mathbf{f}_n^i$
$\boldsymbol{w}_n^i$:	primitive wrench applied on the ith contact, associated with $\mathbf{f}_n^i$
φ_n^i:	the positive coefficient associated with $\boldsymbol{w}_n^i$
d:	the grasp distance
$d_{\boldsymbol{o},\boldsymbol{p}_i^*}$:	distance from the object center $\boldsymbol{o}$ to the projected contact point $\boldsymbol{p}_i^*$
r^l:	radius of the link
r^o:	radius of the spherical object
μ:	coefficient of friction
$\boldsymbol{g}$:	unit vector indicating gravitational direction
θ_j^i:	joint angle of the jth joint on the ith finger
$\boldsymbol{\theta}^i$:	vector of joint angles from the ith finger
$\boldsymbol{\theta}_{actv}^i$:	vector of joint angles parameterizing the contact on the ith finger
$\mathbf{Q}_i$:	set of variables that parameterize $\{\mathbb{C}_i\}$
q^i:	a general element of $\mathbf{Q}_i$
$\mathbf{Q}^o$:	set of variables that parameterize $\{\mathbb{O}\}$

the real Allegro left robotic hand. Both hand models are adapted from (Malvezzi et al. (2013)).

Human hand model

The human right-hand model (20 DOFs) consists of five fingers and palm. The geometric shape of finger phalanges are modeled using cylinders, and the palm is modeled by a cuboid (see Fig. 4.1).

The thumb has four joints, including the trapeziometacarpal junction (TM, 2 joints), metacarpo phalangeal junction (MCP, 1 joint), and interphalangeal (IP, 1 joint). Each one of the fingers has four joints, corresponding to the rotational joints of metacarpophalangeal junction (MCP, 2 joints, orthogonal), proximal interphalangeal junction (PIP, 1 joint), and distal interphalangeal (DIP, 1 joint). The palm connects the bases of all fingers and does not have any degrees of freedom.

We denote each of the hand's articulation (degree of freedom) as *joint* and the segments of the hand (not only finger phalanges but also the palm) as *link*. We number fingers in sequence, from the thumb (F1) to the little finger (F5). Each finger has four phalanges (links) in the anatomical structure. From the wrist to each fingertip, the first finger link (metacarpal phalanx, L1) belongs to the palm, thus is not considered an independent link. The proximal, intermediate, and distal phalanges of each finger are independent links, denoted as L2, L3, and L4, respectively. The palm is also considered as one independent link, represented as L0. Any phalanx in the model can be retrieved by combining the finger name and link name, for example, F2L3 refers to the third link of the second finger, which is the intermediate phalanx of the index finger.

Allegro hand model

The Allegro left-hand model (16 DOFs) has four fingers and palm. Each finger has one ab-/adduction DOF and three extension/flexion DOFs, in the same order as the human hand. We number fingers in sequence, from the thumb (F1) to the ring finger (F4). The geometric shape of finger phalanges are also approximated by cylinders, but differ in sizes from the human hand model (see 4.2). The hand width is measured as the combined width of all fingers (exclude the thumb) close together (i.e. all fingers are fully extended and the ab-/adduction DOF of all fingers are in 0 position). The hand width equals the width of the palm for both hand models.

Table 4.2 – Sizes of hand models (mm).

Hand Type	Finger Length	Hand Width	Hand Length
Human hand	92.0	58.0	120.0
Robotic hand	136.1	139.5	247.7

Definition of reference frame

Grasp synthesis is expressed in the hand's reference frame {ℍ}, whose origin is located at the geometric center of the palm for both hand models (see Fig. 4.1). Each joint' rotation corresponds to the Z axes of each joint's reference frame, and are denoted by blue arrow lines in Fig. 4.1. Axial directions of phalanges correspond to the X-axis of joint reference frames. They are indicated by red arrow lines. The Y axes are determined by following the right-hand convention.

4.2.3   Modeling of objects

We consider spherical and cylindrical shaped objects, as many of the grasping parts of daily life objects can be approximated by such elementary geometries (Hubbard (1996)). The 6-dimensional pose of the object is described by the *object reference frame* {𝕆}, whose origin is located at the geometric center of the object, $\boldsymbol{o}$. For a cylindrical object, we define the Z axis of {𝕆} along the object's central axis. The X and the Y axes are assigned arbitrarily in the orthogonal plane. We represent the orientation of a cylindrical object using a quaternion $\boldsymbol{q}^{\mathcal{O}} = (q_x, q_y, q_z, q_w)$, $||\boldsymbol{q}^{\mathcal{O}}|| = 1$.

As we focus on human-hand dexterous grasps, we assume that the target object is light enough, for the fingers to generate sufficient forces to support the object in a stable grasp configuration. For example, when the force closure property is satisfied. Moreover, we assume that the geometric center of the object and its center of mass coincide. We denote them as the same point $\boldsymbol{o}$.

4.2.4   Modeling of contact

Our algorithm allows the contact point $\boldsymbol{p}_i$ to be assigned to an arbitrary position on the surface of the ith link ($\mathcal{L}_i$). A *contact reference frame* {ℂ$_i$} is defined at each contact point. The contact normal $\boldsymbol{n}_i$ corresponding to the Z-axis of {ℂ$_i$}, perpendicular to the finger surface curvature of the local contact area.

We parameterize a contact point $\boldsymbol{p} \in \mathbf{R}^3$ on a given finger phalanx by a tuple of cylindrical coordinates ρ, ϕ, and α in the local reference frame of the link {𝕃$_i$} (see Appendix A). The palm is modeled as a cuboid and its inner surface region is modeled as a rectangular-shaped surface. Contact points on the palm are parameterized by the Cartesian coordinates of {ℂ$_i$} on the rectangular shaped inner surface of the palm.

Our grasp synthesis algorithm initially considers two contact points, which is the minimum number of contacts necessary to apply opposing forces (Arbib et al. (1985); Iberall (1986)). Grasping using two opposing hand regions is common in human grasping (Iberall (1997)) and can suffice most grasping tasks. Once a successful grasp is generated, the hand can close its free links to establish more contacts with the grasped object (see Sec. 4.4.4).
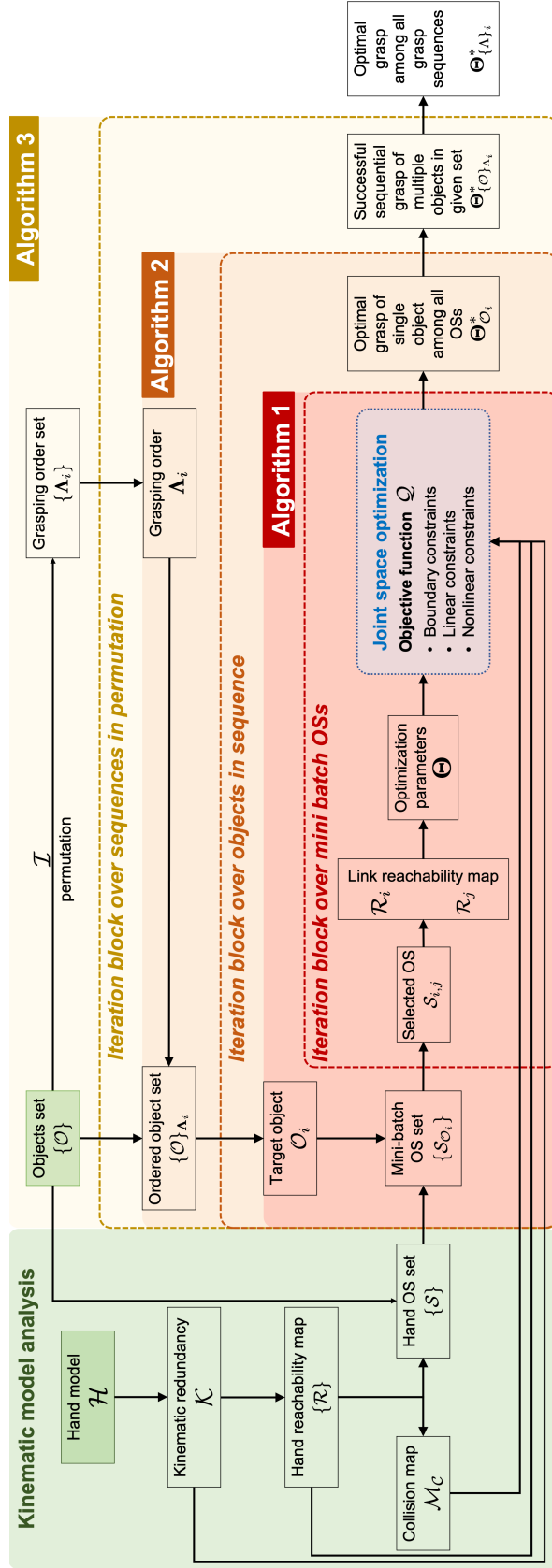
Figure 4.2 – Our proposed framework. It consists of three algorithms: (1) a human-like dexterous grasping algorithm using arbitrary surface regions of the hand (Algorithm 1), (2) an iterative process for sequential grasping of multiple-objects (Algorithm 2), and (3) a greedy grasping algorithm to achieve maximum exploitation of kinematic redundancy in sequential task planning (Algorithm 3).
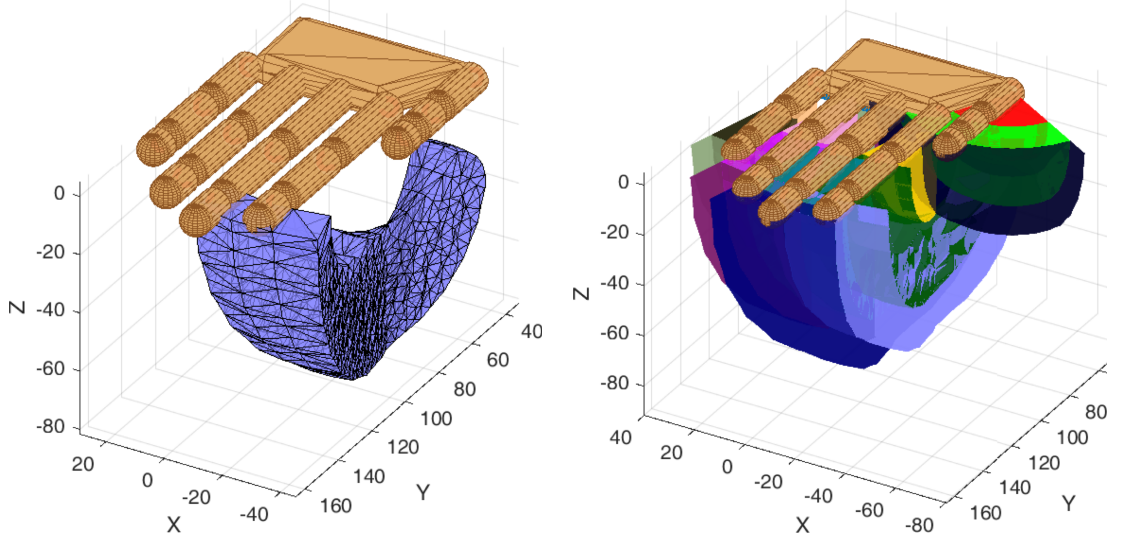
(a) The reachable space of the index finger's distal phalanx (F2L4)

(b) The complete reachability map $\{\mathscr{R}_i\}$ of the hand model

Figure 4.3 – Example of the reachable space $\mathscr{R}$ of one single finger phalanx and the complete reachability map $\{\mathscr{R}\}$ of the human hand model.

As a contact model, we consider each contact area as point contact with friction (coefficient of friction is μ). We approximate the friction cone by a polyhedral convex cone described in Appendix B. We ignore the torsional moments at the point of contact, as we are primarily interested in solving for grasping configurations that ensure the grasped object does not slip off once grasped. In this case, using two contact points to achieve force closure stability condition (Sec. 4.4.3) prevents the happening of translational slip on the contact region rather than the rotational slip, i.e. relative rotation on the contact region with respect to the contact normal.

4.3 Proposed framework

Our proposed framework (Fig. 4.2) consists of three parts: (a) a human-like dexterous grasping algorithm using arbitrary surface regions of the hand (Algorithm 1), (b) an iterative process for sequential grasping of multiple objects (Algorithm 2), and (c) a greedy-grasping algorithm that aims at maximally exploiting the kinematic redundancy in multiple-objects grasping task (Algorithm 3). We introduce each part of the framework in detail in the following sections.

4.4 Human-like dexterous grasping synthesis algorithm

In this section, we introduce our human-like dexterous grasping algorithm that uses arbitrary surface regions of the hand. These grasps are not limited to the fingertips and palms of the

fingers, but include the complete surface of the hand.

4.4.1 Reachability map

The reachability map is a hand-centered representation of the motion feasibility in Cartesian space. To construct the hand's reachability map, we uniformly sample each joint of the hand within its motion range and register the corresponding spatial position of each joint by calculating forward kinematics. Notice that we discriminate the reachable space of one hand link by referring to it as the *reachable space* and the reachable spaces of the entire hand as the *reachability map*. It is a set of the reachable spaces of all links (finger phalanges and the palm). We denote the reachable space of the ith joint as S_i, which is a continuous three-dimensional surface. For each finger phalanx link i, we compute the link's reachable spatial space, $\mathcal{R}_i$, composed of the space enclosed by the reachable set of the joints at both its ends, i.e. S_i and S_{i+1}. The spatial volume bounded in-between is reachable by at least one point on the link, due to the convexity of the link's geometry (finger phalanx is approximated as cylinders) and the continuity of the motion range for each DOF.

The link's reachable space $\mathcal{R}_i$ is represented as a three-dimensional alpha-shape (Edelsbrunner and Mücke (1994)) built up by the sampled Cartesian space points from S_i and S_{i+1}. As an example, Fig. 4.3a shows the reachable space of the distal phalanx of the index finger (F2L4). Since the palm is fixed with respect to {ℍ} and has no DOF, its reachable space is represented as the entire rectangular-shaped palm inner surface. Fig. 4.3b demonstrates the complete reachability map of the hand, {$\mathcal{R}$}.

4.4.2 Opposition space

Given the reachability map of the hand {$\mathcal{R}$}, the complete set of opposition spaces spanned by the hand model can be obtained by analyzing the geometric relationship of reachability maps for each pair of links.

We define a general opposition space (OS) as the convex hull spanned by a group of readability maps:

$$\mathcal{S} = \text{ConvexHull}(\bigcup_i \mathcal{R}_i). \tag{4.1}$$

Within such an OS, opposing forces could be applied at a certain distance. The range value of this distance depends on the spatial relationship of these reachability maps. Each reachability map represents the reachable space of one contact point (on one link), thus to enable a grasp with N contact points, at least N reachability maps are required.

Since our proposed algorithm aims at generating grasps consisting of two contacts, OSs being discussed in this chapter are spanned by a pair of two reachable spaces $\mathcal{R}_i$ and $\mathcal{R}_j$, $i \neq j$, denoted by $\mathcal{S}_{i,j} := \mathcal{R}_i \cup \mathcal{R}_j$. The contact points associated with $\mathcal{S}_{i,j}$ in a grasp are $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$, located on the general hand links $\mathcal{L}_i$ and $\mathcal{L}_j$, respectively. In the following, we drop the

subscripts i, j for simplicity, and denote the OS candidate as $\mathscr{S}$.

Given a pair of spatial points, $\boldsymbol{r}_i \in \mathbf{R}^3$ from $\mathscr{R}_i$ and $\boldsymbol{r}_j \in \mathbf{R}^3$ from $\mathscr{R}_j$, we define respectively the minimum and maximum capacity of $\mathscr{S}$ as:

$$\underline{C}_{\mathscr{S}} = \min||\boldsymbol{r}_i - \boldsymbol{r}_j||, \ \overline{C}_{\mathscr{S}} = \max||\boldsymbol{r}_i - \boldsymbol{r}_j||,$$
$$\boldsymbol{r}_i \in \mathscr{R}_i, \ \boldsymbol{r}_j \in \mathscr{R}_j, \ i \neq j. \tag{4.2}$$

The boundaries of OS capacity constraints the feasible *grasp distance d* inside it, i.e. the distance between contact points on the surface of the object $\mathscr{O}$ being grasped in this OS. For example, for a spherical (or cylindrical) object with radius r^o, the minimum grasp distance enabling a force-closure grasp (see Sec. 4.4.3) is $d = \boldsymbol{d}_{\min}(\boldsymbol{p}_i, \boldsymbol{p}_j) = 2r^o \cos(\arctan(\mu))$.

Given a desired grasp distance d, we say $\mathscr{S}_d$ is *geometrically permissive* for d if the condition holds:

$$\underline{C}_{\mathscr{S}_d} \leq d \leq \overline{C}_{\mathscr{S}_d},$$
$$d := \boldsymbol{d}_{\min}(\boldsymbol{p}_i, \boldsymbol{p}_j), \ \boldsymbol{p}_i \in \mathscr{R}_i, \ \boldsymbol{p}_j \in \mathscr{R}_j, \ i \neq j. \tag{4.3}$$

Fig. 4.4 lists all feasible OSs constructed by the human hand model for a grasp distance $d = 20mm$. $\mathscr{S}_d$ could be used to synthesize a grasp for a target object $\mathscr{O}$ at a grasp distance d. We use $\{\mathscr{S}_d\}$ to represent the set of all geometrically permissive OSs for d.

Notice that such property is related to the object model $\mathscr{O}$ through desired grasp distance d, and being geometrically permissive is a sufficient but not necessary condition for an OS to enable a feasible grasp. In the following, we simply denote $\mathscr{S}_d$ as $\mathscr{S}$, since all OSs being discussed must be geometrically permissive.

4.4.3 Grasp synthesis as a constrained optimization problem

We formulate the grasp synthesis problem as a constrained optimization problem. As soon as an OS candidate $\mathscr{S}$ is chosen, the links used for grasping are determined. We denote the set of all kinematic variables that parametrize the local contact frame $\{\mathbb{C}_i\}$ as $\mathbf{Q}_i$. Such elements include (1) joint angles that precede $\boldsymbol{p}_i$ along the kinematic chain and (2) the cylindrical coordinates that model $\boldsymbol{p}_i$ in local reference frame (see App. A). An element of $\mathbf{Q}_i$ is represented as q.

We consider three types of constraints: boundary constraints, collision constraints (inequality constraints), and contact constraints (equality constraints).

Boundary constraints

Kinematic variables The generalized parameters q in the variable set $\mathbf{Q}_i$ parameterize the corresponding contact $\boldsymbol{p}_i$. It could be a joint angle or a cylindrical coordinate, both subject to
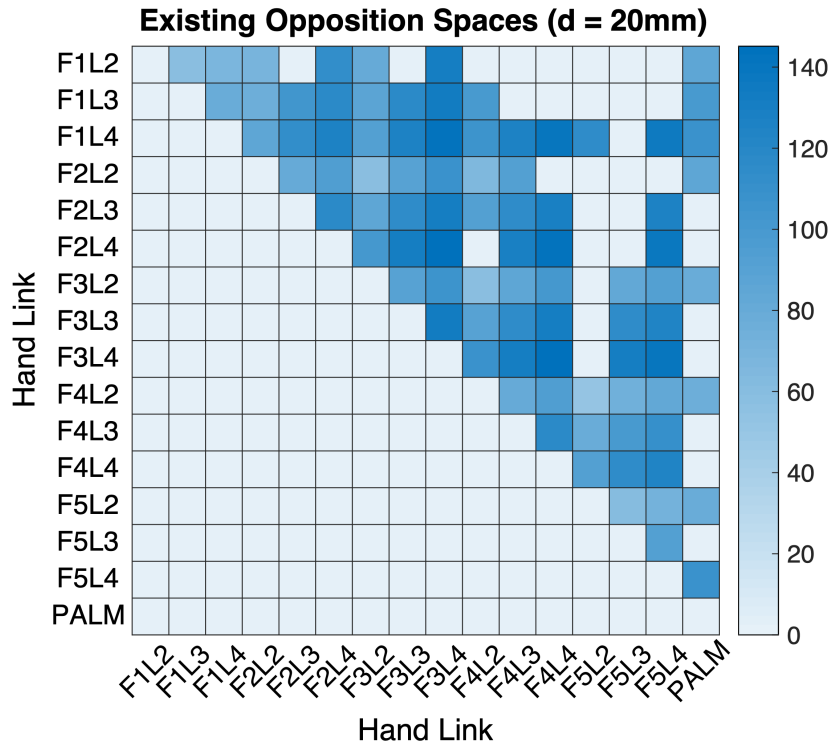
Figure 4.4 – Overview of the geometrically permissive opposition spaces structured inside the human hand model for a grasp distance $d = 20mm$ (coefficient of friction $\mu = 0.5$). Only the upper triangular of the map is visualized due to symmetry. Hand links listed on the label subject to our "finger-link" naming convention. The heatmap of entry values illustrates the largest feasible grasp distance within the corresponding opposition space (all values displayed here are larger than the given desired grasp distance ($d \geq 20mm$)). Empty entries correspond to opposition spaces that are unfeasible (e.g. entries on the diagonal) or not geometrically permissive ($d < 20mm$).

boundary constraints.

$$q \in [\underline{q}, \overline{q}], \ \forall q \in \mathbf{Q}_i, \ i = 1, \dots, N_c \tag{4.4}$$

with $\underline{q}$ and $\overline{q}$ being the corresponding lower and upper bound of q. N_c is the total number of contacts.

Object center Target object is described in $\{\mathbb{H}\}$, and it can only move in the chosen OS to guarantee the existence of contacts. An accurate description of such a bounded region is often difficult to obtain, as the opposition space is an irregular spatial volume in general, and the object's geometry also affects the boundary of this region. We approximate this constraint by restricting the object center $\boldsymbol{o} \in \mathbf{R}^3$ inside the entire opposition space:

$$\boldsymbol{o} \in \mathscr{S}, \tag{4.5}$$

It is possible to approximate the accurate geometric shape description of $\mathscr{S}$ by using the circumscribed spatial cuboid region of $\mathscr{S}$ to simplify the calculation.

Collision constraints

Collision must be avoided for a feasible grasping. We consider three types of collision in planning a grasp: (1) link-object collision, (2) link-link collision, and (3) object-object collision.

Link-object collision To avoid collision between target object and hand link, the spatial distance between the object center $\boldsymbol{o}$ and the center of an arbitrary link $\mathscr{L}_i$ must satisfy:

$$\boldsymbol{d}(\mathscr{O}, \mathscr{L}_i) \geq r_l + d_{\boldsymbol{o}, \boldsymbol{p}_i^*}, \ i = 1, \dots, N_l. \tag{4.6}$$

r_l is the distance from the link center to the contact on the link surface, and $d_{\boldsymbol{o}, \boldsymbol{p}_i^*}$ is the distance from the object center $\boldsymbol{o}$ to $\boldsymbol{p}_i^*$. $\boldsymbol{p}_i^*$ is the projection of $\boldsymbol{p}_i$ onto the object surface along the vector $\overrightarrow{\boldsymbol{p}_i^* \boldsymbol{o}}$.

The distance $\boldsymbol{d}(\mathscr{O}, \mathscr{L}_i)$ can be calculated using sampling based approach (El-Khoury et al. (2013)) by creating N uniformly sampled points on the central axis to represent a cylindrical shape. If $\mathscr{L}_i$ is a finger phalanx and approximated by $N_{\mathscr{L}}$ samples, $\boldsymbol{d}(\mathscr{O}, \mathscr{L}_i)$ can be calculated as $N_{\mathscr{L}}$ distance from the center $\boldsymbol{o}$ of a spherical object to each sample point. r_l is the radius of the phalanx cylinder, and $d_{\boldsymbol{o}, \boldsymbol{p}_i^*}$ is the radius of the object sphere. For a cylindrical object approximated by $N_{\mathscr{O}}$ samples, $\boldsymbol{d}(\mathscr{O}, \mathscr{L}_i)$ has a total number of $N_{\mathscr{L}} \times N_{\mathscr{O}}$ distances, and $d_{\boldsymbol{o}, \boldsymbol{p}_i^*}$ is the radius of the cylinder.

If $\mathscr{L}_i$ represents the palm, r_l is half the thickness of the palm cuboid, and $\boldsymbol{d}(\mathscr{O}, \mathscr{L}_i)$ is calculated as the projection distance of $\boldsymbol{o}$ (for a cylindrical object, it is each one of the samples on its central axis) onto the palm plane, along the palm's surface normal direction.

Link-link collision Constraints in this type ensure that the hand is in a self-collision-free configuration. The overlapping regions of hand reachability map (see Fig. 4.3b) indicate potential collision of links in Cartesian space.

A pair of links $\mathcal{L}_i$ and $\mathcal{L}_j$ may potentially collide, if the minimum spatial distance between their corresponding reachable spaces $\mathcal{R}_i$ and $\mathcal{R}_j$ is smaller than $r_{i,j}^l = r_i^l + r_j^l$. The distance r_i^l or r_j^l is the cylinder radius for a finger phalanx link or half the thickness of the cuboid for the palm link.

We construct a *self-collision map* $\mathcal{M}_{\mathscr{C}}$ to register all potential collisions between links (including palm) in the hand model, by analyzing the reachability map set of the hand $\{\mathcal{R}\}$.

Collisions that exist for all pairs of links $(\mathcal{L}_i, \mathcal{L}_j)$ are checked using sampling based approach to construct $\mathcal{M}_{\mathscr{C}}$. To check the collision between two finger phalanx links, each link has N_k uniformly and orderly sampled points from the head to the tail of the link on its central axis, $\{c_k^i\}$ and $\{c_k^j\}$, $k = 1, 2, \ldots, N_k$, respectively. Then the constraint is formulated as a vector of N_k pairwise distances between sampled points and must satisfy:

$$\boldsymbol{d}(\mathcal{L}_i, \mathcal{L}_j) := [d_{c_1^i, c_1^j}, \ldots, d_{c_k^i, c_k^j}, \ldots, d_{c_{N_k}^i, c_{N_k}^j}]^\mathsf{T} \geq 2\boldsymbol{r}_{i,j}^l,$$
$$\boldsymbol{r}_{i,j}^l \in \mathbf{R}^{N_k \times 1}, \ k = 1, 2, \ldots, N_k, \tag{4.7}$$
$$i = 1, 2, \ldots, N_l, \ j = 1, 2, \ldots, N_l, \ i \neq j.$$

The collision between a finger phalanx link $\mathcal{L}_i$ and the palm is constrained by forcing the projection distance of each sample c_k^i to the palm surface larger than r_i^l.

The self-collision map can be visualized as a heatmap. For example, Fig. 4.5 shows the self-collision map of the human hand model used in this study. The entry values of the map represent the maximum collision depth, i.e. $\max(r_{i,j}^l - \boldsymbol{d}(\boldsymbol{r}_i, \boldsymbol{r}_j))$, $\boldsymbol{r}_i \in \mathcal{R}_i$, $\boldsymbol{r}_j \in \mathcal{R}_j$, $i \neq j$. Only positive values indicate potential collision. The upper bound of this depth is $2r^l$.

Object-object collision In the planning of multiple objects grasping (see Sec. 4.5), any two objects $\mathcal{O}_i$ and $\mathcal{O}_j$ should not intersect. The spatial distance between two objects should satisfy:

$$\boldsymbol{d}(\mathcal{O}_i, \mathcal{O}_j) \geq d_{\boldsymbol{o}_i, \boldsymbol{o}_j^*} + d_{\boldsymbol{o}_j, \boldsymbol{o}_i^*},$$
$$i = 1, 2, \ldots, N_o, \ j = 1, 2, \ldots, N_o, \ i \neq j, \tag{4.8}$$

where $\boldsymbol{o}_j^*$ is the projection of $\boldsymbol{o}_j$ onto the surface of $\mathcal{O}_i$ along the line of $\boldsymbol{o}_i - \boldsymbol{o}_j$, and $d_{\boldsymbol{o}_i, \boldsymbol{o}_j^*}$ is the distance from object center $\boldsymbol{o}_i$ to this projected point. For a spherical object, $d_{\boldsymbol{o}_i, \boldsymbol{o}_j^*}$ equals the radius. A cylindrical object can be represented by a sequence of $N_{\mathcal{O}}$ sample points on its central axis, as introduced previously. In this case, for each one of the $N_{\mathcal{O}}$ samples, $d_{\boldsymbol{o}_i, \boldsymbol{o}_j^*}$ is the radius of the cylinder.
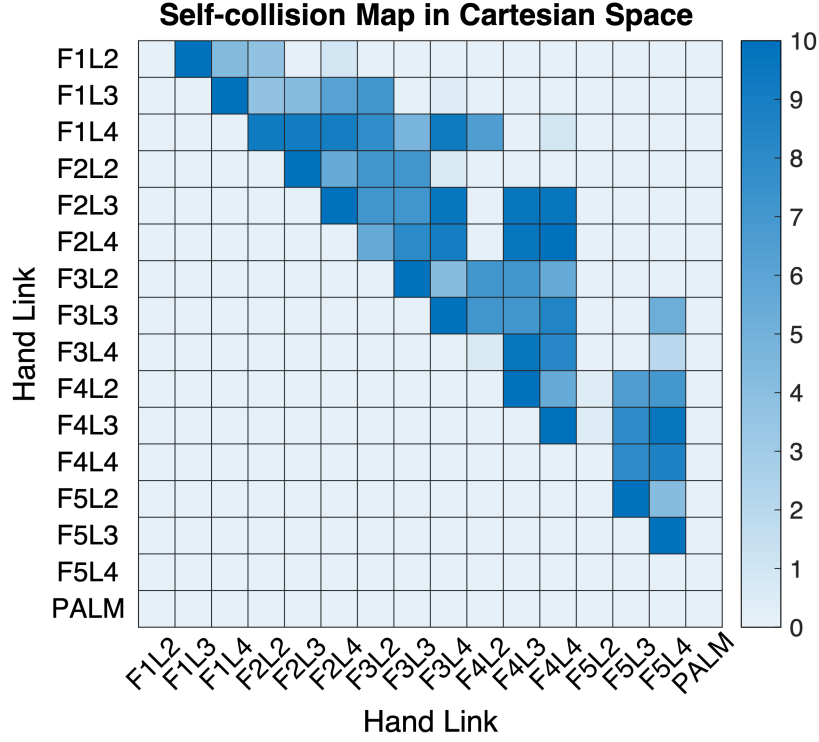
Figure 4.5 – Self-collision map of the human hand model, visualized as heatmap. The radius of all finger phalanges is $r^l = 5mm$. Hand links listed on the label subject to our "finger-link" naming convention. The heatmap of entry values illustrates the maximum potential collision depth in Cartesian space between two links. Empty entries indicate collision-free link pairs.

For irregular-shaped objects, $d_{\boldsymbol{o}_i,\boldsymbol{o}_j^*}$ can be calculated if an explicit surface model is available. Alternatively, the distance can also be estimated based on an implicit representation of the object surface (El-Khoury et al. (2015)). Sampling-based approach could also be applied to calculate distance for irregular shaped objects. For instance, approximate the shape of each object by sampling multiple points on the surface. The minimum distance between samples on the surface of $\mathscr{O}_i$ and samples on $\mathscr{O}_j$ can serve as an estimation of $\boldsymbol{d}(\mathscr{O}_i, \mathscr{O}_j)$.

In this work, we calculate $\boldsymbol{d}(\mathscr{O}_i, \mathscr{O}_j) = r^{\mathscr{O}_i} + r^{\mathscr{O}_j}$ for a pair of spherical objects. Give a cylindrical object and a spherical object, the distance becomes a vector of length $N_{\mathscr{O}}$ with each entry representing the Euclidean distance between the spherical object center and one of the $N_{\mathscr{O}}$ sample points on the central axis of the cylindrical object. Similarly, for a pair of cylindrical objects, the distance vector has $N_{\mathscr{O}_i} \times N_{\mathscr{O}_j}$ entries.

Contact constraints

Contact constraints (1) ensure contacts being established between hand link and the target object, and also (2) provide grasping stability guarantee.

In-contact constraint This equality constraint guarantees that each contact point $\boldsymbol{p}_i$ on the surface of the hand link $\mathscr{L}_i$ must also locate on the object surface:

$$\boldsymbol{d}(\mathscr{O}, \boldsymbol{p}_i) = d_{\boldsymbol{o}, \boldsymbol{p}_i^*}, \; i = 1, 2, \ldots, N_c. \tag{4.9}$$

For a spherical object, $\boldsymbol{d}(\mathscr{O}, \boldsymbol{p}_i) := d_{\boldsymbol{o}, \boldsymbol{p}_i}$, and $\boldsymbol{p}_i^*$ is the projection of $\boldsymbol{p}_i$ onto the sphere surface along the vector $\overrightarrow{\boldsymbol{p}_i \boldsymbol{o}}$. $d_{\boldsymbol{o}, \boldsymbol{p}_i^*}$ is the radius of the object. For a cylindrical object, $\boldsymbol{d}(\mathscr{O}, \boldsymbol{p}_i)$ is the distance from $\boldsymbol{p}_i$ to the line segment of the cylinder central axis, and $\boldsymbol{p}_i^*$ is the projection of $\boldsymbol{p}_i$ onto the cylinder surface along the cylinder radical direction.

Stability constraint We use force closure property as a condition to ensure grasp stability. It states that any external wrenches applied on the grasped object can be counterbalanced by a resultant wrench applied through contacts. In a point contact model with friction, contact force must lie inside the corresponding friction cone, such that the Coulomb's law is satisfied and slip does not occur on the contact point (Kraus et al. (1998)).

Force $\boldsymbol{f}^i$ generated by the ith contact point and lying inside its friction cone can be approximated as a linear combination of all edges of the N_f-sided friction cone, centered about the contact normal $\boldsymbol{n}_i$ (Miller and Allen (1999)):

$$
\begin{aligned}
\boldsymbol{f}^i &\approx \sum_{n=1}^{N_f} \lambda_n^i \mathbf{f}_n^i, \\
\lambda_n^i &\geq 0, \; \sum_{n=1}^{N_f} \lambda_n^i = 1, \; i = 1, \ldots, N_c, \; n = 1, \ldots, N_f.
\end{aligned}
\tag{4.10}
$$

The calculation related to friction cone approximation is explained in Appendix B.

The torque generated by the force component $\mathbf{f}_n^i$ on the object center is:

$$\boldsymbol{\tau}_n^i = \boldsymbol{d}^i \times \mathbf{f}_n^i, \tag{4.11}$$

where $\boldsymbol{d}^i$ is the vector from the object center $\boldsymbol{o}$ to the ith contact point $\boldsymbol{p}_i$, and the associated primitive contact wrench is:

$$\boldsymbol{w}_n^i = \begin{pmatrix} \mathbf{f}_n^i \\ \boldsymbol{\tau}_n^i \end{pmatrix} \tag{4.12}$$

A grasp has force closure if and only if the origin of the wrench space lies inside the convex hull constructed by the contact wrenches (Montana (1991)). This can be formulated as a convex

optimization problem by determining a set of real positive coefficients $\{\varphi_n^i\}$ that satisfy:

$$\exists \varphi_n^i \in \mathbf{R}, \ \varphi_n^i \geq 0, \ \sum_{i,n} \varphi_n^i = 1,$$
$$i = 1, 2, \dots, N_c, \ n = 1, 2, \dots, N_f, \tag{4.13}$$
$$\text{s.t.} \ \sum_{i,n} \varphi_n^i \boldsymbol{w}_n^i = 0.$$

The coefficients $\{\varphi_n^i\}$ are integrated into the grasp synthesis optimization as optimization variables.

Other constraints The quaternion parameterizing the object orientation must be constrained such that its norm is 1:

$$||\boldsymbol{q}^{\mathscr{O}}|| = 1. \tag{4.14}$$

Objective function

We formulate the objective function by considering mainly the grasp quality cost and the movement economy cost.

Grasp quality cost Although satisfying force closure constraint guarantees grasping stability, the force closure property could be in poor quality and is easy to be violated (Zhu and Ding (2004)). This margin of violating force closure property is geometrically equivalent to the minimum distance from the surface of the convex hull to the origin in the grasp wrench space (Borst et al. (2004)), and a handful of metrics have been formulated for grasp quality evaluation (Roa and Suárez (2015); Aleotti and Caselli (2010)).

We use two metrics proposed by (Klein et al. (2020)) as costs in both force and torque aspects. Such metrics have been used in the analysis of human precision grasps of 3D objects and proved to be effective. In a grasp with two contacts, the quality metric can be simply calculated by analyzing the relationship between the force application directions (Nguyen (1988)). In ideal cases, the contact normal should be aligned with the grasp axis (i.e. the line connecting two contact points), such that both applied contact forces lie along the central axes of the corresponding friction cones, and the wrench space has the largest margin under force closure constraint.

The force cost (C_f) penalizes the deviance between grasp axis and contact normal:

$$C_f = \arctan(||\boldsymbol{n}_1 \times (\boldsymbol{p}_2 - \boldsymbol{p}_1)||, \ \boldsymbol{n}_1 \cdot (\boldsymbol{p}_2 - \boldsymbol{p}_1))$$
$$+ \arctan(||\boldsymbol{n}_2 \times (\boldsymbol{p}_1 - \boldsymbol{p}_2)||, \ \boldsymbol{n}_2 \cdot (\boldsymbol{p}_1 - \boldsymbol{p}_2)), \tag{4.15}$$

where $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are the directions of force applications, and $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are the Cartesian coordinates of contacts represented in $\{\mathbb{H}\}$.

The torque cost (C_t) penalizes the magnitude of the total torque to avoid rotation of the grasped object:

$$C_t = ||(\boldsymbol{o} - \boldsymbol{p}_1) \times (-\boldsymbol{g}) + (\boldsymbol{o} - \boldsymbol{p}_2) \times (-\boldsymbol{g})||, \tag{4.16}$$

where $\boldsymbol{g}$ is a unit vector pointing towards the gravitational direction.

Movement economy cost This cost aims to suppress unnecessary motion of fingers, to avoid occupying of adjacent OSs constructed by other links. $\boldsymbol{\theta}_{actv}^i$ is a vector consisting of all joint angles affecting the contact point on the ith finger, and $\Delta\boldsymbol{\theta}_{actv}^i$ indicates the change in joint angles compared to initial position (fully extension or no ab-/adduction). We add movement economy cost to penalize excessive joint angles:

$$C_q = \sum_{i=1}^{N_c} \Delta\boldsymbol{\theta}_{actv}^i{}^\mathsf{T} Q_i \, \Delta\boldsymbol{\theta}_{actv}^i, \ i = 1, \ldots, N_c, \tag{4.17}$$

where N_c is the number of contacts, and Q_i is a positive definite diagonal matrix that prioritizes motion of different finger joints.

Objective function By integrating the above-mentioned costs, we formulate the objective function as:

$$\mathscr{Q} = \lambda \cdot (C_f + C_t) + (1 - \lambda) \cdot C_q, \ \lambda \in [0,1]. \tag{4.18}$$

The weighting parameter λ balances costs in both aspects.

Grasp synthesis formulation

Considering the constraints and objective function described above, we formulate the human-like dexterous grasping synthesis problem as constrained optimization:

$$\Theta^* = \arg\min_{\Theta} \ \mathscr{Q}$$

subject to (4.4), (4.5), (4.6), (4.7), (4.8), (4.9), (4.13), (4.14) $\tag{4.19}$

The optimization variable Θ consists of (1) $\mathbf{Q}^{\mathscr{O}}$, parameters that parameterize the object reference frame $\{\mathbb{O}\}$, including object center position and quaternion (for cylindrical objects), (2) $\mathbf{Q}_i = \{\boldsymbol{\theta}_{actv}^i, \psi_i, \alpha_i\}$, $i = 1, \ldots, N_c$, the kinematic variables that parameterize each contact reference frame $\{\mathbb{C}_i\}$ on the hand link, and (3) $\{\varphi_n^i\}$, $i = 1, \ldots, N_c$, $n = 1, \ldots, N_f$, coefficients associated with the edges of friction cones on contact points.

We summarize our proposed algorithm in Algorithm 1. The kinematic redundancy set $\mathscr{K}$ contains all available degrees of freedoms in the hand model.

Given the hand model $\mathscr{H}$ and a set of its degrees of freedoms $\mathscr{K}$, the first step after ini-

---

Algorithm 1: Human-like grasping using arbitrary hand surface regions

---

Data: hand model $\mathcal{H}$;

kinematic redundancy set $\mathcal{K}$;

target object $\mathcal{O}$;

Result: hand configuration $\mathcal{H}^*$ and object pose $\mathcal{O}^*$ generated by the optimal grasp, described by the corresponding optimization variables Θ^*;

1 initialization;

2 **if** $\mathcal{K} \neq \emptyset$ **then**

3 $\{\mathcal{R}\} \leftarrow$ reachability_map($\mathcal{H}, \mathcal{K}$);

4 $\{\mathcal{S}\} \leftarrow$ opposition_spaces($\{\mathcal{R}\}$);

5 $\{\mathcal{S}_\mathcal{O}\} \leftarrow$ OS_candidates($\{\mathcal{S}\}, \mathcal{O}$);

6 $\{\Theta\} = \emptyset$;

7 $k = 1$;

8   **while** $\{\mathcal{S}_\mathcal{O}\} \neq \emptyset$ & $\mathcal{K} \neq \emptyset$ **do**

9 $\mathcal{S}_{i,j} \leftarrow$ select_OS($\{\mathcal{S}_\mathcal{O}\}$);

10 $\mathbf{Q}_i, \mathbf{Q}_j \leftarrow$ extract_parameters($\mathcal{L}_i, \mathcal{L}_j, \mathcal{S}_{i,j}$);

11 $\Theta_k \leftarrow$ grasp_synthesis($\mathbf{Q}^o, \mathbf{Q}_i, \mathbf{Q}_j, \{\varphi_n^i\}, \{\varphi_n^j\}$);

12 $\{\mathcal{S}_\mathcal{O}\} \leftarrow \{\mathcal{S}_\mathcal{O}\} \setminus \mathcal{S}_{i,j}$;

13 $k = k + 1$;

14   **if** $\Theta_k \neq \emptyset$ **then**

15 $\{\Theta\} \leftarrow \{\Theta\} \bigcup \Theta_k$;

16   **else**

17   continue;

18   **end**

19   **end**

20 $\Theta^* =$ minimum_cost($\{\Theta\}$);

21 $(\mathcal{H}^*, \mathcal{O}^*, \mathcal{K}) =$ update_configuration($\mathcal{H}, \mathcal{O}, \Theta^*$);

22   return $\mathcal{H}^*, \mathcal{O}^*$;

23 **else**

24   return $\mathcal{H}, \mathcal{O}$;

25 **end**

---

tialization is to construct the reachability map set $\{\mathscr{R}\}$ for all hand links. Then, the set of geometrically permissive opposition spaces $\{\mathscr{S}_{\mathscr{O}}\}$ are determined together with the object model $\mathscr{O}$. Any OS from this set, $\mathscr{S}$, can be selected as a candidate for grasp synthesis. Once the candidate OS is selected, the corresponding reachable spaces, $\mathscr{R}_i$ and $\mathscr{R}_j$, along with the associated hand links $\mathscr{L}_i$ and $\mathscr{L}_j$ are also determined. Each link is supposed to establish one contact on the object surface. Kinematic parameters of each contact ($\mathbf{Q}_i$ and $\mathbf{Q}_j$), object pose parameter set $\mathbf{Q}^{\mathscr{O}}$, as well as the coefficients $\{\varphi_n^i\}$ and $\{\varphi_n^j\}$ are added to the optimization variable set.

If multiple opposition space candidates have been selected and more than one OS has resulted in feasible solutions, the optimal grasp is determined as the one leads to the minimum cost among the solution set. Finally, the optimal solution Θ^* is used to update the kinematic configuration of both the hand model and the object model to realize the grasp. After obtaining a successful grasp, the used DOFs in the grasp are excluded from the kinematic redundancy set $\mathscr{K}$. Such DOFs will no longer be available in subsequent tasks.

4.4.4 Increasing contacts

Once an optimal grasp has been determined and $\mathscr{K}$ is not empty, available degrees of freedom in $\mathscr{K}$ can be controlled to move free links and establish more contacts with the grasped object. These extra contacts could be used to either improve grasping quality or to execute desired manipulation motion (Sommer and Billard (2016)).

4.5 Grasping multiple objects

We start by presenting an algorithm to grasp multiple objects in sequence. Then, we propose a kinematic efficiency metric and an associated strategy to explore multiple solutions while leveraging kinematic redundancy of the robot hand.

4.5.1 From 1 to N: sequential grasping of multiple objects

We consider a set of N_o target objects, denoted as $\{\mathscr{O}_i\}$, $i = 1, 2, \ldots, N_o$, and the robotic hand grasps them following the order $\boldsymbol{\Lambda}$, which is an ordered list of objects indices. For example, a total of $N_o = 5$ objects can be grasped in the order of $\mathscr{O}_3, \mathscr{O}_1, \mathscr{O}_5, \mathscr{O}_4, \mathscr{O}_2$, given the grasp order $\boldsymbol{\Lambda} = [3, 1, 5, 4, 2]$.

The *Analysis-Grasp* iterative process

Taking advantage of our proposed grasp synthesis algorithm, we reformulate the problem of grasping multiple objects as a sequence of single object grasping sub-problems. Each sub-problem deals with the grasping of one single object, and it consists of two steps. In the
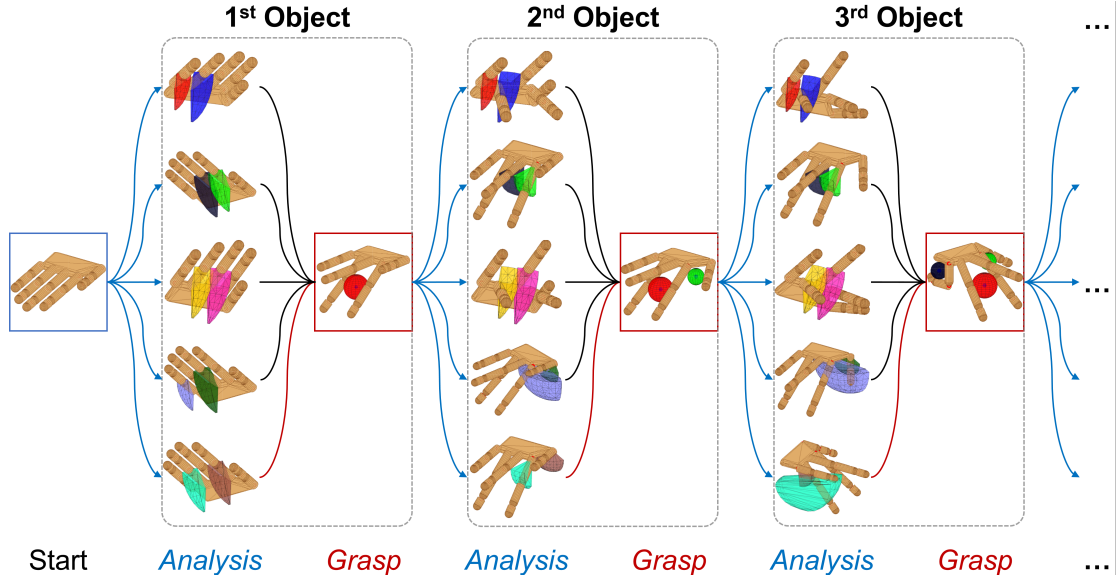
Figure 4.6 – Illustration of the *Analysis-Grasp* iterative process for grasping multiple objects in sequence. In this example, five OSs are selected as the candidates to generate grasp of each object, by analyzing the updated reachability map. The OS that leads to the optimal solution (minimizing the kinematic efficiency metric κ, see Eq. (4.20)) is used in the final grasp.

analysis step, the robotic hand constructs (if this is the first grasp in the sequence) or updates the hand reachability map by resampling all its currently available degrees of freedom, and then selects one or multiple geometrically permissive opposition spaces (OSs) as candidates for subsequent grasping synthesis. In the *grasp* step, each one of the OS candidates is tested for grasping the object. The optimal solution is used to update the configuration of hand and object, and also the kinematic redundancy set $\mathcal{K}$. Fig. 4.6 illustrates this iterative process by demonstrating a three-object grasping problem with five OS candidates have been selected in the grasp of each object. The robotic hand repeats this iterative process for each object, until either all given objects are grasped or no more opposition spaces can be employed to generate a feasible grasp.

Selection of opposition spaces

To determine the optimal grasp for each target object $\mathcal{O}_i$, it is possible to test all feasible OS candidates in the set $\{\mathcal{S}_{\mathcal{O}_i}\}$ using Algorithm 1. However, this inevitably demands large computational cost. It could also result in an optimal but inefficient grasp, e.g. an OS having large space capacity has been used to grasp a tiny object. This may be problematic for the subsequent grasps of potentially large objects.

To alleviate these two potential issues, in each sub-problem, we only select a mini batch of N ($N < |\{\mathcal{S}_{\mathcal{O}_i}\}|$) OS candidates that correspond to the N smallest capacity values $\overline{C}_{\mathcal{S}}$ (see Eq. (4.2)) among all OSs. In case no feasible solutions can be obtained using these N candidates, more

---

Algorithm 2: Sequential grasping of multiple objects

Data: hand model $\mathcal{H}$;
kinematic redundancy set $\mathcal{K}$;
set of target objects $\{\mathcal{O}_i\}$, $i = 1, \ldots, N_o$;
grasping order $\boldsymbol{\Lambda}$;
Result: optimal grasp configuration, including the configuration of hand $\mathcal{H}^*$ and the
configuration of each grasped object $\{\mathcal{O}_i^*\}$, $i = 1, \ldots, N_g$;

1 initialization;
2 $k = 1$;
3 $\{\mathcal{O}_i^*\} = \emptyset$;
4 **while** $\{\mathcal{O}_i\} \neq \emptyset$ & $\mathcal{K} \neq \emptyset$ **do**
5 $\quad$ $\mathcal{O}_k \leftarrow$ select_object($\{\mathcal{O}_i\}$, $\boldsymbol{\Lambda}_k$);
6 $\quad$ $\Theta_k^* \leftarrow$ Algorithm_1($\mathcal{H}$, $\mathcal{K}$, $\mathcal{O}_k$);
7 $\quad$ $\{\mathcal{O}_i\} \leftarrow \{\mathcal{O}_i\} \setminus \mathcal{O}_k$;
8 $\quad$ $k = k + 1$;
9 $\quad$ **if** $\Theta_k^* \neq \emptyset$ **then**
10 $\quad\quad$ $(\mathcal{H}, \mathcal{O}_k^*) \leftarrow$ update_configuration($\mathcal{H}$, $\mathcal{O}_k$, Θ_k^*);
11 $\quad\quad$ $\mathcal{K} \leftarrow$ update($\mathcal{K}$, Θ_k^*);
12 $\quad\quad$ $\{\mathcal{O}_i^*\} \leftarrow \{\mathcal{O}_i^*\} \cup \mathcal{O}_k^*$;
13 $\quad$ **end**
14 **end**
15 $\mathcal{H}^* \leftarrow \mathcal{H}$;
16 return $\mathcal{H}^*$, $\{\mathcal{O}_i^*\}$;

---

feasible OSs can be selected and tested following the ascending order of their capacity values.

Update of hand model

Once the target object has been successfully grasped, the employed OS is excluded from the total OS set $\{\mathscr{S}\}$, and all DOFs affecting the contacted hand links are considered fixed and hence excluded from the kinematic redundancy set $\mathscr{K}$.

Our algorithm enables one link (finger phalanx or the palm) to be reused in multiple grasps, as long as it spans opposition spaces with other links, even though these contacted links are considered to be fixed. We summarize the algorithm for sequential grasping of multiple objects in Algorithm 2.

Notice that Algorithm 2 takes care of testing the mini-batch OS candidates and returns an optimal solution. If no feasible solution is found after testing all OS candidates, indicating the current object cannot be grasped, the hand then continues to grasp the next object. Successfully grasped objects are added to the set $\{\mathscr{O}_i^*\}$. The hand model $\mathscr{H}$ along with its kinematic redundancy set $\mathscr{K}$ are updated at the end of each iteration. Notice that each grasped object potentially brings in novel object-object collision constraints (Eq. (4.8)) to the subsequent grasps.

4.5.2 One step further: greedy grasping of more objects

The proposed sequential grasping algorithm (Algorithm 2) enables the robotic hand to grasp multiple target objects. However, the final grasping configuration may not be "*optimal*" in the sense of number of grasped objects. On the one hand, the solution to each grasp sub-problem is not unique. The selected mini-batch OS candidates do not guarantee the optimal solution among all feasible OSs, and there may also exist multiple OSs resulting in feasible, even optimal grasps. On the other hand, the grasping order $\mathbf{\Lambda}$ affects the hand configuration for each sub-problem. Each grasp is influenced by all its previous grasps, and also affect its subsequent grasps. We propose the following strategies to tackle these issues in a multiple-objects grasping problem.

Kinematic efficiency metric

We measure the kinematic efficiency for each grasp configuration from three aspects: (1) the total number of contacts, $N_c \in \mathscr{N}$, (2) the total number of engaged finger joints, $N_q = \sum_{i=1}^{N_c} |\boldsymbol{\theta}_{actv}^i| \in \mathscr{N}$, and (3) the capacity ratio, $\eta = \overline{C}_{\mathscr{S}}/||\boldsymbol{p}_1 - \boldsymbol{p}_2|| \in [1, \infty)$. $\boldsymbol{\theta}_{actv}^i$ is a vector composed of all joint angles affecting the contact point on i^{th} finger, and $|\boldsymbol{\theta}_{actv}^i|$ represents its cardinality. As stated previously, we plan grasps inside opposition spaces and hence only consider grasps with two contacts. Therefore, $N_c = 2$ remains a constant value in our proposed algorithm.

η measures the *efficiency* of exploiting the selected OS by taking the ratio of its maximum capacity to the actual grasp distance inside this OS. A grasp that fully exploits the OS capacity has a η value close to 1, indicating that the size of the grasped object is slightly smaller or almost equal to the maximum capacity of the OS.

We propose the following kinematic efficiency metric κ by integrating the above-mentioned aspects:

$$\kappa = e^{N_c} \cdot e^{N_q} \cdot e^{\eta}. \tag{4.20}$$

The exponential mapping guarantees the continuity of this metric and the continuity of its derivatives.

It measures the overall consumption of kinematic redundancy in the hand model given a grasp configuration. A more kinematically efficient grasping configuration corresponds to a smaller κ value.

To facilitate the exploitation of kinematic redundancy, we integrate the kinematic efficiency metric into the objective function for single object grasp (see Eq. (4.18)) and use it as the objective function for sub-problems in multiple-objects grasp:

$$\mathscr{Q} = \lambda \cdot (\kappa (C_f + C_t)) + (1 - \lambda) \cdot C_q, \ \eta \in [0, 1]. \tag{4.21}$$

This objective aims at penalizing the excessive use of opposition space capacity and improving the efficiency of utilizing kinematic redundancy for each sub-problem.

Thus, by minimizing the kinematic efficiency metric in the iterative process of grasping multiple objects, the robotic hand (1) prioritizes using the OS candidate that employs fewer DOFs (e.g. the OS closer to the base of the kinematic chain), and (2) when multiple OS candidates demanding the same number of DOFs, prefers the smallest OS possible. It hence preserves as much kinematic redundancy as possible for subsequent tasks during the iterative grasping process.

Notice that for the single object grasping synthesis problem, once a desired OS is given, e^{N_c} and e^{N_q} in κ become constant values. The metric κ is then only a function of η, determined by the optimization variables in contact points. Thus, from a computational point of view, potential abrupt changes in objective function value caused by the change of employed DOFs (N_q) can be avoided when solving the optimization problem for each single object grasp.

We hence integrates the kinematic efficiency metric κ in the single object grasp synthesis problem by replacing the objective function in Algorithm 1 with Eq. (4.21).

Algorithm 3: Greedy grasping of multiple objects

Data: hand model $\mathcal{H}$;
kinematic redundancy set $\mathcal{K}$;
set of target objects $\{\mathcal{O}_i\}$, $i \in \mathcal{I}$;
index set of objects $\mathcal{I} = \{1,\ldots,N_o\}$;
Result: hand configuration $\mathcal{H}^*$ and the set of grasped object poses $\{\mathcal{O}_i^*\}$, $i = 1,\ldots,N_g$
    that associated with the optimal sequential grasping;

1 initialization;
2 $\mathrm{P}_{\mathcal{I}} = \text{generate\_permutation}(\mathcal{I})$;
3 $p = 1$;
4 **while** $p < |\mathrm{P}_{\mathcal{I}}|$ **do**
5 $\quad \Lambda_p \leftarrow \mathrm{P}_{\mathcal{I}}(p)$;
6 $\quad (\mathcal{H}_p, \mathcal{K}, \{\mathcal{O}_i^*\}_p, C_{\Lambda_p}^{seq}) \leftarrow \text{Algorithm\_2}(\mathcal{H}, \mathcal{K}, \{\mathcal{O}_i\}, \Lambda_p)$;
7 $\quad p = p + 1$;
8 **end**
9 $p^* \leftarrow \text{optimal}(\arg\max_p(|\{\mathcal{O}_i^*\}_p|), \arg\min_p(\{C_{\Lambda_p}^{seq}\}))$;
10 $\mathcal{H}^* \leftarrow \mathcal{H}_{p^*}$;
11 $\{\mathcal{O}_i^*\} \leftarrow \{\mathcal{O}_i^*\}_{p^*}$;

Optimal greedy grasping

The solution to a sequential grasping is deterministic, as long as the grasp sequence is defined and the size of the mini-batch OS set for each object is determined. Therefore, to reduce the influence of grasping order on the final solution, we generate a permutation of the objects index set $\mathcal{I}$:

$$\mathrm{P}_{\mathcal{I}} = \{\Lambda_1,\ldots,\Lambda_{N_\Lambda}\}, \; \mathcal{I} = \{1,\ldots,N_o\}, \tag{4.22}$$

where $N_\Lambda = N_o!$ is the number of permutations for N_o objects.

The robot then performs N_Λ independent trials of sequential grasps by following each grasping order Λ in the permutation set P_{Λ}.

The optimal solution to the greedy grasping problem is the sequential grasping having most objects grasped; and for multiple grasping sequences result in the same number of grasped objects, the sequence leads to the minimum *kinematic redundancy cost* is considered optimal. We define the kinematic redundancy cost of a sequential grasping following order Λ_k as:

$$C_{\Lambda_k}^{\text{seq}} = \sum_i^{N_g} \overline{C}_{\mathcal{S}_i}, \; i = 1,\ldots,N_g, \tag{4.23}$$

where $\overline{C}_{\mathcal{S}_i}$ is the maximum capacity of $\mathcal{S}_i$ that is used to grasp the ith object. $C_{\Lambda_k}^{\text{seq}}$ describes the overall consumption of kinematic redundancy that has been used to grasp N_g objects in total. The smaller the cost, the larger the kinematic redundancy remaining in the model. We summarize our greedy grasping algorithm in Algorithm 3.

Nr.	1	2	3	4	5	6	7	8	9	10
Object										
Name	Wooden ball	Ping-pong ball	Soft ball	Tennis ball	Battery	Wine stopper	Pills bottle	Glue stick	Glass bottle	Candy jar
Geometry Type	Sphere	Sphere	Sphere	Sphere	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder
Radius (mm)	17.5	20	29	32	7	12	14.5	8	28.5	37
Height (mm)	N/A	N/A	N/A	N/A	51	45	86	147	105	107

Figure 4.7 – Experimental objects.

Figure 4.8 – Experimental setup: the allegro hand is mounted on a right stand for stability with its fingers in open position. The 10 objects are placed in front of the robotic hand, for ease of visualization of the size of the objects with respect to the hand's size.

(a) F1L4-F2L4   (b) F3L4-PALM   (c) F2L3-F3L3   (d) F2L2-F4L2   (e) F2L2-F2L4   (f) F2L2-F2L4

(g) F1L4-F2L4   (h) F3L4-PALM   (i) F3L4-F4L4   (j) F2L3-F4L3   (k) F2L2-F2L4   (l) F1L4-F2L4
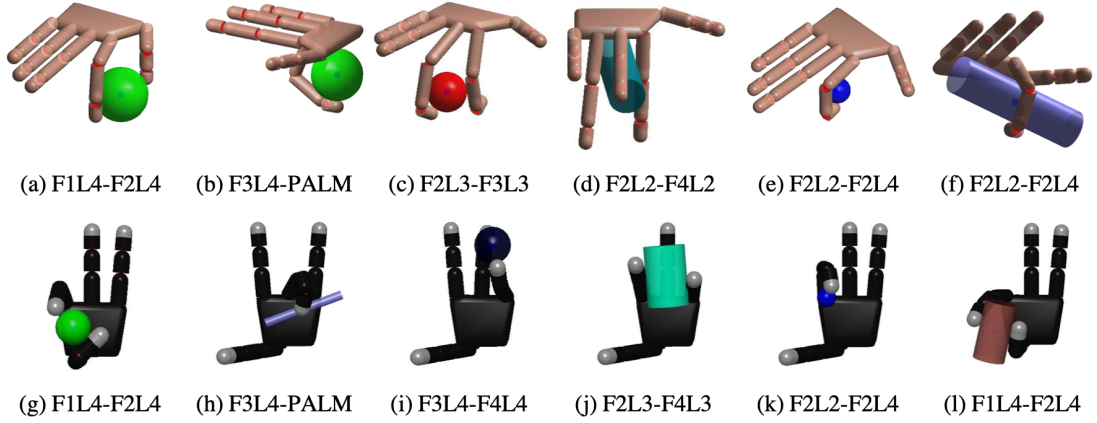
Figure 4.9 – Examples of human-like dexterous grasping of a single object using various opposition spaces, simulated on both human hand model and Allegro hand model. The title of each sub-figure indicates the OS that has been used to form this grasp, represented by its associated links.

4.6 Experiments

We designed three experiments to evaluate our proposed algorithms for (1) human-like dexterous grasping using arbitrary hand surface regions (Algorithm 1), (2) sequential grasping of multiple objects (Algorithm 2), and (3) greedy grasping of multiple objects (Algorithm 3).

4.6.1 Implementation details

Our grasping synthesis algorithms are implemented in MATLAB (The MathWorks®, Inc.). The optimization problem (see Eq. (4.19)) is non-convex and may have multiple local minimums. We use the *GlobalSearch* algorithm together with the sequential quadratic programming (SQP) algorithm in Optimization Toolbox™, in order to find a single global minimum more efficiently. We formulate both objective function and constraints as symbolic expressions using the MATLAB Symbolic Math Toolbox™. Analytical gradients of both objective function and constraints have been calculated to accelerate the optimization.

Prior to solving the optimization problem, we performed offline analysis of the hand kinematic model to construct the reachability map, collision map, and the opposition space set. For simplicity, we used convex hulls (an extreme alpha shape when $\alpha = \infty$) to model the geometry of the reachable space for each finger phalanx.

Each friction cone is approximated by a triangular pyramid. The calculation is explained in Appendix B. We use a coefficient of friction $\mu = 0.5$, typical of common materials (Lide (2004)).

In the objective function, we use a diagonal matrix $Q_i = \mathrm{diag}[10, 50, 25, 10]$ to prioritize the motion of each DOF (first entry corresponds to the ab-/adduction DOF) and $\lambda = 0.5$ to bal-

ance costs in the grasping quality and movement economy aspects. When formulating the constraints, each cylindrical geometry is approximated by a number of $\lceil \frac{h}{r} \rceil$ samples on the central axis, where h is the height of the cylinder and r is its radius. A minimum number of 3 samples are used in case of $h \leq r$.

Hand model moved to the fully open pose as being initialized. This configuration corresponds to the upper bound of all flexion/extension DOFs and the middle value (0 position) of the ab-/adduction DOFs. Variables that parameterize object pose were randomly initialized within corresponding boundaries.

4.6.2   Computational costs

Experiments were conducted in MATLAB R2018b on a laptop with Intel(R) Core™i7-7600 CPU2.80GHz. Constructing a hand model (including computing symbolic expressions of hand kinematics) and offline analyzing kinematic properties takes around one minute. Average time spent on the construction and offline analyses for each hand model in simulation is summarized in Tab. 4.3. The time performance statistics reported here are obtained from ten independent experimental trials. *Modeling* of hand consists of building the kinematic chain and computing the corresponding symbolic expressions of forward kinematics of all links and joints. The *reachability map* $\{\mathcal{R}\}$ and *self-collision* map $\mathcal{M}_{\mathcal{C}}$ are constructed for the entire hand model, with all DOFs being sampled in their corresponding motion ranges.

Table 4.3 – Computational cost on offline construction and analysis of hand models.

Model Type	Modeling	$\{\mathcal{R}\}$	$\mathcal{M}_{\mathcal{C}}$
Human hand	15.12 ± 0.15	2.93 ± 0.16	2.81 ± 0.40
Robotic hand	62.32 ± 1.28	4.90 ± 0.22	3.55 ± 0.04

Formulating a single object grasp synthesis problem takes $10-30$ seconds in most cases, resulting in $20-30$ optimization variables, $100-200$ nonlinear inequality constraints, and around 20 nonlinear equality constraints. Given a complete formulation, such an optimization problem can be solved within seconds (measured for single optimization trial, mean±std: 14.71 ± 12.68 seconds, value range $1.56-45.19$ seconds, statistics obtained from the 50 simulations for the robotic experiments in Fig. 4.10).

4.6.3   Robotic experimental setup

We reproduced our synthesized grasps on a 16-DOF real Allegro robotic left hand (Wonik Robotics Co., Ltd.). The robotic hand was mounted on a metal adapter on a table surface in standing posture (fingertips pointing upwards when fully extended), as Fig. 4.8 shows. The hand was connected to an Ubuntu 18.04 desktop PC through a CAN bus. It was controlled by a joint position PID controller in ROS Melodic at 10 Hz.

We selected 10 everyday objects (see Fig. 4.7) for robotic experiments, and the geometry of each object has been approximated using either sphere or cylinders in various sizes. Opti-Track™motion capture system was used to localize the robotic hand and each target object in the experiment. Reflexive markers have been adhered to the support base of the robotic hand and also the surface of each object to track their positions in space.

4.6.4 Human-like dexterous grasping of a single object

We validated our proposed human-like dexterous grasping Algorithm 1 by generating grasps of a variety of objects using different opposition spaces in the hand model.

Simulation results

We first generated grasps using spherical and cylindrical objects of different sizes on both human right-hand model and Allegro left-hand models in simulation (Fig. 4.9).

For each hand model, we validated our algorithm by generating grasps using different pairs of opposing regions of each hand model. Such grasp poses include (1) the *pinch grasp*, using either the commonly seen thumb-index coupling ((a), (g), and (l)) or the combination of finger and palm ((b), (f), and (h)); (2) the *wrap grasp* that uses phalanges from one single finger ((e) and (k)); (3) the *addiction grasp* by the side faces of finger phalanges, regardless of whether fingers are adjacent ((c) and (i)) or not ((d) and (j)). Except the pinch grasp that employs the thumb and the index finger, most of these grasp poses are atypical for robotic grasping but are common in human grasping, and have been documented in human grasp taxonomy (Feix et al. (2015)). Moreover, it is worth noting that in the pinch grasp formed by the thumb and the index finger ((a) and (g)), the thumb distal phalanx established contact with the object surface on its back side, seems unnatural compared to a human pinch grasp. This is mainly because the distal phalanges of each finger in both hand models have been modeled as cylinder without discriminating the front side and back side faces.

Evaluation on robotic hand

Each grasp was first generated in simulation by applying Algorithm 1, given a desired OS. Once a successful grasp configuration has been determined by solving the optimization problem, the human experimenter held the target object and approached its desired pose, as calculated by the optimization. The actual center position and orientation of the help object was measured by the motion capture system at 10 Hz during this process. As soon as the actual center position of the target object is close to its calculated position (spatial Euclidean distance less than 10 mm), the robotic hand was commanded to move all related fingers simultaneously to their calculated joint angles to grasp the object. After the robotic hand has grasped the object, the human experimenter released the object and checked if the grasp was stable. If the object cannot be grasped by the desired links, or if the object has slipped out of the hand, the grasp

was considered unstable. Then, experimenter controlled the robotic hand to move back to its initial joint configuration (fully open), increased (or decreased) the desired positions of the extension/flexion joints (or ab-/adduction joints, depends on grasp type) that affecting the grasp by 5 degrees, such that the opposing hand links planned for grasp will move closer against each other, hence increasing the grasping force. The experimenter repeated the grasp process several times. If the object cannot be grasped after adjusting desired finger joints, the experimental trial is considered a failure. In a stable grasp configuration, the robotic joint angles and object spatial positions have been registered for evaluation afterwards.

We generated grasps of each selected object using five distinct grasp poses: (1) pinch grasp by a finger and the palm, tested on F3L4-PALM, (2) pinch grasp using the thumb and the index finger, tested on F1L4-F2L4, (3) wrap grasp, tested on F2L2-F2L4, (4) adduction grasp using non-adjacent fingers, tested on F2L3-F4L3, and (5) adduction grip using adjacent fingers, tested on F3L4-F4L4.

Fig. 4.10 lists all grasps performed by the real robotic hand. Notice that for $\mathcal{O}_9$ and $\mathcal{O}_{10}$, the wrap grasp cannot be achieved, since the selected OS F2L2-F2L4 is not geometrically permissive for such large objects. The OS F1L4-F4L4 has been employed instead, resulting in pinch grasp poses using the thumb distal and the ring distal phalanges.

Most objects can be successfully grasped by simply positioning at the calculated grasp configuration. A few grasps were unstable and slight adjustment has been made to re-grasp the object. To evaluate the deviation between simulation results and the grasp using real robotic hand, we quantitatively analyzed registered robotic joint angles and object positions.

We first calculated the deviation in joint angles across all single object grasp trials, as the absolute difference between simulated joint angle values and recorded robot joint angles while performing the same grasp. For most joints (see Fig. 4.11), the deviation value lies within the range of $[0, 5]$ degrees. The largest deviation of 24.56 degree was observed in the second joint (the MCP flexion/extension DOF) of the index finger, when the robotic hand grasped $\mathcal{O}_4$ by wrapping.

We also compared the deviation in object positions between simulation results and robotic experiments. Fig. 4.12 illustrated the absolute deviation in spatial distance, as well as deviation in each Cartesian spatial dimension. Across all experimental trials, objects in the robotic experiments differed from the simulated position by 0.0140 ± 0.0052m (mean $\pm$ std).

Our robotic grasping experimental results demonstrate that the robotic hand is able to grasp the target object using configurations generated by applying Algorithm 1. The deviations in joint angles and object positions are mainly due to the noise and also the inaccuracy in the modeling of the robotic hand. For example, the phalanges of the robotic hand have been approximated by cylinders for simplification, but the actual geometry of the Allegro hand's phalanges is cuboid.

(a) Tested grasps of wooden ball ($\mathcal{O}_1$)

(b) Tested grasps of ping-pong ball ($\mathcal{O}_2$)

(c) Tested grasps of soft ball ($\mathcal{O}_3$)

(d) Tested grasps of tennis ball ($\mathcal{O}_4$)

(e) Tested grasps of battery ($\mathcal{O}_5$)

(f) Tested grasps of wine stopper ($\mathcal{O}_6$)

(g) Tested grasps of pills bottle ($\mathcal{O}_7$)

(h) Tested grasps of glue stick ($\mathcal{O}_8$)

(i) Tested grasps of glass bottle ($\mathcal{O}_9$)

(j) Tested grasps of candy jar ($\mathcal{O}_{10}$)
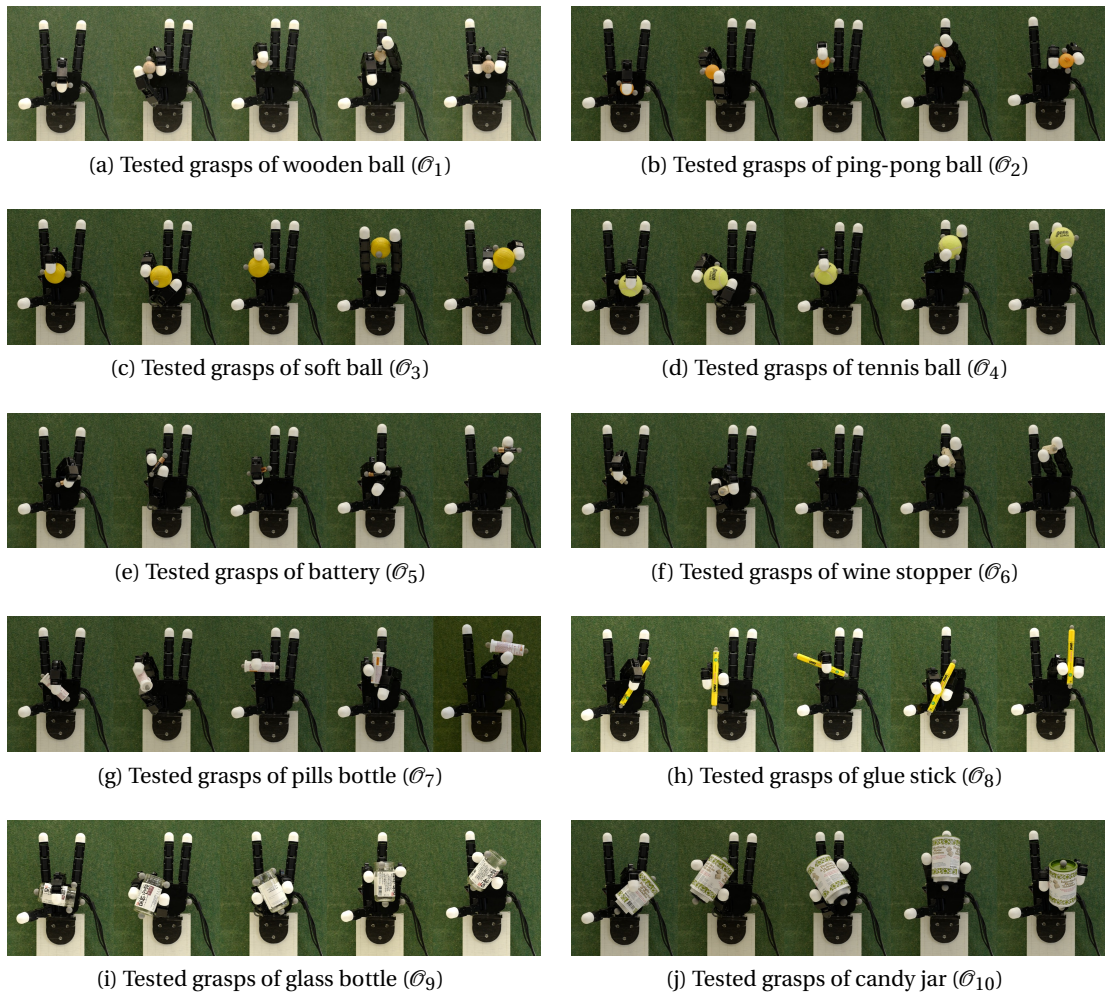
Figure 4.10 – The robotic hand performing human-like dexterous grasps of everyday objects using various opposition spaces spanned by arbitrary opposing surface regions.
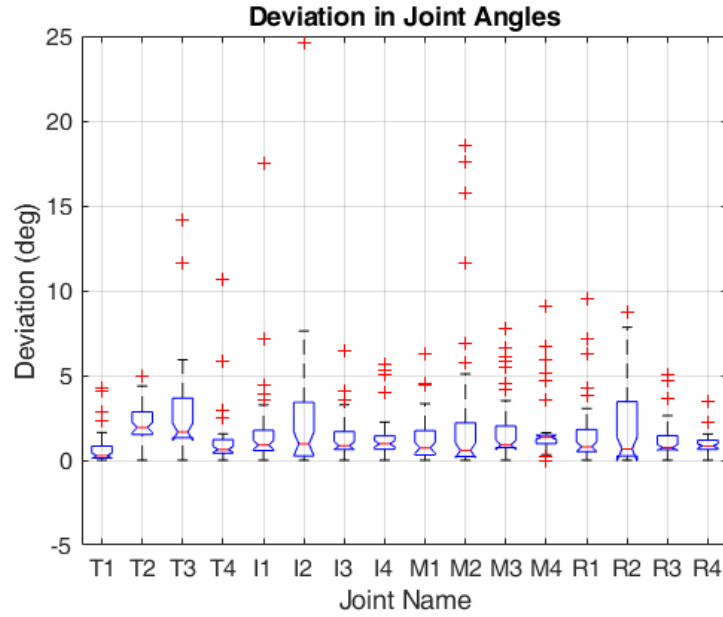
Figure 4.11 – Absolute deviation in robot joint angles (in degree) between simulation and experiment. For each joint name, the first digit indicates the finger name (T: thumb, I: index, M: middle, R: ring), and the second indicates the ordinal number of this joint in the finger. The first joint in each finger corresponds to the ab-/adduction DOF.
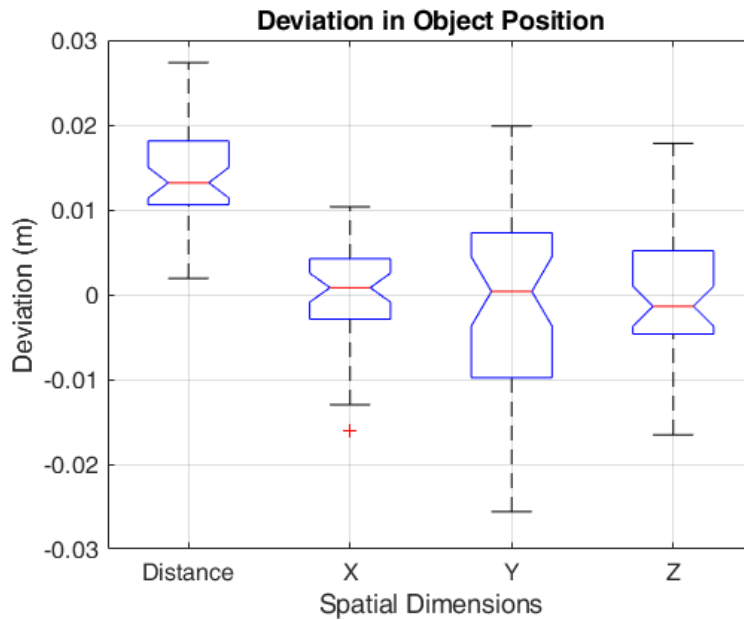
Figure 4.12 – Absolute deviation in grasped object center along each dimension. The first column (*Distance*) represents the deviation in Euclidean distance. The second, third, and forth columns correspond to the deviation in each spatial dimension, respectively.

4.6.5   Sequential grasping of multiple objects

We validated our sequential grasping method in the task of grasping three everyday objects.

We generated sequential grasps in four independent experimental trials. In each trial, three objects have been randomly selected from our object list (Fig. 4.7). For each object, three OS candidates have been tested, and the one that optimizes the objective function (see Eq. (4.21)) has been assigned the final grasp configuration.

Following our proposed Algorithm 2, we first generated the optimal sequential grasping configurations for target objects in ascending order of the object ordinal number. The optimal solution at each step minimizes the use of kinematic capacity of hand among all selected OS candidates for the object.

At each step, once an object has been successfully grasped, the joint configuration and object position have been recorded for reproducing robot grasps afterwards. We see in Fig. 4.13 (left panel) that successful grasping sequences can be generated by solving the problem in all trials.

Then, we reproduced the simulated sequential grasps on the real robotic hand, following the same procedure in Sec. 4.6.4. Experimental results (see Fig. 4.13, right panel) demonstrate that the robotic hand was able to perform successful sequential grasping of objects in all trials.

4.6.6   Greedy grasping of multiple objects

We demonstrated that the robotic hand was able to grasp three objects. Grasping three objects is already challenging. It requires 6 virtual fingers to span three opposition spaces for each object, individually.

In this scenario, we evaluate our proposed greedy grasping method (Algorithm 3) by further exploiting the kinematic redundancy in the hand model and try to grasp more objects. For a four-finger robotic hand used in this study, fingers must be shared in different grasps. It hence demands proper allocation of each single object grasp within the kinematic structure of the hand in this iterative process.

We selected four objects out of the list for evaluation: $\mathcal{O}_2$, $\mathcal{O}_3$, $\mathcal{O}_6$, and $\mathcal{O}_8$. The selection of objects took the geometric size of objects into consideration. For example, large objects ($\mathcal{O}_9$ and $\mathcal{O}_{10}$) may occupy too much space inside hand kinematic structure and thus lead to a lesser chance of successful result.

The permutation of the index set of selected objects $\mathcal{I} = [2, 3, 6, 8]$ results in 24 grasping orders. For each object, three OS candidates were provided for testing. This greedy grasping problem thus required solving a total amount of $24 \times 4 \times 3 = 288$ optimization problems.

Following our proposed algorithm, the robotic hand successfully grasped all four objects in two different poses. Fig. 4.14 shows the generated grasping sequences performed by the

(a) Simulated sequential grasps $\mathcal{O}_1 \to \mathcal{O}_2 \to \mathcal{O}_6$

(b) Robotic sequential grasps $\mathcal{O}_1 \to \mathcal{O}_2 \to \mathcal{O}_6$

(c) Simulated sequential grasps $\mathcal{O}_2 \to \mathcal{O}_5 \to \mathcal{O}_6$

(d) Robotic sequential grasps $\mathcal{O}_2 \to \mathcal{O}_5 \to \mathcal{O}_6$

(e) Simulated sequential grasps $\mathcal{O}_3 \to \mathcal{O}_6 \to \mathcal{O}_9$

(f) Robotic sequential grasps $\mathcal{O}_3 \to \mathcal{O}_6 \to \mathcal{O}_9$

(g) Simulated sequential grasps $\mathcal{O}_3 \to \mathcal{O}_8 \to \mathcal{O}_{10}$
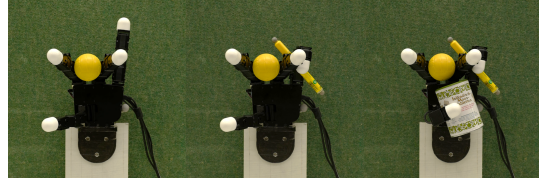
(h) Robotic sequential grasps $\mathcal{O}_3 \to \mathcal{O}_8 \to \mathcal{O}_{10}$

Figure 4.13 – Examples of sequential grasping of three randomly selected objects using robotic hand. Grasp configuration of each step was generated by following proposed Algorithm 2. Each row demonstrates one independent trial of sequential grasping: the left sub-figure shows the simulated grasp configuration, and the right sub-figure shows the actual grasp performed by the robotic hand.

(a) $\mathscr{O}_2 \rightarrow \mathscr{O}_3 \rightarrow \mathscr{O}_6 \rightarrow \mathscr{O}_8$

(b) $\mathscr{O}_2 \rightarrow \mathscr{O}_8 \rightarrow \mathscr{O}_3 \rightarrow \mathscr{O}_6$
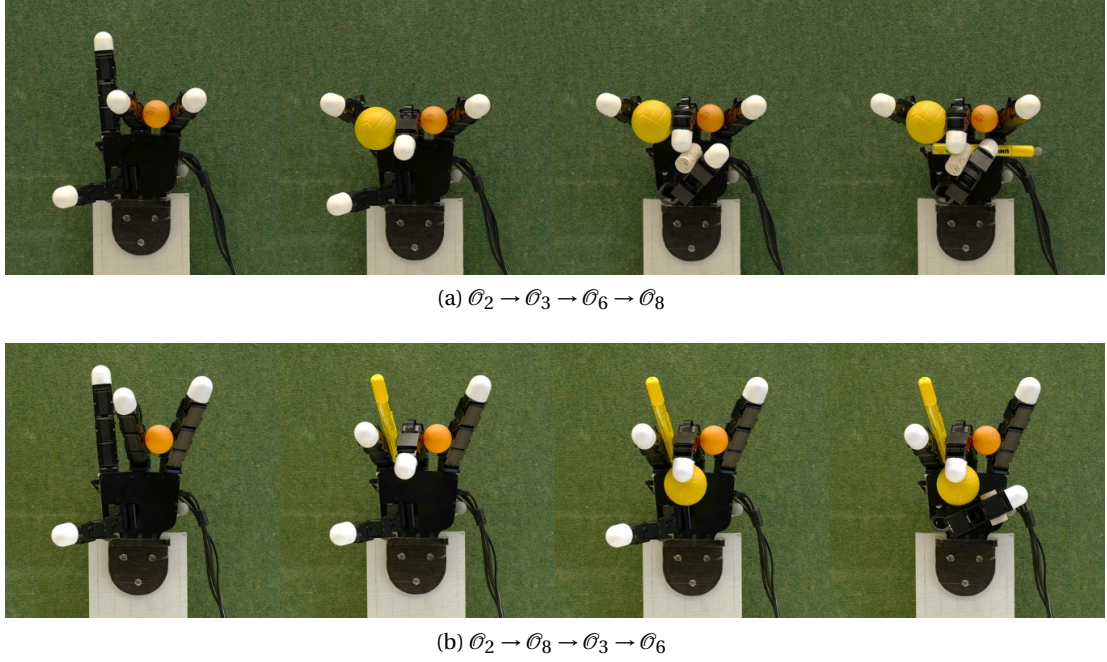
Figure 4.14 – Examples of greedy grasping of multiple objects using robotic hand by applying Algorithm 3. Grasps (a) and (b) followed different grasping orders, both led to successful solutions.

robotic hand. In Fig. 4.14a, the robotic hand followed the same object order as in Fig. 4.13b for the first three objects. When grasping the third object $\mathscr{O}_6$, the robotic hand employed the opposition space F1L3-F3L4, instead of F1L4-PALM that has been used in Fig. 4.13b. This choice spared the space of the palm and thus enabled one extra grasp of $\mathscr{O}_8$, using the space of F1L4-PALM. Grasp in Fig. 4.14b followed a different order, but also resulted in a successful grasping sequence. The most commonly violated constraints observed in failure trials are collision-free constraints and stability constraints.

A total number of 8 contact points (i.e. the minimum number of virtual fingers) are required to span four opposition spaces in total, hence reusing fingers is inevitable. In both successful trials, multiple fingers (including the palm) have participated in more than one grasp. In particular, the middle finger have been shared in three grasps in both cases. The thumb has been reused to grasp $\mathscr{O}_6$ and $\mathscr{O}_8$ in Fig. 4.14a, while the palm surface region has been shared in the grasp of $\mathscr{O}_3$ and $\mathscr{O}_6$ in Fig. 4.14b.

Experimental results demonstrate that our proposed Algorithm 3 enables the hand to grasp more objects than in Sec. 4.6.5 by maximally exploiting its kinematic redundancy.

4.7 Discussion

We have demonstrated that by exploiting the kinematic redundancy, a robotic hand with sufficient DOFs can perform human-like dexterous grasps, and even grasp multiple objects in sequence and hold multiple objects simultaneously. Such grasp types are no longer limited to the common use of only fingertips (pinch grasp) and inner surface of hand (power grasp). Instead, our human-like dexterous grasping synthesis algorithm enables the employment of arbitrary opposing surfaces of hand kinematic structure to achieve atypical dexterous grasp types, such as wrap grasp and adduction grasp, which are common for humans but rarely seen in robotic tasks. Our proposed approach largely increases the dexterity of the robotic hand and hence empowers the hand to grasp multiple objects. We validated our proposed approaches on both a 20-DOF human hand model and a 16-DOF Allegro hand model, and also tested on a real robotic hand. Experimental results validated the efficiency of our proposed algorithms in tasks of both dexterous grasping of a single object and grasping of multiple objects.

Our grasp synthesis algorithm has a hierarchical structure that combines *high-level task-space planning*, i.e. the analysis of kinematic capacity in Cartesian space and determination of opposition space, and *low-level joint-space planning*, i.e. the constrained optimization of search for a feasible solution. This hierarchical formulation has multiple advantages in comparison to the state-of-the-art approaches.

First, it alleviates the computational demands for planning in task space. Constructing the graspability map (Roa et al. (2011)) for a given object demands sampling hand spatial locations and orientations around the target object, with desired properties such as collision-free and stability being validated for every sample. The sampling of such numerous parameters inevitably entails a huge computational cost, and the quality of results largely relies on the number of samples. In contrast, our proposed approach divides the above-mentioned procedure in into task-space planning and joint-space planning. We formulate the problem in a hand-centered instead of object-centered point of view, and only use the task-space information (the reachability map) to roughly restrict the relative position and orientation of the target object with respect to the robotic hand. The optimization in joint space determines the specific grasp configuration and guarantees the constraints are satisfied. Our proposed approach also does not require voxelization of the spatial volume, and precise solution in continuous space can be obtained. Moreover, the properties of hand kinematic model obtained from task-space analysis can be reused in novel tasks or for different target objects.

Second, the commonly used human-like reach-and-grasp motion by simply closing robotic fingers to grasp an object can hardly achieve particular grasp types, such as finger wrap grasp or adduction grasp. Generating these atypical grasp types could be demanded in specific tasks. For example, (El-Khoury et al. (2013)) proposed solution to tackle this problem by generating a library of potential grasps based on sampling. This is, however, inefficient to satisfy a particular task demand, and also requires large computational cost. We provide an alternative solution by exploiting the existing opposing spaces in kinematic structure. The task-space analysis

in our approach efficiently determines the pair of links that can be employed to generate task-demanded grasp type. For example, selecting a pair of links from adjacent fingers will naturally result in an adduction grasp; and generating finger wrap grasp only needs to select an opposition space constructed by links from the same finger. Our human-like dexterous grasping synthesis enables the generation of all opposition grasps that have been discovered in human grasping hand poses (Iberall (1986)), and also avoids the computational costs caused by sampling.

Third, by combining the analysis in both task space and joint space, our proposed approach enables the reuse of analysis results obtained in previous tasks, hence reduces the computation demands. In particular, collisions among hand links and objects must be avoided to achieve a successful grasp. Such constraints are commonly handled in the literature by either checking for all feasible poses of the hand in task space, or searching in entire joint spaces for collision-free configurations. We formulate collision constraints by analyzing the hand's reachability map. Our constructed self-collision map reveals that potential collisions exist only among specific pairs of links. Therefore, hand links that do not collide in Cartesian space are excluded from grasp synthesis. Furthermore, we also reduce the number of involved DOFs by selecting the desired OS candidate prior to joint-space optimization. OSs that are not geometrically permissive for the target object along with their associated finger DOFs are excluded in the problem.

It is worth noting that the proposed geometrically permissive condition is a sufficient but not necessary condition for an OS to result in a feasible grasp configuration. This is because it does not guarantee that the multiple constraints of the problem can be satisfied, especially when the constraints are also related to the specific parameters of the problem, for example, the coefficient of friction. Nevertheless, analysis in task space reduces the dimensionality of optimization variables to a minimum and therefore the complexity of the low-level joint space programming.

Moreover, global optimality is hard to obtain in the planning of multiple objects grasping, mainly due to that multiple geometrically permissive OSs may exist in the high-level planning. The selection of OS affects not only the current grasp, but also subsequent grasps. The only way to guarantee global optimality is to iterate through all feasible OSs for each target object; however, this will result in a significant increase in the dimensionality of the problem, as the number of sub-problems is the cumulative multiplication of feasible OSs for each object. Therefore, we balance the optimality and problem complexity by prioritizing the most kinematically efficient OSs on the one hand, and testing a mini-batch OS candidates on the other hand.

We have validated our proposed approaches on both a human hand model and a robotic hand model, using objects approximated by spherical and cylindrical geometries. Nevertheless, our algorithms are applicable in general cases and the scope of application should not be limited by the models of robots or objects used. As long as kinematic redundancy exists in robot model

structure, our approach can be applied in the planning of grasping tasks. Irregular-shaped object geometries can be simply approximated by a composition of elementary geometries, such as multiple spheres (Hubbard (1996)). In addition, it is also feasible to construct an implicit model of object geometry by sampling data on object's surface offline (El-Khoury et al. (2013)). Implicit geometry models can be represented using learning approaches such as support vector regression model or Gaussian process regression model. This shall be considered in future work as extension to our current approaches.

4.8 Conclusions

This study aimed at improving the dexterity of robotic hands in grasping tasks by making use of all the degrees of freedom of the hand. To this end, we developed a framework that uses all contact areas on fingers and palm as support to simultaneously maintain grasps of multiple objects. Our approach extends state-of-the-art research in grasping in three aspects. First, to our knowledge, this is the first example of an approach to enable sequential grasping of multiple objects by a single hand, while maintaining already grasped objects in the hand. Second, the grasp is no longer restricted to the use of fingertips (pinch grasp) and the inner surfaces of the hand (power grasp). As long as a pair of opposing regions exist in the hand model and span an opposition space, it can be employed to synthesize a grasp. While we have not considered all surfaces, the approach is not limited to the surface considered here and could be extended to consider the upper part of the palm. Third, we proposed the kinematic efficiency metric as a quantitative measure of the kinematic redundancy in the robot model. On this basis, we proposed an associated strategy that facilitates the maximum exploitation of kinematic redundancy in the planning of multiple sequential grasping tasks. Experimental results demonstrated that the robotic hand was able to successfully perform dexterous grasps of a variety of everyday objects and also grasps of multiple objects with only slight deviations from the simulated grasping configuration. Algorithms proposed in this study are not restricted to the hand models or object models used, and can be easily applied to any robotic hands or manipulators having multiple DOFs.

Our approach of exploiting kinematic redundancy has largely improved the dexterity of multi-DOF robotic hands in grasping tasks. It offers the prospect of designing algorithms for robotic dexterous manipulation, and may also inspire the design of novel robotic hands and manipulators.

5 Conclusions

In this last chapter, we review the research work done throughout the thesis, and summarize our main contributions along with limitations. Moreover, we point out potential research directions for future work.

5.1 Contributions

This thesis centers on *dexterity* – both deepening our understanding of human dexterity and improving robot dexterity. We first examined human dexterity in fine-manipulation skills by focusing on the coordination mechanism, which is the core feature of human dexterity. We investigated both the intrinsic and extrinsic coordination in upper limbs and the role distribution across both hands. Then, we took inspiration from the coordination principles of human fingers and developed algorithms that enable robotic hands to achieve human-like dexterity in the grasping of single and multiple objects.

In Chapter 2, we started with the analysis of inter-limb kinematic coordination in upper limbs. The most significant difference between groups revealed in the study lies in the *extrinsic coordination*. Experts manifested a propensity of actively controlling the redundant DOFs in task conditions. This enables them to reduce the energy demand on arms when generating task-required forces, and hence to maintain the upper limbs in a more comfortable posture. In contrast, most novice subjects had to adopt uncomfortable arm postures by rotating their wrists close to joint limits to apply sufficient force. A study of *intrinsic coordination* revealed that hand dominance strongly affects role distribution across both upper limbs.

To the best of our knowledge, this thesis provides the first viewpoint of the acquisition of coordination from the perspective of optimal control theory. This approach in conjunction with a detailed analysis of the different stages of skill acquisition allowed us to observe a modification of the relative importance given by novices as compared to skilled practitioners to different factors inherent to the task, such as comfort and task performance. On the basis of our computational results, we hypothesized that novices tend to privilege satisfying task

demands in the early stage of learning, and make significant efforts at coordinating intrinsic DOFs to this end. As they gain more experience, they improve extrinsic coordination by actively adjusting limb movements to task conditions. This helps to maintain the limbs in more comfortable postures, an advantage when performing tasks that require to stay in the same posture for a long period of time, as is the case in watchmaking. This also alleviates the demand on intrinsic coordination.

In Chapter 3, we shift our attention from upper limbs to hands and fingers. We investigated the role distribution across hands under different task conditions in the task of disassembling a watch screw. Experimental observations indicate that hand dominance strongly influences the role distribution across hands, if more task DOFs need to be controlled (the free-base condition). As the number of DOFs to be controlled in the task decreases, the non-dominant hand contributes more actively to the task by sharing control with the dominant hand (the fixed-base condition). Importantly, our analysis reveals a clear preference for bimanual manipulation strategy over unimanual strategy. We interpret this preference as an effort to both reduce variability caused by biomechanical couplings and to mitigate intrinsic sensorimotor processing burdens, due to the numerous DOFs that need to be controlled in both hands.

In Chapter 4, we took inspiration from the human control principle of finger placements, and proposed algorithms for a multi-fingered robotic hand to achieve human-like dexterous grasping of single and multiple objects. We first proposed a human-like dexterous grasp synthesis algorithm to generate pairwise contacts on arbitrary opposing hand surface regions. Such grasps are no longer limited to the use of fingertips or the inner surface of the hand. Then, we formulated an iterative process that enables the robotic hand to sequentially planning grasps of multiple objects. In addition, we proposed a kinematic efficiency metric and an associated strategy to facilitate the exploitation of kinematic redundancy.

5.2 Implications of human movement studies for robotics

This thesis investigates the principles of human motion in bimanual fine-manipulation tasks, aiming to provide inspiration for designing robot control and learning algorithms. It is difficult to directly apply the discovered human motion principles to robotic systems, mainly due to the differences between the robotic and human motion systems in terms of structure and control mechanisms. Nevertheless, our findings and conclusions from human studies provide inspirations for us to improve the performance of robots.

In Chapter 2, we studied the upper-limb posture in a bimanual fine-manipulation task, and revealed that skillful subjects reduced task difficulty by intentionally modifying the redundant variables in task conditions before performing the task. This conclusion inspires planning and control in robotic tasks, i.e., the introduction of a preparation step may help to reduce task difficulty and facilitate task execution. The robotic system can actively modify and control redundant variables in task conditions during this preparation step.

In grasping tasks, existing studies imitated the human strategy (Chang et al. (2008)) by pre-grasping the target object to enable using a better posture in the subsequent real grasp (Kappler et al. (2012)). Our conclusion is consistent with this particular human strategy and provides inspiration for general robotic tasks. For instance, it may be necessary to introduce a preparatory step in a bimanual lifting task, if the initial pose of the target object does not afford direct lifting or intended subsequent manipulation.

Another typical scenario is collision avoidance. In most robotic tasks that require exploring or interacting with the environment, robotic systems usually needs to generate collision-free trajectories to navigate through obstacles in space. However, in some complex environments, physically feasible trajectories may not exist. In such cases, instead of searching for a collision-free path, an intelligent robotic system may first clear the path by removing or relocating the obstacles before performing the task. For example, mobile robots can remind pedestrians to make more space for its movement by actively making physical contact with them (Shrestha et al. (2015)). Similarly, to pick up an object from a cluster of objects using proper hand poses, the robotic arm-hand system can modify the environment by first picking up and repositioning other objects that are essentially "redundant" in the task.

In Chapter 3, we applied the concept of virtual finger to investigate the principles of role distribution of hands and functional grouping of fingers. Our inspiration from this study is that the abundant degrees of freedom in a robotic system can be systematically grouped into functional units for controlling multiple independent task demands. This inspired our design of robotic grasping synthesis algorithms in Chapter 4.

5.3 Limitations

The work presented in this thesis has limitations mainly in the following aspects.

In Chapter 2, our study provides insight into inter-limb coordination patterns in asymmetric and nonrhythmic two-arm manipulation tasks, and we show that exploiting task redundancy (e.g., the redundant variables in task conditions) can benefit task performance. However, it is unknown how subjects acquire the strategies of actively controlling task redundancy. This question could not be answered in the current study and demands further investigation in future work.

Moreover, our study provides a novel perspective for understanding the human motor learning process. This computational framework can be extended to the analysis of a variety of motor tasks. Notably, this perspective is based on an optimal control framework, and thus as an inherent theoretical underpinning, it assumes that the generated motion is an optimal solution to the underlying optimality criterion. However, it is debatable whether the human motor system actually aims at finding such an optimal solution, or simply adopts a suboptimal solution and relies heavily on sensory feedback (visual or tactile) to achieve online adaptation.

The study conducted in Chapter 3 sheds some further light on the control principles of the human fingers and hands in a bimanual fine-manipulation task. It supports the view that humans prefer bimanual control to unimanual control. However, we cannot conclude on whether this choice is motivated by the economy of efforts, as it is not clear if bimanual control requires fewer computational resources than the highly individualized finger control seen in unimanual case. Taking advantage of functional magnetic resonance imaging (fMRI) techniques to compare the activity of the human brain when using one hand and both hands to perform the same task may provide additional clues to confirm or refute this hypothesis.

In Chapter 4, our proposed algorithms empower, for the first time, a multi-fingered robotic hand with the human-like dexterity to simultaneously grasp multiple objects. Our proposed framework demonstrates the feasibility of advancing dexterity by strategically exploiting kinematic redundancy, and provides a viable and promising solution for robotic hands to increase dexterity in more complex grasping and manipulation tasks. We used an analytical approach to determine the grasping configuration of the robotic hand. However, this algorithm is unlikely to mirror the real human control mechanisms in similar tasks. Furthermore, our imitation of such human dexterity is currently only at the behavioral level. To truly achieve the dexterity of the human hand requires a deeper understanding of the human brain's control mechanism and the structure of the human hand.

5.4   Future work

In this section, based on the contributions and the limitations of the work presented in this thesis, we suggest potential future research topics along with promising directions.

5.4.1   Unraveling the mechanism of human motor learning

Understanding the principles of human motor learning and control is a key research question to answer in human movement science. Although learning of skill can be assessed by analyzing the observed performance, inferring the underlying intrinsic mechanism that drives skill acquisition is difficult and remains open. This is mainly because the control and learning of human motor system is influenced by a variety of factors, such as prior knowledge (e.g., task-related experience), attention, memory, individual difference, and so on. This thesis provides an alternative approach to assess the learning process under the optimal control theoretical framework. Its advantage over other approaches is the ability to quantitatively assess the relative importance of multiple factors that govern the motion generation at a certain skill level.

However, as an inherent limitation of the theoretical framework, the current study is limited in the sense that the selection of candidate factors (i.e., elementary costs or cost components) is not exclusive. Therefore, analysis at multiple skill levels may not be the unique assessment of the learning process. On the one hand, the costs can be formulated at multiple levels —

low-level costs, such as joint velocity and acceleration that appear in the dynamic equation of motion (Mombaur et al. (2010); Berret et al. (2011)), or high-level costs such as kinematic configuration, task compatibility, and biological discomfort that are used in this thesis. On the other hand, it is difficult to determine a set of costs that are both comprehensive and mutually exclusive. A promising solution is to include the most candidates possible, then exclude insignificant and also numerically correlated ones. Alternatively, machine learning approaches can be used to approximate the structure of the optimal criterion, resulting in an *inverse reinforcement learning* problem. This releases the linear structure assumption of the optimal criterion; hence potentially enable a more accurate approximation of the true optimality.

Furthermore, it is also unclear how the learning process is initiated. In other words, what drives the change in the structure of the optimality criterion. These driving factors may be related to the reward mechanisms in the human brain. For example, it is suggested that animals' preference for one action over another is determined by both the reward expectations and the effort cost for pursuing the action (Hull (1943); Charnov (1976)). Although it is not clear whether this mechanism also applies to humans, it is possible that changes in the optimality criterion of human CNS are also driven by specific reward mechanisms in the brain. Such mechanisms may facilitate the human motor system to search for motions that are better adapted to reduce a certain cost or increase a certain reward. Using functional magnetic resonance imaging (fMRI) scans and electroencephalography (EEG) of the human brain, together with the analysis of human movements during the motor learning process may shed more light on this issue.

5.4.2 From static to dynamic: dexterous manipulation by exploiting the kinematic redundancy

We have demonstrated the feasibility of grasping multiple objects by exploiting the kinematic redundancy of the robotic hand. As a direct extension of the current work, our proposed algorithms and framework can be applied to robotic dexterous manipulation tasks. The kinematically redundant finger phalanges can be controlled to establish additional contacts on the surface of grasped objects, hence improving the grasping quality. Similarly, the available fingers can be employed to manipulate the grasped object, e.g., by applying forces to the object. A critical issue to be addressed is to obtain sufficient feedback information on the contact surface, such as grasping force or slip onset. This information can be used to ensure the quality of grasp or to guide the planning of subsequent finger movements. This can be achieved by covering hand surfaces using tactile sensors. In addition, sensory feedback also releases the heavy demands on the accuracy of the model and helps to achieve adaptive control of fingers during manipulation.

5.4.3   From local to global: achieving full-body dexterity

In Chapter 4, we proposed algorithms to enable robotic hands to achieve human-like dexterity in grasping tasks. Although the human hand is the most dexterous body part, virtually the entire human body demonstrates high dexterity, and many body parts exceed their *default* functionality, helping to multitask in everyday activities. For example, humans use only fingers when grasping a ping-pong ball, and naturally extend both hands and arms when it comes to catching a basketball. Moreover, when carrying a large box, it is likely that the entire upper body will coordinate to lift and support it. In the latter cases, the arms and the upper torso play the same role as the fingers in the first case. As another example, we often almost subconsciously use our elbows or shoulders to push and open the door when our both hands are occupied with carried objects. In such scenarios, the involved body parts can be analogized as *virtual fingers* that participate and contribute functionally to the overall task. This strategy of flexibly employing all parts of the body makes it easier for humans to adapt to surroundings, and more efficient in daily life. Our proposed multi-objects grasping approach can be generalized and applied to the task planning and control of robotic systems featuring multiple degrees of freedom, such as robotic arm, mobile robot, and especially to achieve full-body dexterity on humanoid robots.

5.5   Look ahead — towards achieving the true human-like dexterity

Understanding the origins of human dexterity and replicating such dexterity in robotic hands are important research topics in both human movement science and robotics. In this thesis, we investigated the motion principles exhibited by humans in fine-manipulation tasks, and bimanual tasks in particular, from the motor coordination perspective. We then reproduced human-like dexterous grasping abilities on a robotic hand. However, this *dexterity* is only at the behavior level and is not yet comparable to the true human dexterity.

The essential underpinning of the human hand dexterity lies in two aspects — the brain, and the biomechanical structure of the hand. The human brain plays a central role in enabling such dexterity, by controlling all available degrees of freedom to achieve flexible and adaptive motions. The human hand acts as both the actuator and the sensor, afforded by the musculoskeletal system and the skin, executing motor commands while gathering sensory feedback from contact surfaces.

Accordingly, the achievement of human-like dexterity is constrained by the progress of research related to these aspects. On the one hand, our current understanding of human brain is highly limited. Although a large area of the human cerebral cortex is dedicated to the activity of hands (e.g., the *cortical motor homunculus*), how the brain controls these degrees of freedom in the hand remains unclear. As a result, the existing related research work can only stay at the imitation of the human hand behavior, but cannot use the same algorithms as the human brain to control the hand. This is mainly due to the high complexity of the human brain as a system, which makes a complete reconstruction of the human brain a virtually impossible

task in the near future. From the control point of view, it is as if we are trying to model the output pattern of a complex black box system. In recent years, rapid advances in the field of deep learning have enabled robots to perform more complex tasks that are beyond the reach of traditional control and learning methods. For example, learning dexterous in-hand manipulation (OpenAI: Andrychowicz et al. (2020)). Therefore, approximating a *human brain* to control a robot using deep neural networks with a large amount of data from demonstrations seems to be a very promising direction to achieve true human dexterity. At the same time, neurological and physiological studies related to the human brain may provide guidance for designing the architecture of such control systems.

On the other hand, the capabilities of even the most advanced biomimetic robotic hands are still far from those afforded by the musculoskeletal structures of the human hand. The human hand has significant advantages over the most state-of-the-art robotic hands in four essential aspects: number of degrees of freedoms, sensing ability, soft skin, and structural strength.

First, the sufficient number of degrees of freedom (not only the DOFs in the fingers, but also the ones in the palm and the wrist) is a prerequisite for the human hand to achieve flexible motions and perform (multiple) complex tasks. If only the index finger and thumb could be controlled, the function of the hand would be largely limited (analogous to an industrial two-finger gripper). Although most of the degrees of freedom of the palm and wrist cannot be independently controlled, they are the crucial to make the human hand structure more flexible and adaptive. However, these degrees of freedom are usually neglected in most biomimetic robotic hands to reduce the complexity of mechanical structure.

Second, the hand acts not only as an actuator, but also as a sensing equipment. The almost-real-time rich sensory feedback is extremely important for the brain to respond quickly to contact conditions, such as to adjust the positioning of fingers or to regulate the grasping force. For instance, we immediately feel whether a surface is slippery (e.g., steel) or rough (e.g., wood) the moment we touch it. Similarly, we can instantly increase the grasping force once we sense the grasped object is about to slide out of our hand. This ability relies on the sensing feedback from almost the entire surface of the hand, rather than solely the fingertips.

Moreover, the soft skin can largely compensate the model inaccuracy and also increase the robustness against external perturbation on the contact surface. In recent years, more and more soft pneumatic robotic grippers are appearing on industrial assembly lines. However, soft skin is usually ignored in most anthropomorphic robotic hands.

Finally, the human skeletal muscle system can provide a much higher drive load than the motors used in almost all joint-actuated robotic hands. Using tendons to control finger joints is more similar to the control mechanism of the human hand (Kochan (2005); Xu et al. (2013)). The use of high-strength artificial muscles to control the tendons is promising to overcome the limitations of miniature motors that have been widely used to actuate robotic finger joints, thereby largely increasing the payload of the hand.

As more and more robots are involved in the activity of human society, enhancing the dexterity of robots will largely improve their ability to perform complex tasks and delicate operations in industrial production lines as well as in human daily life. Having a highly dexterous prosthesis will also greatly facilitate the convenience of life for amputees. Achieving human-like dexterity on robotic platforms will remain a key topic of research in the future. Importantly, the study of human dexterity and the study of robot dexterity are inextricably linked. Understanding human dexterity will provide direction and ideas for designing mechanical structures and control algorithms for robots, while advancement of robot dexterity will help us humans understand our own motor system from the bottom up.

A Modeling of contact on cylindrical geometry
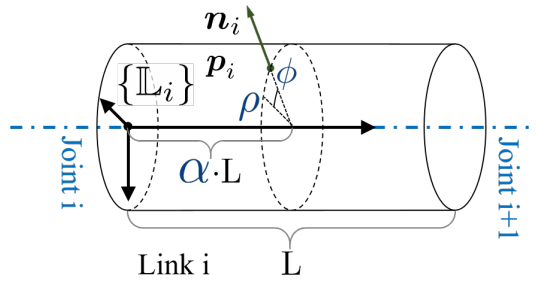
Figure A.1 – Modeling of contact point $\boldsymbol{p}_i$ on a cylindrical finger link geometry $\mathbb{L}_i$ of length L, and its cylindrical coordinates ρ, ϕ, and α.

A contact point on an arbitrary finger phalanx $\boldsymbol{p}_i$ is parameterized by its cylindrical coordinates as defined in Fig. A.1: (1) α, the height from the link base to the contact in the local link reference frame, (2) ϕ, the angular coordinate, and (3) ρ, the radical distance of the contact point to the central axis of the link.

Instead of using absolute values, we represent the height of $\boldsymbol{p}$ as a ratio $\alpha \in [0,1]$ of the total link length L, the height is thus calculated as αL. We enable the contact to happen at any point on the link surface, thus the angular coordinate $\phi \in (-\pi, \pi]$. Since $\boldsymbol{p}_i$ must locate on the surface of finger phalanx, ρ equals the radius of the link.

The Cartesian coordinates of $\boldsymbol{p}_i$ represented in $\{\mathbb{H}\}$ can be obtained by concatenating the local transformation (from the link base reference frame $\{\mathbb{L}_i\}$ to the contact point $\{\mathbb{C}_i\}$) to the finger's hand kinematic chain, and calculated by continually multiplying local homogeneous transformation matrices of each link.

Contact on hand palm can only happen on the inner surface of the palm, which is modeled as a rectangular-shaped surface region. It is parameterized by a vector $\overrightarrow{\boldsymbol{o}_\mathbb{H}\boldsymbol{p}_i} := (v_x, v_y)$ that is located at the palm center $o_\mathbb{H}$, i.e. the origin of $\{\mathbb{H}\}$, and pointing to the contact point $\boldsymbol{p}_i$. The value ranges of v_x and v_y restrict $\boldsymbol{p}_i$ such that it can only move within the palm's inner surface.

B Approximation of friction cone

The friction cone is approximated by a N_f-sided pyramid, defined as:

$$\mathscr{F} = \begin{bmatrix} & \mu\cos(\frac{2\pi i}{n}) & \\ \cdots & 1 & \cdots \\ & \mu\sin(\frac{2\pi i}{n}) & \end{bmatrix}_{3\times N_f} , \; n = 1, 2, \ldots, N_f. \tag{B.1}$$

Each column represents one edge of the approximated friction cone, denoted as $\mathbf{f}_n$, and is normalized such that $||\mathbf{f}_n|| = 1$. This assumes that all finger forces have the same limit. The position of entry 1 in $\mathbf{f}_n$ indicates that the central axis of the pyramid corresponds to the corresponding axis of the local contact reference frame $\{\mathbb{C}\}$.

Bibliography

Agur, A. M. and Dalley, A. F. (2009). *Grant's atlas of anatomy*. Lippincott Williams & Wilkins.

Albrecht, S., Leibold, M., and Ulbrich, M. (2012). A bilevel optimization approach to obtain optimal cost functions for human arm movements. *Numerical Algebra, Control and Optimization*, 2(1):105–127.

Aleotti, J. and Caselli, S. (2010). Interactive teaching of task-oriented robot grasps. *Robotics and Autonomous Systems*, 58(5):539–550.

Andersen, K. W. and Siebner, H. R. (2018). Mapping dexterity and handedness: recent insights and future challenges. *Current Opinion in Behavioral Sciences*, 20:123–129.

Ansuini, C., Giosa, L., Turella, L., Altoè, G., and Castiello, U. (2008). An object for an action, the same object for other actions: effects on hand shaping. *Experimental Brain Research*, 185(1):111–119.

Ansuini, C., Santello, M., Massaccesi, S., and Castiello, U. (2006). Effects of end-goal on hand shaping. *Journal of Neurophysiology*, 95(4):2456–2465.

Arbib, M. and Hoff, B. (1994). Trends in neural modeling for reach to grasp. In *Advances in Psychology*, volume 105, pages 311–344. Elsevier.

Arbib, M. A., Iberall, T., and Lyons, D. (1985). Coordinated control programs for movements of the hand. *Experimental Brain Research*, pages 111–129.

Armbrüster, C. and Spijkers, W. (2006). Movement planning in prehension: do intended actions influence the initial reach and grasp movement? *Motor Control*, 10(4):311–329.

Babik, I. and Michel, G. F. (2016). Development of role-differentiated bimanual manipulation in infancy: Part 1. the emergence of the skill. *Developmental Psychobiology*, 58(2):243–256.

Bagesteiro, L. B. and Sainburg, R. L. (2003). Nondominant arm advantages in load compensation during rapid elbow joint movements. *Journal of Neurophysiology*, 90(3):1503–1513.

Bajaj, N. M., Spiers, A. J., and Dollar, A. M. (2015). State of the art in prosthetic wrists: Commercial and research devices. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 331–338. IEEE.

Bibliography

Baldissera, F., Cavallari, P., and Civaschi, P. (1982). Preferential coupling between voluntary movements of ipsilateral limbs. *Neuroscience Letters*, 34(1):95–100.

Bard, J. F. (2013). *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media.

Baud-Bovy, G. and Soechting, J. F. (2001). Two virtual fingers in the control of the tripod grasp. *Journal of Neurophysiology*, 86(2):604–615.

Bernstein, N. (1967). *The co-ordination and regulation of movements*. Pergamon Press.

Bernstein, N. A., Latash, M. L., and Turvey, M. T. (1996). *Dexterity and Its Development*. Psychology Press.

Berret, B., Chiovetto, E., Nori, F., and Pozzo, T. (2011). Evidence for composite cost functions in arm movement planning: an inverse optimal control approach. *PLoS Computational Biology*, 7(10):e1002183.

Berret, B., Darlot, C., Jean, F., Pozzo, T., Papaxanthis, C., and Gauthier, J. P. (2008). The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. *PLoS Computational Biology*, 4(10):e1000194.

Bhullar, A., Kang, N., Idica, J., Christou, E. A., and Cauraugh, J. H. (2015). Increased visual information gain improves bimanual force coordination. *Neuroscience Letters*, 608:23–27.

Bicchi, A. (1995). On the closure properties of robotic grasping. *The International Journal of Robotics Research*, 14(4):319–334.

Bicchi, A. and Kumar, V. (2000). Robotic grasping and contact: A review. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 348–353. IEEE.

Biegstraaten, M., Smeets, J. B., and Brenner, E. (2006). The relation between force and movement when grasping an object with a precision grip. *Experimental Brain Research*, 171(3):347–357.

Billard, A. and Kragic, D. (2019). Trends and challenges in robot manipulation. *Science*, 364(6446).

Bingham, G. P., Hughes, K., and Mon-Williams, M. (2008). The coordination patterns observed when two hands reach-to-grasp separate objects. *Experimental Brain Research*, 184(3):283–293.

Blinch, J., Flindall, J. W., Smaga, Ł., Jung, K., and Gonzalez, C. L. (2019). The left cerebral hemisphere may be dominant for the control of bimanual symmetric reach-to-grasp movements. *Experimental Brain Research*, 237(12):3297–3311.

Bloomfield, A., Deng, Y., Wampler, J., Rondot, P., Harth, D., McManus, M., and Badler, N. (2003). A taxonomy and comparison of haptic actions for disassembly tasks. In *IEEE Virtual Reality Conference (IEEEVR)*, pages 225–231. IEEE.

Bohg, J., Morales, A., Asfour, T., and Kragic, D. (2013). Data-driven grasp synthesis—a survey. *IEEE Transactions on Robotics*, 30(2):289–309.

Borst, C., Fischer, M., and Hirzinger, G. (2003). Grasping the dice by dicing the grasp. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 4, pages 3692–3697. IEEE.

Borst, C., Fischer, M., and Hirzinger, G. (2004). Grasp planning: How to choose a suitable task wrench space. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 319–325. IEEE.

Bouzit, M. (1996). Design, implementation and testing of a data glove with force feedback for virtual and real objects telemanipulation. *PhD Thesis, Laboratoire de Robotique de Paris, University of Pierre Et Marie Curie*.

Bullock, I. M. and Dollar, A. M. (2011). Classifying human manipulation behavior. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 1–6. IEEE.

Bullock, I. M., Feix, T., and Dollar, A. M. (2013). Finding small, versatile sets of human grasps to span common objects. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1068–1075. IEEE.

Buonomano, D. V. and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annual Review of Neuroscience*, 21(1):149–186.

Butterworth, G. and Hopkins, B. (1988). Hand-mouth coordination in the new-born baby. *British Journal of Developmental Psychology*, 6(4):303–314.

Caldera, S., Rassau, A., and Chai, D. (2018). Review of deep learning methods in robotic grasp detection. *Multimodal Technologies and Interaction*, 2(3):57.

Cannavó, F. (2012). Sensitivity analysis for volcanic source modeling quality assessment and model selection. *Computers & Geosciences*, 44:52–59.

Caplan, B. and Mendoza, J. E. (2011). Edinburgh handedness inventory. *Encyclopedia of Clinical Neuropsychology*, 928.

Cattaert, D., Semjen, A., and Summers, J. (1999). Simulating a neural cross-talk model for between-hand interference during bimanual circle drawing. *Biological Cybernetics*, 81(4):343–358.

Chang, L. Y., Zeglin, G. J., and Pollard, N. S. (2008). Preparatory object rotation as a human-inspired grasping strategy. In *The 8th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 527–534. IEEE.

Bibliography

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9(2):129–136.

Chinn, L., Hoffmann, M., Leed, J., and Lockman, J. (2019). Reaching with one arm to the other: Coordinating touch, proprioception, and action during infancy. *Journal of Experimental Child Psychology*, 183:19–32.

Chiu, S. L. (1987). Control of redundant manipulators for task compatibility. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 4, pages 1718–1724. IEEE.

Chiu, S. L. (1988). Task compatibility of manipulator postures. *The International Journal of Robotics Research*, 7(5):13–21.

Ciocarlie, M. T. and Allen, P. K. (2009). Hand posture subspaces for dexterous robotic grasping. *The International Journal of Robotics Research*, 28(7):851–867.

Cohen, R. and Sternad, D. (2009). Variability in motor learning: relocating, channeling and reducing noise. *Experimental Brain Research*, 193(1):69–83.

Cohen, R. G. and Rosenbaum, D. A. (2004). Where grasps are made reveals how grasps are planned: generation and recall of motor plans. *Experimental Brain Research*, 157(4):486–495.

Colson, B., Marcotte, P., and Savard, G. (2007). An overview of bilevel optimization. *Annals of Operations Research*, 153(1):235–256.

Cover, T. M. (1999). *Elements of Information Theory*. John Wiley & Sons.

Crossman, E. R. (1959). A theory of the acquisition of speed-skill. *Ergonomics*, 2(2):153–166.

Cutkosky, M. and Wright, P. (1986). Modeling manufacturing grips and correlations with the design of robotic hands. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 1533–1539. IEEE.

Cutkosky, M. R. et al. (1989). On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation*, 5(3):269–279.

Daprati, E. and Sirigu, A. (2006). How we interact with objects: learning from brain lesions. *Trends in Cognitive Sciences*, 10(6):265–270.

d'Avella, A. and Bizzi, E. (2005). Shared and specific muscle synergies in natural motor behaviors. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8):3076–3081.

De Souza, R., El-Khoury, S., Santos-Victor, J., and Billard, A. (2015). Recognizing the grasp intention from human demonstration. *Robotics and Autonomous Systems*, 74:108–121.

Deng, Z., Jonetzko, Y., Zhang, L., and Zhang, J. (2020). Grasping force control of multi-fingered robotic hands through tactile sensing for object stabilization. *Sensors*, 20(4):1050.

Deutsch, D. (1983). The generation of two isochronous sequences in parallel. *Perception & Psychophysics*, 34(4):331–337.

Dollar, A. M. (2014). Classifying human hand use and the activities of daily living. In *The Human Hand as an Inspiration for Robot Hand Development*, pages 201–216. Springer.

Donald, B., Gariepy, L., and Rus, D. (2000). Distributed manipulation of multiple objects using ropes. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 450–457. IEEE.

Edelsbrunner, H. and Mücke, E. P. (1994). Three-dimensional alpha shapes. *ACM Transactions on Graphics (TOG)*, 13(1):43–72.

Eizicovits, D. and Berman, S. (2018). Automatic graspability map generation based on shape-primitives for unknown and familiar objects. *Robotics and Autonomous Systems*, 110:1–11.

Eizicovits, D., van Tuijl, B., Berman, S., and Edan, Y. (2016). Integration of perception capabilities in gripper design using graspability maps. *Biosystems Engineering*, 146:98–113.

Ekvall, S. and Kragic, D. (2005). Grasp recognition for programming by demonstration. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 748–753. IEEE.

El-Khoury, S., De Souza, R., and Billard, A. (2015). On computing task-oriented grasps. *Robotics and Autonomous Systems*, 66:145–158.

El-Khoury, S., Li, M., and Billard, A. (2013). On the generation of a variety of grasps. *Robotics and Autonomous Systems*, 61(12):1335–1349.

Endo, Y., Kanai, S., Kishinami, T., Miyata, N., Kouchi, M., and Mochimaru, M. (2007). A computer-aided ergonomic assessment and product design system using digital hands. In *International Conference on Digital Human Modeling*, pages 833–842. Springer.

Engelbrecht, S. E. (2001). Minimum principles in motor control. *Journal of Mathematical Psychology*, 45(3):497–542.

Feix, T., Bullock, I. M., and Dollar, A. M. (2014). Analysis of human grasping behavior: Object characteristics and grasp type. *IEEE Transactions on Haptics*, 7(3):311–323.

Feix, T., Romero, J., Schmiedmayer, H.-B., Dollar, A. M., and Kragic, D. (2015). The grasp taxonomy of human grasp types. *IEEE Transactions on Human-machine Systems*, 46(1):66–77.

Ferguson, G. A. (1956). On transfer and the abilities of man. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 10(3):121.

Bibliography

Ferrari, C. and Canny, J. F. (1992). Planning optimal grasps. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 2290–2295.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381.

Fitts, P. M. and Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology*, 67(2):103.

Flash, T. (1990). The organization of human arm trajectory control. In *Multiple Muscle Systems*, pages 282–301. Springer.

Flash, T. and Hochner, B. (2005). Motor primitives in vertebrates and invertebrates. *Current Opinion in Neurobiology*, 15(6):660–666.

Franz, E. A. (1997). Spatial coupling in the coordination of complex actions. *The Quarterly Journal of Experimental Psychology: Section A*, 50(3):684–704.

Franz, E. A. and Ramachandran, V. (1998). Bimanual coupling in amputees with phantom limbs. *Nature Neuroscience*, 1(6):443.

Friedman, J. and Flash, T. (2007). Task-dependent selection of grasp kinematics and stiffness in human object manipulation. *Cortex*, 43(3):444–460.

Furuya, S., Flanders, M., and Soechting, J. F. (2011). Hand kinematics of piano playing. *Journal of Neurophysiology*, 106(6):2849–2864.

Furuya, S., Nakamura, A., and Nagata, N. (2014). Acquisition of individuated finger movements through musical practice. *Neuroscience*, 275:444–454.

García-Rodríguez, R., Villalva-Lucio, M., and Parra-Vega, V. (2015). Dexterous dynamic optimal grasping of a circular object with pose regulation using redundant robotic soft-fingertips. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6231–6237. IEEE.

Gentilucci, M. (2002). Object motor representation and reaching–grasping control. *Neuropsychologia*, 40(8):1139–1153.

Georgopoulos, A. P., Kalaska, J. F., and Massey, J. T. (1981). Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology*, 46(4):725–743.

Gienger, M., Toussaint, M., and Goerick, C. (2008). Task maps in humanoid robot manipulation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2758–2764. IEEE.

Gilster, R., Hesse, C., and Deubel, H. (2012). Contact points during multidigit grasping of geometric objects. *Experimental Brain Research*, 217(1):137–151.

Goble, D. J., Lewis, C. A., and Brown, S. H. (2006). Upper limb asymmetries in the utilization of proprioceptive feedback. *Experimental Brain Research*, 168(1-2):307–311.

Goldfeder, C. and Allen, P. K. (2011). Data-driven grasping. *Autonomous Robots*, 31(1):1–20.

Gonzalez, F., Gosselin, F., and Bachta, W. (2014). Analysis of hand contact areas and interaction capabilities during manipulation and exploration. *IEEE Transactions on Haptics*, 7(4):415–429.

Gonzalez, S. L. and Nelson, E. L. (2015). Addressing the gap: a blueprint for studying bimanual hand preference in infants. *Frontiers in Psychology*, 6:560.

Grea, H., Desmurget, M., and Prablanc, C. (2000). Postural invariance in three-dimensional reaching and grasping movements. *Experimental Brain Research*, 134(2):155–162.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior*, 19(4):486–517.

Guiard, Y., Olafsdottir, H. B., and Perrault, S. T. (2011). Fitt's law as an explicit time/error trade-off. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1619–1628.

Gutmann, H.-M. (2001). A radial basis function method for global optimization. *Journal of Global Optimization*, 19(3):201–227.

Haken, H., Kelso, J. S., and Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51(5):347–356.

Harada, K. and Kaneko, M. (1998). Kinematics and internal force in grasping multiple objects. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 1, pages 298–303. IEEE.

Heinemann, F., Puhlmann, S., Eppner, C., Élvarez-Ruiz, J., Maertens, M., and Brock, O. (2015). A taxonomy of human grasping behavior suitable for transfer to robotic hands. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 4286–4291. IEEE.

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience*, 8:150.

Heuer, H., Kleinsorge, T., Spijkers, W., and Steglich, C. (2001). Static and phasic cross-talk effects in discrete bimanual reversal movements. *Journal of Motor Behavior*, 33(1):67–85.

Hoff, B. and Arbib, M. A. (1993). Models of trajectory formation and temporal interaction of reach and grasp. *Journal of Motor Behavior*, 25(3):175–192.

Howard, I. S., Ingram, J. N., Körding, K. P., and Wolpert, D. M. (2009). Statistics of natural movements are reflected in motor errors. *Journal of Neurophysiology*, 102(3):1902–1910.

Bibliography

Hu, X. and Newell, K. M. (2011). Modeling constraints to redundancy in bimanual force coordination. *Journal of Neurophysiology*, 105(5):2169–2180.

Hubbard, P. M. (1996). Approximating polyhedra with spheres for time-critical collision detection. *ACM Transactions on Graphics (TOG)*, 15(3):179–210.

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. Appleton-Century.

Iberall, T. (1986). Opposition space as a structuring concept for the analysis of skilled hand movements. *Generation and modulation of action patterns*, 15:158–173.

Iberall, T. (1987a). Grasp planning for human prehension. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence - Volume 2*, pages 1153–1156. Morgan Kaufmann Publishers Inc., San Francisco, CA, United States.

Iberall, T. (1987b). The nature of human prehension: Three dextrous hands in one. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 4, pages 396–401. IEEE.

Iberall, T. (1997). Human prehension and dexterous robot hands. *The International Journal of Robotics Research*, 16(3):285–299.

Jackson, G., Jackson, S., Newport, R., and Harvey, M. (2002). Co-ordination of bimanual movements in a centrally deafferented patient executing open loop reach-to-grasp movements. *Acta Psychologica*, 110(2-3):231–246.

Jacobsen, S., Iversen, E., Knutti, D., Johnson, R., and Biggers, K. (1986). Design of the utah/mit dextrous hand. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 1520–1532. IEEE.

Jäncke, L. (2009). The plastic human brain. *Restorative Neurology and Neuroscience*, 27(5):521–538.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, 18(7):314–320.

Jeannerod, M., Decety, J., and Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia*, 32(4):369–380.

Jerde, T. E., Santello, M., Flanders, M., and Soechting, J. F. (2012). Hand movements and musical performance. In *Music, Motor Control and the Brain*. Oxford University Press.

Jirsa, V. K. and Kelso, S. (2004). *Coordination Dynamics: Issues and Trends*. Springer Science & Business Media.

Johansson, R. S., Theorin, A., Westling, G., Andersson, M., Ohki, Y., and Nyberg, L. (2006). How a lateralized brain supports symmetrical bimanual tasks. *PLoS Biology*, 4(6):e158.

Johansson, R. S., Westling, G., Bäckström, A., and Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, 21(17):6917–6932.

Johnson-Frey, S., McCarty, M., and Keen, R. (2004). Reaching beyond spatial perception: Effects of intended future actions on visually guided prehension. *Visual Cognition*, 11(2-3):371–399.

Kagerer, F. (2016). Nondominant-to-dominant hand interference in bimanual movements is facilitated by gradual visuomotor perturbation. *Neuroscience*, 318:94–103.

Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F., and Miura, Y. (1980). Patterns of static prehension in normal hands. *The American Journal of Occupational Therapy*, 34(7):437–445.

Kang, N. and Cauraugh, J. H. (2018). Coherence and interlimb force control: effects of visual gain. *Neuroscience Letters*, 668:86–91.

Kappler, D., Chang, L. Y., Pollard, N. S., Asfour, T., and Dillmann, R. (2012). Templates for pre-grasp sliding interactions. *Robotics and Autonomous Systems*, 60(3):411–423.

Kelso, J., Southard, D. L., and Goodman, D. (1979a). On the nature of human interlimb coordination. *Science*, 203(4384):1029–1031.

Kelso, J. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(6):R1000–R1004.

Kelso, J. S. (1994). Elementary coordination dynamics. In *Interlimb Coordination*, pages 301–318. Elsevier.

Kelso, J. S. (1997). *Dynamic patterns: The self-organization of brain and behavior*. MIT press.

Kelso, J. S. and Clark, J. E. (1982). *The development of movement control and co-ordination*. New York: Wiley.

Kelso, J. S., Southard, D. L., and Goodman, D. (1979b). On the coordination of two-handed movements. *Journal of Experimental Psychology: Human Perception and Performance*, 5(2):229.

Kim, J.-O. and Khosla, K. (1991). Dexterity measures for design and control of manipulators. In *IEEE/RSJ International Workshop on Intelligent Robots and Systems*, pages 758–763. IEEE.

Kimmerle, M., Ferre, C. L., Kotwica, K. A., and Michel, G. F. (2010). Development of role-differentiated bimanual manipulation during the infant's first year. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 52(2):168–180.

Kimmerle, M., Mick, L. A., and Michel, G. F. (1995). Bimanual role-differentiated toy play during infancy. *Infant Behavior and Development*, 18(3):299–307.

Bibliography

Klapp, S. T. (1979). Doing two things at once: The role of temporal compatibility. *Memory & Cognition*, 7(5):375–381.

Klatzky, R. L., Fikes, T. G., and Pellegrino, J. W. (1995). Planning for hand shape and arm transport when reaching for objects. *Acta Psychologica*, 88(3):209–232.

Klein, L. K., Maiello, G., Paulun, V. C., and Fleming, R. W. (2020). Predicting precision grip grasp locations on three-dimensional objects. *PLoS Computational Biology*, 16(8):e1008081.

Kochan, A. (2005). Shadow delivers first hand. *Industrial Robot.*

Konczak, J., Borutta, M., Topka, H., and Dichgans, J. (1995). The development of goal-directed reaching in infants: hand trajectory formation and joint torque control. *Experimental Brain Research*, 106(1):156–168.

Kraus, P. R., Kumar, V., and Dupont, P. (1998). Analysis of frictional contact models for dynamic simulation. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 976–981. IEEE.

Lamb, P. F. and Bartlett, R. M. (2017). Assessing movement coordination. In *Biomechanical Evaluation of Movement in Sport and Exercise*, pages 22–43. Routledge.

Lang, C. E. and Schieber, M. H. (2004). Human finger independence: limitations due to passive mechanical coupling versus active neuromuscular control. *Journal of Neurophysiology*, 92(5):2802–2810.

Latash, M. L. (2012). The bliss (not the problem) of motor abundance (not redundancy). *Experimental Brain Research*, 217(1):1–5.

Lee, K.-S., Mo, S.-M., Hwang, J.-J., Wang, H., and Jung, M.-C. (2008). Relaxed hand postures. In *Proceedings of the Annual Meeting of Japan Ergonomics Society Japan Ergonomics Society 49th Conference*, pages 436–439. Japan Ergonomics Society.

León, B., Morales, A., and Sancho-Bru, J. (2014). *From robot to human grasping simulation.* Springer.

León, B., Sancho-Bru, J. L., Jarque-Bou, N. J., Morales, A., and Roa, M. A. (2012). Evaluation of human prehension using grasp quality measures. *International Journal of Advanced Robotic Systems*, 9(4):112.

Lepetit, V., Moreno-Noguer, F., and Fua, P. (2009). Epnp: An accurate o (n) solution to the pnp problem. *International Journal of Computer Vision*, 81(2):155.

Li, Z. and Sastry, S. S. (1988). Task-oriented optimal grasping by multifingered robot hands. *IEEE Journal on Robotics and Automation*, 4(1):32–44.

Li, Z.-M., Latash, M., and Zatsiorsky, V. (1998). Force sharing among fingers as a model of the redundancy problem. *Experimental Brain Research*, 119(3):276–286.

Lide, D. R. (2004). *CRC handbook of chemistry and physics*, volume 85. CRC press.

Liu, H., Wu, K., Meusel, P., Seitz, N., Hirzinger, G., Jin, M., Liu, Y., Fan, S., Lan, T., and Chen, Z. (2008). Multisensory five-finger dexterous hand: The dlr/hit hand ii. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3692–3697. IEEE.

Magill, R. and Anderson, D. (2010). *Motor Learning and Control.* McGraw-Hill Publishing New York.

Mainprice, J., Hayne, R., and Berenson, D. (2015). Predicting human reaching motion in collaborative tasks using inverse optimal control and iterative re-planning. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 885–892. IEEE.

Malvezzi, M., Gioioso, G., Salvietti, G., and Prattichizzo, D. (2015). Syngrasp: A matlab toolbox for underactuated and compliant hands. *IEEE Robotics & Automation Magazine*, 22(4):52–68.

Malvezzi, M., Gioioso, G., Salvietti, G., Prattichizzo, D., and Bicchi, A. (2013). Syngrasp: A matlab toolbox for grasp analysis of human and robotic hands. In *IEEE International Conference on Robotics and Automation*, pages 1088–1093. IEEE.

Marteniuk, R., MacKenzie, C., and Baba, D. (1984). Bimanual movement control: Information processing and interaction effects. *The Quarterly Journal of Experimental Psychology Section A*, 36(2):335–365.

Mason, C. R., Gomez, J. E., and Ebner, T. J. (2001). Hand synergies during reach-to-grasp. *Journal of Neurophysiology*, 86(6):2896–2910.

Miller, A. T. and Allen, P. K. (1999). Examples of 3d grasp quality computations. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1240–1246. IEEE.

Mironov, D., Altamirano, M., Zabihifar, H., Liviniuk, A., Liviniuk, V., and Tsetserukou, D. (2018). Haptics of screwing and unscrewing for its application in smart factories for disassembly. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 428–439. Springer.

Mishra, B. (1995). Grasp metrics: Optimality and complexity. In *Algorithmic Foundations of Robotics*, pages 137–166. AK Peters.

Mombaur, K., Truong, A., and Laumond, J.-P. (2010). From human to humanoid locomotion—an inverse optimal control approach. *Autonomous Robots*, 28(3):369–383.

Montana, D. J. (1991). The condition for contact grasp stability. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 412–417.

Morasso, P. (1983). Three dimensional arm trajectories. *Biological Cybernetics*, 48(3):187–194.

Bibliography

Moshtagh, N. et al. (2005). Minimum volume enclosing ellipsoid. *Convex Optimization*, 111(January):1–9.

Mucchiani, C. and Yim, M. (2020). A novel underactuated end-effector for planar sequential grasping of multiple objects. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 8935–8941. IEEE.

Murali, A., Li, Y., Gandhi, D., and Gupta, A. (2018). Learning to grasp without seeing. In *International Symposium on Experimental Robotics*, pages 375–386. Springer.

Murray, R. M., Li, Z., and Sastry, S. S. (2017). *A mathematical introduction to robotic manipulation*. CRC press.

Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics*, 46(2):135–147.

Nguyen, V.-D. (1988). Constructing force-closure grasps. *The International Journal of Robotics Research*, 7(3):3–16.

Oguz, O. S., Zhou, Z., Glasauer, S., and Wollherr, D. (2018). An inverse optimal control approach to explain human arm reaching control based on multiple internal models. *Scientific Reports*, 8(1):5583.

OpenAI: Andrychowicz, M., Baker, B., Chociej, M., Jozefowicz, R., McGrew, B., Pachocki, J., Petron, A., Plappert, M., Powell, G., Ray, A., Schneider, J., Sidor, S., Tobin, J., Welinder, P., Weng, L., and Zaremba, W. (2020). Learning dexterous in-hand manipulation. *The International Journal of Robotics Research*, 39(1):3–20.

Ozawa, R. and Tahara, K. (2017). Grasp and dexterous manipulation of multi-fingered robotic hands: a review from a control view point. *Advanced Robotics*, 31(19-20):1030–1050.

Panagiotopoulos, P. and Al-Fahed, A. (1994). Robot hand grasping and related problems: Optimal control and identification. *The International Journal of Robotics Research*, 13(2):127–136.

Peer, A., Einenkel, S., and Buss, M. (2008). Multi-fingered telemanipulation-mapping of a human hand to a three finger gripper. In *The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 465–470. IEEE.

Peters, M. (1981). Attentional asymmetries during concurrent bimanual performance. *The Quarterly Journal of Experimental Psychology Section A*, 33(1):95–103.

Plamondon, R. and Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences*, 20(2):279–303.

Ponce, J. and Faverjon, B. (1995). On computing three-finger force-closure grasps of polygonal objects. *IEEE Transactions on Robotics and Automation*, 11(6):868–881.

Prattichizzo, D., Malvezzi, M., and Bicchi, A. (2010). On motion and force control of grasping hands with postural synergies. *Robotics: Science and Systems VI*.

Prattichizzo, D., Malvezzi, M., and Bicchi, A. (2011). On motion and force control of grasping hands with postural synergies. In Matsuoka, Y., Durrant-Whyte, H., and J, N., editors, *Robotics: Science and Systems VI*, pages 49–56. MIT Press.

Prattichizzo, D. and Trinkle, J. C. (2016). Grasping. In *Springer Handbook of Robotics*, pages 955–988. Springer.

Preedy, V. R. (2012). *Handbook of anthropometry: physical measures of human form in health and disease*. Springer Science & Business Media.

Prevete, R., Donnarumma, F., d'Avella, A., and Pezzulo, G. (2018). Evidence for sparse synergies in grasping actions. *Scientific Reports*, 8(1):1–16.

Przybyla, A., Coelho, C., Akpinar, S., Kirazci, S., and Sainburg, R. L. (2013). Sensorimotor performance asymmetries predict hand selection. *Neuroscience*, 228:349–360.

Ritter, H. and Haschke, R. (2015). Hands, dexterity, and the brain. *Humanoid Robotics and Neuroscience: Science, Engineering and Society*.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. *Experimental Brain Research*, 71(3):491–507.

Roa, M. A., Hertkorn, K., Zacharias, F., Borst, C., and Hirzinger, G. (2011). Graspability map: A tool for evaluating grasp capabilities. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1768–1774. IEEE.

Roa, M. A. and Suárez, R. (2015). Grasp quality measures: review and performance. *Autonomous Robots*, 38(1):65–88.

Rosenbaum, D. A., Dawson, A. M., and Challis, J. H. (2006). Haptic tracking permits bimanual independence. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5):1266.

Rosenbaum, D. A., Marchk, F., Barnes, H. J., Vaughan, J., Slotta, J. D., and Jorgensen, M. J. (1990). Constraints for action selection: Overhand versus underhand grips. In *Attention and Performance XIII*, pages 321–342. Psychology Press.

Rosenbaum, D. A., Meulenbroek, R. G., and Vaughan, J. (2001). Planning reaching and grasping movements: theoretical premises and practical implications. *Motor Control*, 5(2):99–115.

Ruehl, S. W., Hermann, A., Xue, Z., Kerscher, T., and Dillmann, R. (2011). Graspability: A description of work surfaces for planning of robot manipulation sequences. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 496–502. IEEE.

Bibliography

Sahbani, A., El-Khoury, S., and Bidaud, P. (2012). An overview of 3d object grasp synthesis algorithms. *Robotics and Autonomous Systems*, 60(3):326–336.

Sainburg, R. L. (2002). Evidence for a dynamic-dominance hypothesis of handedness. *Experimental Brain Research*, 142(2):241–258.

Sainburg, R. L. (2020). Lateralization of goal-directed movement. In *Vision and goal-directed movement*. Human Kinetics, Inc.

Sainburg, R. L. and Kalakanis, D. (2000). Differences in control of limb dynamics during dominant and nondominant arm reaching. *Journal of Neurophysiology*, 83(5):2661–2675.

Santello, M., Flanders, M., and Soechting, J. F. (1998). Postural hand synergies for tool use. *Journal of Neuroscience*, 18(23):10105–10115.

Santello, M. and Soechting, J. F. (2000). Force synergies for multifingered grasping. *Experimental Brain Research*, 133(4):457–467.

Scheidt, R. A. and Ghez, C. (2007). Separate adaptive mechanisms for controlling trajectory and final position in reaching. *Journal of Neurophysiology*, 98(6):3600–3613.

Schieber, M. H. and Santello, M. (2004). Hand function: peripheral and central constraints on performance. *Journal of Applied Physiology*, 96(6):2293–2300.

Schneiberg, S., Sveistrup, H., McFadyen, B., McKinley, P., and Levin, M. F. (2002). The development of coordination for reach-to-grasp movements in children. *Experimental Brain Research*, 146(2):142–154.

Scholz, J. P. and Schöner, G. (1999). The uncontrolled manifold concept: identifying control variables for a functional task. *Experimental Brain Research*, 126(3):289–306.

Schöner, G. and Kelso, J. S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, 239(4847):1513–1520.

Schöner, G., Zanone, P. G., and Kelso, J. S. (1992). Learning as change of coordination dynamics: Theory and experiment. *Journal of Motor Behavior*, 24(1):29–48.

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532–545.

Seifert, L., Button, C., and Davids, K. (2013). Key properties of expert movement systems in sport. *Sports Medicine*, 43(3):167–178.

Seifert, L., Wattebled, L., L'Hermette, M., and Herault, R. (2011). Inter-limb coordination variability in ice climbers of different skill level. *Baltic Journal of Sport and Health Sciences*, 1(80).

Semjen, A., Summers, J. J., and Cattaert, D. (1995). Hand coordination in bimanual circle drawing. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5):1139.

148

Sgandurra, G., Cecchi, F., Serio, S. M., Del Maestro, M., Laschi, C., Dario, P., and Cioni, G. (2012). Longitudinal study of unimanual actions and grasping forces during infancy. *Infant Behavior and Development*, 35(2):205–214.

Shimoga, K. B. (1996). Robot grasp synthesis algorithms: A survey. *The International Journal of Robotics Research*, 15(3):230–266.

Shinar, D., Meir, M., and Ben-Shoham, I. (1998). How automatic is manual gear shifting? *Human Factors*, 40(4):647–654.

Shrestha, M. C., Nohisa, Y., Schmitz, A., Hayakawa, S., Uno, E., Yokoyama, Y., Yanagawa, H., Or, K., and Sugano, S. (2015). Using contact-based inducement for efficient navigation in a congested environment. In *The 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 456–461. IEEE.

Siciliano, B. (2012). *Advanced bimanual manipulation: Results from the dexmart project*, volume 80. Springer Science & Business Media.

Smeets, J. B., van der Kooij, K., and Brenner, E. (2019). A review of grasping as the movements of digits in space. *Journal of Neurophysiology*, 122(4):1578–1597.

Smethurst, C. J. and Carson, R. G. (2001). The acquisition of movement skills: Practice enhances the dynamic stability of bimanual coordination. *Human Movement Science*, 20(4-5):499–529.

Sommer, N. and Billard, A. (2016). Multi-contact haptic exploration and grasping with tactile sensors. *Robotics and Autonomous Systems*, 85:48–61.

Sparling, J. W., Van Tol, J., and Chescheir, N. C. (1999). Fetal and neonatal hand movement. *Physical Therapy*, 79(1):24–39.

Sparrow, W., Donovan, E., Van Emmerik, R., and Barry, E. (1987). Using relative motion plots to measure changes in intra-limb and inter-limb coordination. *Journal of Motor Behavior*, 19(1):115–129.

Spijkers, W. and Heuer, H. (1995). Structural constraints on the performance of symmetrical bimanual movements with different amplitudes. *The Quarterly Journal of Experimental Psychology*, 48(3):716–740.

Starke, J., Eichmann, C., Ottenhaus, S., and Asfour, T. (2020). Human-inspired representation of object-specific grasps for anthropomorphic hands. *International Journal of Humanoid Robotics*, 17(02):2050008.

Sternad, D. (2017). Human control of interactions with objects–variability, stability and predictability. In *Geometric and Numerical Foundations of Movements*, pages 301–335. Springer.

Bibliography

Sternad, D., Turvey, M. T., and Schmidt, R. (1992). Average phase difference theory and 1:1 phase entrainment in interlimb coordination. *Biological Cybernetics*, 67(3):223–231.

Summers, J. J., Rosenbaum, D. A., Burns, B. D., and Ford, S. K. (1993). Production of polyrhythms. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2):416.

Sundaralingam, B. and Hermans, T. (2018). Geometric in-hand regrasp planning: Alternating optimization of finger gaits and in-grasp manipulation. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 231–238. IEEE.

Swinnen, S. P. (2002). Intermanual coordination: from behavioural principles to neural-network interactions. *Nature Reviews Neuroscience*, 3(5):348.

Swinnen, S. P. and Gooijers, J. (2015). Bimanual coordination. *Brain Mapping: An Encyclopedic Reference*, 2:475–482.

Swinnen, S. P., Jardin, K., Meulenbroek, R., Dounskaia, N., and Den Brandt, M. H.-V. (1997). Egocentric and allocentric constraints in the expression of patterns of interlimb coordination. *Journal of Cognitive Neuroscience*, 9(3):348–377.

Swinnen, S. P., Jardin, K., Verschueren, S., Meulenbroek, R., Franz, L., Dounskaia, N., and Walter, C. B. (1998). Exploring interlimb constraints during bimanual graphic performance: effects of muscle grouping and direction. *Behavioural Brain Research*, 90(1):79–87.

Swinnen, S. P., Walter, C. B., Lee, T. D., and Serrien, D. J. (1993). Acquiring bimanual skills: contrasting forms of information feedback for interlimb decoupling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(6):1328.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*, volume 89. SIAM.

Thelen, E. (1985). Developmental origins of motor coordination: Leg movements in human infants. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 18(1):1–22.

Theorin, A. and Johansson, R. S. (2007). Zones of bimanual and unimanual preference within human primary sensorimotor cortex during object manipulation. *Neuroimage*, 36:T2–T15.

Theorin, A. and Johansson, R. S. (2010). Selection of prime actor in humans during bimanual object manipulation. *Journal of Neuroscience*, 30(31):10448–10459.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9):907.

Toga, A. W. and Thompson, P. M. (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience*, 4(1):37–48.

Tresch, M. C., Saltiel, P., and Bizzi, E. (1999). The construction of movement by the spinal cord. *Nature Neuroscience*, 2(2):162–167.

Tucker, M. and Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3):830.

Tucker, M. and Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica*, 116(2):185–203.

Turvey, M. T. (1990). Coordination. *American Psychologist*, 45(8):938.

Twitchell, T. E. (1970). Reflex mechanisms and the development of prehension. *Mechanisms of Motor Skill Development*, pages 25–38.

Vainio, L. and Tiainen, M. (2018). Inter-limb coupling of proximal and distal hand actions. *Experimental Brain Research*, 236(1):153–160.

Valyear, K. F., Chapman, C. S., Gallivan, J. P., Mark, R. S., and Culham, J. C. (2011). To use or to move: goal-set modulates priming when grasping real tools. *Experimental Brain Research*, 212(1):125–142.

Vaz, D. V., Kay, B. A., and Turvey, M. T. (2017). Effects of visual and auditory guidance on bimanual coordination complexity. *Human Movement Science*, 54:13–23.

Von Hofsten, C. (1982). Eye-hand coordination in the newborn. *Developmental Psychology*, 18(3):450.

Wang, C., Boyle, J. B., Dai, B., and Shea, C. H. (2017). Do accuracy requirements change bimanual and unimanual control processes similarly? *Experimental Brain Research*, 235(5):1467–1479.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

Wiesendanger, M. (1999). Manual dexterity and the making of tools–an introduction from an evolutionary perspective. *Experimental Brain Research*, 128(1-2):1–5.

Williams, A. M. and Hodges, N. J. (2004). *Skill acquisition in sport: Research, theory and practice*. Routledge.

Winges, S. A., Kornatz, K. W., and Santello, M. (2008). Common input to motor units of intrinsic and extrinsic hand muscles during two-digit object hold. *Journal of Neurophysiology*, 99(3):1119–1126.

Bibliography

Wu, G., Van der Helm, F. C., Veeger, H. D., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X., et al. (2005). Isb recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—part ii: shoulder, elbow, wrist and hand. *Journal of Biomechanics*, 38(5):981–992.

Xu, Z., Kumar, V., and Todorov, E. (2013). A low-cost and modular, 20-dof anthropomorphic robotic hand: design, actuation and modeling. In *The 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 368–375. IEEE.

Xu, Z. and Todorov, E. (2016). Design of a highly biomimetic anthropomorphic robotic hand towards artificial limb regeneration. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3485–3492. IEEE.

Yamada, T., Taki, T., Yamada, M., and Yamamoto, H. (2011). Grasp stability analysis of two objects by considering contact surface geometry in 3d. In *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1108–1115. IEEE.

Yamada, T. and Yamamoto, H. (2013). Grasp stability analysis of multiple objects including contact surface geometry in 3d. In *IEEE International Conference on Mechatronics and Automation (ICRA)*, pages 36–43. IEEE.

Yamada, T. and Yamamoto, H. (2015). Static grasp stability analysis of multiple spatial objects. *Journal of Control Science and Engineering*, 3:118–139.

Yamada, T., Yamanaka, S., Yamada, M., Funahashi, Y., and Yamamoto, H. (2009). Grasp stability analysis of multiple planar objects. In *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1032–1038. IEEE.

Yamanishi, J.-I., Kawato, M., and Suzuki, R. (1979). Studies on human finger tapping neural networks by phase transition curves. *Biological Cybernetics*, 33(4):199–208.

Yao, K. and Billard, A. (2020). An inverse optimization approach to understand human acquisition of kinematic coordination in bimanual fine manipulation tasks. *Biological Cybernetics*, 114(1):63–82.

Yao, K., Sternad, D., and Billard, A. (2021). Hand pose selection in a bimanual fine-manipulation task. *Journal of Neurophysiology*, 126(1):195–212.

Yoshikawa, T. (1984). Analysis and control of robot manipulators with redundancy. In *Robotics Research: The First International Symposium*, pages 735–747. MIT press Cambridge, MA.

Yu, Y. and Fukuda, K. (2013). Analysis of multifinger grasp internal forces for stably grasping multiple polyhedral objects. *International Journal of Mechatronics and Automation*, 3(3):203–216.

Zacharias, F., Borst, C., and Hirzinger, G. (2007). Capturing robot workspace structure: representing robot capabilities. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3229–3236. IEEE.

Zacharias, F., Borst, C., and Hirzinger, G. (2009). Object-specific grasp maps for use in planning manipulation actions. In *Advances in Robotics Research*, pages 203–213. Springer.

Zatsiorsky, V. M., Li, Z.-M., and Latash, M. L. (2000). Enslaving effects in multi-finger force production. *Experimental Brain Research*, 131(2):187–195.

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22.

Zhu, X. and Ding, H. (2004). Planning force-closure grasps on 3-d objects. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1258–1263. IEEE.

Zhuang, K. Z., Sommer, N., Mendez, V., Aryan, S., Formento, E., D'Anna, E., Artoni, F., Petrini, F., Granata, G., Cannaviello, G., Raffoul, W., Billard, A., and Micera, S. (2019). Shared human–robot proportional control of a dexterous myoelectric prosthesis. *Nature Machine Intelligence*, 1(9):400–411.

Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezzetta, E., Scabar, A., and Castiello, U. (2007). Evidence of early development of action planning in the human foetus: a kinematic study. *Experimental Brain Research*, 176(2):217–226.

Kunpeng Yao

CONTACT   EPFL STI IMT LASA, ME A3 394, Station 9   +41 21 693 59 98
CH-1015 Lausanne, Switzerland   kunpeng.yao@epfl.ch

RESEARCH
INTERESTS   Robotic Grasping and Manipulation, Human Motor Control and Learning, Machine Learning in Robotics, Tactile Sensing

EDUCATION   **Ph.D.**, *Robotics, Control, and Intelligent Systems*   Dec. 2017 - Dec. 2021

École Polytechnique Fédérale de Lausanne (EPFL), Learning Algorithms and Systems Laboratory (LASA), Lausanne, Switzerland

Title of thesis: *Enhancing our understanding of human fine manipulation skills and advancing robot dexterity in grasping.*

Thesis advisor: Prof. Aude Billard

M.Sc., *Electrical and Computer Engineering*   Oct. 2014 - Mar. 2017

Technische Universität München (TUM), Munich, Germany

Fakultät für Elektrotechnik und Informationstechnik (*Department of Electrical and Computer Engineering*

Graduation honor: Sehr gut bestanden (*Passed with Distinction*)

B.Sc., *Electrical and Information Engineering*   Oct. 2009 - Jul. 2013

Shanghai Jiao Tong University (SJTU), Shanghai, China

Department of Automation, School of Electrical and Information Engineering

ACADEMIC
EXPERIENCE   **Research Assistant**

Chair of Automatic Control Engineering (LSR), TUM   Jun. 2017 - Nov. 2017

Teaching Assistant

Machine Learning Programming (MICRO-401), EPFL 2019-2020 FS, 2021-2022 FS

Master Semester and Thesis Projects, EPFL   2019-2020 FS, 2021 SS

ACADEMIC
SERVICE   Reviewer for Nature Scientific Reports, IEEE International Conference on Robotics and Automation (ICRA), IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE Robotics and Automation Letters (RA-L).

SKILLS   **Programming** MATLAB, Python, C/C++

Software ROS, Gazebo, Git, LaTeX, Microsoft Office

Language English (proficient), German (intermediate), Chinese (native)

Journal Articles	**Kunpeng Yao**, Dagmar Sternad, and Aude Billard. "Hand Pose Selection in a Bimanual Fine-Manipulation Task." *Journal of Neurophysiology* 126, no. 1 (2021): 195-212
	Kunpeng Yao and Aude Billard. "An inverse optimization approach to understand human acquisition of kinematic coordination in bimanual fine manipulation tasks." *Biological Cybernetics* 114, no. 1 (2020): 63-82
	Konstantinos Chatzilygeroudis, Bernardo Fichera, Ilaria Lauzana, Fanjun Bu, **Kunpeng Yao**, Farshad Khadivar, and Aude Billard. "Benchmark for bimanual robotic manipulation of semi-deformable objects." *IEEE Robotics and Automation Letters* 5, no. 2 (2020): 2443-2450
	Mohsen Kaboli, **Kunpeng Yao**, Di Feng, and Gordon Cheng. "Tactile-based Active Object Recognition and Manifold Learning in an Unknown Workspace." *Autonomous Robots*, 2018/01
	Mohsen Kaboli, Di Feng, **Kunpeng Yao**, Pablo Lanillos, and Gordon Cheng. "A Tactile-based Framework for Active Object Learning and Discrimination using Multi-modal Robotic Skin." *IEEE Robotics and Automation Letters* 2.4 (2017): 2143-2150
Submitted Journal Articles	**Kunpeng Yao** and Aude Billard. "Exploiting Kinematic Redundancy for Robotic Grasping of Multiple Objects." Submitted to *The International Journal of Robotics Research*, 2021.
Conference Proceedings	**Kunpeng Yao** and Aude Billard. "Effect of task conditions on human hand pose selection strategies in bimanual fine manipulation tasks." *Neural Control of Movement (NCM) 2020 Annual Meeting.*
	Kunpeng Yao, Anaïs Haget, and Aude Billard. "Towards Understanding of Human Kinematic Coordination Patterns in Bimanual Fine Manipulation Tasks." *Progress in Motor Control XII: Movement Improvement, 2019.*
	Kunpeng Yao, Mohsen Kaboli, and Gordon Cheng. "Tactile-based Object Center of Mass Exploration and Discrimination." *2017 IEEE-RAS International Conference on Humanoid Robots (Humanoids).*
	Mohsen Kaboli, **Kunpeng Yao**, and Gordon Cheng. "Tactile-based Manipulation of Deformable Objects with Dynamic Center of Mass." *2016 IEEE-RAS International Conference on Humanoid Robots (Humanoids).*
Other Publications	**Kunpeng Yao** and Aude Billard. "Effect of task conditions on human hand control strategies in bimanual fine manipulation." *Workshop on Learning of Manual Skills in Humans and Robots. 2020 IEEE International Conference on Robotics and Automation (ICRA).*
	Kunpeng Yao, Bernardo Fichera, Anais Haget, Ilaria Lauzana, and Aude Billard. "Integrating Multi-sensory Information for Modeling Human Dexterous Bimanual Manipulation Skills." *Workshop: The Intelligence of Touch: Haptics, Tactile, Interaction. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).*