Energy-Optimal Control of Over-Actuated Systems – with Application to a Hybrid Feed Drive

by

Molong Duan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in the University of Michigan
2018

Doctoral Committee:

Associate Professor Chinedum E. Okwudire, Chair
Professor Ilya Kolmanovsky
Associate Professor Kenn Oldham
Professor A. Galip Ulsoy

Molong Duan

molong@umich.edu

ORCID iD: 0000-0001-8624-0133

To my parents, grandparents, and aunts who always support me.

Acknowledgements

It may be the most important decision in my life to start a Ph.D. journey. This journey brought me "back to curiosity", and helped me understand better who I am and what should I pursue in the rest of my life. There are so many people who had helped me, supported me, cared about me during this journey. I would like to express my sincere gratitude to all of them.

My Ph.D. advisor, Dr. Chinedum Okwudire, is definitely at the top of this list. In our collaboration, I learned so many things required for a good researcher, ranging from seeking the academic and practical significances in research activities, to organizing writings and presentations from the audiences' perspectives, and countless others. His encouragements, in both research and life, helped me maintain a positive attitude towards all possible obstacles throughout this journey. He also provided me chances to explore different aspects of academia, including teaching, proposal writing, advising students, etc. I would not be so sure about my academic pursuit without these opportunities and his encouragements.

I would also like to thank my doctoral committee: Dr. A. Galip Ulsoy, Dr. Ilya Kolmanovsky and Dr. Kenn Oldham for their valuable feedback and guidance. Dr. Ulsoy's feedback on control allocation and linear quadratic feedback is very inspirational and greatly helped formulate the research direction. Dr. Kolmanovsky provided very helpful insights in the multi-input, multi-output extension of the proxy-based control, and introduced many potential applications to my work. Dr. Oldham's feedback also helped me shape the research direction in control allocation instead of more detailed research on the hybrid feed drive. All these feedback and guidance are very important to me and are greatly appreciated.

My friends and lab mates also helped me a lot during this journey. Here I would thank Dr. Baoyang Jiang, Dr. Jihyun Lee, Dr. Lingfei Kong, Dr. Xun Liu, Dr. Yihao Zheng, Andrew Edoimioya, Bo Lin, Deokkyun Yoon, Huanyi Shui, Keval Ramani, Oluwami Dosunmu-Ogunbi, Weiyu Cao, Wubing Qin, Xin Dong, Zhengyu Gan, for their kind support in my research and life.

Table of Contents

List of Tables

List of Figures

Abstract

Over-actuated (or input-redundant) systems are characterized by the use of more actuators than the degrees of freedom to be controlled. They are widely used in modern mechanical systems to satisfy various control requirements, such as precision, motion range, fault tolerance, and energy efficiency. This thesis is particularly motivated by an over-actuated hybrid feed drive (HFD) which combines two complementary actuators with the aim to reduce energy consumption without sacrificing positioning accuracy in precision manufacturing.

This work addresses the control challenges in achieving energy optimality without sacrificing control performance in so-called weakly input-redundant systems, which characterize the HFD and most other over-actuated systems used in practice. Using calculus of variations, an optimal control ratio/subspace is derived to specify the optimal relationship among the redundant actuators irrespective of external disturbances, leading to a new technique termed optimal control subspace-based (OCS) control allocation. It is shown that the optimal control ratio/subspace is non-causal; accordingly, a causal approximation is proposed and employed in energy-efficient structured controller design for the HFD. Moreover, the concept of control proxy is proposed as an accurate causal measurement of the deviation from the optimal control ratio/subspace. The proxy enables control allocation for weakly redundant systems to be converted into regulation problems, which can be tackled using standard controller design methodologies. Compared to an existing allocation technique, proxy-based control allocation is shown to dynamically allocate control efforts optimally without sacrificing control performance.

The relationship between the proposed OCS control allocation and the traditional linear quadratic control approach is discussed for weakly input redundant systems. The two approaches are shown to be equivalent given perfect knowledge of disturbances; however, the OCS control allocation approach is shown to be more desirable for practical applications like the HFD, where disturbances are typically unknown. The OCS control allocation approach is validated in simulations and machining experiments on the HFD; significant reductions in control energy without sacrificing positioning accuracy are achieved.

Chapter 1 Introduction and Literature Review

1.1 Background of Over-Actuation

Actuators play a fundamental role in motion control systems, as they provide the necessary forces to regulate the motion. Conventionally, each degree of freedom (DOF) of motion is controlled via a single actuator, resulting in so-called fully-actuated systems [1]. However, fully-actuated systems may be limited in meeting the increasingly multifaceted and stringent control requirements (e.g. fault tolerance, energy efficiency, accuracy) placed on modern motion control systems. Under these circumstances, over-actuation (or input-redundancy) – the use of more actuators than the degrees of freedom to be controlled [2] – is increasingly being employed. To illustrate the wide usage of over-actuation, several representative case studies in aircraft, vehicles, and feed drives are discussed.

1.1.1 Representative Usage of Over-Actuation

1.1.1.1 Over-Actuation in Aircraft

In conventional winged aircraft, roll, pitch and yaw are the three major DOFs to be controlled. These three DOFs are maneuvered through the effectors (i.e., actuators). Among the effectors, two ailerons, two elevators and one rudder formulate the most elemental configuration [3]. They can change the flow of the air around the wings and stabilizer, controlling the aerodynamic forces applied to the aircraft. Note that this elemental configuration yields over-actuation, as more effectors are used compared to the motion DOFs to be controlled. Early aircraft mechanically coupled the motions of the effectors such that roll, pitch and yaw could be controlled separately; this original allocation strategy is referred to as mechanical ganging of effectors [3,4]. It was later discovered that the use of over-actuation could enhance the performance envelope of airplanes, especially during low-speed flight [4]. Over-actuation also allowed aircraft to recover from actuator failure, thus enhancing their fault tolerance [3,5,6]. For example, Figure 1.1 (a) illustrates the use of redundant effectors to compensate the yaw motion

due to a failed effector [7]. This capability of fault tolerance is especially important in the modern fly-by-wire aircraft, where mechanical connections between the pilot's control sticks/pedals to the effectors are replaced by electric signals and distributed actuators [3,7].

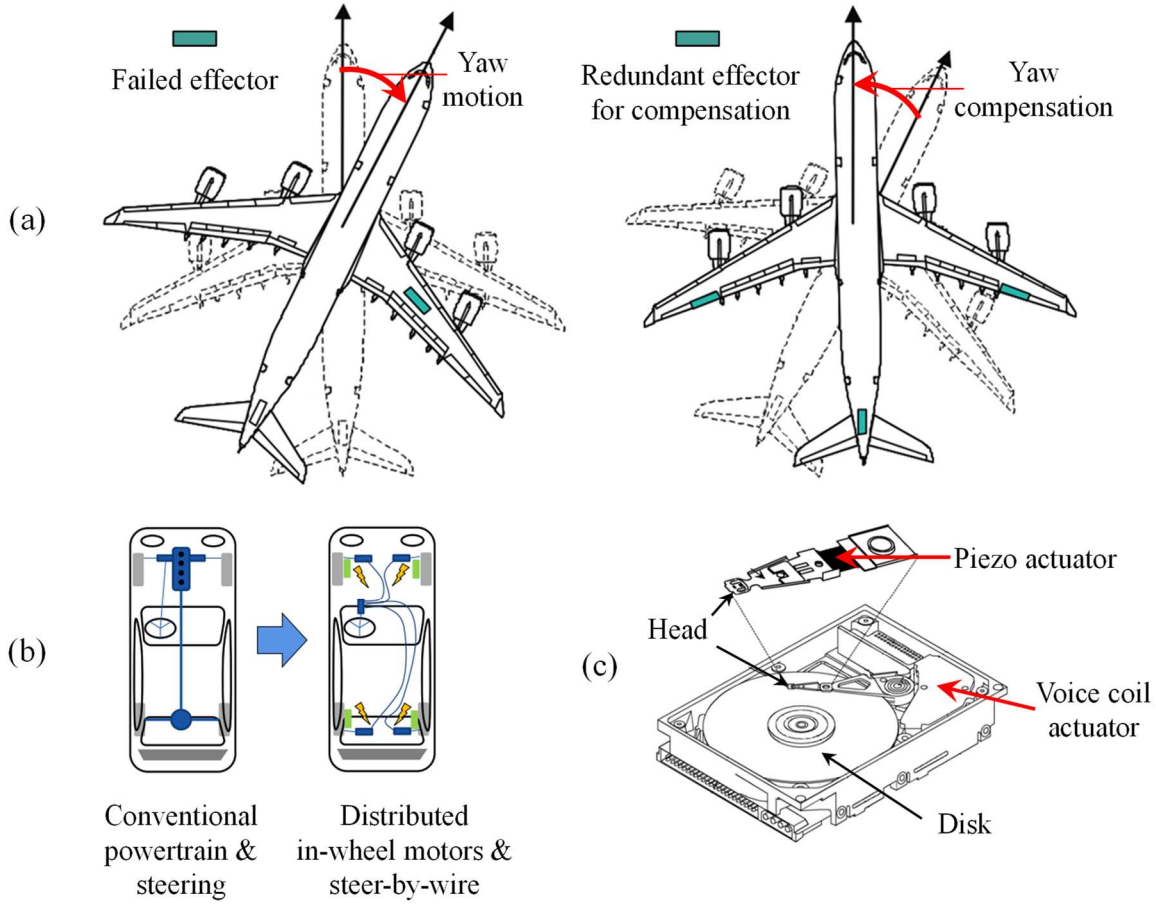

Figure 1.1: (a) The Use of Redundant Effectors to Compensate for a Failed Effector in an Airbus Plane [7], (b) Over-Actuation in the Vehicles using In-Wheel Motors and Steer-by-Wire System to Replace Conventional Powertrain and Steering [8], and (c) Over-Actuated Dual-Stage in Hard Disk Drive [9]

1.1.1.2 Over-Actuation in Vehicles

In rudimentary vehicle designs, the direction of the vehicle is controlled by its steering system while the speed is controlled through gas/brake combination. This combination can be viewed as an early usage of over-actuation: the engine cannot provide abrupt propulsion in reverse direction and brakes are used to overcome this deficiency. Over many decades of automobile development, the use of over-actuation has significantly increased. For instance, at the heart of so-called vehicle stability control, which is mandatory for most modern-day vehicles,

the braking forces to individual wheels of a vehicle are controlled separately to enable enhanced traction and yaw rate control [8,10–12]. Similarly, torque vectoring differentials [13] and multi-axis steering [11,14] are used to enhance the maneuverability of the vehicle; redundant air-pass geometry compressors are used to enhance the response of vehicle turbochargers [15]; a redundant actuator is attached to the steering shaft to reduce the steering torque [16]. Current developments are focused on replacing conventional powertrain and steering system with in-wheel motors and steer-by-wire system (shown in Figure 1.1 (b)) [8,17–19], targeted for improved maneuverability through over-actuation [8]. In steer-by-wire systems, hardware redundancy is usually required to ensure a certain level of fault tolerance [20]. Another important use of over-actuation in automotive applications is in hybrid electric vehicles [21,22] where redundant electric propulsion systems are used in collaboration with internal combustion engines to improve fuel economy.

1.1.1.3 Over-Actuation in Feed Drives

Feed drives are used to generate accurate motion of machine components in various applications like manufacturing, data storage and robotics [23]. In dual-stage feed drives, two feed drives are connected in series or parallel to deliver motion along one DOF. Dual-stage feed drive arrangements formulate over-actuated systems and can be employed to fulfill demanding quality, capacity, productivity, and energy efficiency requirements. One representative case is the combination of a short-range, precise actuator (e.g. piezoelectric actuator) in series with a long-range actuator with limited precision or response time (e.g. linear motor [24], voice coil [25], ball screw feed drive [26,27]). Such a dual-stage configuration simultaneously delivers long-range, precise, and rapid motion, which would have been challenging to achieve with a single actuator. Figure 1.1 (c) illustrates a voice coil actuator connected in series with a piezo actuator for controlling the motion of the read/write head in hard disk drives [9,28–30]. The increased motion range and precision enabled by this dual-stage configuration enables faster read/write speeds (higher throughput) and higher data density (capacity). Apart from series configurations, parallel combinations of redundant actuators are also common in feed drives. For example, gantry structures are commonly actuated through the dual actuators in parallel, which reduces the unwanted yaw vibration [31,32]; in [33], a tilting platform is designed to combine a pneumatic actuator and voice coil motor in parallel to achieve high torque and fast response simultaneously; and in [34], a linear motor drive is connected in parallel with a screw drive to

provide additional damping. The work in this thesis is motivated, from a practical standpoint, by a parallel-configured, dual-stage hybrid feed drive designed to improve energy efficiency in manufacturing, as discussed in the following section.

1.1.2 Hybrid Feed Drive ([35])

In manufacturing machines, such as machine tools, feed drives are responsible for generating accurate motion between tool and workpiece. Hence their positioning accuracy and speed are critical to the quality and productivity of the manufacturing processes [23]. They account for a significant portion of the energy consumption of manufacturing machines [36,37]. It is however recognized that, to achieve truly sustainable manufacturing, improvements in energy efficiency must be achieved without unduly sacrificing quality and productivity [38].

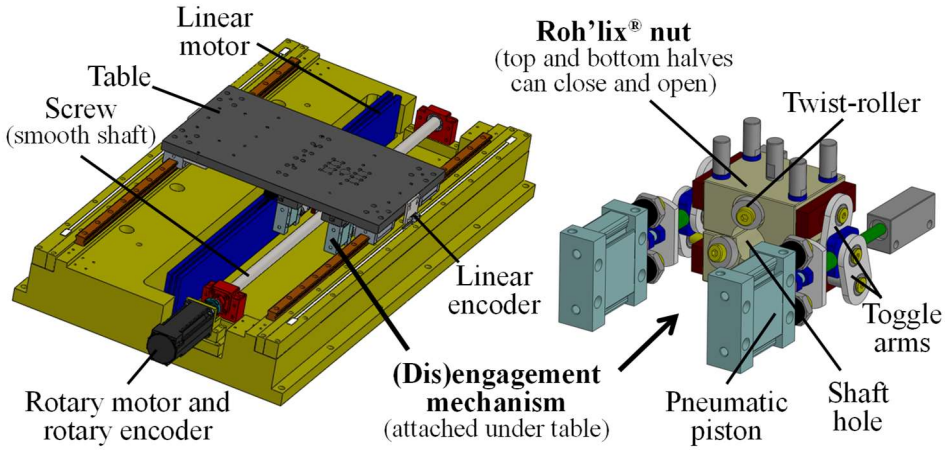

Figure 1.2: Schematic of HFD Prototype Proposed by Okwudire and Rogers [35]

Screw drives (SDs), which convert rotary motion from a rotary motor to linear motion using a screw, are the most common choice for actuating the translatory feed axes of machine tools [23]. The reason is that SDs are cost-effective and have a high mechanical advantage which allows them to support high cutting (i.e., machining) forces with very low energy consumption [39,40]. The speed and accuracy of SDs are however limited because of mechanical issues like vibration, wear, backlash and geometric errors of the screw and associated mechanical components [23,40]. To mitigate these shortcomings, linear motor drives (LMDs) are increasingly being employed [40–45]. LMDs can achieve higher speeds and accelerations than SDs and are not subject to the inaccuracies caused by geometric errors, wear and structural deformations arising from the screw and other mechanical components like bearings, couplers and nuts that are connected to it [23,40]. LMDs are therefore generally more precise than SDs.

However, because they provide no mechanical advantage, they consume a lot more electrical energy to support cutting forces than SDs, thus significantly increasing the energy consumption of the machine [42].

Machining operations typically consist of a combination of two modes – rapid traverse (i.e., high-speed, zero-cutting-force positioning moves) and cutting, usually involving low feed rates and large cutting forces. The idea behind the HFD is to: (*i*) drive the machine table using the LMD during rapid traverse to achieve very high speeds and accelerations with low energy consumption; (*ii*) drive the table using a low-lead SD during cutting to achieve the required cutting speeds with low energy consumption; (*iii*) use the LMD to compensate for vibrations and errors introduced by the SD during cutting; and (*iv*) ensure that the switch between the LMD and SD can be achieved rapidly and energy efficiently at any position of the table within its travel.

These functionalities are realized through a reconfigurable traction SD comprising a rolling helix (Roh'lix®) nut which creates a virtual helix and screw motion when engaged with a smooth shaft [46,47]. The Roh'lix nut consists of two spring-loaded halves that can be separated to disengage the smooth shaft from the nut at any given location. Accordingly, the HFD is operated in two modes – (*i*) the rapid traverse mode, when the nut is disengaged from the shaft and the LMD acts alone; and (*ii*) the cutting mode, when the nut is engaged to the shaft, allowing the SD and LMD to work together in parallel to actuate the table, resulting in an over-actuated system where the goal is to minimize energy consumption without sacrificing positioning accuracy.

1.1.3 Over-Actuated Systems Summary and Control Challenges

Through the discussion of various applications, it can be summarized that the core purpose of adopting over-actuation is to satisfy supplementary control objectives, in addition to the primary control objectives fulfilled by its fully-actuated counterpart. These supplementary objectives can be:

- fault tolerance (e.g. redundant hardware in aircraft and vehicles),
- response time reduction (e.g. gas/brake combination, redundant air-pass geometry compressor, coarse-fine dual stage feed drive),
- enhanced maneuverability (e.g. vehicle stability control system, torque vectoring differentials, multi-axis steering)

- yaw vibration reduction (e.g. redundantly actuated gantry structures)
- enhanced energy efficiency (e.g. hybrid electric vehicle, hybrid feed drive)
- …

Note that these objectives are not exclusive to each other, and an over-actuated system may be designed achieve multiple objectives. For example, the redundant effectors on the aircraft can both enhance the energy efficiency and fault tolerance.

Though the satisfaction of the supplementary control objectives seems desirable, the control of over-actuated systems is not straightforward. All the expected functionalities are not realized through the simple augmentation of actuators, targeted control and sensing algorithms need to be designed accordingly. For example, the fault tolerant control of aircraft shown in Figure 1.1 (a) requires the detection of the effector failure, and a safe transition algorithm towards a different control framework with yaw compensation. In the applications where energy efficiency is required, it is important to ensure that the actuators are minimally counteracting, which otherwise would result in significant energy consumption. Accordingly, enhanced control methods for over-actuated systems are needed, to ensure the satisfaction of both primary and supplementary control objectives.

1.2 Literature Review on Over-Actuated System Control Considering Energy Efficiency

Among the supplementary control objectives, energy efficiency is a fundamental one across different applications. Energy efficiency quantitatively measures the level of control efforts (usually using the two-norm of control signals), and is a major indicator of the degree of collaboration in over-actuated systems. Due to redundancy, a primary control objective can be realized either in a collaborative or an uncooperative combination of actuators, indicating significant potential energy consumption difference. Therefore, minimization of control energy is usually desired, to enhance the synergy among the redundant actuators.

In the early use of over-actuated systems, the avoidance of significant energy consumption was usually inherent. In these scenarios, the control system only needs to guarantee collaboration of actuators on a rudimentary level. Corresponding methods, without quantitatively optimized energy efficiency, are defined as rudimentary methods and discussed in Section 1.2.1. Further development of over-actuated system control optimizes energy efficiency explicitly. This

is usually achieved through two different directions: combining primary objectives with energy efficiency such that existing optimal control methods (e.g. linear quadratic control) can be employed, or decomposing primary and energy efficiency such that controller design follows a two-stage framework. These two different directions are referred as general optimal control methods and control allocation, which are discussed in details in Sections 1.2.2 and 1.2.3.

1.2.1 Rudimentary Methods

In the rudimentary methods, avoidance of significant energy consumption is realized through encouraging collaboration among redundant actuators. The mainstream of these rudimentary methods arises from extensions of existing methods for fully-actuated systems, and embraces a core concept of decomposition. One natural direction is to reduce the coupling among redundant actuators. This decoupling can be achieved in time domain, where the effective time spans of the redundant actuators are disjoint (e.g. gas/brake design), or in frequency domain, where the dominant working frequencies of redundant actuators minimally overlap (e.g. PQ control [48], decoupling network [49], etc.). This concept of decoupling minimizes the common regions where the redundant actuators are active, and thus reduces the probability of actuator counteracting. Another direction of decomposition is to adopt the master-slave framework [50–54], where certain "master" controls are first determined and then used to configure the rest "slave" control inputs. This framework is common in applications where a group of actuators is prioritized (e.g. in daisy chain methods [52,53]), or in applications where certain redundant actuators have limited effect on others (e.g. coarse-fine motion stages [54]). In robotics applications, a special type of master-slave framework is referred as impedance control [55–57], where redundant slave manipulators are controlled as if they are passively connected to the master manipulators. Accordingly, this hierarchy in the master-slave framework enforces actuator collaboration, which inherently benefits energy efficiency. However, these rudimentary methods only avoid significant energy consumption by encouraging actuator collaboration; they do not quantitatively account for energy efficiency and thus could be far from energy optimality.

1.2.2 General Optimal Control Methods

Unlike rudimentary methods, quantitative energy efficiency is considered in general optimal control methods. This optimization of control energy is not unique to over-actuated

systems. General optimal control methods concerns control energy, regardless of over-actuation. The two-norm of the control inputs, i.e.

$$J_u(\mathbf{u}) = \int_0^{+\infty} \mathbf{u}^{\mathrm{T}} \mathbf{R} \mathbf{u} dt, \tag{1.1}$$

is commonly used to define control energy, and is referred to as energy in this dissertation. Without loss of generality, this use of terminology is guided by the hybrid feed drive where the selection of $\mathbf{R}$ is such that J_u represents the Joule heat energy expended by its actuators. The quadratic energy cost has to combine with other control performance metrics to formulate an effective objective function.

For cases where the primary control objectives are also specified in this quadratic format, the optimal control of these two conflicting objectives yields a standard linear quadratic control [29,58]. For standard linear quadratic regulation (LQR) problems with infinite horizon, the solution is specified to be a static state feedback controller, whose gain matrices are specified by Riccati equations [58]. The LQR solutions can be further extended to tracking problems with exogenous inputs [59,60]. This generalization is referred as linear quadratic (LQ) control, which is discussed in detail in Section 5.2. This general optimal control method applies to over-actuated systems, and is shown to strictly enhance controllability with an increased degree of redundancy [61]. In over-actuated systems where the detailed state dynamics is not enforced, the redundancy in the state costs can be used to enhance the transient response of the system [28].

For cases where the primary control objectives are specified using the infinity norm, the optimal control formulate an H_2/H_∞ framework, and can be solved using convex optimization [62,63]. The application of the H_2/H_∞ methods enables a wide range of control objectives, including disturbance rejection, robustness, etc. [24,29]. Nonetheless, the H_2/H_∞ methods usually generate high order controllers even after model reduction, which may not be desirable for practical tuning and implementation [51].

One major critique of these general optimal control methods is that they do not distinguish between over-actuated and fully-actuated systems. Accordingly, they do not capture the characteristics of over-actuation, and thus the primary control objectives and energy are coupled in the design process. This coupling may require trial and error in weight selection and tuning, which is not desirable.

1.2.3 Control Allocation

Unlike the general optimal control methods, control allocation approaches are proposed to take advantage of the structure of the over-actuated systems. The null space in over-actuated systems enables the redistribution of control efforts without affecting the primary control objectives. Accordingly, control allocation usually embraces a two-stage framework, having a high-level nominal controller (aka. 'virtual controller' [10,17,64,65]) which determines the overall control effort required to achieve desired output trajectories, and a control allocator designed to best distribute the control effort to the actuators [10,64–67].

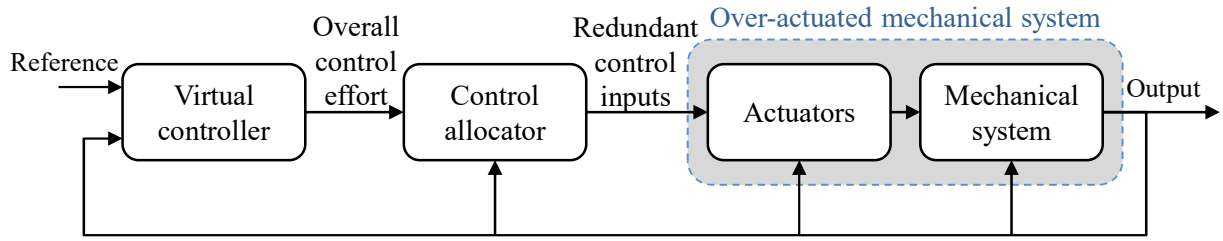

Figure 1.3: Two-Stage Framework in Control Allocation

To facilitate the discussion, two definitions of over-actuated systems (i.e. strong and weak input redundancy), firstly introduced in [68], are discussed. Control allocations methods under each definition are discussed, followed by their connections to the optimal control methods.

1.2.3.1 Definitions of Over-Actuation: Strong and Weak Input Redundancy

Consider a controllable MIMO linear time-invariant (LTI) system

$$\mathbf{y} = \mathbf{G}\mathbf{u} + \mathbf{G}_d\mathbf{d}, \tag{1.2}$$

where $\mathbf{y} \in \mathbb{R}^{n_y}$ is the output of the system, while $\mathbf{u} \in \mathbb{R}^{n_u}$, $\mathbf{d} \in \mathbb{R}^{n_d}$ are the control input and disturbance, respectively. Assume that one minimal state-space realization of the system is given by

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{B}_d\mathbf{d}; \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} + \mathbf{D}_d\mathbf{d}, \end{aligned} \tag{1.3}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$ is the state vector of the system.

Definition 1 (strong and weak input redundancy): The system given by transfer function matrix (1.2) and state-space representation (1.3) is:

(a) strongly input redundant if

$$\mathrm{Ker}\left(\begin{bmatrix}\mathbf{B}\\ \mathbf{D}\end{bmatrix}\right)\neq\varnothing; \tag{1.4}$$

(b) weakly input redundant if

$$\mathrm{Ker}\left(\mathbf{G}(s)\right)\neq\varnothing, \text{ for almost all } s. \tag{1.5}$$

Remark 1: Strong input redundancy in Definition 1 indicates that there exists a family of control inputs that does not affect the internal states, while weak input redundancy relaxes this constraints to output rather than states. Definition 1 arises from [68] but is slightly altered to enable broadband control allocation. The original definition of weakly input redundant systems in [68] focused on only static control allocation (i.e., for $s\rightarrow 0$). This original definition is extended to almost all frequencies as in (1.5), to make it applicable to dynamic (broadband) control allocation.

Remark 2: From the relationship between state-space and transfer function representations of a system, i.e.

$$\mathbf{G}=\mathbf{C}\left(s\mathbf{I}-\mathbf{A}\right)^{-1}\mathbf{B}+\mathbf{D}, \tag{1.6}$$

one can deduce that strongly input redundant systems formulate a subset of weakly input redundant systems, i.e., (1.4) ⇒ (1.5).

Remark 3. The strong and weak input redundancy in Definition 1 can be extended to nonlinear systems. Assume a nonlinear system given by

$$\begin{aligned}\dot{\mathbf{x}}&=\mathbf{f}\left(\mathbf{x},\mathbf{u}\right);\\ \mathbf{y}&=\mathbf{h}\left(\mathbf{x}\right).\end{aligned} \tag{1.7}$$

The states **x** are assumed to have zero initial conditions to simplify the discussion. The output is given by

$$\mathbf{y}(t)=\mathcal{G}\left\{\mathbf{u}(t)\right\}, \tag{1.8}$$

where $\mathcal{G}$ represents an operator from input to output. The system is defined to be strongly input redundant if

$$\forall\mathbf{u},\ \exists\Delta\mathbf{u}\neq 0,\ \text{s.t. }\mathbf{f}\left(\mathbf{x},\mathbf{u}\right)=\mathbf{f}\left(\mathbf{x},\mathbf{u}+\Delta\mathbf{u}\right), \tag{1.9}$$

and to be weakly input redundant if

$$\forall\mathbf{u}(t),\ \exists\Delta\mathbf{u}(t)\neq\mathbf{0}(t),\ \text{s.t. }\mathcal{G}\left\{\mathbf{u}(t)\right\}=\mathcal{G}\left\{\mathbf{u}(t)+\Delta\mathbf{u}(t)\right\}, \tag{1.10}$$

where $\mathbf{0}(t)$ represents a constant zero function.

1.2.3.2 Control Allocation for Strong Input Redundancy

Due to the null space defined in (1.4), there exists a set of control inputs that do not affect system internal states [64–67]. The control effective matrix **B** can be decomposed into

$$\mathbf{B} = \mathbf{B}_v \mathbf{B}_\tau, \tag{1.11}$$

where rank($\mathbf{B}_v$) = rank(**B**) [64,69]. Accordingly, invariant control performance is guaranteed through the linear constraint:

$$\mathbf{B}_\tau \mathbf{u} = \boldsymbol{\tau}, \tag{1.12}$$

where $\boldsymbol{\tau}$ is generally non-redundant (i.e. having the same degree of freedom as **y**), and is specified from the high-level virtual controller in the first stage. Consider other possible linear constraints (e.g. saturation) and the control energy at a particular instance as an objective, i.e.

$$\begin{aligned} &\min \ \mathbf{u}^{\mathrm{T}} \mathbf{R} \mathbf{u}, \\ &\text{subject to } \mathbf{U}^{-} \le \mathbf{u} \le \mathbf{U}^{+}. \end{aligned} \tag{1.13}$$

The problem becomes a quadratic programming problem with linear constraints. The most basic configuration having two control inputs is illustrated in Figure 1.4, where there may exist single or multiple feasible solutions satisfying linear constraints in (1.12) and saturation-type constraints defined in (1.13). The minimization of the two-norm is used to determine the energy-optimal control. There are several methods available in the literature to address this problem, including backstepping [6], redistributive pseudoinverse [69], fixed-point [70], etc. A detailed comparison of these methods is discussed in [67].

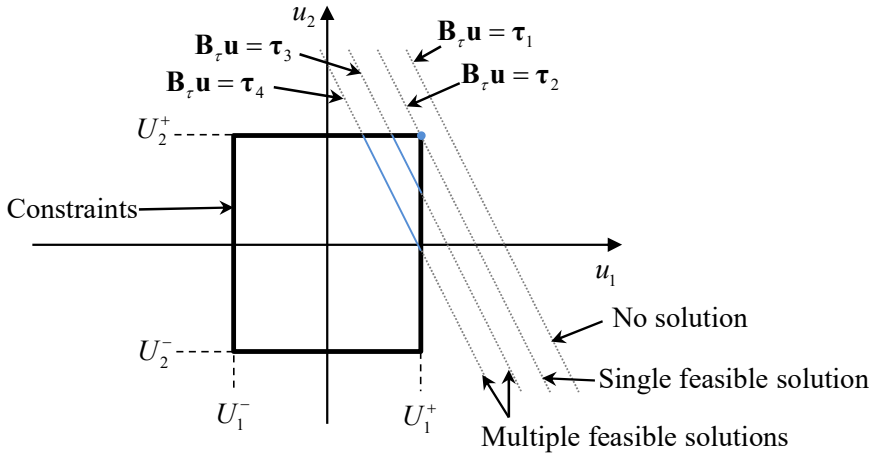

Figure 1.4: Illustration of Control Allocation for Strong Input Redundancy

Although the control allocation methods for strong input redundancy are relatively mature, it should be noted that strong input redundancy is a restrictive condition, which in practice often requires exact collocation of actuators or severe truncation of higher order dynamics [65]. This configuration leads to static allocator structure, fitted for slow allocation and applications where actuator dynamics are similar.

1.2.3.3 Control Allocation for Weak Input Redundancy

Recent advances in over-actuated system research indicate that dynamic allocation is necessary for applications where the fast response is required with unmatched redundant actuator dynamics [65,71]. Preliminary research on weak input redundancy borrow the static actuator model from strong input redundancy, however distribute the control efforts with heuristic dynamic filters [71], or introduce additional allocation memory and penalize actuator rates [65]. Though these dynamic allocation methods enable broadband allocation of control effort, they are still limited since the optimality under broadband disturbance is not guaranteed and the overall control performance may be affected by allocation. More systematic approaches for the weak input redundancy are needed.

To address the dynamic control allocation for weakly input redundant systems, a general control allocation approach for weakly redundant systems is to use a model predictive control framework [15]; however, this often leads to computationally expensive real-time optimization. Using regulator theory, Galeani et al. [72] explored static state feedback structures for optimal control allocation in weakly input redundant systems, by minimizing the infinity norm of the control efforts; a finite dimension relaxation was employed to reduce the computational cost. However, this relaxation leads to a hybrid system, and the associated switching events may introduce undesirable transients. Zaccarian [68] proposed a dynamic allocation method based on a static redundancy model in weakly input redundant systems. The matrix $\mathbf{B}_\perp$ is defined as the orthogonal complement of the DC gain matrix, given by

$$\left(\lim_{s\to 0}\mathbf{G}(s)\right)\mathbf{B}_\perp = \mathbf{0}. \tag{1.14}$$

The static-model-based dynamic allocation in [68] is expressed as

$$\begin{aligned} \dot{\mathbf{w}}_z &= \mathbf{K}_z\mathbf{B}_\perp\mathbf{R}\mathbf{u}; \\ \mathbf{u} &= \boldsymbol{\tau} + \mathbf{B}_\perp\mathbf{w}_z, \end{aligned} \tag{1.15}$$

where $\mathbf{K}_z$ is a tuning parameter of the convergence speed and $\boldsymbol{\tau}$ is the non-redundant overall general efforts generated through the high level virtual controller. This allocator design based on static redundancy is further expanded to nonlinear objectives in [73]. An allocation scheme following similar concept is proposed for first-order nonlinear systems [74]. While these approaches [68,73,74] greatly simplify the problem and reduces computational burden, it is incapable of optimal broadband control allocation. Moreover, given an accurate plant model, the approaches in [68] and [72] both cannot guarantee invariance of the controlled output (i.e., control performance) during broadband control allocation due to their statically-defined null space. Accordingly, computational efficient dynamic control allocation methods with minimal performance variation are needed, for weakly input redundant systems.

1.2.3.4 Connections between Optimal Control and Control Allocation

As two major approaches, the connections between the optimal control approach and control allocation should be explored. Harkegard and Glad [64] discussed the equivalence of the two frameworks in strongly input redundant systems. It is shown that the optimal control problem with a general positive definite performance cost $\mathbf{H}_y$, given by

$$J_{uy}(\mathbf{u}) = \int_0^{+\infty} \left(\mathbf{H}_y + \mathbf{u}^{\mathrm{T}} \mathbf{R} \mathbf{u} \right) dt, \tag{1.16}$$

is equivalent to a control allocation two-stage framework, where the virtual control $\boldsymbol{\tau}$ is optimized through

$$\begin{aligned} &J_{\tau y}(\boldsymbol{\tau}) = \int_0^{+\infty} \left(\mathbf{H}_y + \boldsymbol{\tau}^{\mathrm{T}} \mathbf{R}_\tau \boldsymbol{\tau} \right) dt; \\ &\mathbf{R}_\tau \triangleq \left(\mathbf{B}_\tau \mathbf{R} \mathbf{B}_\tau^{-1} \right)^{-1}, \end{aligned} \tag{1.17}$$

while the allocator decides $\mathbf{u}$ according to the following problem

$$\begin{aligned} &\min \ \mathbf{u}^{\mathrm{T}} \mathbf{R} \mathbf{u}, \\ &\text{subject to } \mathbf{B}_\tau \mathbf{u} = \boldsymbol{\tau}. \end{aligned} \tag{1.18}$$

However, this analysis only applies to strongly input redundant systems, and the relationship between the two approaches for weakly input redundant systems is not explored in previous literature.

1.2.4 Summary of Core Deficiencies in Existing Work

In summary, the following deficiencies are found in existing methods for energy-optimal control of over-actuated system:

- Rudimentary methods avoid significant energy consumption in over-actuated systems through decoupling, master-slave framework, passive connection, etc. They do not explicitly optimize energy consumption and thus could be far from energy optimality.
- General optimal control methods do not exploit the structure introduced by redundancy. Accordingly, the control performance and energy are coupled in the design process, which is not desirable since it may require trial and error in weight selection.
- Control allocation methods for strongly input redundant systems are mature; however, the definition of strong input redundancy is restricted since it often requires exact collocation of actuators or severe truncation of higher order dynamics, which impedes its practical employment in achieving optimal dynamic allocation.
- Control allocation methods for weakly input redundant systems are either computationally expensive or they introduce approximations which result in control performance variations and/or do not guarantee optimality in dynamic (broadband) control allocation.
- The relationship between control allocation and optimal control methods has not been explored for weakly input redundant systems.

1.3 Dissertation Contributions and Outline

To address the deficiencies enumerated above, specifically for weakly input redundant systems, the following contributions are made in this dissertation:

1. An optimal control ratio/subspace for weakly input redundant systems is theoretically derived to specify the optimal relationship among the redundant actuators irrespective of external disturbances, leading to a new approach termed optimal control subspace-based (OCS) control allocation.
2. The optimal control ratio/subspace is shown to be non-causal, which poses challenges for practical implementation. Therefore, a causal approximation is proposed and used to design structured energy-efficient controllers for a weakly input redundant HFD; a relationship quantifying the energy efficiency loss due to the approximation is derived.

3. A causal measurement of the deviation from the optimal control ratio/subspace is introduced as a proxy for deviation from energy optimality. The proxy is used to convert the control allocation problem to a regulation problem compatible with most standard controller design methodologies. Compared to an existing allocation technique [68], the proposed proxy-based control allocation method is shown to dynamically allocate control efforts optimally without sacrificing control performance.
4. The proposed OCS control allocation technique is shown be equivalent to standard LQ control methods, assuming perfect knowledge of the disturbances acting on the controlled system. It is also shown that the OCS control allocation approach is preferable over the LQ control method in practical situations where the disturbance is typically unknown.
5. The OCS control allocation technique is validated in simulation and experiments on the HFD. Significant enhancement in energy efficiency is demonstrated without sacrificing positioning accuracy.

The concept of optimal control ratio, its causal approximation, and its application in HFD have been published in [75,76]. The concept of control proxy, as well as its usage to convert the control allocation problems into regulation problems, have been published in [77–79]. The connections between the proposed OCS control methods and LQ methods are discussed in [80].

The thesis is organized in following order: The concept of optimal control ratio for dual-input, single-output (DISO) weakly input redundant system is introduced in Chapter 2. Using the ratio and its approximation, energy-efficient structured controllers are proposed for the HFD, and validated in simulations and experiments. In Chapter 3, to address the approximation error and implementation challenges of using the optimal control ratio, a proxy signal is established and its usage in OCS control allocation is discussed and validated on the HFD. The concepts of optimal control ratio and proxy for DISO systems is extended to multi-input, multi-output (MIMO) over-actuated systems in Chapter 4. The relationship between the proposed OCS control allocation and the traditional linear quadratic control is then illustrated in Chapter 5, followed by conclusions and future work discussed in Chapter 6.

Chapter 2 Energy-Efficient Control of an Over-Actuated Hybrid Feed Drive via Optimal Control Ratio

2.1 Overview

This chapter explores the energy-efficient control of an over-actuated hybrid feed drive. It starts by deriving an optimal control ratio (OCR) for weakly input redundant dual-input, single-output systems. This OCR helps to determine the energy optimality of controllers in weakly input redundant systems, and is applied in this chapter to the hybrid feed drive (HFD) presented in Chapter 1. Though the OCR is shown to be non-causal (meaning that it cannot be directly evaluated in real time), a causal approximation of OCR is proposed to encourage its practical usage. The energy loss due to the approximation is analyzed and is shown to be small for the HFD. The causal approximation of the OCR is further exploited in structured feedforward and feedback control design and optimization.

This chapter is organized as follows: The derivation of the OCR is provided in Section 2.2. The modeling of HFD, its OCR and causal approximation is discussed in Section 2.3. The application of the causal approximation of the OCR to structured feedforward and feedback design for the HFD is presented in Section 2.4, followed by simulation and experimental validation in Section 2.5, and a chapter summary in Section 2.6.

2.2 Optimal Control Ratio for Dual-Input, Single-Output Systems

Assume a dual-input, single-output (DISO) LTI plant given by

$$y = \mathbf{G}\mathbf{u} + \mathbf{G}_d\mathbf{d}, \tag{2.1}$$

where $\mathbf{G} = [G_1, G_2]$; y, $\mathbf{u} = [u_1, u_2]^{\mathrm{T}}$, and $\mathbf{d}$ are respectively the system output, control inputs and disturbance input; both u_1 and u_2 are assumed to belong to L^2 space and has zero initial values. According to the definition in Section 1.2.3.1, such a system (with more inputs than outputs) is weakly input redundant. In weakly input redundant systems, there exists a family of input trajectories that yield the same output trajectory due to redundancy [15,68]. Assuming a nominal control input $\mathbf{u}_0$ yields a desirable output y_0 under disturbance $\mathbf{d}_0$, i.e.,

$$\mathbf{G}\mathbf{u}_0 = y_0 - \mathbf{G}_d\mathbf{d}_0. \tag{2.2}$$

The family of control signals that replicate y_0 under $\mathbf{d}_0$ defines a set Ω, given by

$$\Omega(\mathbf{u}_0) \triangleq \left\{\mathbf{u} \in \mathbb{R}^2 : \mathbf{G}(\mathbf{u}_0 - \mathbf{u}) = 0\right\}. \tag{2.3}$$

Set Ω is the non-empty optimization space of over-actuated systems without affecting control performance, and can be calculated by establishing the orthogonal complement of system $\mathbf{G}$.

Assume the typical quadratic control energy cost functional

$$J_u \triangleq \int_0^{+\infty} \left(\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}\right) dt = \int_0^{+\infty} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} dt, \tag{2.4}$$

where $\mathbf{R}$ is a positive definite square symmetric matrix which, in general, is a design parameter. However, in certain scenarios, $\mathbf{R}$ can be determined by physical properties of the actuators. To achieve optimality, the variation of J_u should satisfy

$$\delta J_u = 0 \Rightarrow \int_0^{+\infty} \left(\delta\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}\right) dt = 0, \tag{2.5}$$

where $\delta\mathbf{u} = [\delta u_1, \delta u_2]^{\mathrm{T}}$. Moreover, based on (2.3), the variation of every member of set Ω should satisfy the relationship

$$G_1\delta u_1 + G_2\delta u_2 = 0. \tag{2.6}$$

Combining (2.5) and (2.6), we get

$$\begin{aligned} &\int_0^{+\infty} \left(R_{11}u_1\delta u_1 + R_{12}(-u_1(G_2^{-1}G_1\delta u_1) + u_2\delta u_1) - R_{22}u_2G_2^{-1}G_1\delta u_1\right) dt = 0 \\ &\Rightarrow \int_0^{+\infty} \left(R_{11}u_1 + R_{12}(-(G_2^{-1}G_1)^*u_1 + u_2) - R_{22}(G_2^{-1}G_1)^*u_2\right) \delta u_1 dt = 0, \end{aligned} \tag{2.7}$$

where superscript * represents the adjoint operator; the frequency response of the adjoint system is the complex conjugate of the original system at every frequency [81]. According to the fundamental lemma of calculus of variations [82], the integrand of (2.7) must be equal to zero at all time, which yields the energy-optimal control (input) ratio (OCR)

$$\beta^*(s) \triangleq \frac{\hat{u}_1(s)}{\hat{u}_2(s)} = \frac{R_{22}G_1^* - R_{12}G_2^*}{R_{11}G_2^* - R_{12}G_1^*}, \tag{2.8}$$

where the $\hat{}$ accent is used to denote optimality. However, OCR in (2.8) contains adjoint dynamics, which in general is non-causally implementable. Therefore, it requires additional approximation or transformation to facilitate its use in practical applications like the HFD.

2.3 HFD Modeling and Its Optimal Control Ratio

2.3.1 Overview of HFD

As introduced in Section 1.1.2, the HFD is proposed to address the intrinsically different needs of two machining operation modes: high-speed, zero-cutting-force rapid traverse, and low-speed, high-cutting-force cutting motion. According to the machining operation mode, the HFD reconfigurably combines an energy-efficient SD and an energy-costly but more precise LMD to achieve speed, precision and energy efficiency. The detailed schematic of the HFD is illustrated in Figure 1.2 and the prototype is shown in Figure 2.2. An air-core Kollmorgen IL18-100A3 linear motor is placed in the middle while a SD with a smooth shaft is driven by a Kollmorgen AKM33H brushless DC motor on the side. The use of an air-core over an iron-core motor is because the LMD does not need to support large cutting forces. The HFD is equipped with a 0.08 μm resolution optical linear encoder fixed to one side of the table, and a 157 μrad resolution motor-mounted optical rotary encoder.

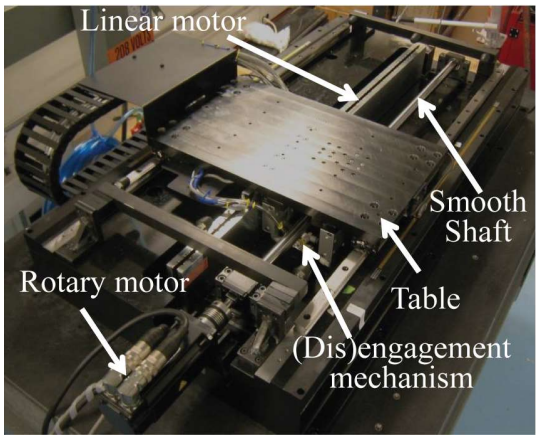

Figure 2.1: HFD Prototype

To facilitate easy engagement and disengagement of the SD from the table, a traction drive SD which uses a rolling helix (or Roh'lix®) nut [46,47] is employed. The Roh'lix nut converts rotary motion to linear motion using rolling element ball bearings that trace a screw motion (of lead l = 5 mm) along a smooth shaft of 25 mm diameter and 965 mm length. It is designed to carry up to 444 N of thrust force without significant slippage between the nut and shaft. The Roh'lix nut consists of two spring-loaded halves that can be separated to disengage the smooth shaft from the nut at any given location. A simple toggle mechanism is designed to separate the two halves of the nut using a pair of fast-acting pneumatic pistons. The HFD is thus

operated in either (*i*) the rapid traverse mode, when LMD actuate the table alone with disengaged nut, or (*ii*) the cutting mode, when the SD and LMD synergistically actuate the table. This dissertation uses the cutting mode control design as a validation tool for the proposed energy-optimal control methods for over-actuated systems.

2.3.2 Two-Mass Model and Identification

A simple two-mass model (like the one shown in Figure 2.2 for the HFD) is commonly used to analyze the mechanical dynamics of SDs for control purposes [83–86]. Even though more advanced SD models are available in the literature (e.g., [87]), the two-mass model is preferred because it is simple and it captures the first axial/torsional vibration mode, which is well known to be the bottleneck for controller design for SDs [83–86]. The LMD does not add any new low-order mechanical dynamics to the HFD shown in Figure 1.2. Therefore, the two-mass model used for SDs can also be used for controller design and analysis of the HFD [34].

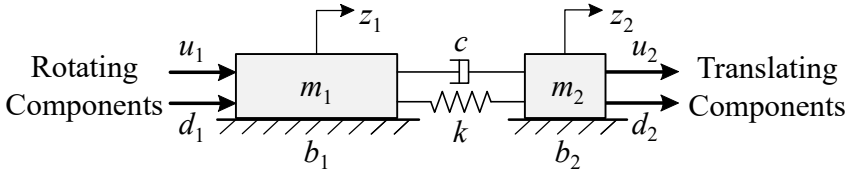

Figure 2.2: Simple Two-Mass Model of HFD

In the two-mass model of Figure 2.2, m_1 and m_2 are the (equivalent) masses of the rotating and translating components of the HFD, while k and c represent the stiffness and viscous damping coefficient of the connecting mechanical components; z_1, z_2 and b_1, b_2 are the (equivalent) displacements and viscous damping coefficients at m_1 and m_2, respectively; u_1 and u_2 are the (equivalent) control forces applied by the rotary and linear motors, while d_1 and d_2 represent external disturbance forces (e.g., non-viscous friction and cutting forces) applied to m_1 and m_2, respectively. The equation of motion of the HFD is given by

$$\begin{aligned} m_1\ddot{z}_1 + b_1\dot{z}_1 + c(\dot{z}_1 - \dot{z}_2) + k(\dot{z}_1 - \dot{z}_2) &= u_1 + d_1, \\ m_2\ddot{z}_2 + b_2\dot{z}_2 + c(\dot{z}_2 - \dot{z}_1) + k(\dot{z}_2 - \dot{z}_1) &= u_2 + d_2. \end{aligned} \tag{2.9}$$

Defining s as the Laplace variable, the plant transfer function matrix, $\mathbf{G}_m$, between the input forces, $\mathbf{u} = \{u_1, u_2\}^{\mathrm{T}}$ and $\mathbf{d} = \{d_1, d_2\}^{\mathrm{T}}$, and the output displacements, $\mathbf{z} = \{z_1, z_2\}^{\mathrm{T}}$, of the HFD is derived as

$$\mathbf{z}(s)=\underbrace{\begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix}}_{\mathbf{G}_m(s)}\left(\mathbf{u}(s)+\mathbf{d}(s)\right), \tag{2.10}$$

where

$$\begin{aligned} G_{11}&=\frac{m_2 s^2+\left(c+b_2\right)s+k}{D_{ol}},\quad G_{22}=\frac{m_1 s^2+\left(c+b_1\right)s+k}{D_{ol}},\\ G_{12}&=G_{21}=\frac{cs+k}{D_{ol}},\quad D_{ol}=a_{ol,4}s^4+a_{ol,3}s^3+a_{ol,2}s^2+a_{ol,1}s;\\ a_{ol,4}&=m_1m_2;\quad a_{ol,3}=b_1m_2+b_2m_1+c\left(m_1+m_2\right)\\ a_{ol,2}&=c\left(b_1+b_2\right)+k\left(m_1+m_2\right)+b_1b_2;\quad a_{ol,1}=\left(b_1+b_2\right)k. \end{aligned} \tag{2.11}$$

The actual HFD system is experimentally measured from the frequency response functions (FRFs) at discrete frequencies (i.e., $\mathbf{G}_a(j\omega)$). The two mass model is identified by curve fitting the two-mass model FRFs (i.e., $\mathbf{G}_m(j\omega)$) with $\mathbf{G}_a(j\omega)$. Figure 2.3 compares the FRFs of the two-mass model (i.e., $\mathbf{G}_m(j\omega)$) with the experimentally measured FRFs (i.e., $\mathbf{G}_a(j\omega)$). As seen, the two FRFs are in good agreement up to 100 Hz, hence $\mathbf{G}_m$ captures the critical axial/torsional mode of the HFD, occurring at 41 Hz. The identified parameters are reported in Table 2.1.

Table 2.1: Identified Parameters of the HFD's Two-Mass Model

m_1 [kg]	616.2	b_1 [kg/s]	44.8	c [kg/s]	5777.2
m_2 [kg]	46.3	b_2 [kg/s]	83.3	k [N/μm]	3.1469

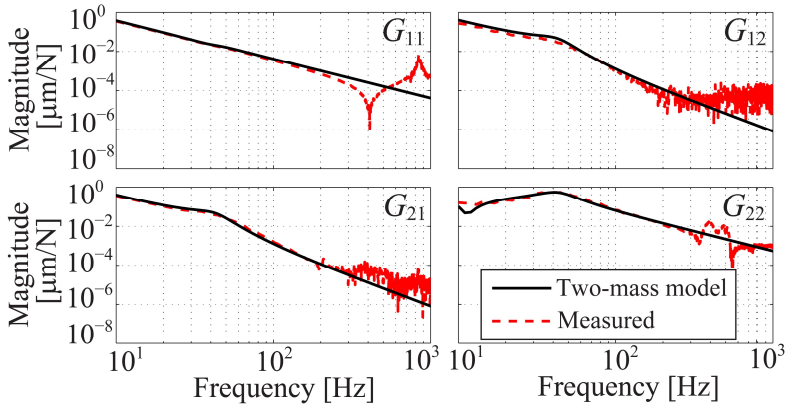

Figure 2.3: Comparison of Two-Mass Model with Experimentally Measured FRF

The energy efficiency of electric motors (linear or rotary) is adversely affected by copper losses, iron losses, friction, and switching losses in the motors and their amplifiers [36,88]. Of these loss mechanisms, copper losses are by far the most dominant, especially at high loads [36], which are typical in machining. Therefore, in this chapter, the heat energy generated as a result of copper losses is used as a measure of actuator efficiency. It is given by the equation [88–93]

$$E_{heat} = \sum_{i=1}^{2} \int_0^{+\infty} \left(\frac{u_i}{K_{mi}} \right)^2 dt. \tag{2.12}$$

The parameter K_{mi} represents the (equivalent) motor constant of the ith actuator in $\mathrm{N}/\sqrt{\mathrm{W}}$. It is often reported in the specification sheets of electric motors. Given the same RMS actuation force, the motor constant provides a measure of the comparative efficiency of each actuator; the larger the motor constant, the more efficient the motor. The motor constants for its rotary and linear motors are $K_{m1} = 380.8\ \mathrm{N}/\sqrt{\mathrm{W}}$ and $K_{m2} = 21\ \mathrm{N}/\sqrt{\mathrm{W}}$, respectively, where K_{m1} has been converted from its nominal value in $\mathrm{Nm}/\sqrt{\mathrm{W}}$ using a factor of $2\pi/l$, with lead $l = 5 \times 10^{-3}$ m. Therefore, actuator efficiency ratio $\gamma = (K_{m1}/K_{m2})^2 = 328.8$, indicating two orders of magnitude difference in the energy efficiencies of the two actuators.

2.3.3 Optimal Control Ratio of HFD

The two-mass model $\mathbf{G}_m$ is fully-actuated. However, the position of the rotary component z_1 does not affect machining accuracy, thus only the position of the table z_2 is the focus. In this regard, the two-mass model of HFD is a DISO system in (2.1) assuming $y = z_2$, $G_1 = G_{21}$, and $G_2 = G_{22}$. The control energy defined in (2.4) relates to the heat generated in HFD through

$$R_{11} = K_{m1}^{-2},\ R_{22} = K_{m2}^{-2},\ R_{12} = 0. \tag{2.13}$$

Accordingly, the OCR, β^*, between the SD force (u_1) and the LMD force (u_2) of the HFD is expressed as

$$u_1 = \underbrace{\gamma \left(G_{22}{}^{-1} G_{21} \right)^*}_{\triangleq \beta^*} u_2. \tag{2.14}$$

Figure 2.4 shows the Bode plot of β^*. Notice, according to (2.14), that β^* combines the actuators' frequency dependent maneuverability with their energy cost. At low frequencies, $\beta^* \to \gamma$, as in static-model-based dynamic allocation methods [68], indicating that the rotary motor is preferred for efficiency. However, at higher frequencies, $\beta^* \to 0$, indicating that it is more efficient to utilize the linear motor for precision positioning, since the rotary motor loses control of the table at high

frequencies due to the mechanical decoupling effect in the two-mass system with non-collocated input and output [94]. This frequency dependent change in OCR requires broadband allocation, which cannot be achieved with static-model-based dynamic allocation methods (e.g., [68]). Notice the positive phase of β^* at resonance and beyond, indicating that it is unstable hence not causally implementable.

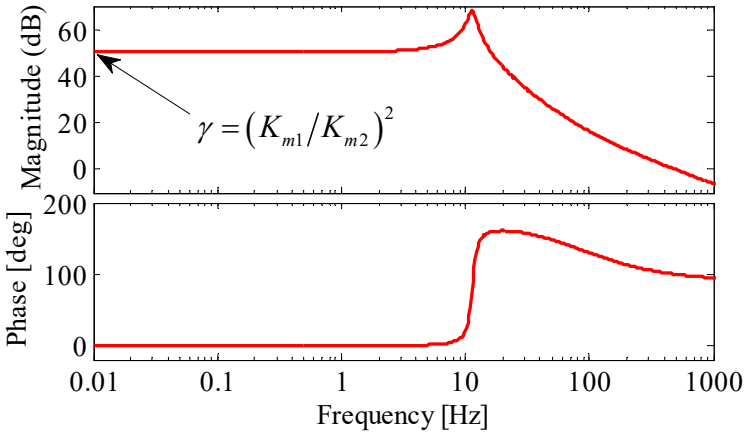

Figure 2.4: Optimal Control Ratio of HFD

2.3.4 Causal Approximation of Optimal Control Ratio of HFD

The OCR, β^*, is not a causal operator. To facilitate causal evaluation of control optimality, a causal function is needed. Such a causal function can be attained by approximating β^* by neglecting the adjoint operation, thus enforcing control ratio between u_1 and u_2 as

$$\beta = \gamma\left(G_{22}^{-1}G_{21}\right). \tag{2.15}$$

The energy efficiency difference caused by using β in place of β^* is analyzed in this section.

Consider η describing an arbitrary ratio (both causal and non-causal) between u_1 and u_2. In accordance with (2.10), the two control signals are uniquely calculated as

$$\begin{cases} u_1 = \eta u_2 \\ y_u = G_{21}u_1 + G_{22}u_2 \end{cases} \Rightarrow \begin{cases} u_1 = \eta\left(G_{21}\eta + G_{22}\right)^{-1} y_u \\ u_2 = \left(G_{21}\eta + G_{22}\right)^{-1} y_u \end{cases}. \tag{2.16}$$

where

$$y_u = y - G_{21}d_1 - G_{22}d_2 \tag{2.17}$$

indicating the pure contribution to the output from u_1 and u_2. Due to Parseval's theorem, the integral of the time domain energy cost function defined in (2.12) can be expressed in the frequency domain. Accordingly, the integrand at each frequency is given by

$$\begin{aligned} E_h(\eta,\omega) &\triangleq \frac{u_1(j\omega)u_1^*(j\omega)}{2\pi K_{m1}{}^2} + \frac{u_2(j\omega)u_2^*(j\omega)}{2\pi K_{m2}{}^2} \\ &= \left.\frac{y_u y_u^*\left(\eta\eta^* + \kappa^2\right)}{2\pi K_{m1}{}^2\left(G_{21}\eta + G_{22}\right)\left(G_{21}\eta + G_{22}\right)^*}\right|_{s=j\omega}. \end{aligned} \tag{2.18}$$

Therefore, the frequency dependent energy efficiency ratio, R_β, between the correct ratio (i.e. $\eta = \beta^*$) and its causal approximation (i.e. $\eta = \beta$) is given by

$$R_\beta(\omega) = \frac{E_h(\beta^*,\omega)}{E_h(\beta,\omega)} = \left.\frac{G_{21}G_{21}{}^*\beta\beta^* + G_{21}G_{22}{}^*\beta + G_{21}{}^*G_{22}\beta^* + G_{22}G_{22}{}^*}{G_{21}G_{21}{}^*\beta\beta^* + G_{21}G_{22}{}^*\beta^* + G_{21}{}^*G_{22}\beta + G_{22}G_{22}{}^*}\right|_{s=j\omega}. \tag{2.19}$$

Note that $R_\beta \approx 1$ when $|\beta(j\omega)| \gg 1$, $|\beta(j\omega)| \ll 1$ or $\angle\beta(j\omega) \approx n\pi$ ($n \in \mathbb{Z}$). Therefore, the deviation of R_β from unity is limited to certain frequency ranges, based on these conditions.

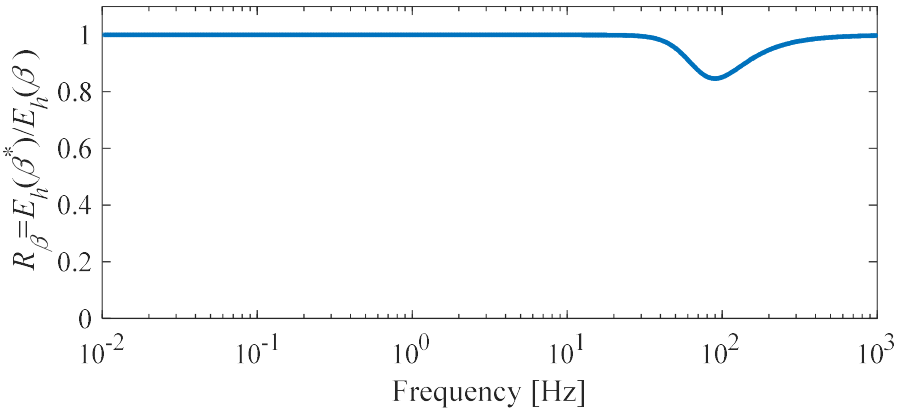

Figure 2.5: Energy Efficiency Ratio R_β as a Function of Frequency

The impact of using β in (2.15) instead of β^* is investigated using R_β, calculated with the hybrid feed drive parameters in Table 2.1; the result is shown in Figure 2.5. At most frequencies $R_\beta \approx 1$; the only exception is the region around 100 Hz, where R_β deviates significantly from unity (because $\beta(j\omega)$ violates the magnitude and phase relationships required to keep $R_\beta \approx 1$). Nonetheless, in the worst case, R_β only drops to 85% at 89 Hz. Therefore, for the HFD, the overall efficiency cost of using the approximate ratio, β, instead of the correct ratio, β^*, is small. Considering the practical benefits gained from its causality combined with its near-optimality,

we conclude that the approximate control ratio β in (2.15) is suitable for energy-efficient controller design for the HFD.

2.4 Energy-Efficient Control of HFD

In motion control applications, it is common to use a so-called two degree-of-freedom (DOF) structure, consisting of a feedforward (FF) and feedback (FB) controller. The FF controller focuses on tracking reference commands while the FB controller handles regulation tasks. The 2-DOF control structure for the HFD is shown in Figure 2.6. The FF controller, $\mathbf{C}_{ff}$, takes the desired position trajectory y_d as its input and generates the FF component, $\mathbf{u}_{ff} = \{u_{ff1}, u_{ff2}\}^{\mathrm{T}}$, of the control signal $\mathbf{u}$, as well as the reference position signal, $\mathbf{z}_r = \{z_{r1}, z_{r2}\}^{\mathrm{T}}$, for the feedback controller. On the other hand, the FB controller, $\mathbf{C}_{fb}$, determines the feedback component, $\mathbf{u}_{fb} = \{u_{fb1}, u_{fb2}\}^{\mathrm{T}}$, of $\mathbf{u}$ based on the error vector, $\mathbf{e}_z = \mathbf{z}_r - \mathbf{z}$. It is of interest to design $\mathbf{C}_{ff}$ and $\mathbf{C}_{fb}$ such that:

(*a*) The best positioning performance is achieved, subject to stability constraints; and

(*b*) Specification (*a*) above is achieved as energy-efficient as possible (i.e., with minimal heat loss from the actuators).

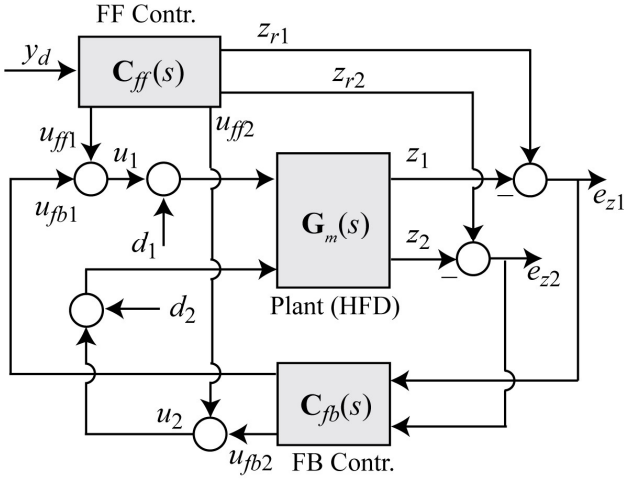

Figure 2.6: Block Diagram of Two-DOF Control Structure for HFD

The underlying assumption in defining the two-tiered specifications above is that, in precision machining, positioning accuracy typically takes precedence over energy efficiency. Therefore, as much as possible, positioning accuracy should not be sacrificed for energy efficiency. It is assumed that, because of the redundant actuation of the HFD, there are multiple combinations of u_1 and u_2 that can achieve the best positioning performance; of those, the most

efficient should be selected. Due to the uncertainty of the disturbance, the FF and FB portions of the control inputs are both expected to align with the OCR β^* or its causal approximation β introduced in Section 2.3.4. An elegant $\mathbf{C}_{ff}$ and $\mathbf{C}_{fb}$ design methodology for achieving the stated goals is proposed in the following two subsections.

2.4.1 Design of Energy-Efficient FF Controller

In the FF controller design $\mathbf{d} = \mathbf{0}$ is assumed, thus perfect tracking can be achieved if

$$y_d = z_2 = G_{21}u_{ff1} + G_{22}u_{ff2}. \tag{2.20}$$

This perfect tracking constraint, combined with the causal approximation of the optimal ratio β, yields a causal energy-efficient FF control law as

$$\begin{aligned} &u_{ff1} = \beta u_{ff2}, \quad u_{ff2} = G_{ff}^{-1} y_d; \\ &G_{ff} \triangleq \beta G_{21} + G_{22}. \end{aligned} \tag{2.21}$$

Accordingly, z_{r1} is given by

$$\begin{aligned} z_{r1} &= G_{11}u_{ff1} + G_{12}u_{ff2} \\ &= \left(\beta G_{11} + G_{12}\right) G_{ff}^{-1} y_d. \end{aligned} \tag{2.22}$$

One major challenge that arises with the energy-optimal FF controller (given by (2.21)) is that, even though each element of $\mathbf{G}$ is minimum phase, the transfer function G_{ff} may have non-minimum phase (NMP) zeros, as such, it may not have a stable inverse. There are several methods in the literature for calculating a stable approximation for the inverse of a transfer function with NMP zeros (e.g., [95–97]). Amongst them, the most notable is Tomizuka's zero phase error tracking control (ZPETC) [95]. ZPETC determines a stable inverse such that there are no errors in phase across all frequencies; however, it gives rise to errors in gain compared to the exact inverse. If the numerator and denominator polynomials of G_{ff} are represented by B_{ff} and A_{ff}, respectively, its stable inverse with zero phase error is given by

$$G_{ff-zpetc}^{-1} = \frac{A_{ff}(s)B_{ffu}(-s)}{B_{ffs}(s){B_{ffu}}^2(0)}, \tag{2.23}$$

where B_{ffs} and B_{ffu} represent the stable and unstable portions of B_{ff}. Note that (2.23) represents the continuous-time equivalent of Tomizuka's discrete-time ZPETC [95].

The problem with the approximate inverse of (2.23) is that it sacrifices energy optimality as well as perfect tracking because it satisfies neither (2.20) nor (2.21); this is also the case for the other similar approximate inversion methods available in the literature (e.g., [96,97]). We

therefore seek an approximation for G_{ff} that can guarantee perfect tracking while maintaining near energy optimality. To determine such a near-energy-optimal (NEO) FF controller, the minimum energy condition (2.5) for HFD, given by

$$\int_0^{+\infty}\left(\frac{2u_1}{K_{m1}{}^2}\delta u_1+\frac{2u_2}{K_{m2}{}^2}\delta u_2\right)dt=0. \tag{2.24}$$

is modified to be

$$\int_0^{+\infty}\left(\left(u_{ff1}-K_{ff}u_{ff2}\right)\delta u_{ff1}+\gamma u_{ff2}\delta u_{ff2}\right)dt=0, \tag{2.25}$$

where K_{ff} is a scalar gain. The implication of (2.25) is that energy optimality is sacrificed by making the variation of the energy functional in (2.24) to be non-zero by introducing K_{ff}. Combining the perfect tracking constraint of (2.20) with (2.25), we get

$$\begin{aligned}&u_{ff1}=\left(K_{ff}+\beta\right)u_{ff2},\qquad u_{ff2}=G_{ff-neo}^{-1}y_d;\\&G_{ff-neo}\triangleq G_{21}K_{ff}+\beta G_{21}+G_{22}.\end{aligned} \tag{2.26}$$

It can be shown that there always exists a $K_{ff}\in[0,\infty)$ that yields minimum phase G_{ff-neo}. To do this, let us re-write G_{ff-neo} as

$$G_{ff-neo}=G_{22}{}^{-1}\left(G_{22}G_{21}K_{ff}+\gamma G_{21}{}^2+G_{22}{}^2\right). \tag{2.27}$$

Therefore, the zeros of G_{ff-neo} can be obtained from the roots of the polynomial

$$\mathrm{num}\left(G_{22}G_{21}\right)K_{ff}+\gamma\,\mathrm{num}\left(G_{21}{}^2\right)+\mathrm{num}\left(G_{22}{}^2\right), \tag{2.28}$$

where num is a function that returns the numerator of the transfer function in its argument. Notice that (2.28) is in the standard form for root locus analysis. When $K_{ff}=0$, the zeros of G_{ff-neo} are the same as those of G_{ff}. However, as $K_{ff}\to\infty$, the zeros of G_{ff-neo} approach the zeros of $G_{22}G_{21}$ which, according to (2.11), are always minimum phase. Since the objective functional of (2.12) is quadratic, near optimality is achieved by selecting the smallest value of K_{ff} that makes the poles of G_{ff-neo}^{-1} stable, and provides sufficient damping to prevent ringing due to poorly damped poles [95]. Note that if G_{ff} is minimum phase, $K_{ff}=0$ becomes the default solution for K_{ff}, in which case G_{ff-neo} equals G_{ff}, which is the energy-optimal solution. Given K_{ff}, z_{r1} can be calculated as

$$\begin{aligned}z_{r1}&=G_{11}u_{ff1}+G_{12}u_{ff2}\\&=\left(G_{11}\left(K_{ff}+\beta\right)+G_{12}\right)G_{ff-neo}^{-1}y_d.\end{aligned} \tag{2.29}$$

2.4.2 Design of Energy-Efficient FB Controller

For the regulation problem, $\mathbf{z}_r$ and $\mathbf{u}_{ff}$ are ignored such that $\mathbf{u} = \mathbf{u}_{fb}$ can be written as

$$\mathbf{u}_{fb} = -\mathbf{K}\mathbf{x}, \tag{2.30}$$

where $\mathbf{x}$ is the extended state vector of the HFD defined as

$$\mathbf{x} = \begin{Bmatrix} z_{2i} & z_1 & z_2 & \dot{z}_1 & \dot{z}_2 \end{Bmatrix}^{\mathrm{T}}; \; z_{2i} \triangleq \int_0^{+\infty} z_2 dt. \tag{2.31}$$

Observe that $\mathbf{x}$ contains all four states of the two-mass model of the HFD, plus an additional state, z_{2i}, representing the integral of the table's position signal, added to ensure zero steady-state regulation of the table's position. Let us assume that $\mathbf{K}$ is a full state FB matrix which can be expressed as

$$\mathbf{K} = \begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} \end{bmatrix}, \tag{2.32}$$

where $K_{11} \ldots K_{25}$ are its gains. This means that the general form of $\mathbf{C}_{fb}$, indicated in Figure 2.6, can be written as

$$\mathbf{C}_{fb}(s) = \begin{bmatrix} K_{12} + K_{14}s & K_{11}/s + K_{13} + K_{15}s \\ K_{22} + K_{24}s & K_{21}/s + K_{23} + K_{25}s \end{bmatrix}. \tag{2.33}$$

In the rest of this section, we will show that there always exists some redundancy in the general form of $\mathbf{C}_{fb}$, in terms of achieving the best FB positioning performance. We therefore propose a systematic method for using the available redundancy to maximize actuator efficiency without sacrificing positioning performance.

2.4.2.1 Redundancy in Achieving Best Positioning Performance

Recall that the precise positioning of the table (i.e., z_2) is of primary concern for the HFD. Therefore, the goal of the FB controller is to minimize the effect of disturbance $\mathbf{d}$ on e_{z2}. This objective can be written as

$$\min J_e \triangleq \left\| \left[W_e \mathbf{S}_d \right] \right\| \text{ s.t. stability constraints,} \tag{2.34}$$

where $\|\cdot\|$ represents a suitable norm (e.g., H_2 or H_∞), W_e is a scalar weighting function of s that indicates the frequencies in e_{z2} that are of utmost importance, and $\mathbf{S}_d$ is defined as a scaled version of the closed loop (CL) disturbance transfer function from $\mathbf{d}$ to e_{z2}, given by the expression

$$\mathbf{S}_d = -\begin{bmatrix}0 & 1\end{bmatrix}\left(\mathbf{I}+\mathbf{G}_m\mathbf{C}_{fb}\right)^{-1}\mathbf{G}_m\mathbf{W}_d, \tag{2.35}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{W}_d = \text{diag}(\{d_{1,\max}, d_{2,\max}\})$ is a matrix that scales d_1 and d_2 by their respective maximum values, $d_{1,\max}$ and $d_{2,\max}$, so that their magnitudes are comparable when calculating the norm of (2.34) [62].

Let $\hat{J}_e$ represent the minimum value of J_e obtained using the full state (i.e., unstructured) FB matrix $\mathbf{K}$ of (2.32); it represents the best positioning performance that can be achieved using $\mathbf{C}_{fb}$ defined in (2.33). The corresponding optimal CL disturbance transfer function, $\hat{\mathbf{S}}_d$, can be expressed as

$$\begin{aligned}\hat{\mathbf{S}}_d &= -\frac{1}{\hat{D}_{cl}}\begin{Bmatrix} s\left[\left(c-\hat{K}_{24}\right)s+\left(k-\hat{K}_{22}\right)\right]d_{1,\max} \\ s\left[m_1s^2+\left(b_1+c+\hat{K}_{14}\right)s+\left(k+\hat{K}_{12}\right)\right]d_{2,\max}\end{Bmatrix}^{\mathrm{T}}; \\ \hat{D}_{cl} &\triangleq \hat{a}_{cl,5}s^5+\hat{a}_{cl,4}s^4+\hat{a}_{cl,3}s^3+\hat{a}_{cl,2}s^2+\hat{a}_{cl,1}s+\hat{a}_{cl,0}.\end{aligned} \tag{2.36}$$

The coefficients of $\hat{D}_{cl}$ are functions of $\hat{K}_{11}\ldots\hat{K}_{25}$ and the plant parameters; their expressions are given by following equations.

$$\begin{aligned}
\hat{a}_{cl,5} &= m_1m_2, \quad \hat{a}_{cl,4} = m_1\left(\hat{K}_{25}+b_2+c\right)+m_2\left(\hat{K}_{14}+b_1+c\right), \\
\hat{a}_{cl,3} &= K_{14}\left(b_2+c+\hat{K}_{25}\right)+\hat{K}_{24}\left(c-\hat{K}_{15}\right)+\hat{K}_{15}c+\hat{K}_{25}\left(b_1+c\right)\cdots \\
&\quad +m_1\left(\hat{K}_{23}+k\right)+m_2\left(\hat{K}_{12}+k\right)+\left(b_1+b_2\right)c+b_1b_2, \\
\hat{a}_{cl,2} &= k\left(\hat{K}_{14}+\hat{K}_{15}+\hat{K}_{24}+\hat{K}_{25}+b_1+b_2\right)+b_1\hat{K}_{23}+b_2\hat{K}_{12}\cdots \\
&\quad +m_1\hat{K}_{21}+c\left(\hat{K}_{12}+\hat{K}_{13}+\hat{K}_{22}+\hat{K}_{23}\right)+\hat{K}_{12}\hat{K}_{25}-\hat{K}_{13}\hat{K}_{24}+\hat{K}_{14}\hat{K}_{23}-\hat{K}_{15}\hat{K}_{22}, \\
\hat{a}_{cl,1} &= k\left(\hat{K}_{12}+\hat{K}_{13}+\hat{K}_{22}+\hat{K}_{23}\right)+c\left(\hat{K}_{11}+\hat{K}_{21}\right)+b_1\hat{K}_{21}\cdots \\
&\quad -\hat{K}_{11}\hat{K}_{24}+\hat{K}_{12}\hat{K}_{23}-\hat{K}_{13}\hat{K}_{22}+\hat{K}_{14}\hat{K}_{21}, \\
\hat{a}_{cl,0} &= k\left(\hat{K}_{11}+\hat{K}_{21}\right)-\hat{K}_{11}\hat{K}_{22}+\hat{K}_{12}\hat{K}_{21}.
\end{aligned} \tag{2.37}$$

Note that the $\hat{}$ accent is used to denote optimality. A sufficient (but not necessary) condition to achieve $\hat{J}_e$, irrespective of the type of norm or weighting function (W_e) used in (2.34), is to match the numerator and denominator coefficients of $\hat{\mathbf{S}}_d$. Matching these coefficients mathematically means that the gain matrix $\mathbf{K}$, with 10 degrees of freedom (i.e., 10 gains to be determined), should satisfy 9 linear/bilinear equations; 4 for the numerators of $\hat{\mathbf{S}}_d$ and 5 for their common denominator. Therefore, at least 1 gain in $\mathbf{K}$ is redundant. As discussed in Section 2.3.2, it is not

uncommon for the condition $d_{1,\max} << d_{2,\max}$ to hold since cutting force (d_2) is dominant and cannot be compensated. Even when the Coulomb friction (d_1) of the bearings is non-negligible, it can be measured and cancelled out reliably through feedforward friction compensation so that the FB controller does not have to deal with it. One can therefore conveniently consider $d_{1,\max} \approx 0$, such that $\hat{\mathbf{S}}_d$ is reduced to a scalar transfer function which can be matched using 7 gains. This leaves at least 3 redundant gains to achieve $\hat{J}_e$.

2.4.2.2 Use of Redundant Gains for Optimizing Efficiency

The set of all FB control inputs, $\mathbf{u}_{fb}$, that yield the same e_{z2} under the influence of a given disturbance input, $\mathbf{d}$, must satisfy the relationship

$$z_2 = G_{21}\left(u_{fb1} + d_1\right) + G_{22}\left(u_{fb2} + d_2\right) \Rightarrow G_{21}\delta u_{fb1} + G_{22}\delta u_{fb2} = 0, \tag{2.38}$$

where $\delta d_1 = \delta d_2 = \delta z_2 = 0$ because they are each specified, even if unknown, functions of time. Accordingly, similar to the FF design, the $\mathbf{u}_{fb}$ is expected to approach the causal approximation of the optimal ratio β, i.e.

$$u_{fb1} = \underbrace{\gamma G_{22}^{-1} G_{21}}_{\beta} u_{fb2}. \tag{2.39}$$

The implication of (2.39) is that to satisfy a given e_{z2} requirement at maximum efficiency, the control inputs are bound relationship given by β. This result is very powerful and can be used to directly or indirectly determine a FB controller that achieves the best performance, $\hat{J}_e$, more efficiently than the optimal full state FB controller.

a. Energy-Efficient FB Controller Design – Direct Approach

The direct approach seeks to structure $\mathbf{K}$ such that (2.39) is satisfied (as much as possible) without compromising $\hat{J}_e$. To do this, note that $\mathbf{u}_{fb}$ can be written as

$$\begin{aligned} &\mathbf{u}_{fb} = \mathbf{S}_u \bar{\mathbf{d}}, \\ &\bar{\mathbf{d}} \triangleq \left\{\bar{d}_1 \quad \bar{d}_2\right\}^{\mathrm{T}}, \quad \mathbf{S}_u \triangleq -\mathbf{C}_{fb}\left(\mathbf{I} + \mathbf{G}_m \mathbf{C}_{fb}\right)^{-1} \mathbf{G}_m \mathbf{W}_d, \end{aligned} \tag{2.40}$$

where $\bar{d}_1, \bar{d}_2 \in [-1, 1]$ represent the scaled values of d_1 and d_2, respectively. Combining (2.40) with (2.39), we get the relationship that must be satisfied by the elements of the re-structured $\mathbf{K}$ in order to maximize efficiency; it is given by

$$
\begin{aligned}
&u_{fb1} - \beta u_{fb2} = 0 \\
&\Rightarrow \begin{cases} \Gamma_1 \triangleq \left(\mathbf{K}_1{}^{\mathrm{T}}\mathbf{Q}_s\mathbf{K}_2 + \boldsymbol{\chi}_2{}^{\mathrm{T}}\mathbf{K}_1 - \beta\boldsymbol{\chi}_2{}^{\mathrm{T}}\mathbf{K}_2\right) d_{1\max} = 0; \\ \Gamma_2 \triangleq \left(\boldsymbol{\chi}_1{}^{\mathrm{T}}\mathbf{K}_1 - \beta\left(\mathbf{K}_1{}^{\mathrm{T}}\mathbf{Q}_s\mathbf{K}_2 + \boldsymbol{\chi}_1{}^{\mathrm{T}}\mathbf{K}_2\right)\right) d_{2\max} = 0, \end{cases}
\end{aligned} \tag{2.41}
$$

where

$$
\boldsymbol{\chi}_1 = \begin{bmatrix} m_1 s^2 + (c + b_1)s + k \\ cs^2 + ks \\ m_1 s^3 + (c + b_1)s^2 + ks \\ cs^3 + ks^2 \\ m_1 s^4 + (c + b_1)s^3 + ks^2 \end{bmatrix}, \boldsymbol{\chi}_2 = \begin{bmatrix} cs + k \\ m_2 s^3 + (c + b_2)s^2 + ks \\ cs^2 + ks \\ m_2 s^4 + (c + b_2)s^3 + ks^2 \\ cs^3 + ks^2 \end{bmatrix}
$$

$$
\mathbf{Q}_s = \begin{bmatrix} 0 & -1 & 0 & -s & 0 \\ 1 & 0 & s & 0 & s^2 \\ 0 & -s & 0 & -s^2 & 0 \\ s & 0 & s^2 & 0 & s^3 \\ 0 & -s^2 & 0 & -s^3 & 0 \end{bmatrix}, \tag{2.42}
$$

$$
\mathbf{K}_i = \begin{bmatrix} K_{i1} & K_{i2} & K_{i3} & K_{i4} & K_{i5} \end{bmatrix}^{\mathrm{T}} \quad (i = 1, 2).
$$

Substituting $s = j\omega$ into $\Gamma_1(s)$ and $\Gamma_2(s)$ in (2.41) (where ω represents the frequency content of **d**), we get a pair of complex-valued equations which are bilinear with respect to the elements of **K**. The real and imaginary parts of both equations must equal zero at every ω contained in **d** in order to attain optimal efficiency. Obviously, **K** may not have enough redundant gains to satisfy the equations for every ω contained in **d**. Therefore, one can maximize efficiency using the available redundancy by determining the **K** that minimizes the least-squares objective

$$
\min J_{u\_d} \triangleq \boldsymbol{\Gamma}(j\omega)^{\mathrm{H}}\boldsymbol{\Gamma}(j\omega) \text{ s.t. } \begin{cases} \text{stability constraints} \\ \left\| \left[W_e \mathbf{S}_d \right] \right\| = \hat{J}_e \end{cases}, \tag{2.43}
$$

where $\boldsymbol{\Gamma} = [\Gamma_1(j\omega), \Gamma_2(j\omega)]^{\mathrm{T}}$ and the superscript H represents the Hermitian transpose. If needed, the least squares objective in (2.43) can be weighted to emphasize some frequencies over others.

b. Energy-Efficient FB Controller Design − Indirect Approach

The traditional approach for enforcing an energy-efficient structure in **K** is to optimize for energy efficiency with positioning performance constraints; i.e.,

$$
\min J_{u-in} \triangleq \left\| \left[\mathbf{W}_u \mathbf{S}_u \right] \right\| \text{ s.t. } \begin{cases} \text{stability constraints} \\ \left\| \left[W_e \mathbf{S}_d \right] \right\| = \hat{J}_e \end{cases}, \tag{2.44}
$$

where $\mathbf{S}_u$ is the transfer function matrix from input $\bar{\mathbf{d}}$ to output $\mathbf{u}_{fb}$ as defined in (2.40), $\mathbf{W}_u$ is a diagonal weighting matrix whose elements reflect knowledge about the relative efficiencies of the actuators. The addition of the constraint related to $\hat{J}_e$ indicates the knowledge that there is redundancy in $\mathbf{K}$ such that the same performance can be achieved in a more efficient manner. The weighting matrix, $\mathbf{W}_u$, is very important because it indirectly determines the structure of $\mathbf{K}$ used in the optimization; a poor choice of $\mathbf{W}_u$ could lead to a sub-optimal structure for $\mathbf{K}$, which cannot attain the best energy efficiency. In practice, $\mathbf{W}_u$ is typically chosen by intuition or trial and error. However, the optimum structure for $\mathbf{K}$ can be obtained, in theory, by selecting $\mathbf{W}_u$ = diag([1, β]), thereby eliminating the sub-optimality that could be created by a poor choice of $\mathbf{W}_u$. Notice that the setup of (2.44) is very similar to (2.43). The only difference is that the objective in (2.43) is designed to directly enforce (as much as possible) the optimal relationship between u_{fb1} and u_{fb2}. On the other hand, (2.44) enforces the optimal relationship indirectly by minimizing the feedback control forces using a suitable norm of $\mathbf{S}_u$, weighted by $\mathbf{W}_u$. A demonstration of the direct and indirect approaches is provided in Section 2.5.

It must be noted before leaving this section that the assumption of a static FB controller in (2.30) is arbitrary. It has been made purely based on its relevance to the desired control method for the HFD (which seeks to avoid high-order controllers). The methods proposed in this section can also be applied to the design of dynamic controllers (e.g., controllers with observers). Moreover, even though the proposed FF and FB controller design methods are presented in the context of the HFD described in Section 2.3, they can be applied more broadly to other over-actuated systems. The methods are valid as long as the plant is dual input, controllable and minimum phase with positioning performance defined on one of its outputs (e.g., z_2 in this work).

2.4.3 Stability Constraints and Higher-Order Dynamics

Thus far, only the low-order dynamics of the HFD, which is assumed to be accurately represented by its two-mass model, $\mathbf{G}_m$, has been considered. Note, however, that stability constraints cannot be enforced in (2.34), (2.43) or (2.44) based on a CL system determined using $\mathbf{G}_m$. This is because $\mathbf{G}_m$ has no NMP zeros, hence the gains of the FB controller can be infinitely high without causing any stability issues [62]. To enforce stability constraints, the actual plant dynamics $\mathbf{G}_a$, which includes un-modeled higher-order dynamics, must be considered. Let us

assume that $\mathbf{G}_a$ is available in the form of a plant model or a frequency response function (FRF) measured at discrete frequencies, ω. Let N_p denote the number of unstable poles in the system's loop transfer function given by $\mathbf{L} = \mathbf{G}_a\mathbf{C}_{fb}$. According to the Generalized Nyquist Theorem [62], the CL system is stable if and only if the Nyquist curve of $\det(\mathbf{I} + \mathbf{L})$ makes N_p counter-clockwise encirclements of the origin without passing the origin.

Most often, it is desired to enforce robust stability, meaning that there are sufficient margins of stability. This is achieved by confining the maximum singular values $\bar{\sigma}$ of the CL sensitivity function, $\mathbf{S} = (\mathbf{I} + \mathbf{L})^{-1}$, below a specified threshold $\bar{\sigma}_{\max}$ at frequencies above the CL bandwidth, ω_{BW}, where instability is likely to occur [62]; i.e.

$$\underbrace{\bar{\sigma}\left([\mathbf{I}+\mathbf{L}(j\omega)]^{-1}\right)}_{\omega>\omega_{BW}} \le \bar{\sigma}_{\max}. \tag{2.45}$$

2.5 Simulations and Experiments

Simulation and experiment results are presented in this section to demonstrate the effectiveness of the proposed FF and FB controller design methods. Simulations are conducted to evaluate the tracking performance and energy efficiency of the near-energy-optimal (NEO) FF control method proposed in Section 2.4.1. FB control is not considered in the simulations because the plant is assumed to be perfectly described by its two-mass model, $\mathbf{G}_m$, and no disturbances are introduced. Under the conditions described, the proposed NEO FF controller is able to achieve perfect tracking with near optimal or optimal efficiency, depending on whether G_{ff} defined in (2.21) has NMP zeros or not. Evaluating G_{ff} using the reported parameters of the HFD reveals that it has a complex conjugate pair of NMP zeros located at $s = 196 \pm 314j$. Therefore, the gain K_{ff} must be selected to ensure a stable inverse of G_{ff-neo} while providing near energy optimality. Figure 2.7 shows the root locus of G_{ff-neo}^{-1} as K_{ff} is varied in the interval $[0, \infty)$. It is found that the unstable pole pair in G_{ff-neo}^{-1} (corresponding to the NMP zero pair in G_{ff-neo}) becomes stable for values of $K_{ff} \ge 30.3$. Note that the point marked $K_{ff} = 1000$ has been highlighted on purpose in the figure; it represents a value of K_{ff} which is utilized for comparative simulations later in this section.

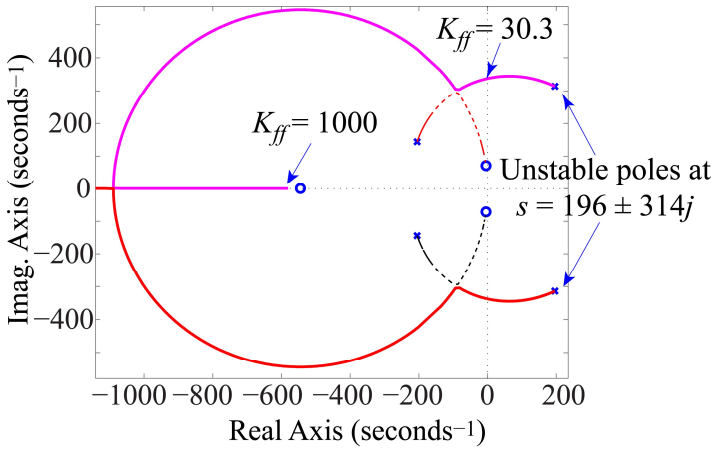

Figure 2.7: Root Locus of G_{ff-neo}^{-1} as a Function of K_{ff}

For tracking tests, a desired position trajectory, y_d, of stroke 15 mm, with kinematic limits of 100 mm/s, 1×10^4 mm/s^2, 1.5×10^6 mm/s^3 and 5×10^8 mm/s^4 for velocity, acceleration, jerk and snap, respectively, is used. Its velocity profile is shown in Figure 2.8 (a); its position, acceleration, jerk and snap profiles are omitted due to space limitations. A FF technique called rigid body FF (RB FF) [98] is used as the benchmark for evaluating the proposed NEO FF controller. RB FF is generated by specifying $z_{r1} = z_{r2} = y_d$ such that

$$\mathbf{u}_{ff} = \mathbf{G}_m^{-1}\begin{bmatrix} y_d & y_d \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} m_1 s^2 + b_1 s \\ m_2 s^2 + b_2 s \end{bmatrix} y_d. \tag{2.46}$$

The implication of (2.46) is that in RB FF each actuator is given the responsibility to move the mass to which it is attached, such that the two-mass system moves as one rigid body. It is very elegant and it guarantees perfect tracking. Note that the zero phase error FF approximation of (2.23) is not selected as the benchmark because it does not meet the perfect tracking requirement stipulated for FF control.

Figure 2.8 (b)-(d) and Table II compare the tracking error (e_{z2}) at the table, and the power/heat wasted by each actuator using RB FF and the proposed NEO FF, with K_{ff} = 30.3 and K_{ff} = 1000. All three controllers achieve perfect tracking, as expected. Notice, however, that RB FF depends heavily on the inefficient LMD. As a result, it is 71.5% less efficient than NEO FF, using K_{ff} = 1000. In theory, NEO FF with K_{ff} = 30.3 should be the most efficient. However, notice from the figure that K_{ff} = 30.3 results in undesirable ringing of the control signals because it gives rise to poorly damped poles in G_{ff-neo}^{-1}. The near-optimal energy formulation of (2.26) does not recognize the effect of poorly damped poles/zeros (much like the near-optimal energy

formulation of (2.21) misses the effect of NMP poles/zeros). Therefore, the onus falls on the controller designer to select a value of K_{ff} that is high enough to avoid poorly damped poles in G_{ff-neo}^{-1}. Figure 2.9 plots the total heat energy of NEO FF in tracking the reference trajectory of Figure 2.8 (a) using values of K_{ff} ranging from 30.3 to 1000. Notice that the energy is initially very high but drops rapidly as the value of K_{ff} increases and the poles become more damped. For this particular case, the energy continues to reduce as K_{ff} is increased but the rate of decrease becomes very low after about K_{ff} = 50. The decrease in energy continues until K_{ff} = 1000 and beyond, but the difference between the energy at K_{ff} = 1000 and that at K_{ff} = 10,000 is just 0.03 J. It must be noted that the relationship between K_{ff} and energy highly depends on the system and desired trajectory.

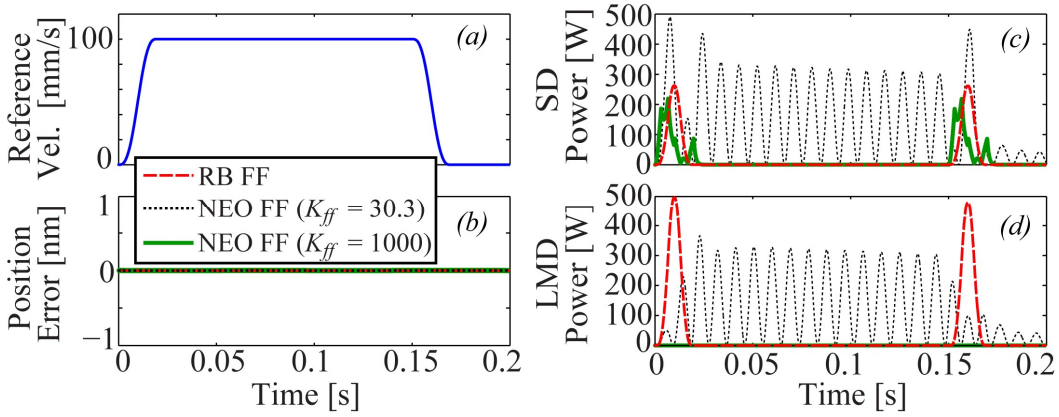

Figure 2.8: Reference Velocity Profile and Comparison of Tracking Errors and Actuator Power of FF Controllers (Simulation)

For instance, Figure 2.10 shows the energy vs. K_{ff} plot generated based on another desired position trajectory, y_d, of stroke 15 mm, with kinematic limits of 100 mm/s, 1×10^3 mm/s^2, 5×10^4 mm/s^3 and 1×10^7 mm/s^4 for velocity, acceleration, jerk and snap, respectively. Just as in Figure 2.9, there is initially a sharp drop in energy as K_{ff} is increased and ringing is reduced. However, in Figure 2.10 there are three local minima at K_{ff} = 66, 200 and 500; beyond K_{ff} = 500, the energy begins to increase slightly but steadily. The difference in the energy vs. K_{ff} profiles of Figure 2.9 and Figure 2.10 is because the trajectory used in Figure 2.10 is less aggressive than that used in Figure 2.9. Therefore, it has is less ringing due to poorly damped poles, and its efficiency is hurt rather than helped by having very high values of K_{ff}. This is because, in theory (i.e., in the absence of ringing), efficiency should decrease as K_{ff} increases. Nonetheless, just as in Figure 2.9, as long as K_{ff} is high enough to provide sufficiently damped poles in G_{ff-neo}^{-1}, the

efficiency of the proposed NEO FF is observed to remain near its optimal value, irrespective of the value of K_{ff}. This arises from the fact that larger K_{ff} encourages the use of more energy-efficient SD.

Table 2.2: Comparison of Tracking Errors and Energy Efficiency of FF Controllers (Simulation)

	RB FF	NEO FF (K_{ff} = 30.3)	NEO FF (K_{ff} = 1000)
Max Tracking Error [nm]	0.000	0.000	0.000
Heat Energy SD [J]	3.967	27.918	3.242
Heat Energy LMD [J]	7.398	24.358	0.001
Total Heat Energy [J]	11.365	52.276	3.243

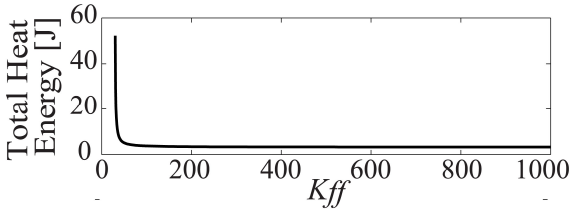

Figure 2.9: Total Heat Energy of Proposed NEO FF as a Function of K_{ff} (based on Reference Trajectory of Figure 2.8 (a)) – Simulation

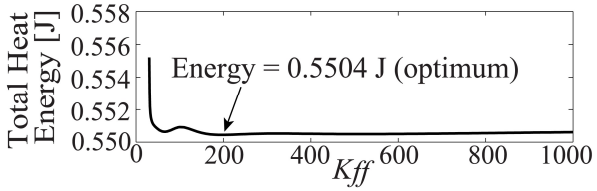

Figure 2.10: Total Heat Energy of Proposed NEO FF as a Function of K_{ff} (Less Aggressive Trajectory)

2.5.1 Evaluation of FB Controllers using Simulations

Simulations are conducted to evaluate EE direct and EE indirect, the direct and indirect methods for using control ratio β derived in Section 2.4.2 for FB controller design. The benchmark is selected as the full (i.e., unstructured) state FB controller discussed in Section 2.4.2. The practical scenario where $d_{1,\max} = 0$ and $d_{2,\max} = 400$ N (which is the cutting force limit of the HFD of Figure 2.1 [35]) is assumed. MATLAB's *hinfstruct* command (which performs H_∞ optimization based on a structured gain matrix) is used to determine the $\hat{\mathbf{S}}_d$ that minimizes the

objective of (2.34) using the H_∞ norm. A weighting function, W_e, suggested by Skogestad and Postlethwaite [62] is adopted; it is given by

$$W_e = \frac{\left(s/M_w^{1/n_w} + \omega_B\right)^{n_w}}{\left(s + \omega_B A_w^{1/n_w}\right)^{n_w}}, \tag{2.47}$$

Equation (2.47) represents a filter of order n_w, designed to weight very low frequencies by a factor of M_w/A_w compared to very high frequencies (i.e., frequencies much higher than ω_B), so as to enforce integral action. We select $n_w = 2$, $M_w = 2$, $A_w = 5 \times 10^{-6}$ and $\omega_B = 200\pi$ rad/s (indicating that we want integral action to taper off beyond 100 Hz).

The HFD is assumed to be perfectly modeled by $\mathbf{G}_m$, and the table-side disturbance force, d_2, is assumed to be a harmonic cutting force at 60 Hz, with an additional DC component; it is given by

$$d_2(t) = 200 + 200\sin(120\pi t) \ \text{[N]}. \tag{2.48}$$

However, as discussed in Section 2.4.3, applying the optimization of (2.34) to $\mathbf{G}_m$ could yield infinitely large gains. Therefore, a 7$^{\text{th}}$ order Butterworth low-pass filter, with a cutoff frequency of 1000 Hz, is applied to each element of $\mathbf{G}_m$ before carrying out the H_∞ optimization in order to have finite stability limits. The default stability settings of MATLAB's *hinfstruct* command are maintained for the optimization. The resulting $\hat{\mathbf{S}}_d(1,2)$ (scaled by $1/d_{2,\max}$) is plotted in Figure 2.11 (a) and the elements of the optimal full state FB gain matrix (i.e., $\widehat{\mathbf{K}}$) are given in Table III. Note that only $\hat{\mathbf{S}}_d(1,2)$ is shown in the figure because $\mathbf{S}_d(1,1) = 0$ based on the assumption that $d_{1,\max} = 0$.

Figure 2.11 (b) shows the optimal transfer function relationship, β, for the HFD. It requires the magnitude of the ratio u_{fb1}/u_{fb2} to be 50.3 dB at $\omega = 0$, and to be 23.5 dB at $\omega = 120\pi$ rad/s. In employing β for energy efficiency optimization of $\mathbf{K}$, we have chosen to use $\hat{\mathbf{S}}_d$ instead of its infinity norm, $\hat{J}_e$, as the performance constraint in (2.43) and (2.44). This is done in order to ensure that the exact same regulation performance as the optimal full state FB controller is achieved at all frequencies. Therefore, using (2.36), we derive 7 constraint equations to ensure that the poles and zeros of $\hat{\mathbf{S}}_d(1,2)$ are matched with 7 out of the 10 gains of $\mathbf{K}$, leaving us with 3 redundant gains. To obtain the energy-efficient controller using the direct method (i.e., EE direct), one of the remaining 3 gains is used to perfectly enforce (2.41) at $\omega = 0$. In theory, the last two gains could be used to satisfy the real and imaginary parts of (2.41) at $\omega = 120\pi$ rad/s,

but the equations are highly nonlinear and difficult to solve analytically. Therefore, the remaining two gains are used to minimize $J_{u\text{-}d}$ (i.e., (2.43)) at $\omega = 120\pi$ rad/s. The elements of the resulting gain matrix are given in Table III.

To determine the energy-efficient controller using the indirect method (i.e., EE indirect), MATLAB's *hinfstruct* command is used to minimize the H_∞ norm of (2.44). A weighting function given by $\mathbf{W}_u = W_\omega \text{diag}([1, \beta])$ is selected, where W_ω is a weighting filter designed to emphasize frequencies $\omega = 0$ and $\omega = 120\pi$ rad/s contained in d_2. It is defined as

$$W_\omega = W_e \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \tag{2.49}$$

Note that W_e is the exact same filter defined in (2.47); it is used in W_ω to place a higher weight on frequencies close to $\omega = 0$. The second-order filter multiplying W_e is designed with $\omega_n = 120\pi$ rad/s and $\zeta = 8.2\times10^{-6}$, such that its magnitude at resonance is equal to $W_e(0)$. The constraint equations that were used to match the poles and zeros of $\hat{\mathbf{S}}_d(1,2)$ for the EE direct method are used to structure $\mathbf{K}$ within MATLAB's *hinfstruct* function so that $\hat{\mathbf{S}}_d(1,2)$ is also matched perfectly in EE indirect. The elements of the resulting FB gain matrix are reported in Table III.

Figure 2.11 (a) compares the magnitudes of $\mathbf{S}_d(1,2)/d_{2,\max}$ for the EE direct and EE indirect controllers to the original $\hat{\mathbf{S}}_d(1,2)/d_{2,\max}$ determined based on the optimal state FB controller; the plots are coincident, indicating that all three methods achieve the best positioning performance at all frequencies. Figure 2.11 (c) and (d) respectively compare $\mathbf{S}_u(1,2)/d_{2,\max}$ and $\mathbf{S}_u(2,2)/d_{2,\max}$ for the three methods. Notice that $\mathbf{S}_u(1,2)/\mathbf{S}_u(2,2) = 50.3$ dB for the EE direct controller, meaning that it perfectly matches β at $\omega = 0$. However, at $\omega = 120\pi$ rad/s, it achieves a $\mathbf{S}_u(1,2)/\mathbf{S}_u(2,2)$ ratio of 32 dB instead of the 23.5 dB stipulated by β, due to errors in the least squares solution applied at that frequency. The EE indirect controller does not perform as well as the EE direct in terms of matching β. This is because, in determining the gain matrix for the EE indirect controller, the *hinfstruct* command was initialized with the gain matrix of the optimal full state FB controller (i.e., $\widehat{\mathbf{K}}$). Comparing the elements of the full state FB controller's gain matrix to those of the EE indirect controller in Table III, one can conclude that the gradient-based algorithm used in *hinfstruct* may have gotten stuck in a local minimum close to its initial value, thus preventing it from reaching a better solution. This is a well-known limitation of gradient-based methods when performing non-convex optimizations. It could be mitigated by testing various initial values in hopes of finding other local minima. Note that the direct method

was also initialized with the elements of the full state FB matrix, when performing the nonlinear least-squares optimization of (2.43). Therefore, its solution could also get stuck in local minima around its starting point, but it appears from Table III that its gains are quite different from those of the full state FB matrix.

The expression

$$\sum_{i=1}^{2}\frac{1}{K_{mi}^2}\left\{\left(200\left|\frac{\mathbf{S}_u^{\mathrm{T}}(2,i)}{d_{2.\max}}\right|_{\omega=0}\right)^2+\frac{1}{2}\left(200\left|\frac{\mathbf{S}_u^{\mathrm{T}}(2,i)}{d_{2.\max}}\right|_{\omega=120\pi}\right)^2\right\}, \tag{2.50}$$

representing the sum of the average power at $\omega = 0$ and $\omega = 120\pi$ rad/s of each actuator's steady-state response to $d_2(t)$ given in (2.48), is used to calculate the average power dissipated in heat by each FB controller. The results are shown in Table 2.. EE direct and EE indirect are respectively 60% and 33% more efficient than the full state FB method. Figure 2.12 compares the ratio of the average power of the EE direct and EE indirect based controllers, relative to that of the full state FB controller, using a disturbance input of $d_2(t) = \sin(2\pi ft)$. The EE direct controller is much more efficient than the full controller from $f = 0$ through $f = 44$ Hz, after which it becomes less efficient. Note that its relative efficiency is 8 dB worse at $f = 60$ Hz, due to the aforementioned errors in the least square solution. The EE indirect controller, on the other hand, is more efficient than the full controller at all frequencies, but only by a very small (0.6 dB) margin (except at $f = 0$ and $f = 60$ Hz where it shows slightly better relative efficiencies of 4.2 dB and 0.8 dB, respectively).

Table 2.3: Comparison of Elements of FB Gain Matrices Calculated using the Full State FB, EE Direct and EE Indirect Methods (Simulation)

	K_{11}	K_{12}	K_{13}	K_{14}	K_{15}
Full State FB	-8.2×10^{5}	-3.1×10^{6}	-3.1×10^{8}	-5.8×10^{3}	-9.3×10^{5}
EE direct	$+9.2\times10^{5}$	-3.1×10^{6}	$+3.8\times10^{9}$	-5.8×10^{3}	$+8.5\times10^{6}$
EE indirect	-4.5×10^{5}	-3.1×10^{6}	-1.8×10^{8}	-5.8×10^{3}	-6.2×10^{5}
	K_{21}	K_{22}	K_{23}	K_{24}	K_{25}
Full State FB	-2.4×10^{10}	$+4.2\times10^{7}$	$+5.2\times10^{7}$	$+9.5\times10^{4}$	$+1.6\times10^{5}$
EE direct	-5.2×10^{7}	-7.6×10^{4}	$+1.1\times10^{8}$	-8.9×10^{1}	$+1.6\times10^{5}$
EE indirect	-2.3×10^{10}	$+7.0\times10^{7}$	$+6.8\times10^{7}$	$+1.3\times10^{5}$	$+1.6\times10^{5}$

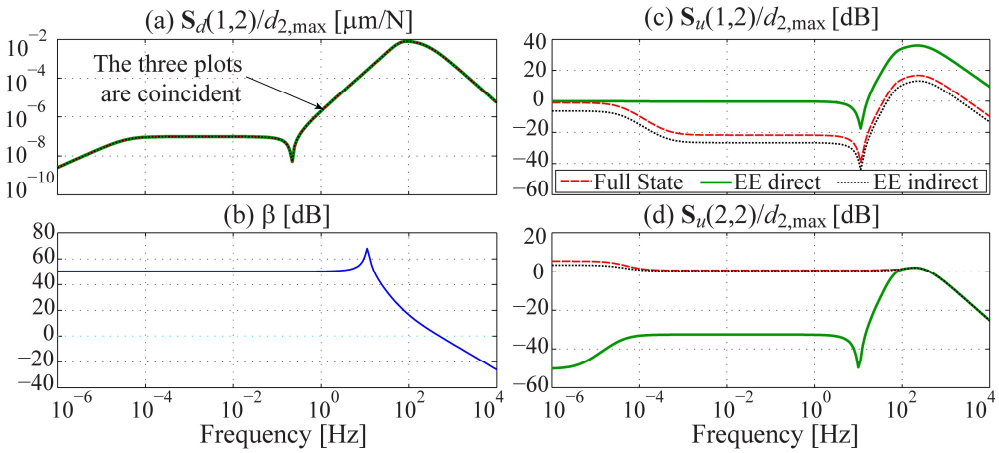

Figure 2.11: Comparison of Positioning Performance and Control Effort of the Full State FB and the Energy-Efficient FB Controllers Designed via the EE Direct and EE Indirect Approaches (Simulation)

Table 2.4: Comparison of Energy Efficiency of FB Controllers Based on Equation (2.50) (Simulation)

Average Power Dissipated in Heat	Full State FB	EE direct	EE indirect
SD [W]	1.36	125.80	0.51
LMD [W]	381.80	26.31	254.44
Total [W]	383.16	152.11	254.95

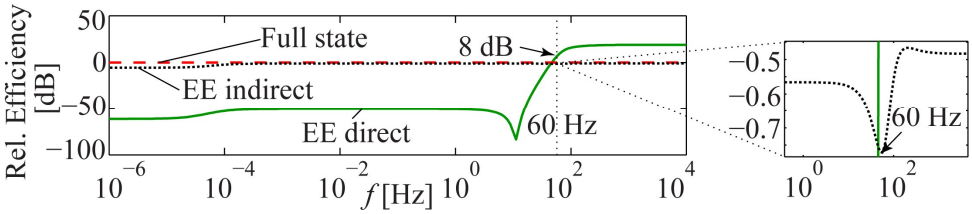

Figure 2.12: Comparison of Efficiencies of Controllers Designed using the EE Direct and EE Indirect Methods (Relative to Efficiency of Full State FB Controller) – Simulation

2.5.2 Machining Experiments

Machining experiments are conducted to evaluate the performance and energy efficiency of the proposed controller design methods. To perform the experiments, the HFD prototype of Figure 2.1 is mounted on the x axis of a FADAL VMC 4020 3-axis milling machine to cut a 15

mm long, 1 mm deep slot in an AISI 1018 steel workpiece. Table V summarizes the key cutting parameters for the operation. A spindle speed of 3600 rpm (i.e., 60 Hz) is selected for the cutting tests. To maintain a reasonable feed per tooth, the desired position trajectory, y_d, is prescribed to travel at a maximum speed of 300 mm/min (i.e., 5 mm/s), with acceleration, jerk and snap limits of 62.5 mm/s^2, 3125 mm/s^3 and 6.25×10^5 mm/s^4, respectively. At such low speeds and accelerations, the theoretical energy contributions of RB FF and NEO FF are negligible (< 6 mJ), and are indistinguishable from noise in the control signals. Consequently, the effect of FF is not considered in the machining tests.

Table 2.5: Cutting Parameters

Spindle speed	3600 rpm
Tool	3/8” dia. HSS end mill
Number of flutes	4
Feed per tooth	0.02 mm/tooth
Feed rate	300 mm/min
Lubrication	None

The optimal full state FB controller is designed by optimizing the H_∞ norm of (2.34) with the actual (i.e., measured) plant dynamics, $\mathbf{G}_a(j\omega)$, substituted for $\mathbf{G}_m$. As mentioned in Section 2.4, the Coulomb friction in the bearings of the rotary motor can be easily cancelled by feedforward action. Therefore, $d_{1,\max} = 0$ is assumed, and $d_{2,\max}$ is selected as 400 N. Notch filters are placed at the 847 Hz resonance peak appearing in $\mathbf{G}_a(1,1)$ (in Figure 2.3), and at the 400 Hz and 487 Hz resonance peaks appearing in $\mathbf{G}_a(2,2)$ in order to increase stability margins. Robust stability constraints are enforced by limiting the maximum singular values of the sensitivity function for frequencies higher than $\omega_{BW} = 628$ rad/s (100 Hz) to $\bar{\sigma}_{\max} = 2.65$.

MATLAB’s *hinfstruct* command, which was used in the simulations, could not be employed for carrying out the optimization in (2.34) using $\mathbf{G}_a$ instead of $\mathbf{G}_m$. This is because $\mathbf{G}_a$ is given as a FRF at discrete frequencies, not as a transfer function modeled in s-domain. Therefore, a Particle Swarm Optimization (PSO) Toolbox in MATLAB [99] is used to carry out

the H_∞ optimization, based on a technique similar to the one reported in [100]. The elements of the resulting optimal full state FB gain matrix are provided in Table 2.6.

Feed forces in milling typically have a DC portion, as well as harmonics at the spindle and tooth passing frequencies [101]. Therefore, for the EE direct controller, we follow the exact same procedure as used in Section 2.5.1. The elements of $\mathbf{K}$ are constrained to match $\hat{\mathbf{S}}_d(1,2)$ using (2.36), and the 3 redundant gains are used for enforcing the β constraint at $\omega = 0$ and $\omega = 120\pi$ rad/s, representing the lowest harmonic in d_2. For the EE indirect method, MATLAB's PSO toolbox is used to carry out the optimization of (2.44) based on $\mathbf{G}_a$, with the elements of $\widehat{\mathbf{K}}$ as its initial values; constraints are added to match $\hat{\mathbf{S}}_d(1,2)$ using (2.36), following the exact same procedure as in the simulations reported in Section 2.5.1. The gains of the EE direct and indirect controllers are reported in Table 2.6.

Table 2.6: Comparison of Elements of FB Gain Matrices Calculated using the Full State FB, EE Direct and EE Indirect Methods (Experiments)

	K_{11}	K_{12}	K_{13}	K_{14}	K_{15}
Full State FB	4.65×10^{7}	4.92×10^{5}	1.06×10^{8}	1.29×10^{4}	6.50×10^{5}
EE Direct	5.48×10^{7}	4.92×10^{5}	8.90×10^{7}	1.29×10^{4}	5.93×10^{5}
EE Indirect	4.80×10^{7}	4.92×10^{5}	8.57×10^{7}	1.29×10^{4}	5.46×10^{5}
	K_{21}	K_{22}	K_{23}	K_{24}	K_{25}
Full State FB	1.52×10^{7}	6.15×10^{5}	1.02×10^{7}	4.85×10^{2}	4.30×10^{4}
EE Direct	2.54×10^{6}	1.59×10^{5}	1.12×10^{7}	1.02×10^{3}	4.30×10^{4}
EE Indirect	6.16×10^{6}	1.57×10^{4}	1.06×10^{7}	-4.25×10^{1}	4.30×10^{4}

Figure 2.13 (a) compares $\mathbf{S}_d(1,2)/d_{2,\max}$ generated using the EE direct and EE indirect controllers to $\hat{\mathbf{S}}_d(1,2)/d_{2,\max}$, generated using the full state FB controller. An almost perfect match is obtained, indicating that the positioning performances of all three controllers are virtually the same. Figure 2.13 (b) compares the maximum singular values of the CL sensitivity function for the three controllers; they are very similar and all satisfy $\bar{\sigma}_{\max} \leq 2.65$ for frequencies above 100 Hz. Figure 2.13 (c) and (d) respectively compare $\mathbf{S}_u(1,2)/d_{2,\max}$ and $\mathbf{S}_u(2,2)/d_{2,\max}$ of all three controllers. In both figures, the two-mass model, $\mathbf{G}_m$, has been used to generate data below 10 Hz which are not contained in the measured FRF, $\mathbf{G}_a$. Again, EE direct is able to match the β

criterion at $\omega = 0$. At $\omega = 120\pi$ rad/s, $\mathbf{S}_u(1,2)/\mathbf{S}_u(2,2) = 21.7$ dB, which is much closer to the target value of 23.5 dB than obtained in the simulations. A plot of the efficiencies of the EE direct and indirect controllers is provided in Figure 2.14, following the same procedure used in the simulation to generate Figure 2.12. The EE direct controller is more efficient than the full state FB controller for frequencies ranging from $f = 0$ to $f = 170$ Hz; and at its worst, it is only 0.3 dB less efficient than the full state FB controller (at 848 Hz). The EE indirect controller is 3.1 dB and 2.8 dB more efficient than the full state FB controller at $f = 0$ and $f = 60$ Hz, respectively; at its worst, it is 3.8 dB less efficient than the full state FB controller (at 0.83 Hz).

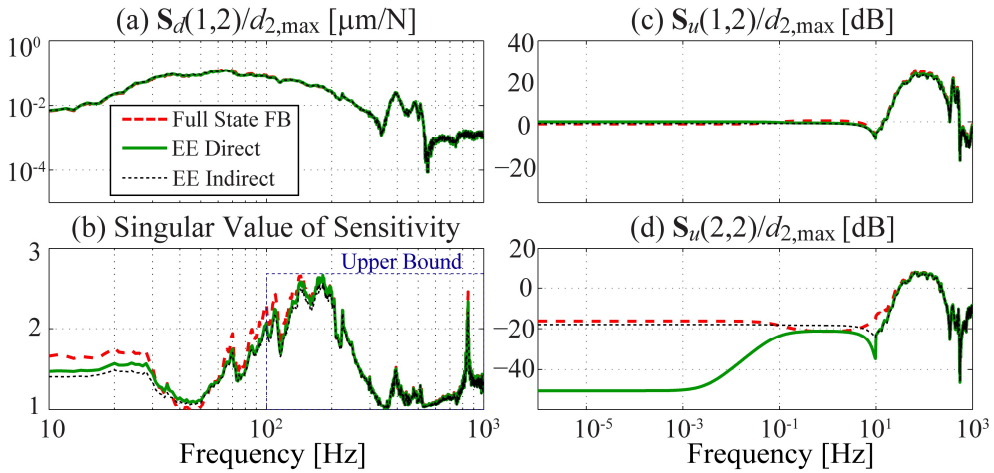

Figure 2.13: Comparison of Positioning Performance, Robustness and Control Effort of the Full State FB and the Energy-Efficient FB Controllers Designed via EE Direct and EE Indirect Approaches (Experiments)

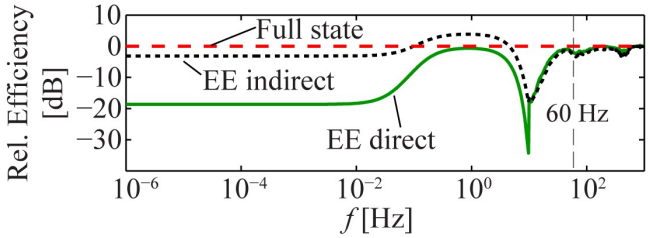

Figure 2.14: Comparison of Efficiencies of Controllers designed using the EE Direct and EE Indirect Methods (Relative to Efficiency of Full State FB Controller) – Experiments

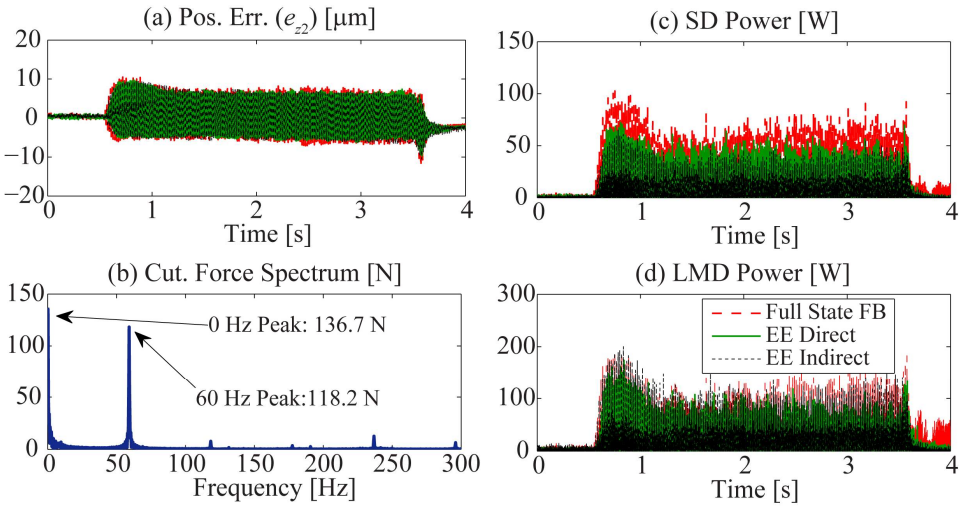

Figure 2.15: Plots of Positioning Error, Cutting Force Spectrum and Actuator Power Dissipated in Heat during Machining Experiment

The controllers are implemented on a dSPACE 1103 motion control platform at a sampling frequency of 10 kHz. A table-mounted high-resolution encoder (see Figure 1.2) is used to measure the position of the table, z_2. However, the rotary position of the motor shaft z_1 is obtained using the observer proposed in [102]. This is because readings from the motor-mounted rotary encoder had an undesirable drift caused by micro slippage, which is known to occur between the Roh'lix nut and smooth shaft [47]. The velocity feedback signals, $\dot{x}_1$ and $\dot{x}_2$, are acquired by taking discrete derivatives of the position signals measured from the rotary and linear encoders, respectively. Figure 2.15 (a) shows a comparison of the positioning errors (e_{z2}), measured from the linear encoder. Notice that all three controllers have nearly identical positioning performance, as also confirmed by the maximum and RMS errors reported in Table 2.7. The disturbance force d_2 during cutting is estimated from the linear motor's current by performing the cut using the LMD only (i.e., by disengaging the Roh'lix nut from the shaft). Figure 2.15 (b) shows the frequency spectrum of the estimated d_2, which clearly shows dominant peaks at 0 and 60 Hz. Figure 2.15 (c) and (d) show the power dissipated as heat in the SD and LMD (calculated using (2.12) without the integral) for each controller; Table 2.7 summarizes the average power dissipated in heat by each actuator during cutting. Observe that even though the positioning performance of all three controllers is almost the same, the EE direct and EE indirect controllers are respectively 26% and 20% more efficient than the full state FB controller.

Table 2.7: Comparison of Tracking Errors and Energy Efficiencies of FB Controllers Based on Machining Experiments

	Max e_{z2} [μm]	RMS e_{z2} [μm]	SD Avg. Pwr. [W]	LMD Avg. Pwr. [W]	Total Avg. Pwr. [W]
Full State FB	12.7	4.3	13.3	24.2	37.5
EE Direct	9.7	3.8	8.8	18.9	27.7
EE Indirect	10.6	4.0	7.9	22.1	30.0

2.6 Summary

This chapter has presented a method for designing a controller that achieves the best positioning accuracy while maximizing the energy efficiency for an over-actuated hybrid feed drive (HFD). An optimal control (input) ratio (OCR) is proposed, for general weak input redundancy, to specify the optimal relationship between the control inputs without sacrificing positioning performance. The OCR is shown to be non-causal, and a causal approximation is proposed for practical real-time implementation. In HFD, the OCR ensures maximum efficiency for any specified positioning accuracy, and the energy efficiency loss due to the causal approximation is shown to be small. The control structure of the HFD is assumed to have a feedforward (FF) controller for tracking and a feedback (FB) controller for regulation. An approach for designing the FF controller using the approximate OCR to achieve perfect tracking while remaining stable and near-energy-optimal has been proposed and verified using comparative simulations. For FB controller design using the approximate OCR, two approaches have been proposed for energy-efficient controller design, without sacrificing positioning accuracy. The first approach directly constrains the control inputs to achieve the requirements of the approximate OCR. The second approach indirectly implements the optimal relationship in the form of a weighting filter for use in traditional H_2/H_∞ controller synthesis. Simulations and machining experiments have been used to demonstrate the benefits of the proposed approaches in significantly improving energy efficiency without sacrificing positioning accuracy for the HFD, under broadband disturbances.

Chapter 3 Proxy-Based Control Allocation for Dual-Input, Single-Output Over-Actuated Systems

3.1 Overview

In the preceding chapter, an optimal control ratio (OCR) was derived for dual-input over actuated systems and shown to be non-causal. Therefore, a causal approximation of the OCR was proposed to enable real-time implementation. However, the approximations involved could lead to significant loss of optimality under certain conditions, as shown in the preceding chapter. Accordingly, in this chapter, the non-causality issue of the OCR is tackled from a different approach that does not require a causal approximation. Through decomposition, a proxy signal is established based on the OCR, and is proven to accurately measure the over-actuated system's deviation from energy optimality. The introduction of the proxy enables the control allocation problem to be converted to a regulation problem, which is solvable using standard control design methodologies. Its resultant optimal control allocation for weak input redundancy is thus achieved through a dynamic system, whose computational load is negligible.

The derivation of the proxy and its relationship to energy optimality are discussed in Section 3.2. The use of the proxy in regulation framework, along with constraints handling techniques, is discussed in Section 3.3. The proposed proxy-based allocation method is validated using hybrid feed drive in simulations and experiments, in comparison with a state-of-the-art control allocation method for weakly redundant systems in Section 3.4, followed by a summary in Section 3.5.

3.2 Control Proxy in Dual-Input, Single-Output Over-Actuated System

3.2.1 Control Allocation Framework

As shown in Figure 3.1, a dual-input, single-output (DISO) system $y = \mathbf{G}\mathbf{u} + \mathbf{G}_d\mathbf{d}$ is defined as in (2.1). This is a weakly input redundant system and has nonzero null space Ker($\mathbf{G}$(s)). The typical two-stage framework of control allocation is shown in Figure 3.1. A nominal control input $\mathbf{u}_0$ from a virtual controller $\mathbf{C}_0$ yields a desirable output y_0 under

disturbance d_0, i.e. the positioning performance is satisfied in this first stage. Set Ω defines the family of control signals that replicate y_0 under $\mathbf{d}_0$. The goal of energy-optimal control allocation is to formulate a mapping $\mathbf{P}$ between $\mathbf{u}_0$ and $\mathbf{u} \in \Omega$ such that control energy is minimized, under the influence of disturbance $\mathbf{d}_0$, without altering $y = y_0$. Accordingly, energy optimality (one supplementary objective) is realized without sacrificing positioning performance (primary control objective).

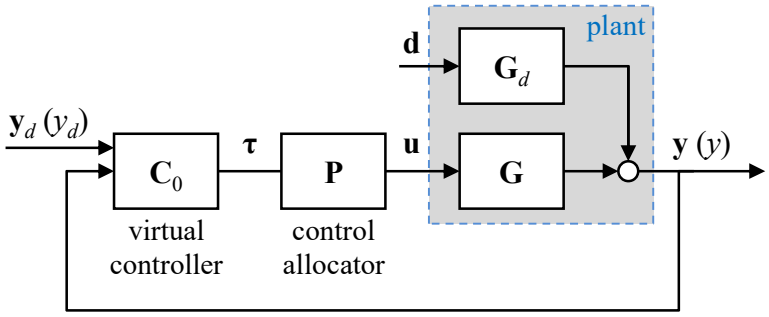

Figure 3.1: Generalized Block Diagram for Control Allocation of Over-Actuated Systems

As discussed in Section 2.2, under quadratic energy cost $J_u = \int(\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u})dt$ defined in (2.4), the control inputs should satisfy the OCR defined as

$$\beta^*(s) \triangleq \frac{\hat{u}_1(s)}{\hat{u}_2(s)} = \frac{R_{22}G_1^* - R_{12}G_2^*}{R_{11}G_2^* - R_{12}G_1^*}, \tag{3.1}$$

to guarantee optimality. However, this relationship was shown in Chapter 2 to be non-causally implementable. A causal stable approximation of β^* (denoted as β) was proposed in Section 2.3.4 however is only valid where the frequency response of β satisfies one of following conditions: $|\beta(j\omega)| \gg 1$, $|\beta(j\omega)| \ll 1$ or $\angle\beta(j\omega) \approx n\pi$ ($n \in \mathbb{Z}$). Significant deviations from energy optimality could result if these conditions are not met, and there is no guarantee that energy optimality is always improved as u_1/u_2 approaches β. To address these issues, a causal proxy signal, representing the deviation from the OCR, is proposed in the following.

3.2.2 Causal Alignment Deviation from Optimal Control Ratio

Given the optimal control ratio (β^*) defined in the preceding section, we seek an allocator $\mathbf{P}$ that enforces the condition $u_1 - \beta^* u_2 = 0$. The operator s has adjoint $s^* = -s$ [81], since the signals it operates on (i.e. u_1 and u_2) belong to the L^2 space and has zero initial and final values. This assumption is valid for control signal $\mathbf{u}$ in almost all practical situations, hence it is adopted here. Accordingly, a causal implementation of $\beta^*(s)$ involves the evaluation of $\beta(-s)$ [103] which

contains unstable poles. We address this challenge by defining u_D, a causal stable measure of the deviation of u_1/u_2 from β^*, as

$$u_D = \beta_2 u_1 - \beta_1 u_2, \tag{3.2}$$

where β_1 and β_2 represent a factorization of β^* defined as

$$\beta^* = \frac{\beta_1}{\beta_2} = \left(\underbrace{\frac{R_{22}G_{1n}^* - R_{12}G_{2n}^*}{\psi}}_{\triangleq \beta_1} \right) \left(\underbrace{\frac{R_{11}G_{2n}^* - R_{12}G_{1n}^*}{\psi}}_{\triangleq \beta_2} \right)^{-1}. \tag{3.3}$$

Here G_{1n} and G_{2n} are respectively coprime numerators of G_1 and G_2 (i.e., $\mathbf{G} = [G_{1n}, G_{2n}]/D_{\text{ol}}$, where D_{ol} is defined such that G_{1n} and G_{2n} do not have shared zeros). On the other hand, ψ is a causal stable denominator, whose selection is discussed in the following section.

The implication of the proposed decomposition of β^* is that its non-causally implementable denominator is replaced by the causal stable denominator, ψ, of β_1 and β_2. Note that $G_{1n}^*(s) = G_{1n}(-s)$ and $G_{2n}^*(s) = G_{2n}(-s)$ can induce non-minimum phase (NMP) zeros in β_1 and β_2, but they do not make β_1 and β_2 unstable. Therefore, β_1 and β_2, unlike β^*, are causally implementable; hence the alignment deviation, u_D, is determinable by a control system in real time.

3.2.3 Relationship between Proxy and Energy Optimality

Based on (3.2), the alignment deviation, u_D, can be viewed as a proxy signal which is usable by a control allocator to indirectly achieve energy optimality. It is therefore instrumental to understand the relationship between u_D and J_u. To do this, for any control input $\mathbf{u} = (\hat{\mathbf{u}} + \delta\mathbf{u}) \in \Omega(\hat{\mathbf{u}})$, J_u can be decomposed as follows:

$$J_u(\mathbf{u}) = \underbrace{\int_0^{+\infty} \hat{\mathbf{u}}^{\mathrm{T}} \mathbf{R} \hat{\mathbf{u}} dt}_{\hat{J}_u} + \underbrace{2\int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}} \mathbf{R} \hat{\mathbf{u}} dt}_{J_{\text{cc}}} + \underbrace{\int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}} \mathbf{R} \delta\mathbf{u} dt}_{\Delta J_u}, \tag{3.4}$$

where $\hat{J}_u$ and ΔJ_u are positive definite terms representing the optimal value of J_u and the energy of deviation $\delta\mathbf{u}$ (i.e. $J_u(\delta\mathbf{u})$), respectively, while J_{cc} is a cross-coupling term between $\delta\mathbf{u}$ and $\hat{\mathbf{u}}$. Based on optimal control ratio $\beta^* = (R_{22}G_1^* - R_{12}G_2^*)(R_{11}G_2^* - R_{12}G_1^*)^{-1}$ (defined in (2.8)), the energy-optimal control, $\hat{\mathbf{u}}$, can be written as

$$\hat{\mathbf{u}} = \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} = \begin{bmatrix} R_{22}G_1^* - R_{12}G_2^* \\ R_{11}G_2^* - R_{12}G_1^* \end{bmatrix} e_0, \tag{3.5}$$

where e_0 is a basis signal. Based on (3.5) and null space relationship

$$G_1\delta u_1 + G_2\delta u_2 = 0, \tag{3.6}$$

the cross-coupling term J_{cc} is written as

$$\begin{aligned} J_{cc} &= 2\int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}}\mathbf{R}\hat{\mathbf{u}}dt \\ &= 2\left(R_{11}R_{22} - {R_{12}}^2\right)\int_0^{+\infty}\left(\delta u_1\left({G_1}^* e_0\right) + \delta u_2\left({G_2}^* e_0\right)\right)dt \\ &= 2\left(R_{11}R_{22} - {R_{12}}^2\right)\int_0^{+\infty}\underbrace{\left(G_1\delta u_1 + G_2\delta u_2\right)}_{\equiv 0} e_0 dt = 0, \end{aligned} \tag{3.7}$$

meaning that $J_u = \widehat{J}_u + \Delta J_u$. Knowing that $u_D = 0$ when $\mathbf{u} = \hat{\mathbf{u}}$, (3.2) can be re-written as

$$u_D = \beta_2\delta u_1 - \beta_1\delta u_2. \tag{3.8}$$

Extracting the numerator of (3.6) and combining it with (3.8), u_D can also be written in terms of δu_1 and δu_2 as

$$\begin{cases} \left(\beta_1 G_{1n} + \beta_2 G_{2n}\right)\delta u_1 = G_{2n}u_D \\ \left(\beta_1 G_{1n} + \beta_2 G_{2n}\right)\delta u_2 = -G_{1n}u_D \end{cases}. \tag{3.9}$$

Based on β_1 and β_2 defined in (3.3) the following relationship is derived

$$\beta_1 G_{1n} + \beta_2 G_{2n} = \frac{\Pi}{\psi}, \tag{3.10}$$

where

$$\Pi \triangleq \begin{bmatrix} G_{2n}^* \\ -G_{1n}^* \end{bmatrix}^{\mathrm{T}} R \begin{bmatrix} G_{2n} \\ -G_{1n} \end{bmatrix}. \tag{3.11}$$

Applying Parseval's theorem and the frequency domain expression of (3.8) to (3.10), the energy increment ΔJ_u can be expressed as

$$\Delta J_u = \int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}}\mathbf{R}\delta\mathbf{u}dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \delta\mathbf{u}^*(\omega)\mathbf{R}\delta\mathbf{u}(\omega)d\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{\psi^*\psi}{\Pi}{u_D}^* u_D d\omega. \tag{3.12}$$

This leads to the relationship:

$$\Delta J_u = \left\|u_D\right\|_2^2, \tag{3.13}$$

if ψ is selected such that $\psi^*\psi = \Pi$. The implication of (3.13) is that the deviation of u_D from zero, measured by $\|u_D\|_2^2$, equals the deviation of J_u from its optimal value of $\widehat{J}_u$. While the selection $\psi^*\psi = \Pi$ may be mathematically convenient, a question that remains is whether it is a causal stable choice for ψ. Notice from (3.11) that Π is self-adjoint (i.e., $\Pi^* = \Pi$); it consists only of zeros, with each zero having a mirror image on the opposite side of the imaginary axis of the s

plane. One can therefore collect all the minimum phase zeros of Π into ψ, leading to a stable and causally implementable ψ. Note that $\mathbf{R}$ is positive definite, indicating based on (3.11) that

$$\Pi(j\omega) = \begin{bmatrix} G_{2n}(-j\omega) \\ -G_{1n}(-j\omega) \end{bmatrix}^{\mathrm{T}} \mathbf{R} \begin{bmatrix} G_{2n}(j\omega) \\ -G_{1n}(j\omega) \end{bmatrix} > 0 \tag{3.14}$$

holds for all $s = j\omega$. Therefore, Π has no zeros on the imaginary axis ($\Pi(j\omega) \neq 0$) as long as G_{1n} and G_{2n} do not share common zeros on the imaginary axis, which is satisfied due to the fact that G_{1n} and G_{2n} are coprime.

3.3 Proxy-Based Control Allocation for DISO Systems

Exploiting the causally implementable proxy u_D, the control allocation problem can be converted into a standard regulation problem. Referring to Figure 3.1, control allocator $\mathbf{P}$ is any mapping between $\mathbf{u}_0$ and $\mathbf{u} \in \Omega(\mathbf{u}_0)$. Based on (2.3), $\mathbf{u}_0$–$\mathbf{u}$ belongs to the orthogonal complement of $\mathbf{G}$ and thus $\mathbf{u}$ is explicitly written as

$$\mathbf{G}(\mathbf{u}_0 - \mathbf{u}) = 0 \Rightarrow \mathbf{u} = \mathbf{u}_0 + \begin{bmatrix} -G_2 \\ G_1 \end{bmatrix} H_v v, \tag{3.15}$$

where v is an arbitrary signal while H_v is a user-defined pre-filter that can be applied to v (e.g., to attenuate high frequency model mismatch). Following the discussions in Sections 4.2.2 and 4.2.3 above, the desire to make $\mathbf{P}$ an energy-optimal dynamic allocator can be indirectly achieved through regulation of u_D. This can be achieved through FF and/or FB control, shown in Figure 3.2 as H_{ff} and H_{fb}, respectively. Note that the Sat(·) block and dotted lines in Figure 3.2 are related to input constraint enforcement discussed in Section 3.3.3; they are ignored in all derivations in this section.

A control designer is free to implement H_{ff} and H_{fb} using a wide range of FF and FB controllers (linear or nonlinear) available in the literature. In the rest of this section, we present a sample FF and FB implementation of the proposed allocation scheme. The presented examples are classical controllers which can easily be implemented in real time, with low computational burden. However, the proposed scheme is also amenable to implementation via advanced controllers like H_2/H_∞ control, sliding mode control, and model predictive control.

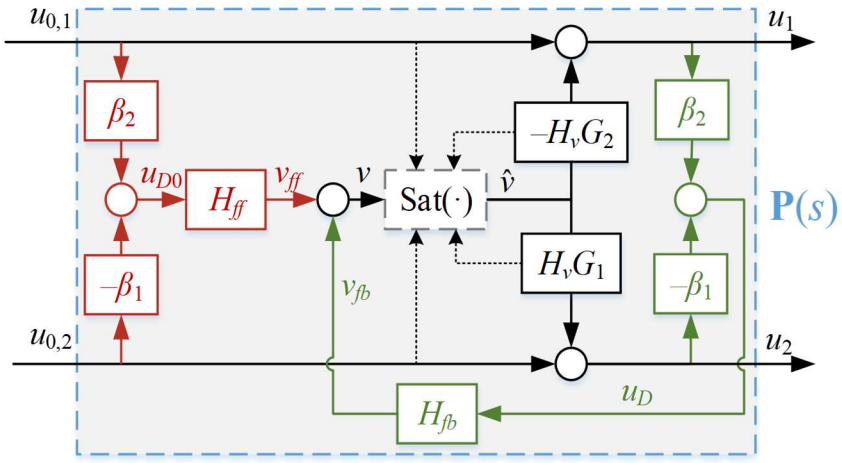

Figure 3.2: General Structure of Proposed Allocation Approach Comprising FF Control (H_{ff}), FB Control (H_{fb}) and Actuator Constraint Handling (Sat(·))

3.3.1 Classical Feedforward Implementation

Focusing on FF control (marked in red in Figure 3.2) by setting $H_{fb} = 0$, $u_D = 0$ requires

$$\begin{aligned} &\beta_2\left(u_{0,1} - H_vG_2v\right) - \beta_1\left(u_{0,2} + H_vG_1v\right) = 0 \\ &\Rightarrow v_{ff} = \left(H_v(\beta_1G_1 + \beta_2G_2)\right)^{-1} u_{D0} = H_{ff}u_{D0}. \end{aligned} \tag{3.16}$$

Note that selecting H_{ff} as defined in (3.16) may not always be feasible in the presence of non-minimum phase (NMP) zeros in $\beta_1G_1 + \beta_2G_2$; moreover, H_{ff} may not be causal. The lack of perfect inversion of H_{ff} indicates that the proxy-based control cannot be energy optimal for every frequency. It also introduces certain level of tradeoff, which can be manipulated by the regulated magnitude of u_D. For practical feedforward implementation, approximate inversion techniques, like zero phase error tracking controller (ZPETC) [95], zero magnitude error tracking controller (ZMETC) [97], filter basis function (FBF) [104,105], etc., may be used to achieve $u_{D0} \approx 0$ over certain frequency or time ranges. Note that these approximate inversion techniques may yield non-proper H_{ff}, which can be addressed via non-causal implementation using future values of u_{D0} (e.g., portions of u_{D0} calculated based on known reference command, y_d) [95]. It can also be addressed by adding fast poles to H_{ff} to make it proper [106], as is demonstrated in Section 3.4.2.2.

3.3.2 Classical Feedback Implementation

Focusing on FB control (marked in green in Figure 3.2) by setting $H_{ff} = 0$, we get

$$v_{fb} = H_{fb} u_D = H_{fb} \begin{bmatrix} \beta_2 & -\beta_1 \end{bmatrix} \mathbf{u}. \tag{3.17}$$

Accordingly, the overall transfer function of **P** is given by

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \underbrace{\frac{\begin{bmatrix} 1+\beta_1 H_{fb} H_v G_1 & \beta_1 H_{fb} H_v G_2 \\ \beta_2 H_{fb} H_v G_1 & 1+\beta_2 H_{fb} H_v G_2 \end{bmatrix}}{1+H_{fb} H_v \left(\beta_1 G_1 + \beta_2 G_2\right)}}_{=\mathbf{P}} \begin{bmatrix} u_{0,1} \\ u_{0,2} \end{bmatrix}. \tag{3.18}$$

The allocator's effect on control alignment with the optimal ratio is evaluated as

$$u_D = \frac{u_{D0}}{1+H_{fb} H_v \left(\beta_1 G_1 + \beta_2 G_2\right)}. \tag{3.19}$$

Here the allocator's characteristic equation

$$\theta(s) \triangleq 1+H_{fb} H_v \left(\beta_1 G_1 + \beta_2 G_2\right) \tag{3.20}$$

is crucial as it determines the allocator performance and stability [77]. FB controller, H_{fb}, can be shaped to deliver a large norm of θ at targeted frequency ranges where efficiency is most desired, while ensuring system stability (e.g., using classical root locus techniques) as demonstrated in Section 3.4.2.2. Note that even though FF and FB implementations have been discussed separately, they can be combined together, as in Figure 3.2, to give

$$v = v_{ff} + v_{fb} = \frac{\left(H_{ff} + H_{fb}\right) u_{D0}}{1+H_{fb} H_v \left(\beta_1 G_1 + \beta_2 G_2\right)}. \tag{3.21}$$

3.3.3 Actuator Constraint Handling

Following the FF and FB design discussed in Section 3.4.2, saturation-type constraints are handled in this section for the DISO system **G** = [G_1, G_2]. Assume that u_1 and u_2 have saturation bounds given by

$$U_1^- \le u_1 \le U_1^+; U_2^- \le u_2 \le U_2^+. \tag{3.22}$$

As shown in Figure 3.2, we seek to enforce these bounds on **u** indirectly through v. This can be done by transferring output bounds to input bounds following the technique discussed in [107,108]. To do this, assume $\mathbf{A}_{c,i}$, $\mathbf{B}_{c,i}$, $\mathbf{C}_{c,i}$, and $\mathbf{D}_{c,i}$ with corresponding state vectors $\mathbf{x}_{c,i}$ (i =1,2), formulate discrete-time realizations of $-H_v G_2$ (for i = 1) and $H_v G_1$ (for i = 2). These two state-space discrete systems are assumed to be strictly proper ($\mathbf{D}_{c,i}$ = **0**) to avoid possible algebraic loops; this can be assured by a proper choice of H_v. The relative degrees of the two systems r_i,

marked by the first nonzero Markov parameters, are each defined as the smallest integer satisfying

$$\mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i-1}\mathbf{B}_{c,i} \neq 0, \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{t_s}\mathbf{B}_{c,i} = 0 \left(t_s = 0,1,\ldots r_i - 2\right). \tag{3.23}$$

where t_s represents the sampled indices of time t in discrete domain. In other words, the relative degree determines the minimum number of time steps needed before a system's input affects its output. With relative degree r_i, only constraints associated with $\mathbf{u}[t_s+r_i]$ and beyond can be enforced by $v[t_s]$ for a given $\mathbf{x}_{c,i}[t_s]$; i.e.,

$$U_i^- \leq \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i}\mathbf{x}_{c,i}\left[t_s\right] + \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i-1}\mathbf{B}_{c,i}v\left[t_s\right] + u_{0,i}\left[t_s + r_i\right] \leq U_i^+. \tag{3.24}$$

Assuming that the fluctuation of u_{0i} is negligible within r_i time steps (i.e., $u_{0,i}[t_s] \approx u_{0,i}[t_s+1]\ldots \approx u_{0,i}[t_s+r_i]$), we can write

$$U_i^- - u_{0,i} - \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i}\mathbf{x}_{c,i} \leq \left(\mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i-1}\mathbf{B}_{c,i}\right)v \leq U_i^+ - u_{0,i} - \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i}\mathbf{x}_{c,i}. \tag{3.25}$$

Accordingly, the limits on $\mathbf{u}$ can be indirectly satisfied by imposing constraints on v as

$$\begin{aligned}
&\underline{v} = \max_{i=1,2}\left(\mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i-1}\mathbf{B}_{c,i}\right)^{-1}\left[U_i^{-\alpha_i} - u_{0,i} - \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i}\mathbf{x}_{c,i}\right];\\
&\overline{v} = \min_{i=1,2}\left(\mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i-1}\mathbf{B}_{c,i}\right)^{-1}\left[U_i^{\alpha_i} - u_{0,i} - \mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i}\mathbf{x}_{c,i}\right];\\
&\alpha_i \triangleq \operatorname{sgn}\left(\mathbf{C}_{c,i}\mathbf{A}_{c,i}^{r_i-1}\mathbf{B}_{c,i}\right),
\end{aligned} \tag{3.26}$$

where $\underline{v}$ and $\overline{v}$ respectively denote the lower and upper bounds on v; α_i is the sign of the first nonzero Markov parameter. Accordingly, the saturation function in Figure 3.2 is defined as [108]

$$\hat{v} = \operatorname{sat}\left(v, \underline{v}, \overline{v}\right) = \max\left(\min\left(\underline{v}, \overline{v}\right), \min\left(v, \max\left(\underline{v}, \overline{v}\right)\right)\right). \tag{3.27}$$

Notice from (3.26) that $\underline{v}$ and $\overline{v}$ are respectively determined by selecting the greater and the smaller value of constraints on v associated with u_1 and u_2. There is however no guarantee that $\underline{v}$ is always smaller than $\overline{v}$, meaning that it may be infeasible for v to lie between $\underline{v}$ and $\overline{v}$. This situation is likely to occur when the dynamics or constraints associated with u_1 and u_2 are very different from each other, especially when high frequency dynamics are considered [108]. High frequency dynamics also make it difficult to satisfy the assumption that $u_{0i}[t_s] \approx u_{0i}[t_s+1]\ldots \approx u_{0i}[t_s+r_i]$. In practice, these issues can be greatly mitigated by using only the low frequency portion of $\mathbf{u}_0$ in (3.26) and selecting pre-filter H_v as a low pass filter to attenuate high frequency dynamics outside the desired allocator bandwidth. The implication is that the Sat(·) block

imposes constraints on low frequency dynamics, but not necessarily on high frequency dynamics, as is illustrated in Section 3.4.3.

3.4 Validation on Over-Actuated Hybrid Feed Drive

3.4.1 Control Proxy for Hybrid Feed Drive

The hybrid feed drive (HFD) [35] introduced in Section 1.1.2 is used to evaluate the proposed control allocation method; the schematic of this dual-input over-actuated system is shown in Section 2.3.

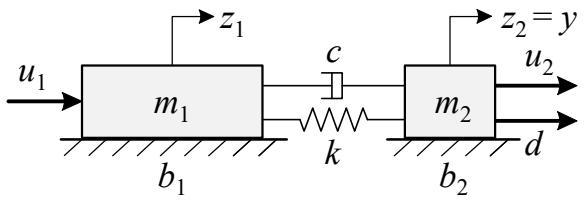

Figure 3.3: Simplified Two-Mass Model of HFD with Cutting Disturbance

Similar to Section 2.3.2, a simplified two-mass model shown in Figure 3.3 is given by

$$\begin{aligned} &m_1\ddot{z}_1 + b_1\dot{z}_1 + c(\dot{z}_1 - \dot{z}_2) + k(\dot{z}_1 - \dot{z}_2) = u_1, \\ &m_2\ddot{z}_2 + b_2\dot{z}_2 + c(\dot{z}_2 - \dot{z}_1) + k(\dot{z}_2 - \dot{z}_1) = u_2 + d. \end{aligned} \tag{3.28}$$

Note that the disturbance on m_1 (mostly friction) is omitted in this formulation (compared to (2.9) and Figure 2.2) since the cutting force on m_2 is dominating and cannot be pre-compensated. The experimentally identified parameters of the HFD's two-mass model are shown in Table 2.1. Since the position of the shaft is not an output to be controlled, over-actuated system

$$y = z_2 = \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^{\mathrm{T}} + G_2 d; \tag{3.29}$$

is formulated. This is a subset of the fully-actuated system defined in (2.10), assuming $G_{21} = G_1$, and $G_{22} = G_2$, $d_1 = 0$ compared to Section 2.3.3. Note that G_1 and G_2 can be decomposed as

$$\begin{aligned} &G_1 = \frac{G_{1n}}{D_{ol}};\ G_2 = \frac{G_{2n}}{D_{ol}}; \\ &G_{1n} = cs + k;\ G_{2n} = m_1 s^2 + (c + b_1)s + k; \end{aligned} \tag{3.30}$$

where D_{ol} is the same as provided in (2.11). The control energy cost functional of the HFD is modeled as the thermal loss (defined in (2.12)) of the HFD, i.e.

$$J_u = \int_0^{+\infty} \left(\left(\frac{u_1}{K_{m1}} \right)^2 + \left(\frac{u_2}{K_{m2}} \right)^2 \right) dt, \tag{3.31}$$

where K_{m1} = 380.8 N/$\sqrt{\text{W}}$ and K_{m2} = 21 N/$\sqrt{\text{W}}$ are the motor constants of the rotary and linear motor, as discussed in Section 2.5. They each represent the efficiency of the corresponding actuator, and ratio $\gamma = (K_{m1}/K_{m2})^2 = 328.8$ represents the energy-optimal static control ratio. Notice that (3.31) follows the general quadratic format in (2.4) with $R_{11} = (1/K_{m1})^2$, $R_{22} = (1/K_{m2})^2$, $R_{12} = 0$. Accordingly, the HFD's optimal control ratio and its factorization are given by

$$\begin{aligned} \beta^* &= \frac{K_{m1}^2}{K_{m2}^2} \frac{k - cs}{m_1 s^2 - (c + b_1)s + k} = \frac{\beta_1}{\beta_2}; \\ \beta_1 &= \frac{k - cs}{K_{m2}^2 \psi}; \beta_2 = \frac{m_1 s^2 - (c + b_1)s + k}{K_{m1}^2 \psi}. \end{aligned} \tag{3.32}$$

Accordingly Π and ψ are determined (based on $\psi^* \psi = \Pi$) as

$$\begin{aligned} \Pi &= \frac{m_1^2 s^4 + \left(2m_1 k - (c + b_1)^2 - \gamma c^2 \right) s^2 + (\gamma + 1) k^2}{K_{m1}^2} \\ &= \underbrace{\left[\frac{m_1 s^2 - a_{\psi 1} s + a_{\psi 0}}{K_{m1}} \right]}_{\psi^*} \underbrace{\left[\frac{m_1 s^2 + a_{\psi 1} s + a_{\psi 0}}{K_{m1}} \right]}_{\psi}, \end{aligned} \tag{3.33}$$

where

$$\begin{aligned} a_{\psi 1} &= \sqrt{2m_1 k \sqrt{(1 + \gamma)} - 2m_1 k + (c + b_1)^2 + \gamma c^2}; \\ a_{\psi 0} &= k \sqrt{(1 + \gamma)}. \end{aligned} \tag{3.34}$$

It can be verified that β_1 and β_2 are causal stable, with poles at –226 ± 204j rad/s.

3.4.2 Design of Nominal Controller and Control Allocator

3.4.2.1 Design of Nominal Controllers

Let us assume that the HFD is required to track a broadband snap-limited reference position signal, y_d, whose velocity profile is shown in Figure 3.4 (a), and kinematic limits are given in Table 3.1. Moreover, as shown in Figure 3.4 (b), during the constant velocity portion of y_d, from 1s to 4.5s, the HFD must reject disturbance force d, given by

$$d = \underbrace{100\cos(20\pi t) - 100}_{\text{Dominant}} + \underbrace{H_d N(0,\sigma^2)}_{\text{Non-dominant}} \quad [\text{N}] \tag{3.35}$$

representing cutting force disturbance with a dominant DC and a harmonic portion (at 10 Hz), combined with a non-dominant broadband portion made up of zero-mean white noise with standard deviation σ = 100 N, low-pass filtered with H_d given by

$$H_d = \frac{\omega_d}{s + \omega_d}. \tag{3.36}$$

Here ω_d = 60π rad/s is employed such that the low pass filter has a cutoff frequency of 30 Hz.

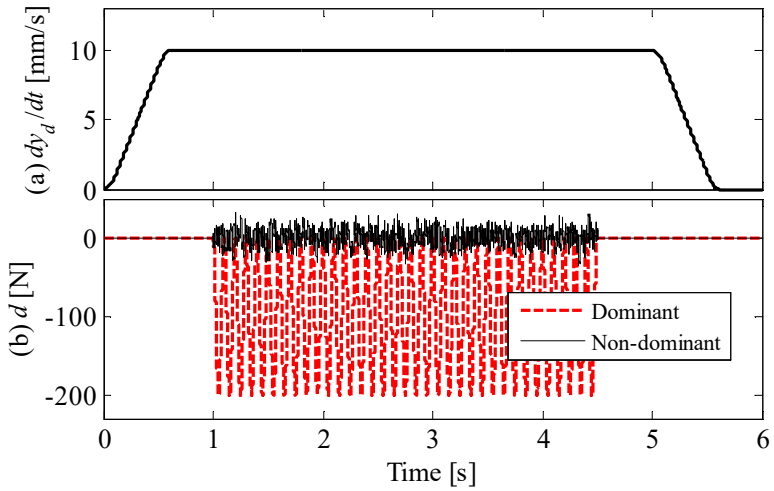

Figure 3.4: (a) Velocity Reference Command and (b) Disturbance Force Profiles

Table 3.1: Kinematic Limits of Reference Trajectory and PID Controller Gains

Pos. [mm]	Vel. [mm·s^{-1}]	Acc. [mm·s^{-2}]	Jerk [mm·s^{-3}]	Snap [mm·s^{-4}]
5	10	20	200	1×10^{4}
K_p [N·m^{-1}]	K_i [N·m^{-1}·s^{-1}]	K_d [N·m^{-1}·s]	T_f [s]	
7.46×10^{6}	4.07×10^{6}	1.01×10^{5}	1.39×10^{-4}	

The nominal controller (i.e., $\mathbf{C}_0$ in Figure 3.1) is designed with a FF component ($\mathbf{C}_{0,ff}$) for reference tracking and a FB component ($\mathbf{C}_{0,fb}$) for regulation such that $\mathbf{u}_0 = \mathbf{u}_{ff0} + \mathbf{u}_{fb0}$.
The FF controller can be designed as

$$\mathbf{u}_{ff0} = \mathbf{C}_{0,ff} y_d = \begin{bmatrix} m_1 s^2 + b_1 s \\ m_2 s^2 + b_2 s \end{bmatrix} y_d, \tag{3.37}$$

which is the simplest FF controller for achieving perfect tracking of y_d; it forces the two-mass HFD to move as one rigid body [109], as discussed in Chapter 2.

The FB controllers used in simulation and the experiment are different. In simulation, the nominal controller is designed to be a PID controller given by

$$\mathbf{u}_{fb0} = \mathbf{C}_{0,fb}\left(y_d - y\right) = \begin{bmatrix} 0 \\ K_p + \dfrac{K_i}{s} + \dfrac{K_d s}{T_f s + 1} \end{bmatrix} \left(y_d - y\right), \tag{3.38}$$

where K_p, K_i, K_d, T_f are calculated using Matlab®'s PID tuning tool targeted at a crossover frequency of 10 Hz, and their values are provided in Table 3.1. Notice that $\mathbf{C}_{0,fb}$ uses only $y = z_2$ for feedback and only u_2 for actuation (i.e., single-input single-output collocated control) which makes for ease of design. In experiments, since the linear motor drive cannot back drive the screw drive, a P/PI-PD controller reported in [98] is used as the feedback controller. The P/PI-PD controller controls the rotary and linear motors using a P/PI and a PD controller, respectively, and is optimized for positioning performance using an H_∞ approach [98]. Even though its design is more involved, it achieves higher bandwidth (107 Hz), and is heuristically more efficient than the PID control approach used in simulations – because it applies the integrator to the rotary motor rather than to the linear motor. Note that neither the FF nor FB controller is designed with efficiency in mind. They both rely on the control allocator to deal with efficiency. This demonstrates the elegance of the two-stage framework upon which the proposed allocator is built.

3.4.2.2 Design of Proposed Energy-Optimal Control Allocator

The proposed allocator is designed using the classical FF and FB implementations discussed in Section 3.3. The FB allocator (H_{fb}) is designed to make $u_D \approx 0$ at 0 and 10 Hz, corresponding to the frequencies of the dominant components of d, which are often known in practical milling scenarios and are most critical to precision and control energy [98]. The FF allocator (H_{ff}) is designed to yield $u_D \approx 0$ in a broadband sense, to provide allocation for portions of u_0 induced by y_d and the broadband portion of d. The pre-filter, H_v, is given by

$$H_v(s) = D_{ol}(s) / D_d(s), \tag{3.39}$$

which replaces the denominator of the original system, D_{ol}, containing undesirable structural modes, with a more desirable denominator, D_d. Here D_d is selected as a third order Butterworth low pass filter with cutoff frequency of 16 Hz, corresponding to the target bandwidth of the allocator (which is greater than the 10 Hz crossover frequency of $\mathbf{C}_{0,fb}$). The numerator of D_d is scaled with stiffness k (which is the DC gain of G_{1n} and G_{2n}) so that the v has force unit [N].

The FB allocator, H_{fb}, is selected as

$$H_{fb}(s) = K_{AL} W_d(s), \tag{3.40}$$

where K_{AL} is a gain which selected to make $\theta(s)$ in (3.20) as high as possible; W_d is a weighting filter that emphasizes frequencies where high $\theta(s)$ is most desirable, since, due to stability constraints, it is impossible to make $\theta(s)$ very high at all frequencies. We select

$$W_d(s) = \frac{1}{s+\varepsilon} \cdot \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \tag{3.41}$$

with $\omega_0 = 20\pi$ rad/s (i.e. 10Hz), so that W_d magnifies $\theta(s)$ at the dominant frequencies contained in d. Note that $\varepsilon = 10^{-5}$ rad/s and $\zeta = 0.01$ are introduced to ensure very high gains at the desired frequencies while keeping the energy of W_d finite. By treating $W_d H_v(\beta_1 G_1 + \beta_2 G_2)$ as an open loop plant, the standard root locus technique is used to determine $K_{AL} = 270\ \mathrm{N}/\sqrt{\mathrm{W}}$ as a high enough FB gain that ensures closed-loop stability. Figure 3.5 shows the root locus and final closed-loop poles of allocator **P** (i.e., the roots of $\theta(s)$).

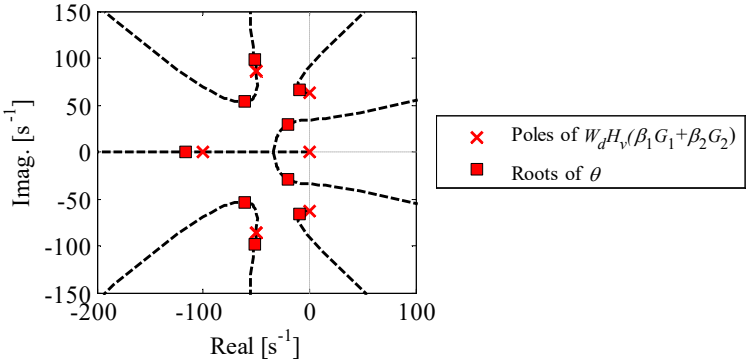

Figure 3.5: Root Locus Plot of FB Allocator

According to (3.16), FF allocator, H_{ff}, is given by

$$H_{ff} = \left(H_v(\beta_1 G_1 + \beta_2 G_2)\right)^{-1} = \frac{D_d(s)}{\psi^*(s)}. \tag{3.42}$$

However, the expression $\psi^*(s) = \psi(-s)$ in the denominator of H_{ff} has unstable poles (see (3.33)). Moreover, H_{ff} is non-proper hence non-causal. An approximation of H_{ff} based on ZPETC [95] is given by

$$H_{ff\_ZPETC} = \frac{\psi(s) D_d(s)}{\psi(0)^2} H_0(s), \tag{3.43}$$

where H_0 is a sixth-order Butterworth low pass filter (similar to the inverse compensation filter [106]) added to make H_{ff_ZPETC} causal.

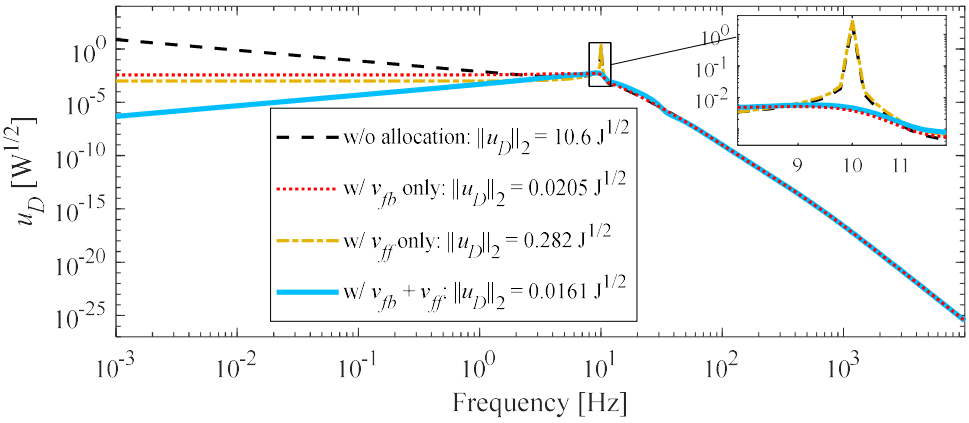

Figure 3.6: Comparison of Estimated Frequency Spectra of u_D Achieved without Allocation and with the Proposed FF and/or FB Allocators

As discussed in Section 3.3, the feedforward and feedback implementations of the proposed allocator can be employed independently, or combined. In Figure 3.6, the frequency spectra of u_D for different implementations are compared by substituting $d(s) \approx W_d(s)$ and $y_d = 0$ in closed loop system transfer functions (i.e., by focusing on the dominant portions of d). Notice that, without allocation, u_D inherits the high DC gain and peak at 10 Hz, corresponding to the dominant portions of d, resulting in $\|u_D\|_2 = 10.6$ $\sqrt{\text{J}}$. The FB allocator eliminates the 10 Hz peak and flattens out the DC gain of u_D such that $\|u_D\|_2 = 0.0205$ $\sqrt{\text{J}}$. The FF allocator reduces u_D (more than achieved by the FB allocator) for frequencies up to 10 Hz, but is not as effective as the FB allocator in suppressing the 10 Hz peak; the result is $\|u_D\|_2 = 0.282$ $\sqrt{\text{J}}$. The combination of both FB and FF allocators does the best job of reducing u_D in a broadband sense, including the 10 Hz peak, resulting in $\|u_D\|_2 = 0.0161$ $\sqrt{\text{J}}$. Therefore, the combination of FF and FB designs is employed in comparing the proposed allocator to an existing allocator in the following section.

3.4.3 Simulation and Experiment Results

In this section, the proposed proxy-based (PB) dynamic allocator designed in Section 3.4.2 above is compared to the static-model-based (SMB) dynamic allocator by Zaccarian [68], which circumvents time-consuming real-time optimization by approximating (1.5) as

$$\mathrm{Ker}\left(\lim_{s\to 0}\mathbf{G}(s)\right)\neq\varnothing, \tag{3.44}$$

In other words, it considers weak redundancy only in a static sense. The SMB allocator is implemented as detailed in [68] based on the parameters of the HFD, and with energy performance matrix $\mathbf{R}$=diag(K_{m1}^{-2}, K_{m2}^{-2}) and allocator speed parameter $\mathbf{K}_z$= 10^3 defined in [68] and in (1.15) as discussed in Section 1.2.3.3.

3.4.3.1 Simulation Results

The nominal FF and FB controllers designed in Section 3.3 are applied to both the PB and the SMB allocator; both allocators are benchmarked against the nominal controller with no allocator, regarding positioning error and control power (i.e. control energy per unit time). Also, in order to illustrate the constraint handling approach proposed in Section 3.3.3, an additional case with saturation limits $U_1^{\pm} = U_2^{\pm} = \pm$ 100 N applied to the PB allocator is simulated. The inputs to the simulation are y_d and d given in Figure 3.4.

Table 3.2: Positioning Performance and Average Control Power Consumption Comparison (Simulation)

	Pos. Err. [μm]		Control Power [W]		
	Max.	RMS	P_1	P_2	P_{total}
No Allocator	22.73	8.96	0	21.982	21.982
SMB Allocator [68]	23.08	9.41	0.038	11.092	11.130
Proposed PB Allocator (w/o constraints)	22.73	8.96	0.047	1.108	1.155
Proposed PB Allocator (w/ constraints)	22.73	8.96	0.034	1.396	1.430

As shown in the tracking results of Figure 3.7, and statistics reported in Table 3.2, relative to the no allocator case, the SMB allocator introduces 2% and 5% percent more maximum and RMS tracking errors, respectively. However, the proposed PB allocator (with and without constraints) preserves the tracking performance perfectly. Notice from Figure 3.7 that, as

expected, the PB and SMB allocators both re-assign more of the static control effort to the more-efficient rotary motor, compared to the no allocator case which relies heavily on the less-efficient linear motor. The major difference between the SMB allocator and the PB allocator lies at higher frequencies. For instance, in keeping with β^* (see Figure 2.4), the PB allocator assigns much more regulation of the 10 Hz disturbance signal to the rotary motor, as shown in Figure 3.7 and Figure 3.8; the SMB allocator does not. As a result, without constraints imposed, the PB allocator respectively consumes 95% and 90% less average control power than the no allocator and the SMB allocator cases. Figure 3.7 also illustrates the effectiveness of the constraint handling approach discussed in Section 3.3.3, applied to the PB allocator; notice that it truncates the portions of u_1 that exceed the imposed limits and re-assigns them to u_2. This re-assignment does not affect positioning performance, because, as shown in Figure 3.2, it occurs in the null space of **G**. However, the re-assignment violates the requirements of β^* resulting in 24% more average control power consumption compared to the unconstrained case. Nonetheless, even with constraints imposed, the proposed PB allocator still significantly outperforms the no allocator and SMB allocator cases.

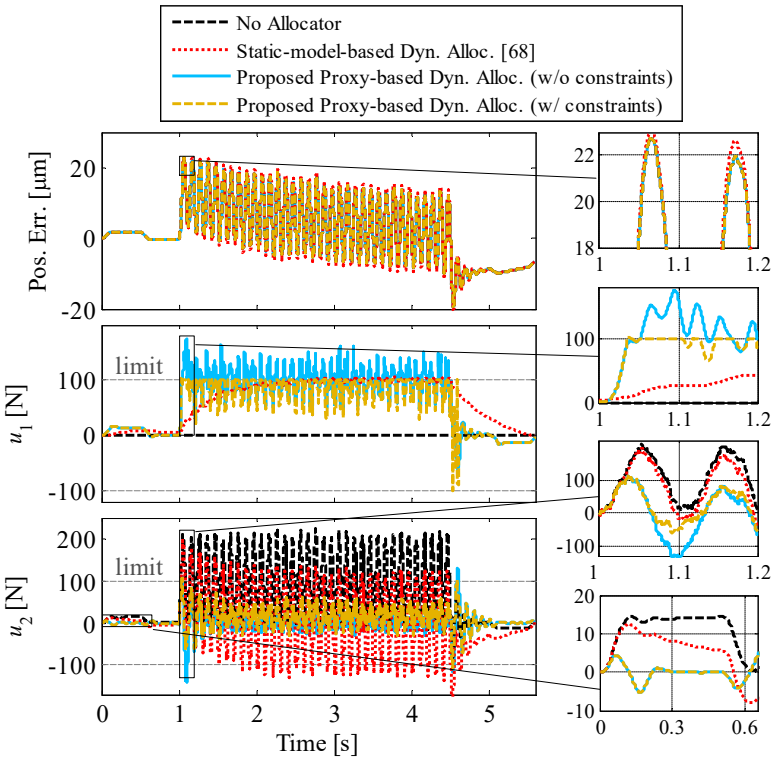

Figure 3.7: Time Domain Comparison of Positioning Performance and Control Efforts of Allocation Methods (Simulation).

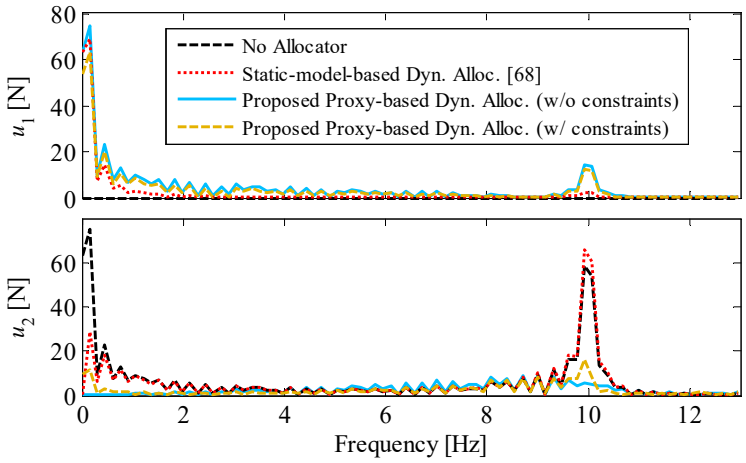

Figure 3.8: Comparison of Control Effort Frequency Spectra of Allocation Methods (Simulation)

3.4.3.2 Experiment Results

The proposed PB allocation approach is evaluated in experiments conducted on an in-house built prototype of the HFD (see Section 1.1.2 for details), controlled at 10 kHz sampling frequency using dSPACE DS1103 control board. The same FF controller used in the simulations, combined with a P/PI-PD controller reported in [98], is used as the nominal FB controller in the experiments. The exact same reference command, disturbance forces, SMB allocator [68], and FF+FB implementations of proposed PB allocator as used in simulations are adopted in the experiments; the disturbance forces are applied to the table via the linear motor. An additional case with saturation limits of $U_1^{\pm} = U_2^{\pm} = \pm$ 150 N is considered. The control signals are low-pass filtered with a cut-off frequency of 100 Hz before passing them to the constraint handler to avoid issues with high-frequency control signals, as discussed in Section 3.3.3.

The time domain and frequency domain results from the experiments are shown in Figure 3.9 and Figure 3.10, and the tracking error and average control power statistics are reported in Table 3.3. Unlike in the simulations, in the experiments, the tracking errors of the proposed PB allocator are slightly higher compared to those of the nominal controller. This is because the PB allocator, being model-based, can only guarantee invariance of control performance in the absence of modeling errors, which is not the case in the experiments. From the energy point of view, the SMB allocator [68] consumes only 2% less average control power than the no allocator case, due to the more-efficient nominal controller adopted in the experiments. However, for the

same reasons as discussed in the simulations, without constraints, the proposed PB allocator consumes 45% and 24% less control power than the SMB allocator, without and with constraints imposed, respectively. Notice from Figure 3.9 that even with constraints imposed, u_1 appears to violate the constraints. The reason is that, as explained in Section 3.3.3, the constraint handler does not impose constraints on signals above the allocator bandwidth regulated by H_v (16 Hz). Figure 3.11 compares the control signals of Figure 3.9 for the proposed allocator with and without constraints, filtered using a low-pass filter with 16 Hz cut off frequency. Notice that, as expected, the constraint handler redistributes the control effort from u_1 to u_2 to satisfy the constraints. The constraint handler is thus suitable for many practical applications where constraint violation by high-frequency portions of signals is not a major concern.

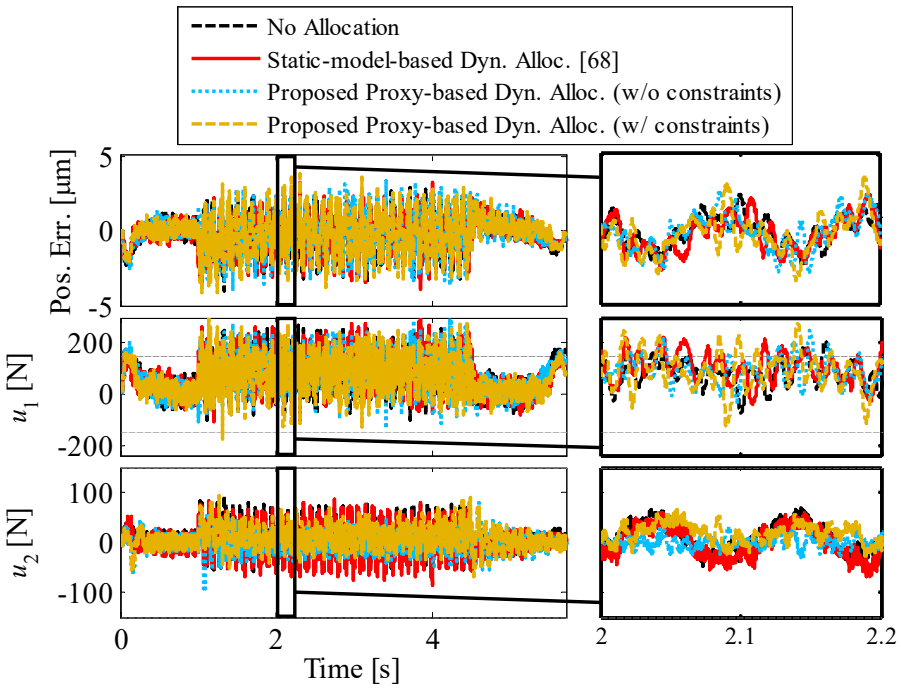

Figure 3.9: Time Domain Comparison of Positioning Performance and Control Efforts of Allocation Methods (Experiment).

Table 3.3: Positioning Performance and Average Control Power Consumption Comparison (Experiments)

	Pos. Err. [μm]		Control Power [W]		
	Max.	RMS	P_1	P_2	P_{total}
No Allocator	4.02	1.04	0.065	2.016	2.081
SMB Allocator [68]	3.94	1.02	0.068	1.964	2.031
Proposed PB Allocator (w/o constraints)	3.98	1.20	0.072	1.055	1.127
Proposed PB Allocator (w/ constraints)	4.19	1.12	0.062	1.472	1.534

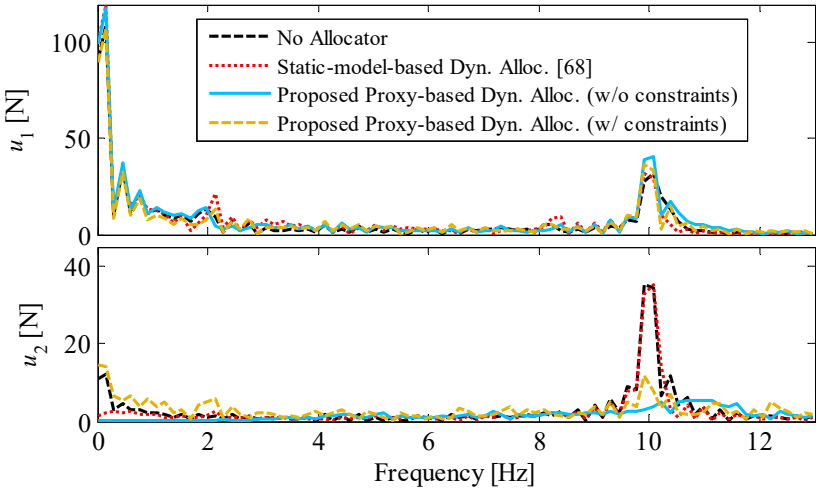

Figure 3.10: Comparison of Control Effort Frequency Spectra of Allocation Methods (Experiment).

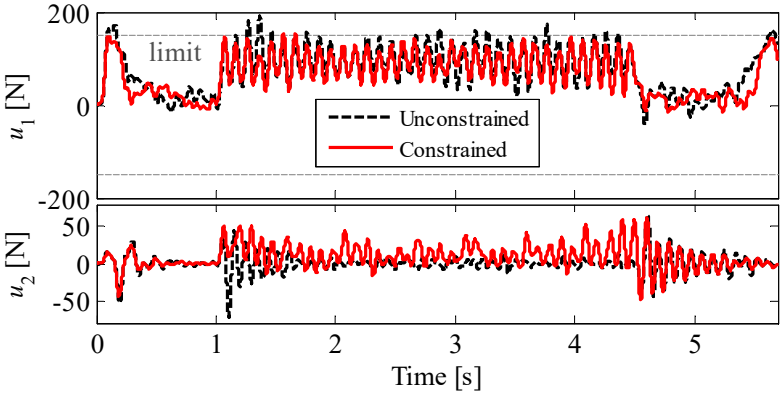

Figure 3.11: Comparison of Control Efforts of Proposed Proxy-Based Dynamic Allocator with and without Constraints, Low-Pass Filtered with 16 Hz Cut Off

3.5 Summary

In this chapter, an elegant method is proposed for resolving the non-causality problem of the optimal control (input) ratio derived in Chapter 2 for energy-optimal dynamic control allocation of dual-input single-output over actuated systems. Through factorization, a causal and stable deviation measure from the optimal control ratio is derived. The deviation measure is shown to be an accurate proxy for deviation from control energy optimality for weak input redundancy. Hence a proxy-based dynamic allocation approach is proposed to drive the control system to control energy optimality by regulating the deviation measure using classical or advanced feedforward (FF) or feedback (FB) controllers. The resultant control allocator is simple and computationally efficient since it does not require real-time optimization. A method for handling actuator constraints without sacrificing performance is also presented. The proposed proxy-based dynamic allocation approach is compared to an existing static-model-based dynamic allocator in simulations and experiments on an over-actuated hybrid feed drive. Large improvements in efficiency without sacrificing performance are demonstrated. The effectiveness of the constraint handler on imposing constraints on low frequency portions of control signals is also shown.

Chapter 4 Proxy-Based Control Allocation for Multi-Input, Multi-Output Over-Actuated Systems

4.1 Overview

In this chapter, the proxy-based control allocation for dual-input, single-output (DISO) system is extended to multi-input, multi-output (MIMO) over-actuated systems. The concept of optimal control ratio is extended to optimal control subspace in Section 4.2, where the control proxy is also realized through matrix fraction description of MIMO systems. The proxy-based control allocation for MIMO system is discussed with feedback design through H_∞ synthesis in Section 4.3, followed by simulation examples in Section 4.4 and a summary in Section 4.5.

4.2 Control Proxy in Multi-input, Multi-output Over-Actuated System

4.2.1 Optimal Control Subspace

The concept of optimal control subspace (OCS) is a natural extension of the OCR in DISO systems. Consider a general MIMO weakly input redundant system given by

$$\mathbf{y} = \mathbf{G}\mathbf{u} + \mathbf{G}_d\mathbf{d}, \tag{4.1}$$

where $\mathbf{y} \in \mathbb{R}^{n_y}$ is the output of the system, while $\mathbf{u} \in \mathbb{R}^{n_u}$, $\mathbf{d} \in \mathbb{R}^{n_d}$ are the control input and disturbance, respectively. All signals in $\mathbf{u}$ are assumed to belong to L^2 space and have zero initial values. According to Section 1.2.3.1, such a system with more inputs than outputs is defined as weakly input redundant system since $\mathrm{Ker}(\mathbf{G}(s)) \neq \emptyset$.

For weakly input redundant system with $n_u > n_y$, define redundancy degree n_r as

$$n_r = n_u - n_y. \tag{4.2}$$

Without loss of generality, the first n_y control inputs are assumed to formulate a non-redundant control input set, i.e. the overall system transfer function matrix is divided as

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_p & \mathbf{G}_r \end{bmatrix}; \mathbf{u} = \begin{bmatrix} \mathbf{u}_p^{\mathrm{T}} & \mathbf{u}_r^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \tag{4.3}$$

where $\mathbf{G}_p$ is a nonsingular square transfer function matrix from the first n_y principal control inputs $\mathbf{u}_p$ to the outputs, and $\mathbf{G}_r$ ($n_y \times n_r$) is a transfer function matrix from the remaining n_r redundant control inputs $\mathbf{u}_r$ to the outputs. Due to the nature of over-actuated systems, there exists null space within which the control inputs' variations would not affect the system outputs, i.e.

$$\begin{bmatrix}\mathbf{G}_p & \mathbf{G}_r\end{bmatrix}\begin{bmatrix}\delta\mathbf{u}_p \\ \delta\mathbf{u}_r\end{bmatrix} = \mathbf{0} \Rightarrow \delta\mathbf{u}_p = -\mathbf{G}_p^{-1}\mathbf{G}_r\delta\mathbf{u}_r. \tag{4.4}$$

A quadratic cost J_u is defined for the MIMO system as

$$J_u(\mathbf{u}) = \int_0^{+\infty} \mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}dt = \int_0^{+\infty} \begin{bmatrix}\mathbf{u}_p \\ \mathbf{u}_r\end{bmatrix}^{\mathrm{T}} \begin{bmatrix}\mathbf{R}_p & \mathbf{R}_{pr} \\ \mathbf{R}_{pr}^{\mathrm{T}} & \mathbf{R}_r\end{bmatrix}\begin{bmatrix}\mathbf{u}_p \\ \mathbf{u}_r\end{bmatrix} dt, \tag{4.5}$$

where $\mathbf{R}$ is a positive definite symmetric matrix divided according to the first n_y principal control inputs and n_r redundant inputs. To achieve energy optimality, the variation of J_u should satisfy

$$\delta J_u = \int_0^{+\infty} \begin{bmatrix}\delta\mathbf{u}_p \\ \delta\mathbf{u}_r\end{bmatrix}^{\mathrm{T}} \begin{bmatrix}\mathbf{R}_p & \mathbf{R}_{pr} \\ \mathbf{R}_{pr}^{\mathrm{T}} & \mathbf{R}_r\end{bmatrix}\begin{bmatrix}\mathbf{u}_p \\ \mathbf{u}_r\end{bmatrix} dt = 0. \tag{4.6}$$

Consider the variation within the null space specified by (4.4),

$$\int_0^{+\infty} \left(\left(-\mathbf{G}_p^{-1}\mathbf{G}_r\delta\mathbf{u}_r\right)^{\mathrm{T}} \mathbf{R}_p\mathbf{u}_p + \left(-\mathbf{G}_p^{-1}\mathbf{G}_r\delta\mathbf{u}_r\right)^{\mathrm{T}} \mathbf{R}_{pr}\mathbf{u}_r + \delta\mathbf{u}_r^{\mathrm{T}}\mathbf{R}_{pr}^{\mathrm{T}}\mathbf{u}_p + \delta\mathbf{u}_r^{\mathrm{T}}\mathbf{R}_r\mathbf{u}_r \right) dt = \mathbf{0}. \tag{4.7}$$

Taking the adjoint operation of transfer function matrix $\mathbf{G}_p^{-1}\mathbf{G}_r$

$$\int_0^{+\infty} \delta\mathbf{u}_r^{\mathrm{T}} \left(\left((\mathbf{G}_p^{-1}\mathbf{G}_r)^* \mathbf{R}_p - \mathbf{R}_{pr}^{\mathrm{T}} \right)\mathbf{u}_p + \left((\mathbf{G}_p^{-1}\mathbf{G}_r)^* \mathbf{R}_{pr} - \mathbf{R}_r \right)\mathbf{u}_r \right) dt = \mathbf{0}, \tag{4.8}$$

where * indicate the adjoint operation as mentioned in Section 2.2. Based on the fundamental lemma of calculus of variation for multivariable systems [82], following condition holds for the energy-optimal control inputs.

$$\left[\left(\mathbf{G}_p^{-1}\mathbf{G}_r\right)^* \mathbf{R}_p - \mathbf{R}_{pr}^{\mathrm{T}} \right]\hat{\mathbf{u}}_p = \left[\mathbf{R}_r - \left(\mathbf{G}_p^{-1}\mathbf{G}_r\right)^* \mathbf{R}_{pr} \right]\hat{\mathbf{u}}_r, \tag{4.9}$$

where $\hat{}$ accent is used to denote optimality as defined in Section 2.2. Note that $(\mathbf{G}_p^{-1}\mathbf{G}_r)^*$ is non-causal (because of the adjoint operator), and the relation in (4.9) formulates optimal control subspace.

4.2.2 Proxy as a Deviation from OCS

To design an energy-optimal allocator **P**, as discussed in Section 3.3, it is desirable to enforce the optimal control subspace defined in (4.9). The operator s has adjoint $s^* = -s$ [81], since the signals it operates on (i.e. **u**) belong to the L^2 space and have zero initial and final values. This assumption is valid for control signal **u** in almost all practical situations, hence it is adopted here. Here we propose a causal alignment deviation measure through decomposition of non-causal operators. First, matrix fraction description [110] is used to decompose $\mathbf{G}_p^{-1}\mathbf{G}_r$, whose adjoint operation in (4.9) is the source of the non-causality. The matrix fraction description is given by

$$\mathbf{G}_p^{-1}(s)\mathbf{G}_r(s) = \mathbf{N}(s)\mathbf{D}^{-1}(s), \tag{4.10}$$

where $\mathbf{N}$ ($n_y \times n_r$) and $\mathbf{D}$ ($n_r \times n_r$) are coprime transfer function polynomials without denominators, such that (4.10) formulates a right coprime fraction of $\mathbf{G}_p^{-1}\mathbf{G}_r$. Note that here $\mathbf{N}$ and $\mathbf{D}$ are not unique: all possible coprime $\mathbf{N}$ and $\mathbf{D}$ pairs are related to each other through unimodular transformation matrices [110]. The importance of this right coprime fraction is that $\mathbf{N}$ and $\mathbf{D}$ only contain the numerator polynomials, whose adjoint

$$\left(\mathbf{N}(s)\right)^* = \mathbf{N}^{\mathrm{T}}(-s);\ \left(\mathbf{D}(s)\right)^* = \mathbf{D}^{\mathrm{T}}(-s) \tag{4.11}$$

are also numerator polynomials due to $s^* = -s$ [81], which does not lead to instability. Deviation from (4.9), which originally cannot be causally evaluated, is measured by proxy $\mathbf{u}_D$ defined as

$$\mathbf{u}_D = \boldsymbol{\beta}_p\mathbf{u}_p - \boldsymbol{\beta}_r\mathbf{u}_r, \tag{4.12}$$

where

$$\begin{aligned}\boldsymbol{\beta}_p &= \boldsymbol{\psi}^{-1}\left(\mathbf{N}^*\mathbf{R}_p - \mathbf{D}^*\mathbf{R}_{pr}^{\mathrm{T}}\right);\\ \boldsymbol{\beta}_r &= \boldsymbol{\psi}^{-1}\left(\mathbf{D}^*\mathbf{R}_r - \mathbf{N}^*\mathbf{R}_{pr}\right).\end{aligned} \tag{4.13}$$

Here $\boldsymbol{\psi}$ is defined as a square ($n_r \times n_r$) nonsingular numerator polynomial with minimum phase zeros such that both $\boldsymbol{\psi}^{-1}\mathbf{D}^*$ and $\boldsymbol{\psi}^{-1}\mathbf{N}^*$ are proper and stable transfer function matrices, and can thus be evaluated in real time. Comparing the results from the MIMO and DISO cases, one observes that the $\mathbf{D}$ and $\mathbf{N}$ matrices in (4.10) are generalizations of G_{1n} and G_{2n} defined in (3.3).

4.2.3 Relationship between Proxy and Control Energy Optimality

It is clear from (4.9) and (4.12) that enforcing $\mathbf{u}_D = \mathbf{0}$ ensures energy optimality. However this perfect condition may not be always satisfied due to various reasons such as limited control

bandwidth, non-minimum phase (NMP) zeros, etc., as discussed in Section 3.3. Therefore, it is instrumental to understand how nonzero $\mathbf{u}_D$ is related to the energy cost J_u. For any control signal $\mathbf{u} = \hat{\mathbf{u}} + \delta\mathbf{u}$ which belongs to $\Omega(\hat{\mathbf{u}})$, its energy cost J_u is decomposed into

$$J_u(\mathbf{u}) = \underbrace{\int_0^{+\infty} \hat{\mathbf{u}}^{\mathrm{T}}\mathbf{R}\hat{\mathbf{u}}dt}_{J_u(\hat{\mathbf{u}})} + \underbrace{2\int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}}\mathbf{R}\hat{\mathbf{u}}dt}_{J_{\mathrm{cc}}} + \underbrace{\int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}}\mathbf{R}\delta\mathbf{u}dt}_{J_u(\delta\mathbf{u})}, \tag{4.14}$$

where $J_u(\hat{\mathbf{u}})$ and $J_u(\delta\mathbf{u})$ are positive definite terms representing the optimal energy cost and energy cost of $\delta\mathbf{u}$. Note that $\delta\mathbf{u}$ is assumed to satisfy the null space condition specified by (4.4), such that the cross-coupling term $J_{\mathrm{cc}} = 0$, i.e.

$$\begin{aligned} J_{\mathrm{cc}} &= 2\int_0^{+\infty} \begin{bmatrix} \delta\mathbf{u}_p \\ \delta\mathbf{u}_r \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \mathbf{R}_p & \mathbf{R}_{pr} \\ \mathbf{R}_{pr}^{\mathrm{T}} & \mathbf{R}_r \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_p \\ \hat{\mathbf{u}}_r \end{bmatrix} dt \\ &= 2\int_0^{+\infty} \left(\delta\mathbf{u}_p^{\mathrm{T}}\mathbf{R}_p\hat{\mathbf{u}}_p + \delta\mathbf{u}_p^{\mathrm{T}}\mathbf{R}_{pr}\hat{\mathbf{u}}_r + \delta\mathbf{u}_r^{\mathrm{T}}\mathbf{R}_{pr}^{\mathrm{T}}\hat{\mathbf{u}}_p + \delta\mathbf{u}_r^{\mathrm{T}}\mathbf{R}_r\hat{\mathbf{u}}_r \right) dt \\ &= 2\int_0^{+\infty} \delta\mathbf{u}_r^{\mathrm{T}} \left(\underbrace{\left((\mathbf{G}_p^{-1}\mathbf{G}_r)^*\mathbf{R}_p - \mathbf{R}_{pr}^{\mathrm{T}} \right)\hat{\mathbf{u}}_p + \left((\mathbf{G}_p^{-1}\mathbf{G}_r)^*\mathbf{R}_{pr} - \mathbf{R}_r \right)\hat{\mathbf{u}}_r}_{=0} \right) dt = 0. \end{aligned} \tag{4.15}$$

The integrand in (4.15) is always zero due to the optimal control subspace condition in (4.9). This zero cross-coupling term indicates that the $J_u(\delta\mathbf{u})$ comprises the positive definite energy increment $J_u(\mathbf{u}) = J_u(\hat{\mathbf{u}}) + J_u(\delta\mathbf{u})$ from the optimal control inputs. Knowing that $\mathbf{u}_D = \mathbf{0}$ when $\mathbf{u} = \hat{\mathbf{u}}$, (4.12) can be re-written as

$$\mathbf{u}_D = \boldsymbol{\beta}_p\delta\mathbf{u}_p - \boldsymbol{\beta}_r\delta\mathbf{u}_r. \tag{4.16}$$

Combined with (4.4), $\delta\mathbf{u}$ is related to the proxy $\mathbf{u}_D$ as

$$\begin{aligned} \delta\mathbf{u}_r &= -\left[\boldsymbol{\beta}_p \left(\mathbf{G}_p^{-1}\mathbf{G}_r \right) + \boldsymbol{\beta}_r \right]^{-1} \mathbf{u}_D = -\mathbf{D}\boldsymbol{\Pi}^{-1}\boldsymbol{\psi}\mathbf{u}_D; \\ \delta\mathbf{u}_p &= \mathbf{G}_p^{-1}\mathbf{G}_r \left[\boldsymbol{\beta}_p \left(\mathbf{G}_p^{-1}\mathbf{G}_r \right) + \boldsymbol{\beta}_r \right]^{-1} \mathbf{u}_D = \mathbf{N}\boldsymbol{\Pi}^{-1}\boldsymbol{\psi}\mathbf{u}_D, \end{aligned} \tag{4.17}$$

where $\boldsymbol{\Pi}$ is defined to be a self-adjoint system given by

$$\boldsymbol{\Pi} = \begin{bmatrix} \mathbf{N} \\ -\mathbf{D} \end{bmatrix}^* \mathbf{R} \begin{bmatrix} \mathbf{N} \\ -\mathbf{D} \end{bmatrix} \tag{4.18}$$

and according to (4.13),

$$\boldsymbol{\beta}_p \left(\mathbf{G}_p^{-1}\mathbf{G}_r \right) + \boldsymbol{\beta}_r = \boldsymbol{\psi}^{-1}\boldsymbol{\Pi}\mathbf{D}^{-1}. \tag{4.19}$$

Applying Parseval's theorem and the frequency domain expression from (4.17) to (4.19), the energy increment due to deviation from optimal subspace $\Delta J_u = J_u(\mathbf{u}) - J_u(\hat{\mathbf{u}}) = J_u(\delta\mathbf{u})$ is given by

$$\Delta J_u = \int_0^{+\infty} \delta\mathbf{u}^{\mathrm{T}}\mathbf{R}\delta\mathbf{u}dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \delta\mathbf{u}^*(\omega)\mathbf{R}\delta\mathbf{u}(\omega)d\omega$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathbf{u}_D^*\boldsymbol{\psi}^*\left(\boldsymbol{\Pi}^{-1}\right)^* \underbrace{\begin{bmatrix}\mathbf{N}\\ -\mathbf{D}\end{bmatrix}^* \mathbf{R}\begin{bmatrix}\mathbf{N}\\ -\mathbf{D}\end{bmatrix}}_{=\boldsymbol{\Pi}} \boldsymbol{\Pi}^{-1}\boldsymbol{\psi}\mathbf{u}_D d\omega \tag{4.20}$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathbf{u}_D^*\boldsymbol{\psi}^*\left(\boldsymbol{\Pi}^{-1}\right)^* \boldsymbol{\psi}\mathbf{u}_D d\omega.$$

The energy increment ΔJ_u is the square of the two-norm of the proxy $\mathbf{u}_D$, i.e.

$$\Delta J_u = \left\|\mathbf{u}_D\right\|_2^2, \tag{4.21}$$

given

$$\boldsymbol{\Pi} = \boldsymbol{\psi}\boldsymbol{\psi}^*. \tag{4.22}$$

The implication of (4.21) is that the deviation of $\mathbf{u}_D$ from zero is directly proportional to the deviation of $J_u(\mathbf{u})$ from its optimal value of $J_u(\hat{\mathbf{u}})$. Accordingly, decreasing the 2-norm of $\mathbf{u}_D$ via regulation of the proxy $\mathbf{u}_D$ strictly enhances the energy efficiency.

Notice from (4.18) that the self-adjoint $\boldsymbol{\Pi}$ consists only of zeros symmetrically placed about the imaginary axis. One can therefore collect all the minimum phase zeros of $\boldsymbol{\Pi}$ into a stable and causally implementable $\boldsymbol{\psi}$; this decomposition in (4.22) is referred as left spectral factorization [111]. Note that $\mathbf{N}$ and $\mathbf{D}$ are coprime (share no common zeros) and $\mathbf{R}$ is positive definite, indicating $\boldsymbol{\Pi}$ has no zeros on the imaginary axis because

$$\boldsymbol{\Pi}(j\omega) = \begin{bmatrix}\mathbf{N}(-j\omega)\\ -\mathbf{D}(-j\omega)\end{bmatrix}^{\mathrm{T}} \mathbf{R}\begin{bmatrix}\mathbf{N}(j\omega)\\ -\mathbf{D}(j\omega)\end{bmatrix} > 0 \tag{4.23}$$

holds for all $s = j\omega$. Under this condition, stable and minimum phase solutions $\boldsymbol{\psi}$ to (4.22) always exist [111].

4.3 Proxy-Based Control Allocation Design for MIMO System

As Figure 3.1 illustrates, the control allocator $\mathbf{P}$ aims at redistributing control efforts $\mathbf{u}_0$ within $\Omega(\mathbf{u}_0)$ such that the control performance is preserved. Therefore, according to (4.4), it is assumed that

$$\mathbf{u} = \mathbf{u}_0 + \begin{bmatrix}-\mathbf{N}\boldsymbol{\psi}_0^{-1}\\ \mathbf{D}\boldsymbol{\psi}_0^{-1}\end{bmatrix}\mathbf{v}, \tag{4.24}$$

where $\mathbf{v} \in \mathbb{R}^{n_r}$ is an arbitrary signal while $\boldsymbol{\psi}_0$ is a user-defined nonsingular square polynomial transfer function numerator matrix with minimum phase zeros; $\boldsymbol{\psi}_0$ is a pre-filter that helps to eliminate in undesirable pole dynamics the original system. Note that (4.18) satisfies the null space defined in (4.4) as

$$\begin{bmatrix} \mathbf{G}_p & \mathbf{G}_r \end{bmatrix} \begin{bmatrix} -\mathbf{N}\boldsymbol{\psi}_0^{-1} \\ \mathbf{D}\boldsymbol{\psi}_0^{-1} \end{bmatrix} \mathbf{v} = \mathbf{0}. \tag{4.25}$$

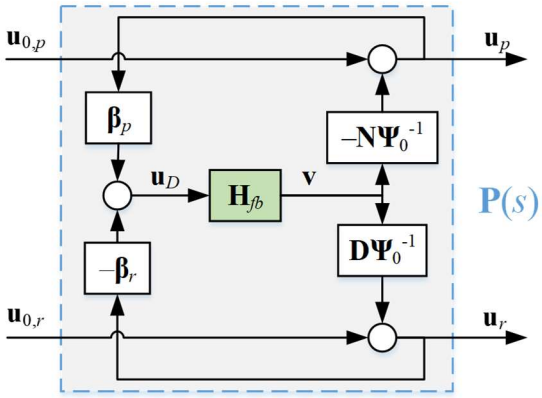

Figure 4.1: Proposed Control Allocator based on Feedback Design

Assume that the proxy of the nominal controller $\mathbf{u}_0$ (before allocation) is defined as

$$\mathbf{u}_{D0} = \begin{bmatrix} \boldsymbol{\beta}_p & -\boldsymbol{\beta}_r \end{bmatrix} \mathbf{u}_0. \tag{4.26}$$

Then the proxy after allocation $\mathbf{u}_D$ is given by

$$\mathbf{u}_D = \mathbf{u}_{D0} + \Delta\mathbf{u}_D = \mathbf{u}_{D0} + \underbrace{\begin{bmatrix} \boldsymbol{\beta}_p & -\boldsymbol{\beta}_r \end{bmatrix} \begin{bmatrix} -\mathbf{N}\boldsymbol{\psi}_0^{-1} \\ \mathbf{D}\boldsymbol{\psi}_0^{-1} \end{bmatrix}}_{\triangleq \mathbf{H}_v} \mathbf{v}, \tag{4.27}$$

where $\mathbf{H}_v$ is a square ($n_r \times n_r$) dynamic system, mapping $\mathbf{v}$ to its manipulation of the proxy $\Delta\mathbf{u}_D$. Notice from (4.27) that one possibility is to make $\mathbf{u}_D = \mathbf{0}$ via feedforward control using $\mathbf{v} = -\mathbf{H}_v^{-1}\mathbf{u}_{D0}$, which ideally would yield the energy-optimal control input. However, such a feedforward design is not robust and is limited by possible NMP zeros within $\mathbf{H}_v$ [77]. Therefore a feedback relationship

$$\mathbf{v} = \mathbf{H}_{fb}\mathbf{u}_D \tag{4.28}$$

is assumed as shown in Figure 4.1. With this assumption, $\mathbf{H}_{fb}$ can be designed using various MIMO controller synthesis methods. Here a representative H_∞ controller synthesis framework

[112] is illustrated in Figure 4.2, since infinity system norm marks the upper bound on the signal 2-norm gain. In Figure 4.2, $\mathbf{W}_d$ is a weighting filter describing the disturbance profile; $\mathbf{v}_w$ is signal $\mathbf{v}$ filtered by $\mathbf{W}_v$ to penalize the high frequency component; $\mathbf{U}_{0,d}$ represents the transfer function matrix from disturbance to the nominal control $\mathbf{u}_0$ given by

$$\mathbf{U}_{0,d} = \left(\mathbf{I} - \mathbf{C}_0\mathbf{G}\right)^{-1}\mathbf{C}_0\mathbf{G}_d. \tag{4.29}$$

Note that $\mathbf{U}_{0,d}$ is invariant with respect to both $\mathbf{P}$ and $\mathbf{H}_{fb}$, since the allocation process works in the null space and output is invariant. The equivalent plant marked in the blue box in Figure 4.2 is referred to as $\mathbf{L}_d$, i.e.

$$\begin{bmatrix} \mathbf{v}_w \\ \mathbf{u}_D \end{bmatrix} = \mathbf{L}_d\mathbf{d}. \tag{4.30}$$

Through feedback $\mathbf{H}_{fb}$, the infinity norm of $\mathbf{L}_d$ is minimized with standard H_∞ solvers. In cases where the allocator is expected to be designed without knowledge of the nominal controller $\mathbf{C}_0$, the frequency profile of $\mathbf{u}_{D0}$ can be assumed to be correlated with the disturbance frequency profile, i.e. the two blocks with dashed lines in Figure 4.2 are omitted and the sizes of $\mathbf{W}_d$ and $\mathbf{d}$ are adjusted to conform to the dimension of $\mathbf{u}_{D0}$.

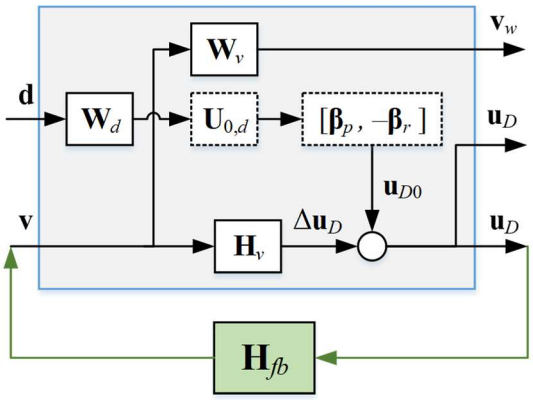

Figure 4.2: H_∞ Controller Synthesis of Control Allocator

4.4 Simulation Examples

Simulation examples from [68] are used to illustrate the proposed proxy-based control allocation method. System $\mathbf{G}$ is defined to be a 3-input, 1-output system with state-space realization

$$\left[\begin{array}{c|c}\mathbf{A} & \mathbf{B}\\ \hline \mathbf{C} & \mathbf{D}\end{array}\right]=\left[\begin{array}{cc|ccc} -0.157 & -0.094 & 0.87 & 0.253 & 0.743\\ -0.416 & -0.45 & 0.39 & 0.354 & 0.65\\ \hline 0 & 1 & 0 & 0 & 0\end{array}\right]. \tag{4.31}$$

As expressed in (4.1) and (4.3), the first $n_y = 1$ input channel formulates a non-redundant set and the transfer function matrix is divided as

$$\mathbf{G}=\left[\mathbf{G}_p \;\vdots\; \mathbf{G}_r\right]=\frac{\left[\begin{array}{c}0.39(s-0.77)\\ \hline 0.354(s-0.1403)\\ 0.65(s-0.3185)\end{array}\right]^{\mathrm{T}}}{s^2+0.607s+0.1098}. \tag{4.32}$$

Note that the system **G** is a stable system with NMP zeros in each control channel. With this matrix partition, $\mathbf{G}_p^{-1}\mathbf{G}_r$ is achieved and corresponding right coprime matrix fraction (**N** and **D**) defined in (4.10) is calculated as

$$\begin{aligned}\mathbf{G}_p^{-1}\mathbf{G}_r &= \left[\begin{array}{cc}\dfrac{0.90769(s-0.1403)}{s-0.77} & \dfrac{1.6667(s-0.3185)}{s-0.77}\end{array}\right]\\ &= \underbrace{\left[\begin{array}{cc}0.4793 & 0.90769(s-0.1403)\end{array}\right]}_{\mathbf{N}}\underbrace{\left[\begin{array}{cc}-1.3173 & s-0.77\\ 1 & 0\end{array}\right]^{-1}}_{\mathbf{D}}.\end{aligned} \tag{4.33}$$

With same $\mathbf{R} = \mathrm{diag}(100,1,1)$ defined in [68], the self-adjoint system is given by

$$\mathbf{\Pi}=\begin{bmatrix}\mathbf{N}\\ -\mathbf{D}\end{bmatrix}^{*}\mathbf{R}\begin{bmatrix}\mathbf{N}\\ -\mathbf{D}\end{bmatrix}=\begin{bmatrix}24.913 & 41.429(s-0.1203)\\ -41.429(s+0.1203) & -83.39(s+0.163)(s-0.163)\end{bmatrix} \tag{4.34}$$

and its spectral factorization is calculated as

$$\mathbf{\psi}=\begin{bmatrix}4.9913 & 0\\ -8.3002(s+0.1203) & 3.8075(s+0.2901)\end{bmatrix}. \tag{4.35}$$

Note that although the original system **G** contains NMP zeros, **Π** is always self-adjoint and a minimum phase spectral factor **ψ** can always be found to satisfy (4.22). Accordingly, $\boldsymbol{\beta}_p$ and $\boldsymbol{\beta}_r$ defined in (4.13) are given by

$$\begin{aligned}\boldsymbol{\beta}_p &= \left[\begin{array}{cc}9.4351 & -\dfrac{3.272s+0.8715}{s+0.2901}\end{array}\right]^{\mathrm{T}};\\ \boldsymbol{\beta}_r &= \begin{bmatrix}-0.2639 & 0.2\\ -\dfrac{0.838s+0.2717}{s+0.2901} & \dfrac{0.4368s+0.05252}{s+0.2901}\end{bmatrix},\end{aligned} \tag{4.36}$$

which are causal and stable, thus $\mathbf{u}_D$ is evaluated and regulated in real time. The same LQG controller as in [68] is used as the nominal controller $\mathbf{C}_0$. Without loss of generality, disturbance **d** defined in Figure 3.1 is assumed to be affecting output from the first control channel, i.e. $\mathbf{G}_d = \mathbf{G}(:,1)$. Here two nominal disturbance profiles are considered:

(a) Step unit disturbance: $\mathbf{d} = 1$ ($t > 0$);

(b) Sinusoidal disturbance at 1 Hz: $\mathbf{d} = 100\sin(2\pi t)$;

Accordingly, $\mathbf{W}_d$ is defined as an integrator with resonance term

$$\mathbf{W}_d(s) \triangleq \frac{1}{s+\varepsilon}\frac{\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2}, \tag{4.37}$$

where $\omega_0 = 2\pi$ rad/s (i.e. 1Hz), $\zeta = 0.1$ and $\varepsilon = 10^{-5}$ rad/s are used. In other words, the internal model principle [113] is employed in $\mathbf{W}_d$ to magnify the targeted frequency ranges. Weighting filter $\mathbf{W}_v$ in Figure 4.1 is defined as a high pass filter:

$$\mathbf{W}_v(s) \triangleq \frac{\omega_2}{\omega_1}\frac{s+\omega_1}{s+\omega_2}, \tag{4.38}$$

where $\omega_1 = 100\pi$ rad/s (i.e. 50 Hz) and $\omega_2 = 1000\pi$ rad/s (i.e. 500 Hz). In the same vein, $\boldsymbol{\psi}_0^{-1}$ in the allocator structure (Figure 4.2) is designed to be a diagonal transfer function consisting of identical third order low pass Butterworth filters each with a 10 Hz cutoff frequency. The designs of both $\mathbf{W}_v$ and $\boldsymbol{\psi}_0$ ensure that only the low frequency contents of the control efforts are redistributed within the allocator. This arises from practical robustness concerns as the model tends to be less accurate at higher frequencies.

Following the H_∞ design framework introduced in Figure 4.2, $\mathbf{H}_{fb}$ is designed without considering the controller dynamics, such that $\mathbf{L}_d$ maps the input $\mathbf{u}_{D0}$ to the output $[\mathbf{v}_w^{\mathrm{T}}, \mathbf{u}_D^{\mathrm{T}}]^{\mathrm{T}}$. This synthesis calculation is conducted with MATLAB 8® H_∞ synthesis tool. The open loop (OL) and closed loop (CL) Bode plots of $\mathbf{L}_d$ are illustrated in Figure 4.3. Note that, in open loop, input $\mathbf{u}_{D0}$ does not affect $\mathbf{v}_w$, and it only diagonally contributes to the proxy $\mathbf{u}_D$. These diagonal components are inherited from $\mathbf{W}_d$ defined in (4.37), and are the major contributors to $\mathbf{L}_d$ in open loop. The large open loop DC gains of $\mathbf{L}_d$ are flattened and the resonance is smoothed with closed-loop feedback $\mathbf{H}_{fb}$, indicating that the low frequency components as well as the resonant peak in $\mathbf{u}_{D0}$ are regulated.

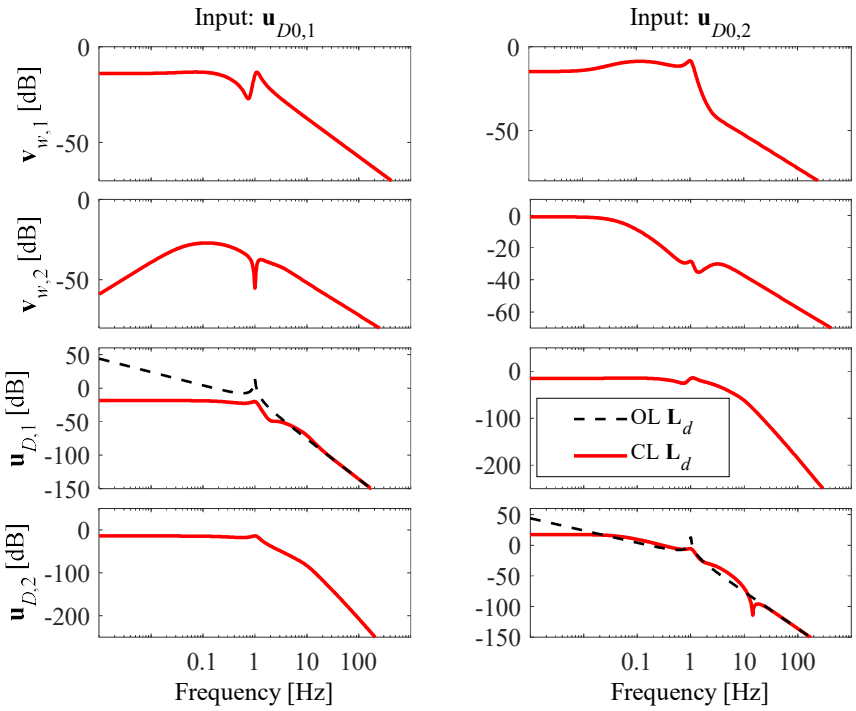

Figure 4.3: Equivalent System $\mathbf{L}_d$ in H_∞ Synthesis of $\mathbf{H}_{fb}$

The output and control inputs of the constant step disturbance (i.e. Case (a)) are shown in Figure 4.4. It is noticeable that the original LQG control input $\mathbf{u}_0$ has conflicting control inputs and does not yield an energy-efficient combination. Both the SMB allocator in [68] and the proposed proxy-based allocation method converge to an optimal solution where the heavily penalized u_1 is avoided. Notice that, compared to the SMB dynamic allocation proposed in [68], the proxy-based allocation method does not change the original system output while the SMB allocator in [68] introduces large deviations from the original system output. This difference fundamentally arises from the fact that the SMB allocation method in [68] employs statically defined null space while the proxy-based method both defines the null space and optimally allocates control efforts in broadband. This difference is further illustrated in Case (b) with sinusoidal disturbance, whose output and control inputs are shown in Figure 4.5. The heavily penalized u_1 is minimally redistributed by the SMB allocator in [68], while it is almost fully cancelled in the proposed proxy-based allocation scheme. The three input channels of the proposed proxy-based method work synergistically and thus greatly reduce the energy cost J_u without altering the controlled output.

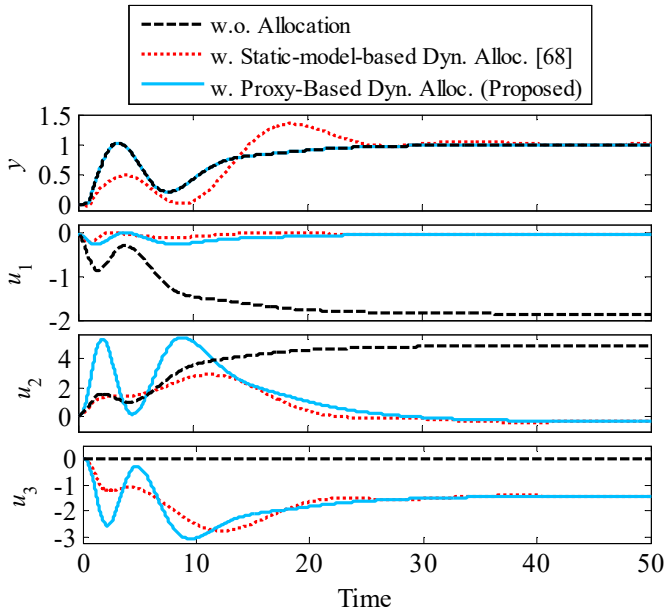

Figure 4.4: Control Allocation Results with Step Disturbance (Case (a))

To further quantify the performance of the allocators, signal $\Delta\mathbf{y}$ is defined as the output deviation from the case without allocation. Also, the steady state power consumption for each input P_1, P_2, P_3 and their combination P_{total} are defined by their contributions to the total J in unit time. $\overline{P}_{\text{total}}$ is the overall power averaged over the evaluated time horizon (i.e., including transients). The statistics for the deviations and power consumptions for Cases (a) and (b) are listed in Table 4.1. It is shown that the SMB allocator in [68] can introduce severe output deviation, especially for the step disturbance in Case (a) where the disturbance is not continuous, while the proxy-based method's output deviations are negligible for both cases. In Case (a), both allocators consume less than 1% of the overall average power consumption of the nominal controller. Note that the overall average power consumption ($\overline{P}_{\text{total}}$) of the proposed allocator is a bit higher than that of the SMB allocator in [68] because of the additional effort it takes to keep the output unaltered during transients; the SMB allocator in [68] is unable to maintain the desired output during transients hence it consumes less power. However, both of the allocators converge to same static optimal control as shown in Figure 4.4, resulting less than 2% steady state power difference (as observed from P_{total}). The benefit of the proposed allocator is more pronounced in Case (b), where the disturbance signal is more dynamic. The steady state power consumption of the SMB allocator in [68] over the no allocator case is 20% less, while the proposed proxy based allocator provides 99% less steady state power consumption than the no allocator case, due to its capability to optimally redistribute dynamic control efforts (at non-zero frequencies). Note that

in Case (b) the proposed allocator introduces more significant transients in some of the control efforts. This arises from the presence of NMP zeros in the controlled system, coupled with the relatively fast allocator dynamics. Even with these transients, the proposed allocator is capable of achieving 58% less overall average power consumption than the SMB allocator in [68], based on $\overline{P}_{\text{total}}$.

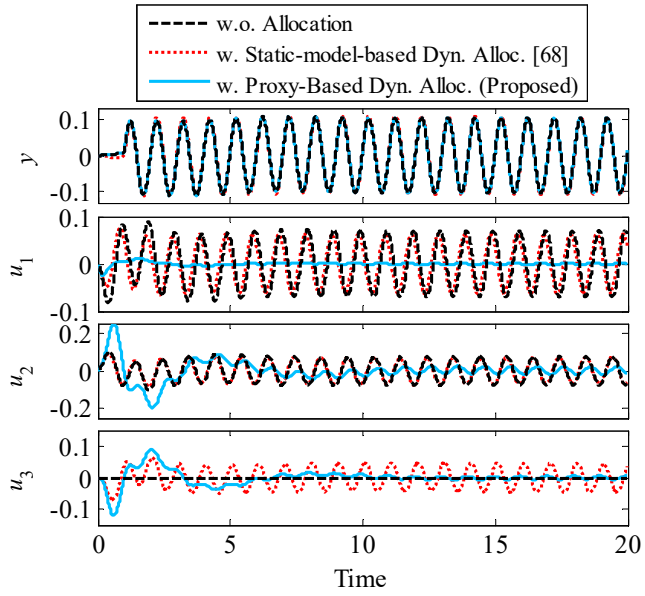

Figure 4.5: Control Allocation Results with Sinusoidal Disturbance (Case (b))

Table 4.1: Output Deviation and Steady State Power Consumption Comparison

		max $\lvert\Delta \mathbf{y}\rvert$	$\overline{P}_{\text{total}}$	Steady State Power			
				P_1	P_2	P_3	P_{total}
(a)	Without allocation	0	263	309	21	0	331
	SMB Allocator in [68]	1.00	5.2	0.04	0.08	1.92	2.04
	PB allocator	3×10^{-5}	8.88	0.07	0.05	1.94	2.06
(b)	Without allocation	0	14.87	9.90	0.11	0	10.02
	SMB Allocator in [68]	0.46	8.52	7.68	0.20	0.08	7.97
	PB allocator	1×10^{-5}	3.56	0.01	0.06	0.01	0.08

4.5 Summary

The concept of optimal control ratio for dual-input, single-output systems introduced in Chapter 2 is extended to optimal control subspace for multi-input, multi-output (MIMO) weakly input redundant systems in this chapter. The proxy-based control allocation introduced in Chapter 3 is also extended to MIMO systems using matrix fraction description and spectral factorization. The causality and accurate measurement of deviation from optimality are maintained in the proposed proxy signal for MIMO systems. The control allocation problem is converted into a regulation problem, and is solved with standard H_∞ synthesis tools. The proposed proxy-based control allocation is compared with a state-of-art static-model-based dynamic allocation method in simulation studies. Significant improvements in energy efficiency without affecting system outputs are observed, especially under the influence of broadband disturbances.

Chapter 5 Connections between Energy-Optimal Control Allocation and Linear Quadratic Control

5.1 Overview

In this chapter, the relationship between energy-optimal control allocation developed in Chapter 3 and Chapter 4, and the traditional linear quadratic (LQ) control approach discussed in Chapter 1 is explored under weak input redundancy. In Section 5.2, the energy-optimal control allocation approach developed in the preceding chapters is formalized under the name of optimal control subspace-based (OCS) control allocation, and some background on LQ control are provided. In Section 5.3, it is shown that LQ control is a special case of energy-optimal control allocation, and yield identical solutions, under perfectly known exogenous disturbance and reference inputs. However, for most practical cases in which the disturbance is unknown, it is shown that both methods need a certain level of approximation. A Kalman filter is assumed for the LQ approach while zero magnitude/phase approximation of control allocation approach is assumed. In Section 5.4, the OCS control allocation method is shown to be superior to the LQ control because it is more explicit about its approximations, and a numerical example is provided to justify this position. A brief summary is provided in Section 5.5.

5.2 Background

Following the definition in (1.2) and (4.1), a redundant MIMO over-actuated system is defined as $\mathbf{y}=\mathbf{G}\mathbf{u} + \mathbf{G}_d\mathbf{d}$, where n_y, n_u, n_d represent the dimension of the output, control input and disturbance, respectively. In this section we assume the system to be strictly proper, such that one minimal state-space realization of the system is given by

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{B}_d\mathbf{d}; \\ \mathbf{y} &= \mathbf{C}\mathbf{x}, \end{aligned} \tag{5.1}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$ is the state vector of the system. Note that the assumption of a strictly proper system eliminates the direct feedthrough terms in the state-space representation (i.e. $\mathbf{D} = \mathbf{D}_d = \mathbf{0}$ compared with (1.3)). This assumption simplifies the expression, and is consistent with the later-

discussed linear quadratic (LQ) tracking controller. Accordingly, the condition of strong input redundancy becomes

$$\mathrm{Ker}\left(\mathbf{B}\right) \neq \varnothing; \tag{5.2}$$

while the condition for weakly input redundancy is same as in (1.5). The relationship between the state-space representation to the transfer function is given by

$$\mathbf{G} = \mathbf{C}\left(s\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{B};\ \mathbf{G}_d = \mathbf{C}\left(s\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{B}_d, \tag{5.3}$$

Same as defined in (4.2), a positive integer $n_r = n_u - n_y$, is used to indicate the system's redundancy degree for weakly input redundant system.

5.2.1 Optimal Control Subspace-Based (OCS) Control Allocation

The methods developed in the preceding chapters for energy-optimal control allocation for DISO then MIMO systems can be referred to as optimal control subspace-based (OCS) control allocation. The overarching idea in OCS control allocation is to decompose a weakly input redundant system **G** and its quadratic energy cost **R** according to n_y-dimension principal component and n_r-dimension redundant component, i.e.

$$\mathbf{G} = \begin{bmatrix}\mathbf{G}_p & \mathbf{G}_r\end{bmatrix};\ \mathbf{u} = \begin{bmatrix}\mathbf{u}_p^{\mathrm{T}} & \mathbf{u}_r^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}},\ \mathbf{R} = \begin{bmatrix}\mathbf{R}_p & \mathbf{R}_{pr} \\ \mathbf{R}_{pr}^{\mathrm{T}} & \mathbf{R}_r\end{bmatrix}. \tag{5.4}$$

In order to achieve energy optimality, the optimal control should lie within the OCS, given by

$$\begin{aligned}&\left[\left(\mathbf{G}_p^{-1}\mathbf{G}_r\right)^{*}\mathbf{R}_p - \mathbf{R}_{pr}^{\mathrm{T}}\right]\hat{\mathbf{u}}_p = \left[\mathbf{R}_r - \left(\mathbf{G}_p^{-1}\mathbf{G}_r\right)^{*}\mathbf{R}_{pr}\right]\hat{\mathbf{u}}_r \\ &\Leftrightarrow \left[\left(\mathbf{G}_p^{-1}\mathbf{G}_r\right)^{*} \quad -\mathbf{I}\right]\mathbf{R}\hat{\mathbf{u}} = \mathbf{0},\end{aligned} \tag{5.5}$$

where * represents the adjoint operation [114] and $\hat{}$ indicates optimality. The methods based on causal approximation of the OCR (Chapter 2) or proxy-based allocation (Chapters 3 and 4) can therefore be considered as specific realizations of OCS control allocation.

5.2.2 LQ controller (with reference tracking and disturbance rejection)

Besides control allocation, another natural thought in handling over-actuated system is to formulate the problem under the LQ control framework.

Problem 1. Assume that $\mathbf{y}_d$ represents the desired output of the system defined in (1.2) and (1.3), find optimal **u** such that the quadratic objective on tracking performance and control energy, given by

$$J_{uy} = \frac{1}{2}\int_0^{+\infty} \left(\left(\mathbf{y} - \mathbf{y}_d\right)^{\mathrm{T}} \mathbf{Q}_y \left(\mathbf{y} - \mathbf{y}_d\right) + \mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u} \right) dt, \tag{5.6}$$

is minimized.

Remark 4. The formulation of Problem 1 agrees with the framework of weak input redundancy since it does not involve the internal states in the objective. However, the results from the general LQ formulation based on system states are amenable to Problem 1 by defining

$$\mathbf{Q} \triangleq \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\mathbf{C}, \tag{5.7}$$

such that $\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}=\mathbf{y}^{\mathrm{T}}\mathbf{Q}_y\mathbf{y}$.

Theorem 1. Assume the disturbance $\mathbf{d}$ and desired output $\mathbf{y}_d$ are generated from exo-systems

$$\dot{\mathbf{d}} = \mathbf{E}\mathbf{d};\ \ \dot{\mathbf{y}}_d = \mathbf{F}\mathbf{y}_d. \tag{5.8}$$

where $\mathbf{E}$ and $\mathbf{F}$ are square matrices with no eigenvalue on the right half-plane. Assume matrix $\mathbf{A}$ is Hurwitz, the optimal solution $\hat{\mathbf{u}}$ to Problem 1 satisfies following feedback and feedforward structure:

$$\begin{aligned} \hat{\mathbf{u}} &= -\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{p}; \\ \mathbf{p} &= \mathbf{P}_1\mathbf{x} + \mathbf{P}_2\mathbf{d} + \mathbf{P}_3\mathbf{y}_d, \end{aligned} \tag{5.9}$$

where the $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ satisfy the Riccati equation (5.10) and Sylvester equations (5.11)(5.12) as

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}_1 + \mathbf{P}_1\mathbf{A} - \mathbf{P}_1\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}_1 + \mathbf{Q} = \mathbf{0}. \tag{5.10}$$

$$\mathbf{A}_{cl}^{\mathrm{T}}\mathbf{P}_2 + \mathbf{P}_2\mathbf{E} = -\mathbf{P}_1\mathbf{B}_d; \tag{5.11}$$

$$\mathbf{A}_{cl}^{\mathrm{T}}\mathbf{P}_3 + \mathbf{P}_3\mathbf{F} = \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y. \tag{5.12}$$

Matrix $\mathbf{A}_{cl}$ is part of the closed loop state-space model, i.e.

$$\mathbf{A}_{cl} \triangleq \mathbf{A} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}_1. \tag{5.13}$$

The solution to Sylvester equations (5.11) and (5.12) are unique if $\mathbf{A}_{cl}$ shares no common eigenvalues with $\mathbf{E}$ and $\mathbf{F}$.

Remark 5. Theorem 1, which applies to infinite length $\mathbf{d}$ and $\mathbf{y}_d$ in continuous systems, is a natural extension of the optimal control of finite length $\mathbf{d}$ and $\mathbf{y}_d$, as discussed in [60,115].

5.3 Equivalence of OCS Control Allocation and LQ Control with Known Disturbance Dynamics

In this section, it is shown that the LQ solution for a weakly input redundant system, including general external disturbances and desired output tracking, formulates a two-stage stage structure as in control allocation. It is also shown that the optimal state feedback in (5.9) also yields the two-stage control allocation structure.

Theorem 2. The optimal solution $\hat{\mathbf{u}}$ to Problem 1 satisfies the following two-stage control allocation structure:

$$\hat{\mathbf{u}} = \mathbf{R}^{-1}\mathbf{G}^{*}\boldsymbol{\tau}; \tag{5.14}$$

$$\boldsymbol{\tau} = \mathbf{Q}_y\left(\mathbf{y}_d - \mathbf{y}\right). \tag{5.15}$$

Proof. According to Pontryagin's minimum principle [116], the Hamiltonian of the system is written as

$$H = \frac{1}{2}\left(\left(\mathbf{x}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} - \mathbf{y}_d^{\mathrm{T}}\right)\mathbf{Q}_y\left(\mathbf{C}\mathbf{x} - \mathbf{y}_d\right) + \mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}\right) + \mathbf{p}^{\mathrm{T}}\left(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{B}_d\mathbf{d}\right), \tag{5.16}$$

where $\mathbf{p}$ is a vector of co-states. The co-states are crucial to the optimal solution because

$$\frac{\partial H}{\partial \mathbf{u}}\left(\hat{\mathbf{u}}\right) = 0 \Rightarrow \mathbf{R}\hat{\mathbf{u}} + \mathbf{B}^{\mathrm{T}}\mathbf{p} = \mathbf{0} \Rightarrow \hat{\mathbf{u}} = -\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{p}. \tag{5.17}$$

The combination of state and co-state dynamics is written as

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix}\mathbf{x}\\ \mathbf{p}\end{bmatrix} &= \begin{bmatrix} \dfrac{\partial H}{\partial \mathbf{p}} \\ -\dfrac{\partial H}{\partial \mathbf{x}} \end{bmatrix} = \mathbf{M}\begin{bmatrix}\mathbf{x}\\ \mathbf{p}\end{bmatrix} + \begin{bmatrix}\mathbf{B}_d\\ \mathbf{0}\end{bmatrix}\mathbf{d} + \begin{bmatrix}\mathbf{0}\\ \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\end{bmatrix}\mathbf{y}_d, \\ \mathbf{M} &\triangleq \begin{bmatrix} \mathbf{A} & -\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}} \\ -\underbrace{\mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\mathbf{C}}_{\mathbf{Q}} & -\mathbf{A}^{\mathrm{T}} \end{bmatrix}. \end{aligned} \tag{5.18}$$

Using Laplace transform, both $\mathbf{x}$ and $\mathbf{p}$ can be expressed as outputs relative to $\mathbf{d}$ and $\mathbf{y}_d$. The co-state is expressed as

$$\mathbf{p} = \begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{p}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{N}\left(\begin{bmatrix}\mathbf{B}_d\\ \mathbf{0}\end{bmatrix}\mathbf{d} + \begin{bmatrix}\mathbf{0}\\ \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\end{bmatrix}\mathbf{y}_d\right), \tag{5.19}$$

where $\mathbf{N} \triangleq (s\mathbf{I}-\mathbf{M})^{-1}$. In $\mathbf{N}$, only N_{21} and N_{22} contribute to the co-states, i.e.

$$\mathbf{p} = N_{21}\mathbf{B}_2\mathbf{d} + N_{22}\mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\mathbf{y}_d, \tag{5.20}$$

which can be analytically determined by the block-wise matrix inversion formula [117] as

$$\begin{aligned} N_{21} &= -\left(s\mathbf{I} + \mathbf{A}^{\mathrm{T}}\right)^{-1}\mathbf{Q}\left(s\mathbf{I} - \mathbf{A} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\left(s\mathbf{I} + \mathbf{A}^{\mathrm{T}}\right)^{-1}\mathbf{Q}\right)^{-1}; \\ N_{22} &= \left(s\mathbf{I} + \mathbf{A}^{\mathrm{T}} - \mathbf{Q}\left(s\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}. \end{aligned} \tag{5.21}$$

Using the relationship in (5.5) and the definition of the adjoint of system dynamics [103]

$$\mathbf{G}^{*} = -\mathbf{B}^{\mathrm{T}}\left(s\mathbf{I} + \mathbf{A}^{\mathrm{T}}\right)^{-1}\mathbf{C}^{\mathrm{T}}, \tag{5.22}$$

the optimal control in (5.17) yields

$$\begin{aligned} \hat{\mathbf{u}} &= -\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\left(s\mathbf{I} + \mathbf{A}^{\mathrm{T}} - \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\mathbf{G}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\mathbf{y}_d \\ &\quad -\mathbf{R}^{-1}\mathbf{G}^{*}\mathbf{Q}_y\mathbf{C}\left(s\mathbf{I} - \mathbf{A} + \mathbf{B}\mathbf{R}^{-1}\mathbf{G}^{*}\mathbf{Q}_y\mathbf{C}\right)^{-1}\mathbf{B}_d\mathbf{d}. \end{aligned} \tag{5.23}$$

Note that the inversions in (5.23) can be simplified using the Woodbury matrix identity formula, i.e.

$$\begin{aligned} \hat{\mathbf{u}} &= \mathbf{R}^{-1}\mathbf{G}^{*}\boldsymbol{\tau}; \\ \boldsymbol{\tau} &= -\mathbf{Q}_y\left(\mathbf{I} + \mathbf{G}_\tau\mathbf{Q}_y\right)^{-1}\mathbf{G}_d\mathbf{d} + \left(\mathbf{I} + \mathbf{Q}_y\mathbf{G}_\tau\right)^{-1}\mathbf{Q}_y\mathbf{y}_d, \end{aligned} \tag{5.24}$$

where $\boldsymbol{\tau} \in \mathbb{R}^{n_y}$ is defined as the virtual control input and

$$\mathbf{G}_\tau \triangleq \mathbf{G}(s)\mathbf{R}^{-1}\mathbf{G}^{*}(s). \tag{5.25}$$

Replacing $\mathbf{G}_d\mathbf{d}$ with $\mathbf{y}–\mathbf{G}_\tau\boldsymbol{\tau}$ in (5.24) and applying Woodbury matrix identity formula

$$\begin{aligned} &\left(\mathbf{I} - \mathbf{Q}_y\left(\mathbf{I} + \mathbf{G}_\tau\mathbf{Q}_y\right)^{-1}\mathbf{G}_\tau\right)\boldsymbol{\tau} = \left(\mathbf{I} + \mathbf{Q}_y\mathbf{G}_\tau\right)^{-1}\boldsymbol{\tau} = -\mathbf{Q}_y\left(\mathbf{I} + \mathbf{G}_\tau\mathbf{Q}_y\right)^{-1}\mathbf{y} + \left(\mathbf{I} + \mathbf{Q}_y\mathbf{G}_\tau\right)^{-1}\mathbf{Q}_y\mathbf{y}_d \\ &\Rightarrow \boldsymbol{\tau} = -\left(\mathbf{I} + \mathbf{Q}_y\mathbf{G}_\tau\right)\mathbf{Q}_y\left(\mathbf{I} + \mathbf{G}_\tau\mathbf{Q}_y\right)^{-1}\mathbf{y} + \mathbf{Q}_y\mathbf{y}_d = \mathbf{Q}_y\left(\mathbf{y}_d - \mathbf{y}\right). \end{aligned} \tag{5.26}$$

This is identical to the expression shown in (5.14) and (5.15). ■

Theorem 3. The feedback/feedforward solution specified in Theorem 1 satisfies the control allocation structure in Theorem 2.

Proof. The state dynamics of the corresponding system in closed loop is given by

$$\dot{\mathbf{x}} = \mathbf{A}_{cl}\mathbf{x} + \mathbf{B}_{cl,d}\mathbf{d} + \mathbf{B}_{cl,y}\mathbf{y}_{\mathrm{d}}, \tag{5.27}$$

where

$$\begin{aligned}\mathbf{B}_{cl,d} &= \mathbf{B}_d - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}_2;\\ \mathbf{B}_{cl,y} &= -\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}_3.\end{aligned} \tag{5.28}$$

The state dynamics of this system is expressed as

$$\begin{aligned}\mathbf{x}(t) &= \int_0^t e^{\mathbf{A}_{cl}(t-\tau)}\left(\mathbf{B}_{cl,d}\mathbf{d}(\tau) + \mathbf{B}_{cl,y}\mathbf{y}_d(\tau)\right)d\tau\\ &= \left(\mathbf{P}_4 e^{\mathbf{E}t} - e^{\mathbf{A}_{cl}t}\mathbf{P}_4\right)\mathbf{d}_0 + \left(\mathbf{P}_5 e^{\mathbf{F}t} - e^{\mathbf{A}_{cl}t}\mathbf{P}_5\right)\mathbf{y}_{d0},\end{aligned} \tag{5.29}$$

where $\mathbf{d}_0$ and $\mathbf{y}_{d0}$ are the initial values of $\mathbf{d}$ and $\mathbf{y}_d$, respectively; $\mathbf{P}_4$ and $\mathbf{P}_5$ are matrices satisfying following Sylvester equations

$$\mathbf{P}_4\mathbf{E} - \mathbf{A}_{cl}\mathbf{P}_4 = \mathbf{B}_{cl,d}; \tag{5.30}$$

$$\mathbf{P}_5\mathbf{F} - \mathbf{A}_{cl}\mathbf{P}_5 = \mathbf{B}_{cl,y}. \tag{5.31}$$

Accordingly the optimal control is expressed as

$$\mathbf{u} = -\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\left(\left((\mathbf{P}_1\mathbf{P}_4 + \mathbf{P}_2)e^{\mathbf{E}t} - \mathbf{P}_1 e^{\mathbf{A}_{cl}t}\mathbf{P}_4\right)\mathbf{d}_0 + \left((\mathbf{P}_1\mathbf{P}_5 + \mathbf{P}_3)e^{\mathbf{F}t} - \mathbf{P}_1 e^{\mathbf{A}_{cl}t}\mathbf{P}_5\right)\mathbf{y}_{d0}\right). \tag{5.32}$$

From the control allocation structure, the adjoint system can be evaluated through convolution as

$$\begin{aligned}&\mathbf{R}^{-1}\mathbf{G}^{*}\mathbf{Q}_y\left(\mathbf{y}_d - \mathbf{y}\right)\\ &= \mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}e^{-\mathbf{A}^{\mathrm{T}}t}\int_t^{+\infty} e^{\mathbf{A}^{\mathrm{T}}\sigma}\mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\left(\mathbf{y}_d(\sigma) - \mathbf{C}\mathbf{x}(\sigma)\right)d\sigma\\ &= \mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}e^{-\mathbf{A}^{\mathrm{T}}t}\{\int_t^{+\infty} e^{\mathbf{A}^{\mathrm{T}}\sigma}\left(\mathbf{Q}e^{\mathbf{A}_{cl}\sigma}\mathbf{P}_4 - \mathbf{Q}\mathbf{P}_4 e^{\mathbf{E}\sigma}\right)d\sigma\mathbf{d}_0\\ &\quad + \int_t^{+\infty} e^{\mathbf{A}^{\mathrm{T}}\sigma}\left(\left(\mathbf{Q}\mathbf{P}_5 - \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y\right)e^{\mathbf{F}s} - \mathbf{Q}e^{\mathbf{A}_{cl}s}\mathbf{P}_5\right)d\sigma\mathbf{y}_{d0}\}.\end{aligned} \tag{5.33}$$

This integration can be simplified with two equations

$$\mathbf{A}^{\mathrm{T}}\left(\mathbf{P}_1\mathbf{P}_4 + \mathbf{P}_2\right) + \left(\mathbf{P}_1\mathbf{P}_4 + \mathbf{P}_2\right)\mathbf{E} = -\mathbf{Q}\mathbf{P}_4; \tag{5.34}$$

$$\mathbf{A}^{\mathrm{T}}\left(\mathbf{P}_1\mathbf{P}_5 + \mathbf{P}_3\right) + \left(\mathbf{P}_1\mathbf{P}_5 + \mathbf{P}_3\right)\mathbf{F} = -\mathbf{Q}\mathbf{P}_5 + \mathbf{C}^{\mathrm{T}}\mathbf{Q}_y, \tag{5.35}$$

which can be conveniently shown by the Riccati equation (5.10) and Sylvester equations (5.11), (5.12), (5.30), and (5.31). Accordingly,

$$\frac{d}{dt}\left(e^{\mathbf{A}^{\mathrm{T}}t}\left(\mathbf{P}_1\mathbf{P}_4 + \mathbf{P}_2\right)e^{\mathbf{E}t}\right) = -e^{\mathbf{A}^{\mathrm{T}}t}\mathbf{Q}\mathbf{P}_4 e^{\mathbf{E}t}; \tag{5.36}$$

$$\frac{d}{dt}\left(e^{\mathbf{A}^{\mathrm{T}}t}\left(\mathbf{P}_1\mathbf{P}_5 + \mathbf{P}_3\right)e^{\mathbf{F}t}\right) = e^{\mathbf{A}^{\mathrm{T}}t}\left(\mathbf{C}^{\mathrm{T}}\mathbf{Q}_y - \mathbf{Q}\mathbf{P}_5\right)e^{\mathbf{F}t}. \tag{5.37}$$

These relationships simplify the integration in (5.33) and lead to

$$\hat{\mathbf{u}} = \mathbf{R}^{-1}\mathbf{G}^{*}\mathbf{Q}_y\left(\mathbf{y}_d - \mathbf{y}\right), \tag{5.38}$$

which is the structure specified in Theorem 2. ■

Remark 6. The expression for $\hat{\mathbf{u}}$ in (5.14) satisfies the optimal control subspace specified in (5.5). Accordingly, Theorem 3 indicates that that the LQ solution with perfectly known disturbance and reference models satisfies the two-stage structure of OCS control allocation. The only difference between the two frameworks is that the LQ control technique has a specific way of generating $\boldsymbol{\tau}$, while in OCS control allocation (and energy-optimal control allocation, in general) is open to many different ways of generating $\boldsymbol{\tau}$, according to different positioning objectives. Accordingly, the LQ control is a specific subset of OCS control allocation where $\boldsymbol{\tau}$ is generated as in (5.15) under quadratic tracking cost.

Remark 7. The OCS control allocation technique can be regarded as the solution to a more general objective J'_{uy} with incremental cost H_y on $\mathbf{y}$ and $\mathbf{y}_d$, i.e.

$$J'_{uy} = \frac{1}{2}\int_0^{+\infty} \left(H_y\left(\mathbf{y}, \mathbf{y}_d\right) + \mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u} \right) dt. \tag{5.39}$$

Similar to derivations in Theorem 2, the two-stage framework is preserved in its optimal solution structure as

$$\begin{aligned} \hat{\mathbf{u}} &= \mathbf{R}^{-1}\mathbf{G}^{*}\boldsymbol{\tau}; \\ \boldsymbol{\tau} &= -\frac{\partial H_y}{\partial \mathbf{y}}. \end{aligned} \tag{5.40}$$

This generalization of the linear case in (5.14) and (5.15) indicates that, in OCS control allocation, the virtual control input $\boldsymbol{\tau}$ can be generated in various ways to meet different control performance requirements.

5.4 Differences between LQ and OCS Control Allocation with Unknown Disturbance

While Theorem 3 has shown that $\hat{\mathbf{u}}$ for LQ control (5.9) is identical to $\hat{\mathbf{u}}$ for OCS control allocation (5.14) and (5.15), they approach the computation of $\hat{\mathbf{u}}$ from different perspectives. The LQ control technique depends on perfect knowledge/model of internal states $\mathbf{x}$, disturbance $\mathbf{d}$ and reference $\mathbf{y}_d$ through an exo-system. On the other hand, OCS control allocation requires the determination of $\hat{\mathbf{u}}$ from $\mathbf{G}^{*}$. Both of these approaches come with practical challenges. The operator $\mathbf{G}^{*}$ is in general non-causal, hence it cannot be evaluated going forward in time. Therefore, in practice, it is typically approximated in one form or the other. The Static-model-

based dynamic control allocation method in [68] can be interpreted as an approximation guaranteeing convergence to the DC gain of $\mathbf{G}^*$. Alternatively, OCS control allocation strives to approximate $\mathbf{G}^*$ in broadband using, for instance, a causal approximation of $\mathbf{G}^*$ in Chapter 2 and a proxy-based approximation of $\mathbf{G}^*$ through feedforward and/or feedback regulation as discussed in Section 3.3 and Section 4.3. Similarly, in practice, the internal states of system, disturbance and reference may not be available for use in the LQ technique. In such a scenario, observers are utilized to estimate (approximate) these states.

Remark 8. Even though both frameworks require some sort of approximation, the closeness of the approximate solutions to the optimal solution depends on how well they each approximate $\mathbf{G}^*$. In this regard, one can argue that the OCS control allocation framework is superior to the LQ framework because it is explicit about $\mathbf{G}^*$ as the target of approximation while LQ indirectly approximates $\mathbf{G}^*$ through observers.

Remark 9. While accurately approximating $\mathbf{G}^*$ is an obvious measure of nearness to optimality, it is shown in Section 3.3 and Section 4.3 that deviation from the optimal subspace defined in (5.5) is a more-straightforward measure of nearness to optimality in the OCS control allocation framework; specifically, it is shown that the two-norm squared of the deviation from the optimal subspace is directly proportional to the deviation of J_u from its optimal value. However, the optimal subspace is also non-causal, in general, so it needs to be approximated.

The superiority of the OCS control allocation technique relative to the LQ control technique in accurately approximating the optimal solution (i.e., the deviation from the optimal subspace) is demonstrated in the following example.

Example 1. Consider a dual-input, single-output (DISO) system derived from a truncation of the 3-input system in [68]. The state space representation is

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right] = \left[\begin{array}{cc|cc} -0.157 & -0.094 & 0.87 & 0.253 \\ -0.416 & -0.45 & 0.39 & 0.354 \\ \hline 0 & 1 & 0 & 0 \end{array}\right] \tag{5.41}$$

with disturbance term $\mathbf{B}_d\mathbf{d}$ being zero-mean white noise with covariance $\mathbf{Q}_n$. Assume that $\mathbf{y}_d = \mathbf{0}$ and the output $\mathbf{y}$ also suffers from zero mean white noise with covariance $\mathbf{R}_n$; also assume that $\mathbf{R}$ = diag[1,2] and $\mathbf{Q}_y$ = 1. In this context, the Kalman filter with feedback gain $\mathbf{L}_{kf}$ is usually adopted as an optimal observer for LQ control because it minimizes the covariance of state estimation error [58], i.e.

$$\dot{\mathbf{x}}_o = \left(\mathbf{A} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}_1\right)\mathbf{x}_o + \mathbf{L}_{kf}\left(\mathbf{y} - \mathbf{C}\mathbf{x}_o\right), \tag{5.42}$$

where $\mathbf{x}_o$ is the observed state $\mathbf{x}$. This observer essentially establishes a relationship from $\mathbf{y}$ to $\mathbf{u}$, given by the transfer function

$$\mathbf{G}_{kf} \triangleq -\mathbf{B}^{\mathrm{T}}\mathbf{P}_1\left(s\mathbf{I} - \mathbf{A} + \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}_1 + \mathbf{L}_{kf}\mathbf{C}\right)^{-1}\mathbf{L}_{kf}\mathbf{Q}_y^{-1}, \tag{5.43}$$

such that $\mathbf{u}(s) = -\mathbf{R}^{-1}\mathbf{G}_{kf}\mathbf{Q}_y\mathbf{y}(s)$, which is in the same form as the control allocation structure specified in Theorem 2. Accordingly, $\mathbf{G}_{kf}$ is an approximation of $\mathbf{G}^*$ from the realization through an optimal state observer.

From the OCS control allocation viewpoint, one obvious way to approximate $\mathbf{G}^*$ is to use $\mathbf{G}^{\mathrm{T}}$ (as in Section 2.3.4). This is an approximation which guarantees zero magnitude error in the approximation of $\mathbf{G}^*$; thus it is referred to, here, as the zero magnitude error approximation of $\mathbf{G}^*$ (denoted by $\mathbf{G}_{zm}$); i.e.,

$$\left|\mathbf{G}_{zm}(j\omega)\right| = \left|\mathbf{G}^*(j\omega)\right|. \tag{5.44}$$

Another sensible approximation of $\mathbf{G}^*$ arises from the zero phase error tracking controller [95], commonly used for approximating the inversion of non-minimum phase systems. It has the form

$$\mathbf{G}_{zp}(s) = \frac{\mathbf{\Pi}_G^{\mathrm{T}}(-s)\,d_G(s)\,d_G^{-2}(0)}{d_{lpf}(s)}, \tag{5.45}$$

where d_{lpf} is a low-pass type denominator while pair ($\mathbf{\Pi}_G$, d_G) is a Smith-Mcmillan form of $\mathbf{G}$ [110] satisfying

$$\mathbf{G}(s) = \frac{\mathbf{\Pi}_G(s)}{d_G(s)}, \tag{5.46}$$

such that $\mathbf{G}_{zp}(0) = \mathbf{G}^*(0)$, and $\angle\mathbf{G}_{zp}(j\omega) \approx \angle\mathbf{G}^*(j\omega)$ where the phase distortion introduced by the low-pass type denominator d_{lpf} is minimal.

Remark 10. The zero magnitude approximation, $\mathbf{G}_{zm}$, and the zero phase approximation, $\mathbf{G}_{zp}$, represent two general approximation methods of $\mathbf{G}^*$ in the absence of any knowledge of $\mathbf{d}$ or $\mathbf{y}_d$. More tailored approximations of $\mathbf{G}^*$ can be achieved, given more information on these signals, as in the proxy-based approach to OCS control allocation discussed in Chapters 3 and 4. Note that these approximations $\mathbf{G}_{zm}$ and $\mathbf{G}_{zp}$ do not guarantee that the closed loop system following original structure in (5.38) is causally stable. Therefore in practical implementations, additional virtual control components are needed.

Note from Remark 9 that the deviation from the optimal control subspace, compared to the absolute approximation error of $\mathbf{G}^*$, is a more straightforward way of determining nearness to optimality. In this DISO system example, the deviation from the optimal control subspace can be evaluated from the control input ratio between the two redundant inputs; i.e., $\mathbf{u}(1)/\mathbf{u}(2)$. Figure 5.1 compares the control ratio calculated based on $\mathbf{G}_{kf}$, $\mathbf{G}_{zm}$, and $\mathbf{G}_{zp}$ to that calculated using $\mathbf{G}^*$. For $\mathbf{G}_{kf}$, $\mathbf{R}_n = \mathbf{I}$ is used, while $\mathbf{Q}_n$ varies from $0.01\mathbf{I}$ to $100\mathbf{I}$.

Notice that control ratio with $\mathbf{G}_{zm}$ matches the gain of $\mathbf{G}^*$ exactly, while the phase is reversed, while $\mathbf{G}_{zp}$ matches both magnitude and phase of $\mathbf{G}^*$ exactly. The perfect match with $\mathbf{G}_{zp}$ arises from the fact the magnitude deviation of $\mathbf{G}^*$ is cancelled in calculating the ratio, i.e. $|\mathbf{G}_{zp}(1)|/|\mathbf{G}_{zp}(2)| = |\mathbf{G}^*(1)|/|\mathbf{G}^*(2)|$ even though $|\mathbf{G}_{zp}| \neq |\mathbf{G}^*|$. However, notice that the control ratios in all instances of $\mathbf{G}_{kf}$ deviate significantly from those of $\mathbf{G}^*$, in both magnitude and phase. The implication of such large deviations in terms of optimal control allocation is huge. For instance, a phase error of around 180° is observed at most frequencies of the ratio calculated using $\mathbf{G}_{kf}$. This implies that the two actuators would be in conflict with each other and ruin the energy performance at most frequencies. Therefore, while the Kalman filter is an optimal state observer, in this example, acts against the optimal alignment indicated by $\mathbf{G}^*$, hence yields far-from-optimal control allocation.

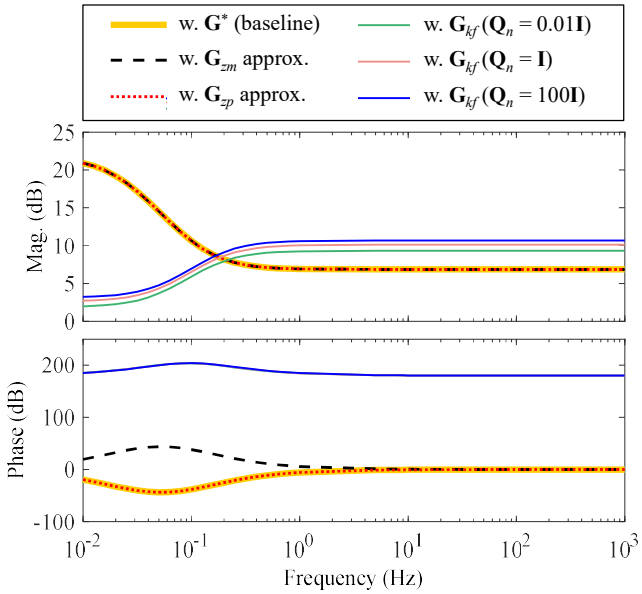

Figure 5.1: Comparison of Control Ratio ($\mathbf{u}(1)/\mathbf{u}(2)$) Calculated using $\mathbf{G}^*$, and its Zero-Magnitude ($\mathbf{G}_{zm}$), Zero-Phase ($\mathbf{G}_{zp}$) and Kalman Filter ($\mathbf{G}_{kf}$) Approximations.

5.5 Summary

The connections between the control allocation and linear quadratic (LQ) control frameworks for optimally distributing control inputs in weakly input redundant systems are explored. It is analytically shown that, for a representative class of exogenous disturbance and references, the LQ control technique is identical to the optimal control subspace (OCS) based control allocation technique. However, for this equivalence to hold, the OCS based control allocation requires evaluation of a (generally) non-causal relationship between control inputs, while the LQ control technique requires perfect knowledge of the system, disturbance and reference states. Neither of these conditions is practically tenable, therefore, approximations are needed. In this regard, the OCS based control allocation technique is superior because it is explicit about the relationship that must be accurately approximated to attain optimality; therefore, desirable approximations of the optimal relationship can be achieved. On the other hand, the LQ control technique implicitly approximates the optimal relationship via estimation of states, disturbances and/or reference signals. Using a classical example based on the Kalman filter, it is shown that the implicit approach of the LQ control technique may yield very poor alignment among the redundant control inputs. The key takeaway from this study is that OCS control allocation technique maintains control alignment better for weakly redundant systems compared to the LQ framework when future disturbance or trajectory is unknown.

Chapter 6 Conclusion and Future Work

6.1 Conclusions

In this dissertation, novel energy-optimal control methods for over-actuated systems are proposed. They apply to the generally defined over-actuated systems (i.e. weak input redundancy), and address the challenges of existing methods for weak input redundancy, including computational efficiency and broadband suboptimality. Also, the connections between the optimal control and control allocation framework, which have not been elucidated for weakly input redundant systems in existing work, are demonstrated in this dissertation. The equivalence of control allocation framework and linear quadratic control under perfect knowledge of disturbance is proved, indicating that the two-stage framework in control allocation does not sacrifice optimality. The advantage of control allocation in preserving control inputs alignment is also shown, when the disturbance is not known a priori.

Specifically, the theoretical contributions of this dissertation have been summarized as follows:

1. The optimal control ratio/subspace for weakly input redundant systems is theoretically derived. It specifies the optimal relationship among the redundant actuators irrespective of external disturbances. The optimal relationship is broadband, and thus enables new dynamic control allocation approaches with negligible computational load.
2. The optimal control ratio/subspace, shown to be non-causal, poses challenges for practical real-time implementation. A causal approximation is proposed and the corresponding energy efficiency loss due to the approximation is analyzed. Using this causal approximation, structured energy-efficient controllers for weakly input redundancy are designed and validated.
3. A causal proxy signal is proposed to accurately measure the deviation from the optimal control ratio/subspace. It converts the control allocation problem to a regulation problem, which is compatible with most existing controller design methodologies. Its capability in

optimal dynamic control allocation without sacrificing control performance is shown, in comparison with a state-of-art control allocation method for weak input redundancy [68].

4. The equivalence between the OCS control allocation and standard LQ control methods is shown, with perfect knowledge of external disturbance. The OCS control allocation is shown to be more advantageous in maintaining redundant control inputs alignments, in practical situations where the disturbance is typically unknown.

In addition to pursuing advancements in control theory on over-actuated systems, the study also focuses on its application on a hybrid feed drive (HFD) for machining applications. The HFD is an over-actuated system, since it adopts a precise but energy costly linear motor drive along with an energy efficient screw drive. The HFD is weakly input redundant since the linear motor drive and screw drives are non-collocated; it also requires dynamic control allocation to reject the broadband and unknown cutting disturbances. The proposed control allocation methods are shown to significantly reduce the energy consumption without sacrificing the positioning accuracy, under broadband cutting forces. Though the time-domain cutting forces are unknown, the dominant cutting frequencies, estimated from the type of cutting tool and the spindle speed, are shown to be helpful in the allocator design. Through this study on HFD, it is shown that the hardware redundancy in over-actuated systems cannot be fully exploited without advanced control allocation methods.

The mentioned theoretical and practical significances are from the published journal articles ([75,76,78]), submitted journal article ([80]), and published conference articles ([77,79]).

6.2 Future Research Directions

6.2.1 Robust Control Allocation

The proposed method is shown to allocate the control efforts of over-actuated systems effectively and efficiently with a reasonably accurate model of the system. However this accurate model of systems may not be always available and model mismatch usually exists. For example the two-mass model discussed in this dissertation introduces modeling errors, especially at high frequencies (shown in Figure 2.3). Also, considering possible actuator degradation, malfunction or failure, the probability of system dynamics variation is higher in over-actuated systems due to a larger quantity of actuators. To understand the effect of modeling error, on both primary control performance (e.g. positioning accuracy in HFD) and energy efficiency, the robustness of

current control allocation methods needs to be evaluated. For cases where model uncertainty is not negligible, robust control allocation methods are needed.

The existing work on robust control allocation mainly focuses on strong input redundancy. For instance, a robust high-level virtual controller is proposed in [18] to avoid performance degradation. Damped least square approaches [118] are introduced to mitigate possible singularity introduced by model uncertainty. A robust least square approach is proposed in [119] to convert robust allocation problems into the H_2/H_∞ framework. Formulating a min-max optimization, a robust control allocation method is presented in [120] such that both model uncertainty and actuator failure/degradation can be addressed in a unified framework. However all these methods apply to strong input redundancy and do not generalize to weak input redundancy. Existing methods for robust allocation in weak input redundancy only enforce a robust stable condition [121], which does not explicitly consider the structure of over-actuation and thus may introduce unnecessary tradeoff. Accordingly, robust control allocation for weak input redundancy is a meaningful future research direction.

6.2.2 Nonlinear Control Allocation

The proposed methods and validation applies to linear time-invariant systems. However in practical applications (e.g. aircraft and automobile discussed in Section 1.1.1), nonlinearity is frequently encountered. Most existing methods for strong input nonlinear redundancy linearize the system in certain working condition to simplify the discussion [10], or adopt a control-Lyapunov approach to facilitate effective allocation [122]. The control allocation with weak input nonlinear redundancy usually embraces gradient-based methods [73,74]. However, they are limited to certain types of nonlinear systems [74], or may introduce performance tradeoff [73].

Current OCS control allocation has potential to be extended to nonlinear systems from following aspects: (i) The proxy signal established in Chapter 3 is shown to accurately measure the deviation from optimality since its squared two norm equals the energy increment. This condition may be relaxed into a general Lyapunov-like function for general nonlinear systems, or a partial set of nonlinearities. (ii) The adjoint dynamics $\mathbf{G}^*$ is not a concept limited to linear systems. The adjoint dynamics of the nonlinear system may specify the optimal allocation structure, which may yield control allocation methods for nonlinear systems. Therefore, an extension of current framework to nonlinear systems is worth looking into.

6.2.3 Sensor Fusion

Control and sensing are dual problems. Therefore, energy-optimal control allocation of over-actuated systems has a counterpart in the world of sensing: optimal estimation of signals from redundant sensors. This problem is referred to as sensor fusion, and has broad applications in machine fault diagnosis, sensor noise reduction, pattern recognition, etc. Most existing sensor fusion algorithms are based on linear minimum variance objective or maximum likelihood principle [123–126]. However, these methods assume identical state dynamics, such that the different measurements in redundant sensors arise from the output noise or different output matrices. This concept is similar to strong input redundancy in the over-actuated system. A generalization similar to weak input redundancy is needed to facilitate a computationally efficient and broadband optimal sensor fusion algorithm.

Bibliography

[1] Xin, X., and Liu, Y., 2014, Control design and analysis for underactuated robotic systems, Springer London, London.

[2] Schneiders, M. G. E., Molengraft, M. J. G. Van De, and Steinbuch, M., 2004, "Benefits of over-actuation in motion systems," Proceedings of the 2004 American Control Conference, pp. 505–510.

[3] Oppenheimer, M. W., Doman, D. B., and Bolender, M. A., 2010, "Control allocation," The Control Handbook, W.S. Levine, ed., CRC Press, Boca Raton, pp. 8-1–24.

[4] Durham, W., Bordignon, K. A., and Beck, R., 2016, Aircraft control allocation, John Wiley & Sons, Ltd, Chichester, UK.

[5] Edwards, C., Lombaerts, T., and Smaili, H., 2010, Fault tolerant flight control, Springer.

[6] Härkegård, O., 2003, "Backstepping and control allocation with applications to flight control," Linköping University.

[7] Goupil, P., 2011, "AIRBUS state of the art and practices on FDI and FTC in flight control system," Control Eng. Pract., **19**(6), pp. 524–539.

[8] Edrén, J., 2014, "Motion modelling and control strategies of over-actuated vehicles," KTH Royal Institute of Technology.

[9] Numasato, H., and Tomizuka, M., 2003, "Settling control and performance of a dual-actuator system for hard disk drives," IEEE/ASME Trans. Mechatronics, **8**(4), pp. 431–438.

[10] Johansen, T. A., and Fossen, T. I., 2013, "Control allocation—a survey," Automatica, **49**(5), pp. 1087–1103.

[11] Jonasson, M., Andreasson, J., Jacobson, B., and Trigell, A. S., 2010, "Global force potential of over-actuated vehicles," Veh. Syst. Dyn., **48**(9), pp. 983–998.

[12] Jonasson, M., and Thor, M., 2018, "Steering redundancy for self-driving vehicles using differential braking," Veh. Syst. Dyn., **56**(5), pp. 791–809.

[13] De Novellis, L., Sorniotti, A., and Gruber, P., 2014, "Wheel torque distribution criteria for

electric vehicles with torque-vectoring differentials," IEEE Trans. Veh. Technol., **63**(4), pp. 1593–1602.

[14] Tagesson, K., Sundstrom, P., Laine, L., and Dela, N., 2009, "Real-time performance of control allocation for actuator coordination in heavy vehicles," 2009 IEEE Intelligent Vehicles Symposium, IEEE, pp. 685–690.

[15] Zhou, J., Canova, M., and Serrani, A., 2016, "Predictive inverse model allocation for constrained over-actuated linear systems," Automatica, **67**, pp. 267–276.

[16] Zaremba, A. T., Liubakka, M. K., and Stuntz, R. M., 1998, "Control and steering feel issues in the design of an electric power steering system," Proceedings of the American Control Conference, pp. 36–40.

[17] Chen, Y., and Wang, J., 2014, "Adaptive energy-efficient control allocation for planar motion control of over-actuated electric ground vehicles," IEEE Trans. Control Syst. Technol., **22**(4), pp. 1362–1373.

[18] Chen, Y., and Wang, J., 2012, "Fast and global optimal energy-efficient control allocation with applications to over-actuated electric ground vehicles," IEEE Trans. Control Syst. Technol., **20**(5), pp. 1202–1211.

[19] Chen, Y., and Wang, J., 2014, "Design and experimental evaluations on energy efficient control allocation methods for overactuated electric vehicles: longitudinal motion case," IEEE/ASME Trans. Mechatronics, **19**(2), pp. 538–548.

[20] Hayama, R., Higashi, M., Kawahara, S., Nakano, S., and Kumamoto, H., 2010, "Fault-tolerant automobile steering based on diversity of steer-by-wire, braking and acceleration," Reliab. Eng. Syst. Saf., **95**(1), pp. 10–17.

[21] Kim, N., Cha, S., and Peng, H., 2011, "Optimal control of hybrid electric vehicles based on Pontryagin's minimum principle," IEEE Trans. Control Syst. Technol., **19**(5), pp. 1279–1287.

[22] Sciarretta, A., Back, M., and Guzzella, L., 2004, "Optimal control of parallel hybrid electric vehicles," IEEE Trans. Control Syst. Technol., **12**(3), pp. 352–363.

[23] Altintas, Y., Verl, A., Brecher, C., Uriarte, L., and Pritschow, G., 2011, "Machine tool feed drives," CIRP Ann. - Manuf. Technol., **60**(2), pp. 779–796.

[24] Zheng, J., Su, W., and Fu, M., 2010, "Dual-stage actuator control design using a doubly coprime factorization approach," IEEE/ASME Trans. Mechatronics, **15**(3), pp. 339–348.

[25] Dong, W., Tang, J., and ElDeeb, Y., 2009, "Design of a linear-motion dual-stage actuation system for precision control," Smart Mater. Struct., **18**(9), p. 095035.

[26] Fujita, T., Matsubara, A., Kono, D., and Yamaji, I., 2010, "Dynamic characteristics and dual control of a ball screw drive with integrated piezoelectric actuator," Precis. Eng., **34**(1), pp. 34–42.

[27] Pahk, H. J., Lee, D. S., and Park, J. H., 2001, "Ultra precision positioning system for servo motor–piezo actuator using the dual servo loop and digital filter implementation," Int. J. Mach. Tools Manuf., **41**(1), pp. 51–63.

[28] Suh, S., Chung, C. C., and Lee, S., 2002, "Design and analysis of dual-stage servo system for high track density HDDs," Microsyst. Technol., **8**(2–3), pp. 161–168.

[29] de Callafon, R. A., Nagamune, R., and Horowitz, R., 2006, "Robust dynamic modeling and control of dual-stage actuators," IEEE Trans. Magn., **42**(2), pp. 247–254.

[30] Peng, K., Chen, B. M., Lee, T. H., and Venkataramanan, V., 2004, "Design and implementation of a dual-stage actuated HDD servo system via composite nonlinear control approach," Mechatronics, **14**(9), pp. 965–988.

[31] Gordon, D. J., and Erkorkmaz, K., 2012, "Precision control of a T-type gantry using sensor/actuator averaging and active vibration damping," Precis. Eng., **36**(2), pp. 299–314.

[32] Giam, T. S., Tan, K. K., and Huang, S., 2007, "Precision coordinated control of multi-axis gantry stages.," ISA Trans., **46**(3), pp. 399–409.

[33] Shinno, H., Yoshioka, H., and Hayashi, M., 2009, "A high performance tilting platform driven by hybrid actuator," CIRP Ann. - Manuf. Technol., **58**(1), pp. 363–366.

[34] Frey, S., Groh, K., and Verl, A., 2012, "Semi-active damping of drive systems," J. Vib. Control, **19**(5), pp. 742–754.

[35] Okwudire, C., and Rodgers, J., 2013, "Design and control of a novel hybrid feed drive for high performance and energy efficient machining," CIRP Ann. - Manuf. Technol., **62**(1), pp. 391–394.

[36] Saidur, R., 2010, "A review on electrical motors energy use and energy savings," Renew. Sustain. Energy Rev., **14**(3), pp. 877–898.

[37] Vijayaraghavan, A., and Dornfeld, D., 2010, "Automated energy monitoring of machine tools," CIRP Ann. - Manuf. Technol., **59**(1), pp. 21–24.

[38] Helu, M., Behmann, B., Meier, H., Dornfeld, D., Lanza, G., and Schulze, V., 2012, "Impact of green machining strategies on achieved surface quality," CIRP Ann. - Manuf. Technol., **61**(1), pp. 55–58.

[39] Dahmus, J. B., and Gutowski, T. G., 2004, "An Environmental Analysis of Machining," Manufacturing Engineering and Materials Handling Engineering, ASME, pp. 643–652.

[40] Pritschow, G., and Philipp, W., 1990, "Direct drives for high-dynamic machine tool axes," CIRP Ann. - Manuf. Technol., **39**(1), pp. 413–416.

[41] Robert, A., 1997, "Attack of the linear motors," Manuf. Eng., **5**, pp. 60--64.

[42] Pritschow, G., 1998, "A comparison of linear and conventional electromechanical dives," CIRP Ann. - Manuf. Technol., **47**(2), pp. 541–548.

[43] Xu, L., and Yao, B., 2001, "Adaptive robust precision motion control of linear motors with negligible electrical dynamics: theory and experiments," IEEE/ASME Trans. Mechatronics, **6**(4), pp. 444–452.

[44] Butcher, M., and Karimi, A., 2010, "Linear parameter-varying iterative learning control with application to a linear motor system," IEEE/ASME Trans. Mechatronics, **15**(3), pp. 412–420.

[45] Sato, K., Katori, M., and Shimokohbe, A., 2013, "Ultrahigh-acceleration moving-permanent-magnet linear synchronous motor with a long working range," IEEE/ASME Trans. Mechatronics, **18**(1), pp. 307–315.

[46] "Roh'Lix® linear actuators" [Online]. Available: https://www.zero-max.com/ce-rohlix-linear-actuators. [Accessed: 08-Feb-2018].

[47] Buice, E. S., Otten, D., Yang, R. H., Smith, S. T., Hocken, R. J., and Trumper, D. L., 2009, "Design evaluation of a single-axis precision controlled positioning stage," Precis. Eng., **33**(4), pp. 418–424.

[48] Schroeck, S. J., Messner, W. C., and McNab, R. J., 2001, "On compensator design for linear time-invariant dual-input single-output systems," IEEE/ASME Trans. Mechatronics, **6**(1), pp. 50–57.

[49] Juhász, L., and Maas, J., 2013, "Control of hybrid nanopositioning systems for trajectory-tracking applications," Mechatronics, **23**(6), pp. 617–629.

[50] Glöß, R., 2006, "Nanometer precise hybrid actuator in positioning mechanism with long travel range," International Conference and Exhibition on New Actuators and Drive

Systems, Bremen, Germany, pp. 668–671.

[51] Ding, J., Tomizukas, M., and Numasato, H., 2000, "Design and robustness analysis of dual stage servo system," Proceedings of the 2000 American Control Conference, IEEE, pp. 2605–2609.

[52] Bordignon, K. A., 1996, "Constrained control allocation for systems with redundant control effectors," Virginia Polytechnic Institute and State University.

[53] Buffington, J. M., and Enns, D. F., 1996, "Lyapunov stability analysis of daisy chain control allocation," J. Guid. Control. Dyn., **19**(6), pp. 1226–1230.

[54] Hirano, T., Fan, L.-S., Lee, W. Y., Hong, J., Imaino, W., Patanaik, S., Chan, S., Webb, P., Horowitz, R., Aggarwal, S., and Horsley, D. A., 1998, "High-bandwidth high-accuracy rotary microactuators for magnetic hard disk drive tracking servos," IEEE/ASME Trans. Mechatronics, **3**(3), pp. 156–165.

[55] Oh, Y., Chung, W. K., Youm, Y., and Suh, I.-H., 1998, "Experiments on extended impedance control of redundant manipulator," Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, pp. 1320–1325.

[56] Lin, Z. C., Patel, R. V, and Balafoutis, C. A., 1995, "Impact reduction for redundant manipulators using augmented impedance control," J. Robot. Syst., **12**(5), pp. 301–313.

[57] Saglia, J. A., Tsagarakis, N. G., Dai, J. S., and Caldwell, D. G., 2009, "A high-performance redundantly actuated parallel mechanism for ankle rehabilitation," Int. J. Rob. Res., **28**(9), pp. 1216–1227.

[58] Anderson, B. D. O., and Moore, J. B., 1971, Linear optimal control, Prentice-Hall, Inc., Englewood Cliffs, NJ.

[59] Geering, H., 2007, Optimal control with engineering applications.

[60] Singh, A. K., and Pal, B. C., 2017, "An extended linear quadratic regulator for LTI systems with exogenous inputs," Automatica, **76**, pp. 10–16.

[61] Duan, Z., Huang, L., Yao, Y., and Jiang, Z. P., 2012, "On the effects of redundant control inputs," Automatica, **48**(9), pp. 2168–2174.

[62] Skogestad, S., and Postlethwaite, I., 2007, Multivariable Feedback Control: Analysis and Design, Wiley New York.

[63] Khargonekar, P. P., and Rotea, M. A., 1991, "Mixed H2/H∞ control: a convex optimization approach," IEEE Trans. Automat. Contr., **36**(7), pp. 824–837.

[64] Härkegård, O., and Glad, S. T., 2005, "Resolving actuator redundancy—optimal control vs. control allocation," Automatica, **41**(1), pp. 137–144.

[65] Härkegård, O., 2004, "Dynamic control allocation using constrained quadratic programming," J. Guid. Control. Dyn., **27**(6), pp. 1028–1034.

[66] Petersen, J. A. M., and Bodson, M., 2006, "Constrained quadratic programming techniques for control allocation," IEEE Trans. Control Syst. Technol., **14**(1), pp. 91–98.

[67] Bodson, M., 2002, "Evaluation of optimization methods for control allocation," J. Guid. Control. Dyn., **25**(4), pp. 703–711.

[68] Zaccarian, L., 2009, "Dynamic allocation for input redundant control systems," Automatica, **45**(6), pp. 1431–1438.

[69] Kirchengast, M., Steinberger, M., and Horn, M., 2018, "Input matrix factorizations for constrained control allocation," IEEE Trans. Automat. Contr., **63**(4), pp. 1163–1170.

[70] Naskar, A. K., Patra, S., and Sen, S., 2017, "New control allocation algorithms in fixed point framework for overactuated systems with actuator saturation," Int. J. Control, **90**(2), pp. 348–356.

[71] Lallman, F., Davidson, J., and Bundick, W., 2001, "Integrated reconfigurable control allocation," AIAA Guidance, Navigation, and Control Conference and Exhibit, American Institute of Aeronautics and Astronautics, Reston, Virigina.

[72] Galeani, S., Serrani, A., Varano, G., and Zaccarian, L., 2015, "On input allocation-based regulation for linear over-actuated systems," Automatica, **52**(2015), pp. 346–354.

[73] De Tommasi, G., Galeani, S., Pironti, A., Varano, G., and Zaccarian, L., 2011, "Nonlinear dynamic allocator for optimal input/output performance trade-off: Application to the JET tokamak shape controller," Automatica, **47**(5), pp. 981–987.

[74] Passenbrunner, T. E., Sassano, M., and Zaccarian, L., 2016, "Optimality-based dynamic allocation with nonlinear first-order redundant actuators," Eur. J. Control, **31**, pp. 33–40.

[75] Duan, M., and Okwudire, C. E., 2016, "Energy-efficient controller design for a redundantly-actuated hybrid feed drive with application to machining," IEEE/ASME Trans. Mechatronics, **21**(4), pp. 1822–1834.

[76] Duan, M., and Okwudire, C. E., 2016, "Correction to 'Energy-efficient controller design for a redundantly-actuated hybrid feed drive with application to machining,'" IEEE/ASME Trans. Mechatronics, **21**(6), pp. 2999–3000.

[77] Duan, M., and Okwudire, C. E., 2016, “Near energy optimal control allocation for dual-input over-actuated systems,” Proceedings of the ASME 2016 Dynamic Systems and Control Conference, Minneapolis, p. V001T01A011.

[78] Duan, M., and Okwudire, C. E., 2018, “Proxy-based optimal control allocation for dual-input over-actuated systems,” IEEE/ASME Trans. Mechatronics, **23**(2), pp. 895–905.

[79] Duan, M., and Okwudire, C., 2017, “Proxy-based optimal dynamic control allocation for multi-input, multi-output over-actuated systems,” Proceedings of the ASME 2017 Dynamic Systems and Control Conference, ASME, Tyson, VA, p. V001T03A005.

[80] Duan, M., and Okwudire, C. E., “Connections between control allocation and linear quadratic control for weakly redundant systems,” Automatica (under Revision).

[81] Curtain, R. F., and Zwart, H., 2012, An introduction to infinite-dimensional linear systems theory, Springer Science & Business Media.

[82] Komzsik, L., 2014, Applied calculus of variations for engineers, CRC Press.

[83] Varanasi, K., and Nayfeh, S., 2004, “The dynamics of lead-screw drives: low-order modeling and experiments,” J. Dyn. Syst. Meas. Control, **126**, pp. 388--396.

[84] Okwudire, C., and Altintas, Y., 2009, “Minimum tracking error control of flexible ball screw drives using a discrete-time sliding mode controller,” J. Dyn. Syst. Meas. Control, **131**(5), pp. 051006-1–12.

[85] Gordon, D. J., and Erkorkmaz, K., 2013, “Accurate control of ball screw drives using pole-placement vibration damping and a novel trajectory prefilter,” Precis. Eng., **37**(2), pp. 308–322.

[86] Sepasi, D., Nagamune, R., and Sassani, F., 2012, “Tracking control of flexible ball screw drives with runout effect and mass variation,” IEEE Trans. Ind. Electron., **59**(2), pp. 1248–1256.

[87] Okwudire, C. E., and Altintas, Y., 2009, “Hybrid modeling of ball screw drives with coupled axial, torsional, and lateral dynamics,” J. Mech. Des., **131**(7), p. 071002.

[88] Wang, Y., Ueda, K., and Bortoff, S. A., 2013, “A Hamiltonian approach to compute an energy efficient trajectory for a servomotor system,” Automatica, **49**(12), pp. 3550–3561.

[89] Halevi, Y., Carpanzano, E., and Montalbano, G., 2014, “Minimum energy control of redundant linear manipulators,” J. Dyn. Syst. Meas. Control, **136**(5), p. 051016.

[90] Huang, M. S., Hsu, Y. L., and Fung, R. F., 2012, “Minimum-energy point-to-point

trajectory planning for a motor-toggle servomechanism," IEEE/ASME Trans. Mechatronics, **17**(2), pp. 337–344.

[91] Yoon, D., and Okwudire, C. E., 2015, "Magnet assisted stage for vibration and heat reduction in wafer scanning," CIRP Ann. - Manuf. Technol., **64**(1), pp. 381–384.

[92] Park, J., 1996, "Motion profile planning of repetitive point-to-point control for maximum energy conversion efficiency under acceleration conditions," Mechatronics, **6**(6), pp. 649–663.

[93] Wang, Y., Zhao, Y., Bortoff, S. A., and Ueda, K., 2015, "A real-time energy-optimal trajectory generation method for a servomotor system," IEEE Trans. Ind. Electron., **62**(2), pp. 1175–1188.

[94] Schmidt, R. M., Schitter, G., and Rankers, A., 2014, The design of high performance mechatronics: high-tech functionality by multidisciplinary system integration, IOS Press, Amsterdam, Netherlands.

[95] Tomizuka, M., 1987, "Zero phase error tracking algorithm for digital control," J. Dyn. Syst. Meas. Control, **109**(1), p. 65.

[96] Rigney, B. P., Pao, L. Y., and Lawrence, D. A., 2009, "Nonminimum phase dynamic inversion for settle time applications," IEEE Trans. Control Syst. Technol., **17**(5), pp. 989–1005.

[97] Butterworth, J. A., Pao, L. Y., and Abramovitch, D. Y., 2012, "Analysis and comparison of three discrete-time feedforward model-inverse control techniques for nonminimum-phase systems," Mechatronics, **22**(5), pp. 577–587.

[98] Duan, M., and Okwudire, C. E., 2015, "Energy efficiency and performance optimized control of a hybrid feed drive," Proceedings of ASME 2015 International Manufacturing Science and Engineering Conference, ASME, Charlotte, NC, p. V002T05A007.

[99] Chen, S., 2014, "Another Particle Swarm Toolbox," MATLAB Cent. File Exch. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/25986-another-particle-swarm-toolbox/content/psopt/pso.m.

[100] Kim, T.-H., Maruta, I., and Sugie, T., 2008, "Robust PID controller tuning based on the constrained particle swarm optimization," Automatica, **44**(4), pp. 1104–1110.

[101] Altintas, Y., 2012, Manufacturing automation: metal cutting mechanics, machine tool vibrations, and CNC design, Cambridge university press.

[102] Darouach, M., Zasadzinski, M., and Xu, S. J., 1994, "Full-order observers for linear systems with unknown inputs," IEEE Trans. Automat. Contr., **39**(3), pp. 606--609.

[103] Kothare, M. V., and Morari, M., 1996, Multiplier theory for stability analysis of anti-windup control systems, California Institute of Technology.

[104] Ramani, K. S., Duan, M., Okwudire, C. E., and Ulsoy, A. G., 2017, "Tracking control of linear time-invariant nonminimum phase systems using filtered basis functions," J. Dyn. Syst. Meas. Control, **139**(1), pp. 011001-1–11.

[105] Duan, M., Yoon, D., and Okwudire, C. E., 2017, "A limited-preview filtered B-spline approach to tracking control – With application to vibration-induced error compensation of a 3D printer," Mechatronics.

[106] Weck, M., and Ye, G., 1990, "Sharp corner tracking using the IKF control strategy," CIRP Ann. - Manuf. Technol., **39**(1), pp. 437–441.

[107] Rojas, O. J., and Goodwin, G. C., 2002, "A simple antiwindup strategy for state constrained linear control," IFAC World Congress, Elsevier, Barcelona, Spain, pp. 109–114.

[108] Chambon, E., Burlion, L., and Apkarian, P., 2018, "Time-response shaping using output to input saturation transformation," Int. J. Control, **91**(3), pp. 534–553.

[109] Lambrechts, P., Boerlage, M., and Steinbuch, M., 2005, "Trajectory planning and feedforward design for electromechanical motion systems," Control Eng. Pract., **13**(2), pp. 145–157.

[110] Kailath, T., 1980, Linear systems, Prentice-Hall Englewood Cliffs, NJ.

[111] Sebek, M., 2015, "Spectral factorization," Encyclopedia of Systems and Control, Springer London, London, pp. 1289–1295.

[112] Doyle, J. C., Glover, K., Khargonekar, P. P., and Francis, B. A., 1989, "State-space solutions to standard H2 and H∞ control problems," IEEE Trans. Automat. Contr., **34**(8), pp. 831–847.

[113] Francis, B. A., and Wonham, W. M., 1976, "The internal model principle of control theory," Automatica, **12**(5), pp. 457–465.

[114] Beard, R. W., 2002, "Linear operator equations with applications in control and signal processing," IEEE Control Syst. Mag., **22**(2), pp. 69–79.

[115] Bryson, A., and Ho, Y., 1975, Applied optimal control: optimization, estimation, and

control, Blaisdell Publishing Company.

[116] Kirk, D., 2012, Optimal control theory: an introduction, Dover Publications, Inc., Mineola, New York.

[117] Petersen, K. B., and Pedersen, M. S., 2008, "The matrix cookbook," Tech. Univ. Denmark.

[118] Berge, S. P., and Fossen, T. I., 1997, "Robust control allocation of overactuated ships; experiments with a model ship," IFAC Proc. Vol., **30**(22), pp. 193–198.

[119] Cui, L., and Yang, Y., 2011, "Disturbance rejection and robust least-squares control allocation in flight control system," J. Guid. Control. Dyn., **34**(6), pp. 1632–1643.

[120] Shen, Q., Wang, D., Zhu, S., and Poh, E. K., 2017, "Robust control allocation for spacecraft attitude tracking under actuator faults," IEEE Trans. Control Syst. Technol., **25**(3), pp. 1068–1075.

[121] Cocetti, M., Serrani, A., and Zaccarian, L., 2016, "Dynamic input allocation for uncertain linear over-actuated systems," Proceedings of the 2016 American Control Conference, IEEE, pp. 2906–2911.

[122] Johansen, T. a, 2004, "Optimizing nonlinear control allocation," Decision and Control, 2004. CDC. 43rd IEEE Conference on, p. 3435–3440 Vol.4.

[123] Saha, R. K., 1996, "Track-to-track fusion with dissimilar sensors," IEEE Trans. Aerosp. Electron. Syst., **32**(3), pp. 1021–1029.

[124] Sun, S. L., and Deng, Z. L., 2004, "Multi-sensor optimal information fusion Kalman filter," Automatica, **40**(6), pp. 1017–1023.

[125] Kim, K. H., 1994, "Development of track to track fusion algorithms," Proceedings of 1994 American Control Conference - ACC '94, IEEE, pp. 1037–1041.

[126] Chong, C., 2017, "Forty years of distributed estimation: A review of noteworthy developments," 2017 Sensor Data Fusion: Trends, Solutions, Applications (SDF), IEEE, pp. 1–10.