

# Stability and Transparency in Mixed-Reality Bilateral Human Teleoperation

David G. Black , Member, IEEE, and Septimiu E. Salcudean , Life Fellow, IEEE

**Abstract**—Recent work introduced the concept of human teleoperation (HT), where the remote robot typically considered in conventional bilateral teleoperation is replaced by a novice person wearing a mixed-reality head-mounted display and tracking the motion of a virtual tool controlled by an expert. HT has advantages in cost, complexity, and patient acceptance for telemedicine in low-resource communities or remote locations. However, the stability, transparency, and performance of bilateral HT are unexplored. In this article, we, therefore, develop a mathematical model of the HT system using test data. We then analyze various control architectures with this model and implement them with the HT system, testing volunteer operators and a virtual fixture-based simulated patient to find the achievable performance, investigate stability, and determine the most promising teleoperation scheme in the presence of time delays. We show that instability in HT, while not destructive or dangerous, makes the system impossible to use. However, stable and transparent teleoperation is possible with small time delays ( $< 200$  ms) through three-channel teleoperation, or with large time delays through model-mediated teleoperation with local pose and force feedback for the novice.

**Index Terms**—Augmented reality, force feedback, haptics, human–computer interaction, stability, teleoperation, transparency.

## I. INTRODUCTION

MANY remote and underresourced communities experience severe challenges in accessing qualified medical care. For example, ultrasound imaging is important, widely used, and much lower in cost than other modalities such as computed tomography or magnetic resonance. However, capturing and interpreting ultrasound images requires a high degree of expertise that is not commonly present in many small communities. As a result, a sonographer or radiologist must be transported to the town on a regular basis, or patients must be sent to a major medical center. Either case leads to long wait times and difficulty

Received 16 April 2025; revised 13 August 2025; accepted 5 September 2025. Date of publication 23 September 2025; date of current version 10 October 2025. This work was supported by the Vanier Canada Graduate Scholarship, C.A. Laszlo Chair in Biomedical Engineering, Mitacs, and Rogers Communications, Inc. This article was recommended for publication by Associate Editor Farokh Atashzar and Editor Dana Kulic upon evaluation of the reviewers' comments. (Corresponding author: David G. Black.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of British Columbia Behavioural Research Ethics Board (BREB), under Application No. H22-01195. Informed consent was given by all participants.

The authors are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver V6T 1Z4, Canada (e-mail: dgbblack@ece.ubc.ca).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TRO.2025.3613464>, provided by the authors.

Digital Object Identifier 10.1109/TRO.2025.3613464

handling urgent cases. In communities across Canada, patients are flown hundreds of kilometers for standard ultrasound exams. This takes up to three days and exerts a high social and financial cost on the community.

Therefore, teleultrasound is an important and growing field. However, current commercially available technologies are often impractical. Video teleguidance is simple, low-cost, and accessible to anyone but is highly inefficient and imprecise if the person being guided does not already have ultrasound experience [1]. On the other hand, robotic teleultrasound gives the physician complete and precise control but is expensive and complex to set up and maintain. We thus recently introduced a novel teleguidance method called human teleoperation (HT) to address the shortcomings of both existing approaches [1]. This method is also useful in any other application that could benefit from hand-over-hand guidance, including other telehealth, remote maintenance, inspection, teaching, and more.

## A. Background on Human Teleoperation

HT is described in detail in [2]. Diagrams of a prototype system are shown in Figs. 1 and 6.

In HT, a local novice person performs an ultrasound exam on a patient while being guided by a remote operator, the sonographer or radiologist. The novice wears a mixed-reality (MR) head-mounted display (HMD), which projects a virtual ultrasound probe into their field of view. The pose of this virtual probe is controlled in real time by the operator, who manipulates a haptic device with an ultrasound probe-shaped end effector. As the operator moves the virtual probe, the novice aligns their real probe to the virtual one and follows it as it moves, thus achieving the operator's desired motion. The ultrasound image and video from the novice's HMD are streamed to the operator so they can carry out the procedure. The two sides are also in verbal communication. Following the MR virtual probe is very intuitive, and no prior ultrasound experience is required for the novice. Indeed, the ability of the novice to track the MR input was previously characterized [3], [4], showing tracking lags of 200–350 ms and steady-state error of less than 3 mm. The results of these tests are used to inform the mathematical modeling in the following sections. Relative to audio/video guidance, HT is very precise and tightly coupled, although the accuracy and speed are slightly less than a robot's [1]. Such tight coupling of the operator and novice enables this teleguidance system to be designed and analyzed such as a robotic teleoperation system in which the follower robot is replaced by the novice person, leading to the name HT.

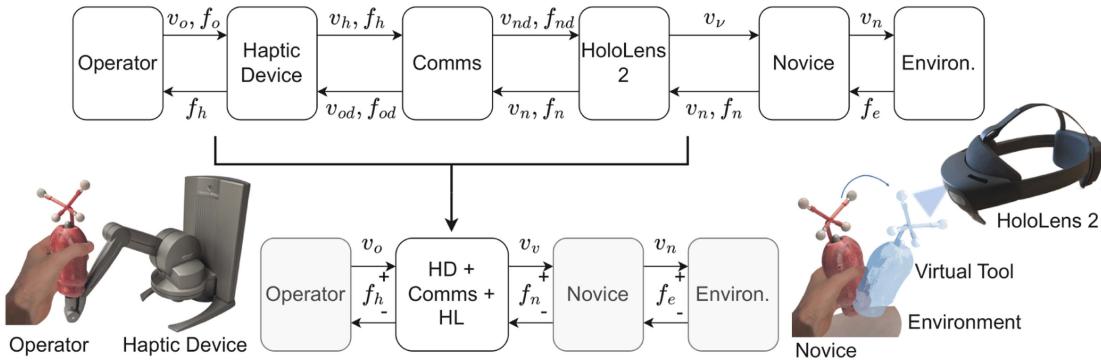


Fig. 1. Conceptual overview of the teleoperation system. The gray boxes of operator, novice, and environment have to be modeled, while the haptic device, communications, and HoloLens 2 represent the part of the system we can directly control. In general HT, the novice moves a tool in space, as shown here, but for our tests, this tool was connected to a haptic device to record the motion and simulate a known environment (see Fig. 6).

While the position and orientation of the ultrasound probe are key during procedures, and are tracked by the HMD using infrared reflective markers [5], the applied force is also an important factor. Applying the correct level of force avoids excessive deformation of structures, such as blood vessels, displaces gas that otherwise obstructs the image, and allows imaging below the ribs, for example. Therefore, the ultrasound probe in HT is instrumented with force sensing, as previously described [6], [7]. Moreover, the operator and novice sides communicate over the Internet using WebRTC, a fast peer-to-peer framework [4]. Together, these enable fast feedback of pose and force to the operator.

The haptic device of the operator has three actuated joints, allowing it to apply forces to the operator’s hand. This way, it is possible to render a realistic haptic sense of the patient to the operator, which is essential for three primary reasons. First, sonographers are accustomed to resting their hand on the patient and exerting relatively large forces up to approximately 20 N [8]. Without this support, the scanning by sonographers feels unrealistic and unintuitive. More importantly, it is very difficult to hold the probe at exactly the height of the patient surface or to make precise motions if the operator’s hand is floating in space instead of resting on the surface. Finally, the operator uses their sense of touch to perform the exam, for example, by palpating tissue or by feeling the ribs so as to image between them. Beyond teleultrasound, other contact tasks also require haptic feedback, making it essential in HT.

### B. Stability and Transparency in Robotic Teleoperation

Haptic feedback has been extensively studied in the context of robotic teleoperation. In robot-assisted minimally invasive surgery [9] and surgical training [10], haptic feedback has been shown to increase task performance and decrease tissue damage [11], [12]. Similarly, in teleoperation of mobile robots [13], uncrewed aerial vehicles [14], and robots for micromanipulation and assembly [15], haptic feedback has shown utility.

As shown in the telerobotics control literature, haptic feedback increases the transparency of the teleoperation. Transparency describes how well the follower and operator sides match each other in terms of force and position (velocity) [16]. In

a perfectly transparent system, the velocity and force of the operator and follower (often referred to in prior literature as master and slave, respectively) would be exactly equal at all times. Thus, the follower would perfectly track the operator’s motion and the operator would feel exactly as if they were touching the remote environment directly. This is a primary goal in teleoperation that has been addressed using many approaches including direct force feedback [17], position feedback [18], combined position and force feedback [19], impedance feedback [20], adaptive control [21], local force feedback at the follower [22], and more.

The other essential objective in teleoperation is stability. Especially in medical applications in which the novice robot interacts with patients, instability in the system can constitute a health hazard. Unfortunately, even a small amount of time delay can destabilize nominally stable bilateral teleoperation systems [23]. Since any remote teleoperation system inevitably includes communication at a distance, leading to delays, this is an important problem that has been studied extensively. Due to the complex, nonlinear nature of teleoperators, system passivity is commonly used to show stability [24]. For example, it has been shown that transmitting wave variables rather than the values themselves guarantees the passivity of the system under arbitrary time delays [25], [26]. However, this can degrade tracking accuracy and transparency. Thus, others have used Smith predictors [8], [27], [28] and model predictive control [29] to deal with the time delay by predicting the feedback before it arrives. Similarly, model-mediated control allows instantaneous feedback by keeping a local model of the remote environment on the operator side [30]. As these methods are affected by modeling accuracy, however, others have studied transmitting adaptive combinations of velocity and force [31] and robust control approaches [32]. A particularly successful method called time-domain passivity was proposed by Hannaford and Ryu [33], [34], in which an energylike quantity of the system is observed and dissipative elements are added to eliminate only the net energy output, thus maintaining passivity with less degradation of performance.

### C. Stability and Transparency in Human Teleoperation

Despite extensive research on robotic teleoperation, there is no research that examines the stability and transparency of

HT. In HT, the follower robot is replaced by a human, leading to fundamentally different system behavior. In particular, the human novice may be considered passive [3], [35], [36], but their reaction time and imperfect tracking lead to potentially larger delays and inaccurate motions and forces. The novice can intrinsically respond to changing environmental conditions and does not require trajectory planning around joint limits and singularities, nor initial calibration, safe deceleration, or compliant controllers. However, they must be guided effectively according to the capabilities of their vision and perception, without providing too much input so as to cause cognitive overload or rapid fatigue. Thus, HT is influenced by other factors than robotic teleoperation, but stability and transparency remain paramount. To achieve efficient, performant teleoperation, it is thus important to study the stability and transparency of the HT system.

#### D. Contributions

This article presents an initial application of the concepts of stability and transparency from telerobotics to bilateral HT. Although the initial motivation is teleultrasound, this analysis holds for any application of HT. In particular, the following contributions are included.

- 1) A physical model of the human-in-the-loop system is derived and used to create a hybrid matrix formulation of the teleoperation with time delays in Section II-B.
- 2) This model is used with several candidate controllers to investigate their respective transparency and expected performance for HT in Section II-C.
- 3) In Section II-E, the stability of bilateral HT is investigated.
- 4) In Section III-D, the robustness of the derived models to parametric and dynamic uncertainty is explored using the structured singular value.
- 5) Tests with the real system and volunteer novices were performed to validate the modeling and characterize the novice behavior under different controllers. The environment/patient was simulated using a haptic device virtual fixture. The setup is described in Sections II-F and II-G, and the results are presented in Section III.
- 6) It is shown that instability in HT manifests itself differently than in telerobotics and primarily degrades system performance.

HT, despite many parallels with telerobotics, constitutes a novel system whose responses to control inputs have not yet been studied. While this article does not constitute a comprehensive analysis or test of all possible HT approaches, it lays the foundations both for future practical tests with human volunteers and for further algorithmic development, for example, in predictive or robust control schemes.

## II. METHODS

### A. Notation

In the following,  $f$  is force,  $x$  position,  $\dot{x} = v$  velocity,  $\tau$  time delay,  $m$  mass,  $b$  damping, and  $k$  a spring constant or gain. The subscripts are  $n$  for novice,  $o$  for operator,  $e$  for

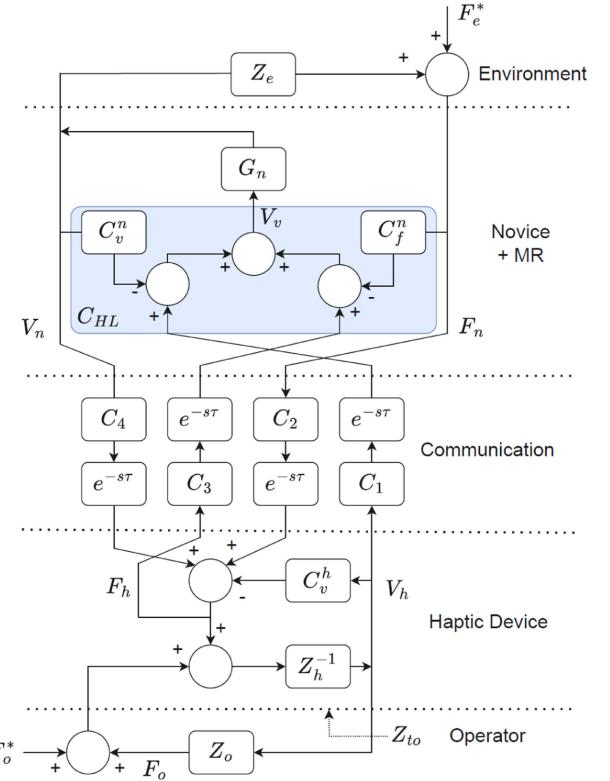


Fig. 2. General model of HT. The operator (subscript  $o$ ) interacts with a haptic device (subscript  $h$ ), while the novice ( $n$ ) interacts with the environment ( $e$ ). The communication channel induces time delays of  $\tau$  on the force and velocity, which are transmitted bilaterally. The MR headset creates a visual control output potentially using all four channels of force and velocity, denoted  $C_{HL}$ . The novice responds to the MR input according to the novice transfer function,  $G_n$ . The controllers  $C_j^i$  are on the novice or haptic device if the superscript  $i = f$  or  $h$ , respectively, and involve force or velocity, if the subscript  $j = f$  or  $v$ , respectively. Most haptic devices are not equipped with force sensing, so there is no  $C_f^h$ .

environment,  $v$  for the virtual tool, and  $h$  for haptic device. Note, novice and follower are equivalent here, but we use novice to avoid confusion with subscript  $f$  for force. Scalar variables are lowercase, Laplace domain variables are capitalized, column vectors are bolded lowercase, and matrices are bolded uppercase.

### B. Modeling

A general linear-time invariant model of HT is shown in Fig. 2. Although the overall structure is similar to previous robotic teleoperation models and has the same expert components, the novice side is different. We assume that the novice is seated or otherwise supported, so we model only the relationship between the novice's MR input (pose of virtual ultrasound transducer in the headset) and the motion of the novice's hand (pose of the actual ultrasound transducer in the hand). In following the virtual tool, the human novice is controlled through position commands and implicitly accounts for his/her own arm dynamics. We, therefore, assume that the novice's arm dynamics in tracking the MR input is linear, time-invariant, and passive, with minimal coupling between axes. This is justified by detailed analysis of MR-guided human response [3], [35], building on widely used

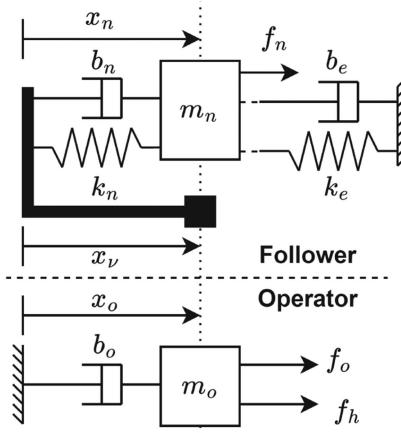


Fig. 3. Models of the operator and haptic device (bottom), and the novice potentially in contact with the environment (top). The damping of the operator arm and haptic device together is represented by  $b_o$ . The black fixture attached to the novice mass is rigid and massless. The square end at position  $x_\nu$  represents the virtual tool. The novice force  $f_n = -k_ex_n - b_e\dot{x}_n$  cancels the environment dynamics at all times. Thus, when the novice motion has converged (i.e.,  $b_n$  and  $k_n$  are at equilibrium), the novice position  $x_n$  should equal  $x_\nu$ .

prior work on human arm models [36]. We further assume the novice sees only the virtual tool pose and the operator receives feedback only by forces applied by the haptic device. This does not take into account the operator's visual feedback from video or other sensor streams such as ultrasound images, nor the verbal communication between operator and novice.

Beyond the virtual tool pose, it is possible to give additional input to the novice such as changing the tool color or displaying an error bar, arrow, or visual renderings in the MR headset [3]. These options could be added in the blue  $C_{HL}$  box in Fig. 2. However, human response to color or error bars may be distracting and may lead to cognitive overload for the novice [3]. Hence, this article examines only controlling the virtual tool pose and other renderings are left for future work.

Given these assumptions about the HT system, we develop models of the novice and operator in contact with the environment and haptic device, respectively, and study the tradeoffs between stability and performance for different choices of the control parameters in Fig. 2.

1) *Follower*: The model was derived from measured response data of 11 volunteer novices [3], [4]. It was found that a second-order system with two poles and one zero matched the measured behavior well. This is equivalent to attaching the novice's hand and tool, which have a certain mass, to the virtual tool by a spring and damper, as shown in Fig. 3. The novice either moves freely in space or is in contact with the environment, represented by an impedance  $Z_e = b_e + \frac{1}{s}k_e$ , which is taken to be constant.

We assume the novice applies whatever force is necessary to match the virtual tool, whether or not they are in contact with the environment, so there is no steady-state error. In other words,  $f_n = -k_ex_n - b_e\dot{x}_n$  at all times, so it perfectly cancels the environment contact dynamics. This way, intermittent contact with the environment has no effect on the novice's stability or tracking, which is as expected from experience; the human hand

TABLE I  
APPROXIMATE MODEL PARAMETER VALUES DETERMINED THROUGH SIMULATION AND SYSTEM IDENTIFICATION BASED ON PREVIOUS DATA [3]

Environment	Follower		Operator		
$k_e$	10	$k_n$	1	$k_a$	100
$b_e$	1	$b_n$	0.275	$b_o$	0.1
$m_p$	0.02	$m_n$	0.02	$m_o$	0.1
				$k_{o1}$	0.01
				$k_{o2}$	80

All units are SI - N, m, s, kg.

does not become unstable, even in high-frequency contact with a stiff surface. The novice can thus be represented by the passive (strictly passive if  $b_n > 0$ ) transfer function

$$X_n = \left( \frac{b_ns + k_n}{m_ns^2 + b_ns + k_n} \right) x_\nu = G_n x_\nu \quad (1)$$

$$F_n = -(b_es + k_e)X_n \quad (2)$$

where  $k_e$  and  $b_e$  are zero when not in contact with the environment. For this to be linear time invariant (LTI), we assume the environment is unmoving and its impedance is constant. This is approximately true, for example, when ultrasound scanning a specific region of a patient such as the abdomen. In this case,  $F_e^* = 0$  in Fig. 1.

2) *Operator and Haptic Device*: The operator hand holding the haptic device is modeled as a mass and damper system, which is subject to the applied force of the operator and the haptic device, as shown in Fig. 3.

$$m_o\ddot{x}_o + b_o\dot{x}_o = f_h + f_o. \quad (3)$$

The operator changes their applied force depending on the haptic device force to track a desired trajectory. In particular, to choose a representative response, we use a proportional-integral-derivative-inspired controller

$$f_o = -k_{o1}x_o - k_{o2}\dot{x}_o + k_a \int (x_o^* - x_o) dt \quad (4)$$

where  $x_o^*$  is the operator's desired motion. The gains in (4) can be tuned to obtain a response that resembles the recorded operator data. We ignore the operator's desired velocity for simplicity as it introduces a zero eigenvalue and the operator is primarily interested in position. We thus find the following Laplace-domain representation of the operator:

$$X_o = \frac{F_h + \frac{k_a}{s}X_o^*}{m_osc^2 + (b_o + k_{o2})s + k_{o1} + \frac{k_a}{s}} \quad (5)$$

$$F_o = \frac{k_a}{s}X_o^* - \left( k_{o2}s + k_{o1} + \frac{k_a}{s} \right) X_o. \quad (6)$$

The state-space formulations of the operator and novice are derived in the Appendix.

3) *Parameter Identification*: To check the modeling of Sections II-B1 and II-B2, the operator and novice parameters were independently estimated to obtain responses similar to previous measurements [3], [4]. An effective set of parameters is shown in Table I. The chosen model parameter values were obtained by fitting the models to the previously measured data using the MATLAB linear system identification toolbox and

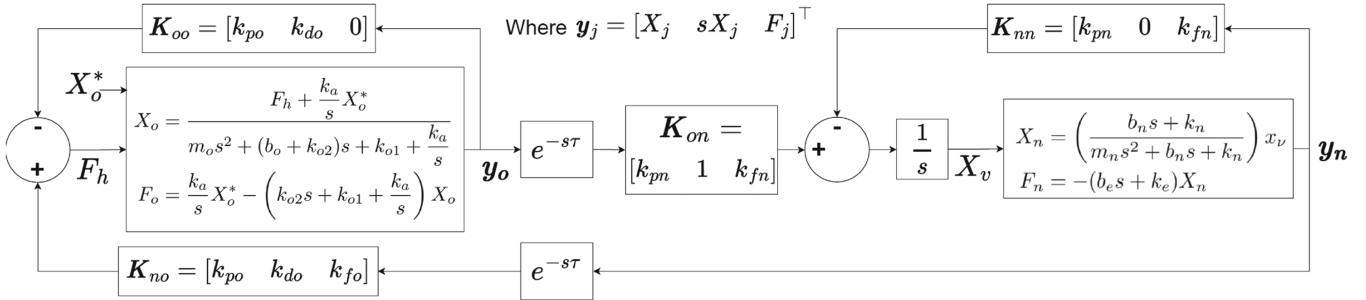


Fig. 4. Interconnection of the derived models in the closed-loop teleoperation system. Different behavior is achieved by changing the feedback and feed-forward gains of the positions, velocities, and forces. The operator-side model is given in (5), while the novice side is in (1).

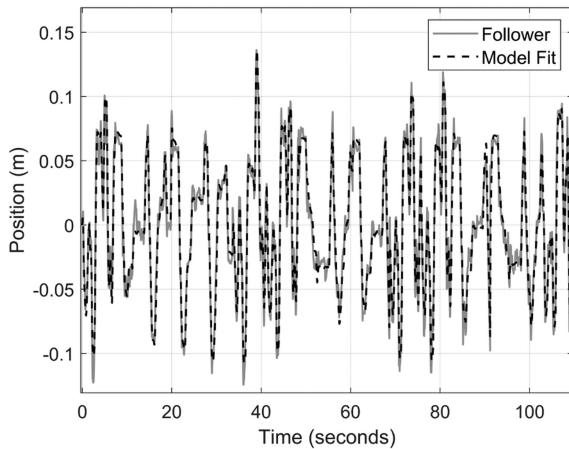


Fig. 5. Derived model with parameters fitted to the measured data from previous tests [3], [4]. The MSE between the two is 0.67 mm.

some manual tuning. Specifically, linear gray-box estimation was utilized with adaptive subspace Gauss–Newton search to fit parameters to the derived ordinary differential equations, with nonnegativity constraints and regularization on all parameters, and a stability constraint on the system. The final system fit the measured dataset with a mean-squared error (MSE) of 0.67 mm in the novice motion. The fit accuracy is relatively insensitive to changes in the parameters, which shows robustness in the model. The fit MSE has a partial derivative of 0.125 mm/kg for novice mass with all other parameters fixed at their chosen values. All other parameters have slopes of less than 0.0625 mm/unit in magnitude. The measured and fitted data are shown in Fig. 5.

This way, approximate numerical values could be assigned to the various environment, novice, and operator parameters, as shown in Table I. These values are useful for simulating and testing controllers before implementing them on physical hardware and for numerically assessing the stability of a controller.

4) *Teleoperation System:* The full teleoperation system is shown in Fig. 4, given the models derived previously. Let  $\mathbf{y}_j = [x_j \ \dot{x}_j \ f_j]^\top$   $j = \{o, n\}$  be the output of the operator and novice, respectively. In this case, we control the virtual tool pose,  $x_\nu$  for the novice, and the haptic device force,  $f_h$  for the operator. Each can be a function of the position, velocity, and force of the operator and novice as shown in the diagram. Some gains are never used, so they are set to zero. For example, the haptic device is not equipped with a force sensor, so its actual

force cannot be fed back. Thus, the general inputs to the operator and novice are

$$f_h = k_{fo} f_n + k_{po}(x_n - x_o) + k_{do}(\dot{x}_n - \dot{x}_o) \quad (7)$$

$$\dot{x}_\nu = \dot{x}_o + k_{pn}(x_o - x_n) + k_{fn}(f_h - f_n). \quad (8)$$

Let  $\mathbf{K}_{oo} = [k_{po} \ k_{do} \ 0]$ ,  $\mathbf{K}_{on} = [k_{pn} \ 1 \ k_{fn}]$ ,  $\mathbf{K}_{nn} = [k_{pn} \ 0 \ k_{fn}]$ , and  $\mathbf{K}_{no} = [k_{po} \ k_{do} \ k_{fo}]$ . Then, in terms of the individual model outputs, these expressions become

$$f_h = \mathbf{K}_{no} \mathbf{y}_n - \mathbf{K}_{oo} \mathbf{y}_o \quad (9)$$

$$\dot{x}_\nu = \mathbf{K}_{on} \mathbf{y}_o - \mathbf{K}_{nn} \mathbf{y}_n. \quad (10)$$

For nonzero time delay, the equations are

$$f_h = e^{-s\tau} \mathbf{K}_{no} \mathbf{y}_n - \mathbf{K}_{oo} \mathbf{y}_o \quad (11)$$

$$\dot{x}_\nu = e^{-s\tau} \mathbf{K}_{on} \mathbf{y}_o - \mathbf{K}_{nn} \mathbf{y}_n. \quad (12)$$

The full state-space notation of the closed-loop teleoperation system is shown in the Appendix. In addition, (13) shows how the model of Fig. 4 relates to the generic teleoperation model of Fig. 2

$$\begin{aligned} C_1 &= 1 + \frac{1}{s} k_{pn} & C_2 &= k_{fo} \\ C_3 &= k_{fn} & C_4 &= k_{do} + \frac{1}{s} k_{po} \\ C_v^n &= \frac{1}{s} k_{pn} & C_f^n &= k_{fn} \\ C_v^h &= k_{do} + \frac{1}{s} k_{po} & Z_e &= b_e + \frac{1}{s} k_e \\ Z_o &= m_o s + b_o + k_{o2} + \frac{1}{s} k_{o1} + \frac{1}{s^2} k_a. \end{aligned} \quad (13)$$

And  $Z_h$  depends on the haptic device.

### C. Transparency Conditions

Bilateral teleoperation systems are commonly expressed using a hybrid representation as follows:

$$\begin{bmatrix} f_h \\ -\dot{x}_n \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} \dot{x}_o \\ f_n \end{bmatrix} = \mathbf{H}(s) \begin{bmatrix} \dot{x}_o \\ f_n \end{bmatrix} \quad (14)$$

For perfect transparency,  $\dot{x}_o = \dot{x}_n$  and  $f_h = f_n$ , so the ideal hybrid matrix consists of elements  $h_{11} = h_{22} = 0$ ,  $h_{12} = 1$ , and  $h_{21} = -1$ . Equations (11) and (12) together with the operator

and novice models can be expressed as a hybrid matrix with the following elements:

$$h_{11} = \frac{e^{-s^2\tau}(k_{po} + k_{do}s)(1 - G_n^{-1})}{G_n^{-1}s + k_{pn} - e^{-s^2\tau}k_{fn}(k_{po} + k_{do}s)} \quad (15)$$

$$h_{12} = \frac{(G_n^{-1}s + k_{pn})k_{fo} - k_{fn}(k_{po} + k_{do}s)}{G_n^{-1}s + k_{pn} - e^{-s^2\tau}k_{fn}(k_{po} + k_{do}s)}e^{-s\tau} \quad (16)$$

$$h_{21} = -\frac{s + k_{pn} - e^{-s\tau}k_{fn}(k_{po} + k_{do}s)}{G_n^{-1}s + k_{pn} - e^{-s^2\tau}k_{fn}(k_{po} + k_{do}s)} \quad (17)$$

$$h_{22} = \frac{k_{fn}(1 - e^{-s^2\tau}k_{fo})s}{G_n^{-1}s + k_{pn} - e^{-s^2\tau}k_{fn}(k_{po} + k_{do}s)}. \quad (18)$$

Notice that all elements of the  $\mathbf{H}$  matrix share the same denominator. Consider the case where time delay approaches zero ( $\tau \rightarrow 0$ ) and we use unity-gain force feedback ( $k_{fo} = 1$ ). Then,  $h_{22} = 0$  and  $h_{12} = 1$ . The terms  $h_{11}$  and  $h_{21}$  achieve their ideal values of 0 and  $-1$ , respectively, if and only if  $G_n = 1$ , which would occur if the novice were infinitely fast and perfectly accurate. Of course, this is not the case and the transient response will never be exactly unity. Thus, with no time delay and  $k_{fo} = 1$ , the hybrid matrix becomes

$$\mathbf{H}(s) = \begin{bmatrix} \frac{(k_{po} + k_{do}s)(1 - G_n^{-1})}{G_n^{-1}s + k_{pn} - k_{fn}(k_{po} + k_{do}s)} & 1 \\ -\frac{s(1 - G_n^{-1})}{G_n^{-1}s + k_{pn} - k_{fn}(k_{po} + k_{do}s)} - 1 & 0 \end{bmatrix}. \quad (19)$$

However, at steady state, if there is no remaining error,  $G_n$  tends to 1, so  $h_{11} = 0$  and  $h_{21} = -1$ . Similarly, if the operator velocity is very small, for example, when in contact with an environment during an ultrasound,  $v_o \rightarrow 0$ , so transparency is achieved at steady state.

The novice's response,  $G_n$ , can only be altered through practice, not external control, but  $k_{po}$ ,  $k_{do}$ ,  $k_{pn}$ , and  $k_{fn}$  can be tuned to bring  $h_{11}$  closer to 0. Indeed, taking either novice gain,  $k_{pn}$  or  $k_{fn}$ , to infinity drives  $h_{11}$  and  $h_{21}$  to their ideal values. With very large operator gains,  $k_{po}$  and  $k_{do}$ ,  $h_{21} \rightarrow 0$  and  $h_{11} \rightarrow \frac{G_n - 1}{k_{fn}G_n}$ . However, such large gains are prevented by practical considerations such as haptic device joint torque limits and stability. Adding time delay further degrades the transparency.

#### D. Control Architectures

In this subsection, various possible teleoperation architectures based on this model are described. Their stability is analyzed in Section II-E.

*1) Two-Channel Teleoperation:* Three two-channel teleoperation approaches were implemented, differing in which variable is reflected from novice to operator and whether the novice receives local position feedback. In each case, the operator motion is sent to the novice.

*a) Position-position (PP) ( $C_2 = C_3 = 0$ ):* In this architecture, the novice position is fed back to the operator, and a force proportional to the position error and its time derivative is applied to the haptic device. This is one of the first bilateral teleoperation schemes to have been developed and has the advantage of not requiring a force sensor at the novice. In this

case, however,  $k_{fo} = k_{pn} = k_{fn} = 0$ , which leads to problems with the transparency. In Fig. 2, this is equivalent to setting  $C_2 = C_3 = C_v^n = C_f^n = 0$ . The hybrid matrix becomes

$$\mathbf{H}(s) = \begin{bmatrix} e^{-s^2\tau} \left( \frac{k_{po}}{s} + k_{do} \right) (G_n - 1) & 0 \\ -G_n & 0 \end{bmatrix}. \quad (20)$$

The obvious problem is that the haptic device force has no relation to the novice force. In practice, the bigger problem is that the novice lags the operator and the returned force is further delayed, so the operator always feels forces resisting their motion, which makes operation difficult and frustrating. This is evident in  $h_{11}$ , which becomes further from 0 as the novice response  $G_n$  departs from unity and as the round-trip time of the communication ( $2T$ ) increases.

*b) Position-force (PF) ( $C_3 = C_4 = 0$ ):* This teleoperation scheme is the naïve method of providing force feedback, with position being sent from operator to novice and measured forces being fed back to the operator. In Fig. 2, this is equivalent to setting  $C_3 = C_4 = C_v^n = C_f^n = C_v^h = 0$ , or, with our particular implementation,  $k_{po} = k_{do} = k_{pn} = k_{fn} = 0$ . The hybrid matrix then becomes

$$\mathbf{H}(s) = \begin{bmatrix} 0 & e^{-s\tau}k_{fo} \\ -G_n & 0 \end{bmatrix}. \quad (21)$$

The time-delayed force feedback is seen in  $h_{12}$ , and motion tracking is affected by the novice's response. If the tracking is not fast and accurate, it can lead to oscillation due to the delay. Often in practice,  $k_{fo} < 1$  is used to improve stability. Intuitively, this reduces the jerk on the operator when the novice side changes suddenly, e.g., when first touching the environment. Similarly, low-pass filtering the feed-back force injects damping into the system, greatly reducing the oscillatory behavior. Both, however, decrease transparency.

In addition, in HT, the novice can easily misjudge the alignment of the virtual and real tools due to imperfect depth perception as well as partial occlusion of the real tool by the virtual one. With this controller, there is no possibility to correct the novice's pose to account for such steady-state errors.

*c) Position-force with local position feedback (PFP) ( $C_3 = C_4 = 0$ ):* To correct for these steady-state errors, the pose of the novice can be measured, and the tracking error fed back locally. In particular, the control law for the novice is then

$$\dot{x}_v = \dot{x}_o + k_{pn}(x_o - x_n). \quad (22)$$

In Fig. 2, this is equivalent to setting  $C_3 = C_4 = C_f^n = C_v^h = 0$ . To render this on the HoloLens 2, the desired velocity is numerically integrated using the time step between successive frames, and the resulting pose is applied to the virtual probe. In practice, a deadband is used for the velocity error before integration, and the output  $x_v$  is saturated to avoid losing track of the virtual probe.

By adding the position error to the virtual probe signal, the tracking error is effectively amplified with an integrator. If the novice is not well aligned in one axis, the virtual probe starts moving further away in that axis, causing the novice to realize their error and better align the probe.

This control scheme creates the following hybrid matrix, with  $k_{po} = k_{do} = k_{fn} = 0$ :

$$\mathbf{H}(s) = \begin{bmatrix} 0 & e^{-s\tau}k_{fo} \\ -\frac{s+k_{pn}}{G_n^{-1}s+k_{pn}} & 0 \end{bmatrix}. \quad (23)$$

As expected, adding position error as local force-feedback provides an opportunity to modify  $h_{21}$ . As  $k_{pn}$  increases, if stability is maintained, we achieve kinematic correspondence.

2) *Three-Channel Teleoperation ( $C_4 = 0$ )*: As in the two-channel position–force teleoperation, only the novice’s force is reflected to the operator. However, both pose and force are sent from operator to novice, enabling the novice control law to include a local force feedback term

$$\dot{x}_v = \dot{x}_o + k_{pn}(x_o - x_n) + k_{fn}(f_h - f_n). \quad (24)$$

This is useful during contact phases of the teleoperation, in which the novice is to apply a certain desired force. With a relatively stiff environment, even a large change in force will not cause a change in the MR virtual display, so employing only position and velocity in the control law in (24) leads to relatively poor force tracking.

In Fig. 2, this is equivalent to setting  $C_4 = C_v^h = 0$ . The hybrid matrix with this controller becomes

$$\mathbf{H}(s) = \begin{bmatrix} 0 & e^{-s\tau}k_{fo} \\ -\frac{s+k_{pn}}{G_n^{-1}s+k_{pn}} & \frac{k_{fn}(1-e^{-s2\tau})s}{G_n^{-1}s+k_{pn}} \end{bmatrix}. \quad (25)$$

From this, it is apparent that the novice force now has a bearing on novice velocity, as desired. Numerically, this decreases transparency. However, in practice, the novice is relatively insensitive to small changes in virtual probe pose and the human hand has limited motion resolution. These effects are not well modeled and make the added local force feedback term desirable.

3) *Model-mediated teleoperation*: One problem with the aforementioned methods is that the operator’s sensation depends directly on the novice’s actions. While this gives a true representation of the system on the novice side, it can lead to an unsteady experience for the operator who feels every small jolt and inadvertent movement of the novice. This effect can be reduced using low-pass filtering, but only at the cost of response speed. Furthermore, the performance decreases with increasing time delays.

One method to potentially overcome both problems is to render to the operator a local virtual haptic environment that is a replica of the novice’s real environment. This is known as model-mediated teleoperation [37]. In our present system, a point cloud or mesh of the patient is captured and sent to the operator. This mesh is used as a virtual fixture or keep-out volume by the haptic device [1]. An outward force dependent on an estimated impedance and the novice’s tool pose within the mesh can be applied, as long as the impedance estimation is significantly faster than the change in the environment (i.e., the patient moves slowly).

With estimated parameters  $\hat{b}_e, \hat{k}_e$ , the haptic device and novice forces are given by

$$f_h = -\hat{b}_e \dot{x}_o - \hat{k}_e x_o \quad (26)$$

$$f_n = -b_e \dot{x}_n - k_e x_n. \quad (27)$$

Plugging in the novice model of (1) gives

$$F_h = -(\hat{b}_e s + \hat{k}_e) X_o \quad (28)$$

$$F_n = -(b_e s + k_e) G_n x_\nu. \quad (29)$$

For the virtual probe pose, we can substitute this into (8)

$$x_\nu = e^{-s\tau} X_o + k_{pn}(e^{-s\tau} X_o - G_n x_\nu) + k_{fn}(e^{-s\tau} F_h - f_n) \\ = \frac{1 + k_{pn} - k_{fn}(\hat{b}_e s + \hat{k}_e)}{1 + [k_{pn} - k_{fn}(b_e s + k_e)] G_n} e^{-s\tau} X_o. \quad (30)$$

Thus, we obtain an expression for the novice velocity as follows:

$$V_n = \frac{1 + k_{pn} - k_{fn}(\hat{b}_e s + \hat{k}_e)}{1 + [k_{pn} - k_{fn}(b_e s + k_e)] G_n} G_n e^{-s\tau} V_o. \quad (31)$$

Similarly, from (26) and (27), the novice force is

$$F_n = \left( \frac{b_e s + k_e}{\hat{b}_e s + \hat{k}_e} \right) \frac{1 + k_{pn} - k_{fn}(\hat{b}_e s + \hat{k}_e)}{1 + k_{pn} - k_{fn}(b_e s + k_e)} G_n e^{-s\tau} F_h. \quad (32)$$

Here, it is clear that if  $G_n = 1$  and the impedance estimate is perfect, the novice motion and force will correspond precisely to the delayed operator motion and force, so the hybrid matrix will be

$$\mathbf{H}(s) = \begin{bmatrix} 0 & e^{sT} \\ -e^{-s\tau} & 0 \end{bmatrix}. \quad (33)$$

Notice how the time delay in the force is negative. Depending on the communication delay, the operator feels the force feedback from the local model before the novice has applied it. This way, the model-mediated method gives predictive haptic feedback.

In general, even with imperfect impedance estimation, the damping is very small, and for many tasks, the steady-state velocity is zero, so damping is negligible. Assuming  $G_n$  is constant, the steady-state gain is

$$f_n = \left( \frac{k_e}{\hat{k}_e} \right) \frac{k_{pn} - k_{fn}\hat{k}_e}{k_{pn} - k_{fn}k_e} f_h. \quad (34)$$

If  $k_{pn} - k_{fn}k_e > 0$  and the environment impedance is underestimated, i.e.,  $k_e > \hat{k}_e$ , the operator moves too far into the environment for a given force, so the novice must press harder to match the virtual tool and  $f_n > f_h$ . Conversely, overestimation leads to  $f_n < f_h$ .

If the impedance model is inaccurate, a small  $k_{fn}$  reduces the pose tracking error due to the mismatch, but the force experienced by the operator remains proportionally inaccurate. If the novice response,  $G_n$  is poor, large  $k_{fn}$  can suppress the error. Thus, every effort should be made to have an accurate and fast impedance estimation and environment model while using a reasonably large  $k_{fn}$  to compensate for the novice response. Making  $k_{fn}$  too large, however, becomes unintuitive for the novice and amplifies modeling errors. In addition, taking  $k_{pn} = 0$  ensures a steady-state gain of 1 but may lead to poor position tracking. Thus, the best strategy may be to have large

$k_{fn}$  and zero  $k_{pn}$  in the normal direction to the tissue (the direction of the force), and nonzero  $k_{pn}$  in the tangent directions.

### E. Stability

Since the novice is a person, not a robot, instability does not lead to violent motions that cause harm or damage. However, it does render the teleoperation unusable, so some analysis of stability can give insight into how well the system will perform. In this section, we use a state-space formulation of the model that is derived in the [Appendix](#).

For the zero time-delay case, the system stability is given by the eigenvalues of the state-space  $\mathbf{A}$  matrix in [\(44\)](#). These depend on the choice of the feedback and feedforward gains from [\(7\)](#) and [\(8\)](#).

For PF control with no local feedback loops, only  $k_{fo}$  is nonzero. In this case,  $\mathbf{A}$  is stable. However, it has one eigenvalue close to zero. Increasing  $k_{fo}$  makes the novice slower and more oscillatory and in turn pushes this eigenvalue closer to zero. At large enough  $k_{fo}$ , the system becomes unstable. Thus, we expect that with poor tracking from the novice, the PF feedback will be oscillatory.

Conversely, if we add local position and force feedback for the novice, the hybrid matrix is given in [\(25\)](#). By the Routh–Hurwitz criterion for a third degree polynomial,  $h_{21}$  from the hybrid matrix,  $\mathbf{H}$ , is stable if

$$b_n(k_n + k_{pn}b_n) - m_nk_{pn}k_n > 0. \quad (35)$$

This is clearly true if  $k_{pn} = 0$ , meaning no local position feedback. With the feedback, and using the parameter values from Table I, the system is stable for any  $k_{pn} \geq 0$ . Depending on the novice's behavior ( $k_n$  and  $b_n$ ), however, this may change. Notice also that with no time delay, the stability is independent of the local force feedback,  $k_{fn}$ .

An alternative criterion to determine the stability of the teleoperation system in Fig. 1 uses the scattering matrix

$$\mathbf{S}(s) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} (\mathbf{H}(s) - I)(\mathbf{H}(s) + I)^{-1}. \quad (36)$$

If  $\|\mathbf{S}(s)\|_2 \leq 1$ , the two-port network in Fig. 1, terminated by  $Z_e$  and  $Z_o$ , is passive if  $Z_e$  and  $Z_o$  are passive, and therefore, the overall system is stable [\[38\]](#). For example, the hybrid matrix in [\(33\)](#) leads to a scattering operator norm of 1, so stability is maintained irrespective of delay. However, this is more difficult to show when the novice response or parameter estimation are nonideal. In reality, the impedance estimation is time-varying, delayed, and not perfectly accurate. Furthermore, passivity is a conservative method of guaranteeing stability, so a system may be stable even if  $\|\mathbf{S}(s)\|_2 > 1$ . Less conservative criteria were developed for passive  $Z_e$  and  $Z_o$  using Llewellyn's absolute stability criterion or the structured singular value of  $S$ , and the conservatism can be further reduced when bounded impedances are considered [\[32\]](#).

### F. Experimental Setup

Section II thus far has shown how the developed models enable predictions of how different controllers will affect the system. For example, local pose and force feedback will likely improve the tracking and speed of the position control, while model-mediated teleoperation will be more stable with time delays but dependent on the model's fidelity. Adding local feedback will decrease this dependence and improve accuracy. The following sections describe experiments to determine if the modeling is accurate, the controllers effective, and the assumptions justified, thus enabling future development of more complex controllers using the same methods. Moreover, we will show empirically that stability is much less of a concern in HT than in robotic teleoperation, and is in this case secondary to performance and transparency.

To test the HT architectures in a controlled and close-to-ideal environment, we developed the following system. In order to measure the novice's motions and forces while presenting them with a known environment, instead of interacting with a patient or a tissue phantom, their tool was connected to the end effector of a second haptic device. The novice and operator sides thus both consisted of a host PC running Windows 11 and a haptic device (Touch X, 3D Systems, Rock Hill, SC, USA). The haptic devices were controlled via C++ programs using OpenHaptics, with a graphical user interface written in Qt. The two clients communicated over a fast WebRTC connection, exchanging position, velocity, and force data, as well as occasional synchronization messages to measure the communication time delay [\[4\]](#). The time delay was controlled by handling all sending and receiving of WebRTC messages on the operator side on a separate thread, which delayed every message by a set time before sending it or forwarding it to the main graphics and haptics threads. The novice side also had a Microsoft HoloLens 2 MR HMD which communicated with the novice side PC application over a local WebSocket, receiving messages forwarded from the WebRTC connection.

The virtual tool used in the tests was an ultrasound probe viewed through the HoloLens 2. An identical ultrasound probe shape was 3-D printed and attached to the novice's Touch X stylus. Four infrared (IR) reflective spheres were also attached to the dummy probe and were tracked by the HoloLens 2, allowing the HoloLens to compute the haptic device end effector pose in the HoloLens frame [\[5\]](#). This was used to perform a registration before every test. Suppose the operator's haptic device has base frame represented by the homogeneous transform  ${}^w\mathbf{T}_o$  relative to the world while the novice's has base frame  ${}^w\mathbf{T}_n$  and the HoloLens is at  ${}^w\mathbf{T}_h$ , as shown in Fig. 6. When the operator moves their haptic device end effector to a pose  ${}^w\mathbf{T}_o {}^o\mathbf{T}_e$ , the virtual probe is rendered in the HoloLens at pose  ${}^w\mathbf{T}_h {}^h\mathbf{T}_v$ , and the novice moves their handle to pose  ${}^w\mathbf{T}_n {}^n\mathbf{T}_e$ . The virtual probe should be rendered such that when the novice aligns perfectly to the virtual probe (i.e.,  ${}^w\mathbf{T}_n {}^n\mathbf{T}_e = {}^w\mathbf{T}_h {}^h\mathbf{T}_v$ ), their handle is at  ${}^n\mathbf{T}_e = {}^o\mathbf{T}_e$ . This way, if the novice matches the operator, the poses output by the haptic device control software, OpenHaptics, are equal. Note that the novice pose is  ${}^n\mathbf{T}_e = ({}^w\mathbf{T}_n)^{-1} {}^w\mathbf{T}_h {}^h\mathbf{T}_v$ , so to achieve  ${}^n\mathbf{T}_e = {}^o\mathbf{T}_e$ , the virtual probe

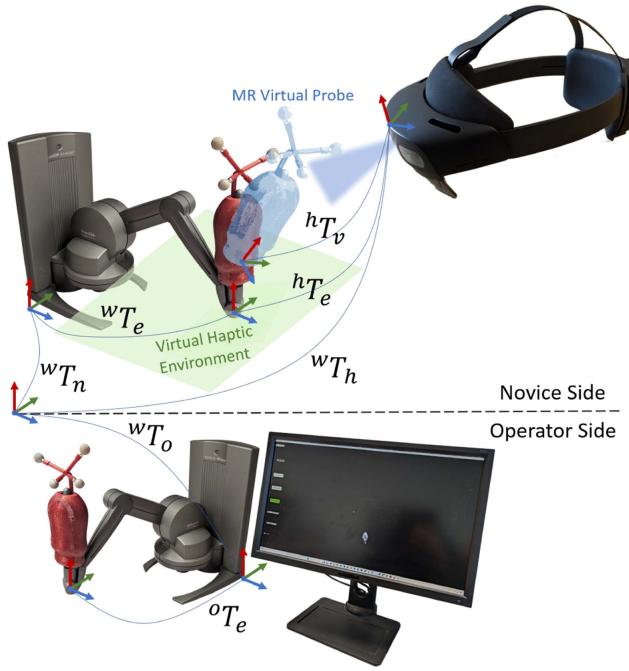


Fig. 6. Experimental setup showing the operator and novice sides, which shared one host computer but were separated by a screen. The virtual probe rendered by the HoloLens 2 is shown in semitransparent blue. The red objects on the haptic devices are the ultrasound probe-shaped end effectors with retro-reflective IR markers. The novice-side haptic device is used only to provide a precisely known novice environment during the tests, and to measure the novice's pose and force.

pose must be set to  $^h\mathbf{T}_v = (^w\mathbf{T}_h)^{-1} ^w\mathbf{T}_n \circ ^o\mathbf{T}_e$ . For this, we must determine the novice haptic device base frame relative to the HoloLens:  $^h\mathbf{T}_n = (^w\mathbf{T}_h)^{-1} ^w\mathbf{T}_n$ .

Thus, at the start of each test, the novice haptic device end effector was moved to several different positions and orientations while the HoloLens recorded the measured pose ( $^h\mathbf{T}_e$ ) and received the actual pose from OpenHaptics ( $^n\mathbf{T}_e$ ) over the WebSocket. These are related by  $^w\mathbf{T}_h \cdot ^h\mathbf{T}_e = ^w\mathbf{T}_n \cdot ^n\mathbf{T}_e$ , so  $^h\mathbf{T}_n = (^w\mathbf{T}_h)^{-1} ^w\mathbf{T}_n = ^h\mathbf{T}_e \cdot (^n\mathbf{T}_e)^{-1}$ . After recording several hundred samples of  $(^h\mathbf{T}_e, ^n\mathbf{T}_e)$ , the least-squares-optimal value of  $^h\mathbf{T}_n$  was computed using the Kabsch–Umeyama algorithm, which ensures a valid homogeneous transform using singular value decomposition [39].

During tests, the operator performed arbitrary motions with their haptic device. The motion and force were sent to the novice PC over WebRTC, and forwarded to the HoloLens over the WebSocket. Here, they were transformed into HoloLens coordinates and rendered to the novice through the virtual probe pose. The novice grasped the dummy ultrasound probe on their haptic device end effector and aligned it with the virtual probe to follow the desired motion. A virtual environment in the form of a flat surface of specified stiffness and damping was rendered haptically to the novice, leading to forces generated due to the motions. The applied forces and novice motions were sent to the HoloLens for rendering local feedback, as described previously, and to the operator for haptic feedback. Depending on the architecture being tested, the operator's haptic device either applied forces based on this feedback and/or simply rendered

an identical virtual environment to the novice one to test the model-mediated feedback.

The operator and novice PC clocks were synchronized by sending timestamps along with every data message, and all forces, positions, orientations, and velocities were in the same coordinate frame since both sides used identical haptic devices. All data received on the operator side was recorded to a file for later analysis in MATLAB (Mathworks Inc., Natick, MA, USA).

### G. Tests

Using this setup, we performed preliminary tests of the various teleoperation architectures. These tests provide a validation of our mathematical modeling under controlled conditions. The architectures we considered are listed as follows and described in the previous sections.

- 1) Position–position (PP):  $\dot{x}_v = \dot{x}_o$  and  $f_h = k_{po}(x_n - x_o) - k_{do}\dot{x}_o$ .
- 2) Position–force (PF):  $\dot{x}_v = \dot{x}_o$  and  $f_h = k_{fo}f_n - k_{do}\dot{x}_o$ .
- 3) Position–force with local position feedback (PFPF):  $\dot{x}_v = \dot{x}_o + k_{pn}(x_o - x_n)$  and  $f_h = k_{fo}f_n - k_{do}\dot{x}_o$ .
- 4) Position–force with local pose and force feedback (PFPF):  $\dot{x}_v = \dot{x}_o + k_{pn}(x_o - x_n) + k_{fn}(f_h - f_n)$  and  $f_h = k_{fo}f_n - k_{do}\dot{x}_o$ .
- 5) Model-mediated (M):  $\dot{x}_v = \dot{x}_o$  and  $f_h = \begin{cases} k_e(x_0 - x_o) - b_e\dot{x}_o & \text{if } x_o \text{ inside mesh} \\ -b_e\dot{x}_o & \text{otherwise} \end{cases}$  where  $x_0$  is the position of the virtual mesh surface.
- 6) Model-mediated with local position feedback (MP): Same as (M) but with  $\dot{x}_v = \dot{x}_o + k_{pn}(x_o - x_n)$ .
- 7) Model-mediated with local position and force feedback (MFP): Same as M and MP but with  $\dot{x}_v = \dot{x}_o + k_{pn}(x_o - x_n) + k_{fn}(f_h - f_n)$ .

Note, in MFP, the force feedback was given normal to the surface and the position feedback was tangent to the surface. In PFPF and PFPF, a deadband of an acceptably small distance error,  $d_{err}$ , is applied to the integrator since the error is never exactly zero. The threshold was chosen based on the steady-state tracking error from step response tests [4]. Moreover, when integrating  $\dot{x}_v$  to render the virtual probe position, the offset from the actual desired position was limited to be less than  $d_{max}$ .

The controller parameters were tuned by hand before testing to give good nominal (no time delay) performance while remaining stable according to Section II-E. This was also attempted with time delays, but the performance remained relatively poor, as shown below, except for the model-mediated approaches. Thus, similar parameter values were used for a given controller for all tests at different time delays.

For each teleoperation architecture, the operator performed arbitrary smooth motions for 1–4 min with the novice tracking. The position and force root-mean-squared (rms) tracking errors were analyzed, as well as the steady-state error for steplike motions. Next, the two most promising methods of each type (force feedback and model-based feedback) were selected and tested again five times, each time with increasing communication time delay (0, 50, 250, 500, and 1000 ms). Then, the same methods were tested with three different environment stiffness values

TABLE II  
TELEOPERATION PERFORMANCE FOR DIFFERENT CONTROL SCHEMES AT ZERO TIME DELAY

	$e_p$ (mm)	$e_{ps}$ (mm)	$e_f$ (N)	$e_{fs}$ (N)	$\tau$ (ms)
PF	7.92	6.57	0.18	0.12	440
PFP	5.91	3.83	<b>0.16</b>	0.11	38
PFPF	6.18	<b>1.52</b>	0.33	<b>0.07</b>	48
M	6.75	5.12	2.66	2.45	254
MF	6.91	6.23	0.94	0.9	284
MFP	<b>3.86</b>	3.38	0.58	0.47	<b>36</b>

The RMS position tracking error,  $e_p$ , steady-state position error,  $e_{ps}$ , RMS force tracking error,  $e_f$ , steady state force error,  $e_{fs}$ , and approximate time delay,  $\tau$ , between novice and operator are shown. The bold values are the best (lowest) error values.

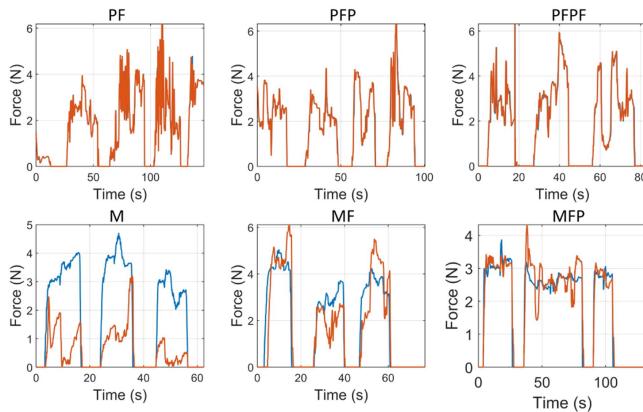


Fig. 7. Force tracking with no time delay using various teleoperation schemes.

(0.2, 0.4, and 0.8 N/mm) and no time delay, with and without low-pass filtering.

In the tests, the operator was one of the authors, and the novice was a volunteer. Five novices were tested, including three males and two females with age ranging from 22 to 57, with various backgrounds. None were experienced in ultrasound or MR. The volunteers were instructed to follow the virtual probe closely, focusing on matching its pose as accurately and quickly as possible. Ethics approval for the tests was obtained from the University of British Columbia Behavioural Research Ethics Board (BREB), approval number H22-01195. Informed consent was given by all participants.

### III. RESULTS

#### A. Tracking Tests With No Delay

The average results of the first test are shown in Table II, and one set of tests is plotted in Figs. 7 and 8. For these nominally zero time delay tests, the communication round trip time was measured to be  $3.69 \pm 2.67$  ms (average  $\pm$  standard deviation) using the method described in [4]. For later nonzero delay tests, the nominal time delay is also reported but may vary similarly.

Even with no delay, it was immediately apparent that PP teleoperation was impractical. The lag of the novice in tracking the input signal made it difficult for the operator to move at all if the stiffness was high, while low stiffness gave the operator very poor feeling of where the surface was. In addition, the novice, being human, does not track very smoothly, so the operator continually received unwanted and jarring feedback.

Low-pass filtering helped but increased the delay, which again made it harder for the operator to move. Thus, while the position tracking was by definition good, the teleoperation experience was intolerable and the method was removed from the tests.

Direct force feedback (PF controller) performed much better, although the lack of explicit position coordination on the operator side was apparent, with relatively large tracking errors. In addition, contact was oscillatory: due to the tracking lag of the novice, the operator initially moved well below the surface before receiving force feedback, which was stronger than anticipated due to the novice's attempt to reach the virtual tool inside the surface. This jerked the operator's hand upward, which the novice tried to track, thus decreasing the force, so the operator's hand moved down again. This leads to periodic and uncomfortable motion when in intermittent contact. This is similar to the chattering effect in robots, but at a lower rate and larger amplitude. This back-and-forth reflection of waves is very apparent in Fig. 7. Both behaviors match what was expected from (21).

Adding local position feedback for the novice (PFP controller) helped decrease position error and substantially reduced tracking lag. By effectively magnifying any tracking error, the local feedback has a predictive effect, moving the probe further than the operator has yet moved and encouraging the novice to react immediately. This reduced the oscillation during contact. The feedback also led to much more accurate position tracking, as shown in Fig. 8. With local position and force feedback (PFPF controller), position tracking suffered slightly, since the virtual probe no longer exclusively represents the desired pose, but aims to control the force as well. In (25), this is reflected by the nonzero  $h_{22}$  element, while  $k_{pn}$  enables the more responsive position tracking. Interestingly, rms force tracking also became slightly worse since the increased feedback to the novice sometimes caused initial overshoot. However, in steady state, the force quickly converged to be more accurate than without local feedback. The addition of force additionally had a damping effect during contact, discouraging sudden changes in force, and thus, leading to much less oscillation.

The teleoperation architectures based on a local model or mesh showed different behavior. Contact oscillation was completely eliminated due to the stable and unchanging mesh. This led to a much more comfortable and intuitive experience. With local position feedback, the position tracking was the best of any of the methods. On the other hand, force tracking was completely lacking unless explicitly enforced with local feedback (MFP controller), as shown in Fig. 7. Even so, the force experienced by the operator was not as accurate as when the measured force was fed back directly, and is affected by modeling errors in the surface shape and impedance. Again, adding local position feedback (MP and MFP controllers) greatly reduced the tracking lag.

#### B. Tracking Tests With Delay

The effect of communication delays on the force and position tracking is shown in Figs. 9 and 10 for the PFPF and MFP controllers. Due to the delay, discussing rms error is not meaningful. However, several important results are apparent in the

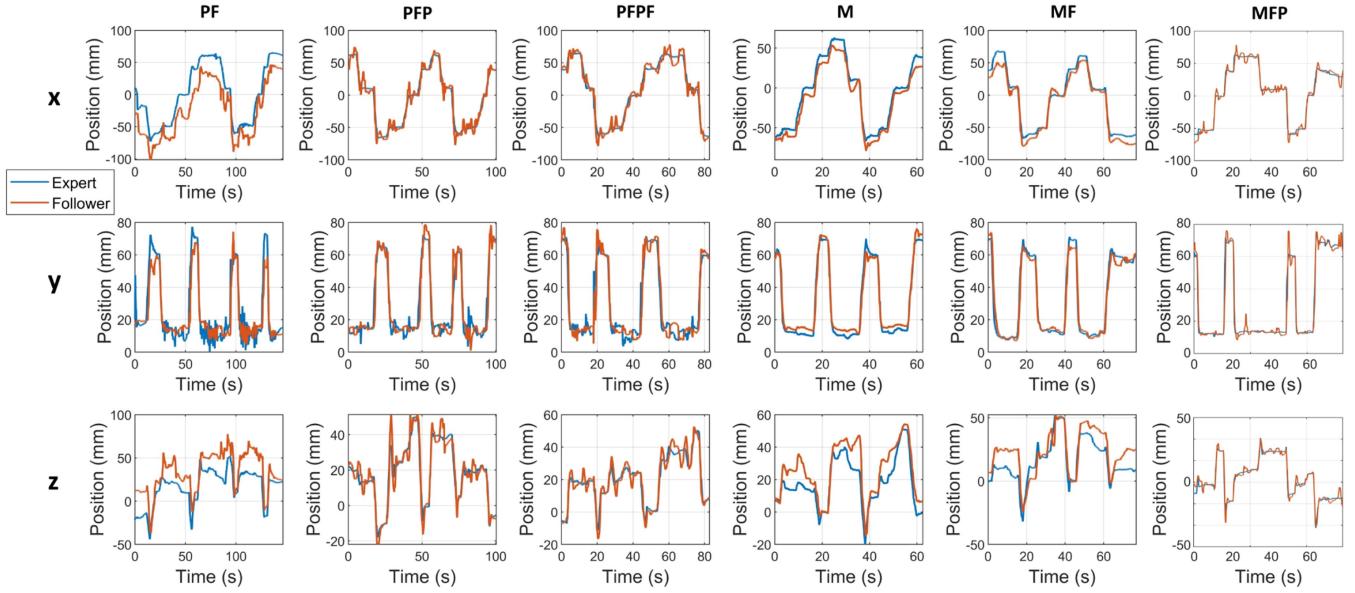


Fig. 8. Position tracking with no time delay for the six teleoperation architectures. PF shows significant oscillation during contact, while PF, M, and MF have relatively poor position tracking. Adding explicit local position feedback for the novice greatly improves the tracking accuracy and speed.

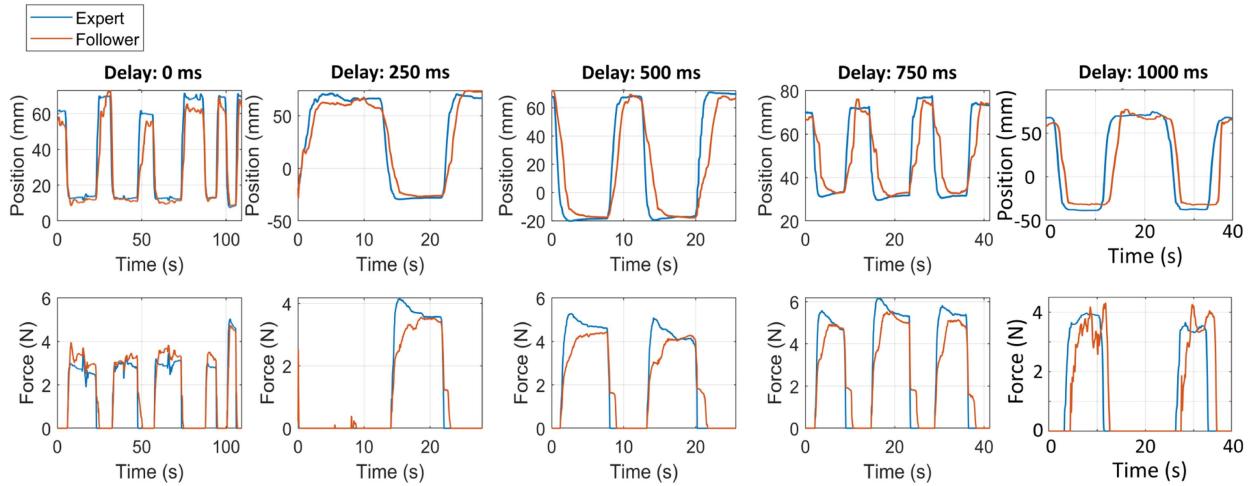


Fig. 9. Position and force tracking (both normal to the virtual surface) in the presence of time delays with the MFP controller (mesh local model with local position and force feedback). Time delay has little effect on the force tracking and only delays the position tracking.

plots. With the three-channel architecture (PFPF), the character of the response underwent two distinct phases. Initially, up to between 250 and 500 ms delay, the response became increasingly oscillatory as the wave reflections were exacerbated by the larger delays. At 500 ms, however, the operator's approach changed and they started moving more slowly and carefully, anticipating the unexpected application of force feedback. As a result, the oscillations ceased. However, as the delay continued increasing, the operator was deeper and deeper within the surface before the novice's force was applied, measured, and fed back. This led to excessive offsets between the virtual probe and the novice, making tracking impossible.

Conversely, with model-mediated MFP teleoperation, the operator always received immediate haptic feedback with no delay. Consequently, their motion was unaffected and the novice was able to track without problems, simply delayed by the given

amount. Thus, while the delay made teleoperation impractical with direct force feedback, it had little effect when operating with a local model. Due to the local position and force feedback on the novice side, the tracking was good for both. This reflects (33) with perfect modeling, but will degrade with an imperfect local model.

### C. Environment Stiffness and Low-Pass Filtering

Finally, the effect of environment stiffness on the teleoperation performance was tested. At zero delay, the stiffness was varied from 0.4 N/mm (the value used for all other tests) down to 0.2 N/mm to simulate a very soft environment and up to 0.8 N/mm for a stiff environment. The damping was kept constant and was not noticeable. These tests were performed with the PFPF architecture since the stiffness has little effect on

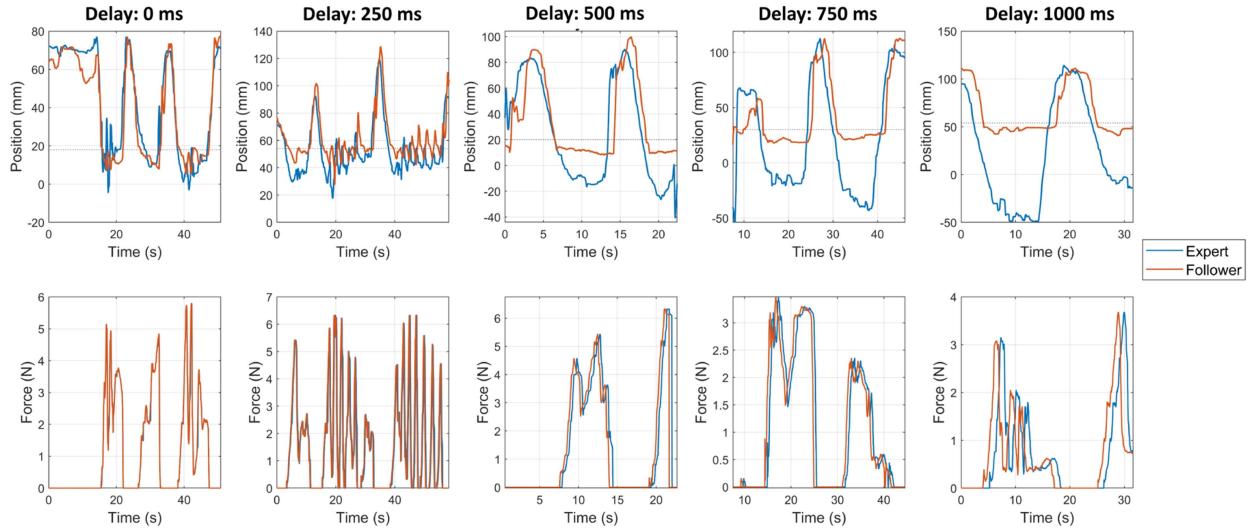


Fig. 10. Position and force tracking (both normal to the virtual surface) in the presence of time delays with the PFPF controller (position–force teleoperation with local position and force feedback for the novice). With delays over approximately 500 ms, tracking becomes infeasible.

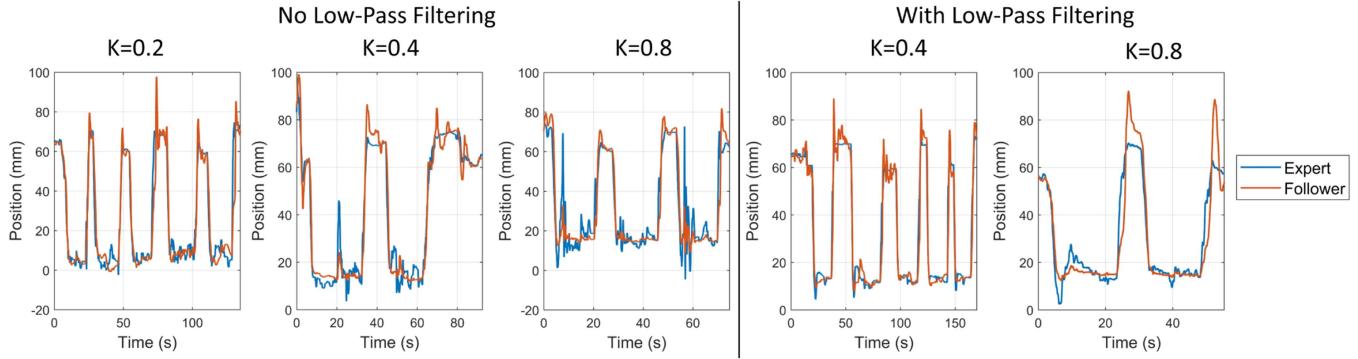


Fig. 11. Position tracking at different environment stiffnesses with and without low-pass filtering, with the PFPF teleoperator. In contact (position close to 0 mm), the oscillations increase with stiffness but decrease with low-pass filtering.

TABLE III  
TELEOPERATION PERFORMANCE (RMS TRACKING ERROR IN MILLIMETERS)  
FOR STIFF AND SOFT ENVIRONMENTS, WITH AND WITHOUT LOW-PASS  
FILTERING (LP)

Stiffness (N/mm)	0.2	0.4	0.8
PFPF	6.48	6.18	7.53
PFPF + LP	5.04	4.78	6.31

the local mesh feedback as long as it is approximately equal for the operator and novice.

It was already noted in previous tests that the direct force feedback approaches had oscillatory behavior in contact due to the delay of the novice. Therefore, low-pass filtering of the feedback force was also tested, using a single-pole infinite impulse response filter with a time constant of approximately 0.1 s. This injects damping into the system and is expected to decrease the oscillation.

The results of these tests are shown in Table III and Fig. 11. The oscillatory behavior increases with stiffness, as expected, but strongly decreases in the presence of low-pass filtering, without greatly increasing the phase lag. This leads to better rms error, but more importantly substantially improved the feel

for the operator. The feedback was much more stable and less disruptive than without the filtering. The lower environment stiffness slightly increased tracking error compared to normal. This is because on the softer virtual surface, the operator and novice could move more for small changes in force, making it harder to maintain a precise position.

#### D. Robustness to Parameter Variation

Throughout this article, we have made a number of simplifying assumptions about the system. In particular, the novice and operator models are taken to be constant. However, they may be affected by external, time-varying influences such as distractions or fatigue and, as described in Section II-D1b, the novice may have imperfect depth perception. In addition, in ultrasound exams, for example, the novice is often hesitant or unable to apply large forces on a patient, so may fail to follow a virtual tool placed too far within the patient. In this case,  $f_n \neq -k_e x_n - b_e \dot{x}_n$ , as assumed in Fig. 3, so additional dynamics are introduced. The linear time-invariant assumption in the model is thus not always accurate, and there is both parametric and dynamic uncertainty in the novice and operator models. Thus, the mathematical

TABLE IV  
ACCEPTABLE PERCENT VARIATION OF PARAMETER VALUES FROM NOMINAL TO MAINTAIN STABILITY FOR UNCERTAINTY ON ONLY THE OPERATOR, ONLY THE NOVICE, ONLY THE ENVIRONMENT, OR ALL THREE

Parameter	Uncertainty on			
	Operator	Novice	Environment	All
$k_e$			58%	33%
$b_e$			58%	33%
$b_n$		41%		39%
$m_n$		85%		0.4%
$m_o$	92%			83%
$b_o$	9.2%			8.3%
$k_a$	18%			17%
$k_{o1}$	100%+			100%+
$k_{o2}$	18%			17%

modeling can be improved through stochastic approaches as both users' actions are not deterministic and their imperfect sensory perception and motor skills lead to random fluctuations in the motions. These imperfections could be treated as disturbances in the existing model in future work. However, as shown in Fig. 5 and the related fit error, the models with constant parameters were able to capture complex recorded motions for an extended period of time without any drift due to parameters changing. Indeed, with controller gains  $\mathbf{K}_{no}$ ,  $\mathbf{K}_{on}$ ,  $\mathbf{K}_{oo}$ , and  $\mathbf{K}_{nn}$  chosen to give realistic performance at the parameter values identified in Table I, the acceptable ranges of the model parameters are shown in Table IV. To compute these values, the structured singular value,  $\mu$ , of the closed-loop system was computed in MATLAB using the robstab command and the state-space model in the Appendix, with uncertainty on different combinations of parameters. On average, all parameters simultaneously can vary up to approximately 37% without causing instability. Thus, the LTI assumption gives a reasonable representation of the system.

In addition, although it is assumed that the novice can handle intermittent environment contact implicitly without affecting their tracking, switching contact affects the force fed back to the operator and may thus destabilize the closed-loop system. This is shown in Fig. 10, where the tracking is unstable at 250-ms delay during intermittent contact, leading to very ineffective teleoperation. This shows that instability in HT does not constitute a health or equipment hazard but makes guidance impossible. However, at higher delays, this instability disappears due to an apparent change in strategy. The instability is also eliminated when using model-mediated teleoperation.

#### IV. DISCUSSION

This article has introduced the concept of bilateral HT as a control problem analogous to robotic bilateral teleoperation but with some unique characteristics. The modeling and simulation of the system were described, and various control architectures were explored in terms of their implementation, theoretical transparency and stability, simulation, and practical performance on a test system. The results show that high performance, stable, transparent bilateral HT is possible but that time delays pose a challenge. With less than approximately 200-ms communication latency, a three channel teleoperation system performed best. In

this scheme, the novice's set point velocity is a function of the operator's velocity and the novice's position and force tracking error, and the novice's force is fed back to the operator with significant low-pass filtering. With time delays, however, direct feedback becomes impractical and alternate approaches are required. One such approach of model-mediated teleoperation, in which the operator interacts with a local virtual model of the novice's environment, had excellent performance irrespective of time delay, although it is dependent on an accurate model. Unlike a follower robot, the human novice will always maintain stability. However, the novice's response is slower and less accurate than most robots', with stochastically varying performance based on focus and other factors. Given that the human novice's physical capabilities can be matched quite well to those of the expert, whether this limits achievable transparency remains the subject of future research.

This article represents an initial exploration of a novel topic and as a result has some limitations. First, the tests were performed with relatively few subjects and in idealized conditions. They will have to be augmented by practical tests with numerous subjects. For example, to achieve the performance of the ideal model-mediated teleoperation in this article, future work will have to capture an accurate geometric model of the environment. This is feasible through RGB-D cameras, LiDAR, time-of-flight (ToF) depth cameras, stereo reconstruction, or even monocular 3-D reconstruction. Almost all robotic or HT systems are equipped with at least one of the aforementioned sensing modalities. For example, the HoloLens 2 has stereo cameras and a ToF depth sensor, and can easily be augmented with a higher performance external RGB-D camera. Therefore, the assumptions in this article are not inaccurate. However, very fine details such as the patient's ribs during an ultrasound procedure may have a substantial effect on the measured force and will likely not be captured well in the mesh.

Similarly, the controllers presented here use pose and force, which have to be measured in real scenarios where the novice is not moving a haptic device. For HT, the force sensing must be low-profile and have sufficient load capability [6], [7], while pose tracking presents another set of challenges. Optical trackers are accurate but susceptible to occlusion, while electromagnetic trackers are affected by external disturbances. Extensive research has been performed in this direction [5], but the solutions are imperfect and future work will have to test practical bilateral teleoperation with all of these challenges.

Other future work includes investigation of further control architectures, for example, predictive control using the mathematical models to extrapolate trajectories, or robust control by defining uncertainties on the parameters in the derived state-space model. This article also did not consider architectures with nonzero  $k_{po}$ , i.e., four-channel approaches where the novice position is fed back to the operator. It has been shown that using all four channels of velocity and force is important for optimal transparency [19]. However, with the novice's unpredictable motions, this may not be true for HT. The hand-tuned controller parameters worked for these controlled tests, but more advanced and potentially adaptive methods to tune the controller

parameters based on individual novice dynamics and varying communication delays should also be explored.

Traditional passivity-based methods such as wave variables [25] for achieving time-delay robust stability were not used in this article because they tradeoff performance to maintain stability for the sake of safety. As instability in HT degrades performance rather than safety, this tradeoff is not helpful. The time-domain passivity approach (TDPA) [33] is less conservative than wave variables and may lead to better performance. Future work will use TDPA to insert passivity observers and controllers in the most promising teleoperation controllers determined in this article.

Finally, orientation tracking was not explicitly evaluated in this article but can be taken to be analogous to position tracking, i.e., consisting of similar lag and accuracy. This was the case in previous tests of HT [3]. Moreover, in [3], different rendering schemes were tested. These tests could be extended to find the best rendering method for improved depth perception. During the steplike motions in this article, the distribution of the novice position was approximately Gaussian about the desired location. However, in the depth direction from the novice's point of view (approximately  $z$  in Fig. 8), the position variance was significantly greater than in the other directions, pointing to possible depth perception limitations.

In model-mediated teleoperation, the environment impedance model affects the stability. We assumed in the tests that the environment model is fixed and known, with no delay, no error in the parameter estimation, and  $F_e^* = 0$ . Of course, in a real ultrasound exam, for example, the patient breathes and moves, and the impedance varies when examining different anatomies. When scanning the liver and kidneys, the probe moves from ribs to soft tissue and back. Practical impedance estimation schemes based on pose and force measurement require several measurement samples to converge to an estimate. Thus, the impedance model is delayed, varying, and not perfectly accurate. However, as shown in (34), the novice in the model-mediated teleoperation can converge to the correct normal force and tangent position despite modeling errors by having local force feedback as well as local position feedback in directions orthogonal to the force. In this case, only the operator feels a slightly incorrect force that is only transient while the impedance estimation converges. In practice, the ultrasound probe is moved slowly and contact is maintained, so the impedance is approximately constant. Thus, the test results are relatively realistic and show the effectiveness of model-mediated teleoperation in this system. Nonetheless, future work should validate the presented results with the real force sensing, pose tracking, environment reconstruction, and impedance estimation.

## V. CONCLUSION

This article is a first exploration of bilateral HT from a controls perspective. The results showed that three-channel teleoperation was effective at small time delays, while model-based teleoperation with local force and pose feedback achieved good performance even with large latency and imperfect impedance

estimation. It was found that instability in HT manifests differently than in robotics by substantially degrading system performance rather than causing hazard or harm. This is important to keep in mind when developing controllers. With the new system model, simulation, and practical setup, future work can build off the presented methods and tests to achieve high-performance, transparent, stable HT despite time delays. This will in turn enable effective remote guidance and execution of important tasks such as ultrasound exams with the relatively simple, low-cost mixed reality HT system.

## APPENDIX

In this section, we derive a state-space model of the full teleoperation system. This formulation is used in this article to assess stability and robustness.

From Section II-B, the novice is given by

$$\begin{aligned} \dot{\mathbf{x}}_f &= \begin{bmatrix} \dot{x}_n \\ \ddot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k_n}{m_n} & -\frac{b_n}{m_n} \end{bmatrix} \begin{bmatrix} x_n \\ \dot{x}_n \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{k_n}{m_n} & \frac{b_n}{m_n} \end{bmatrix} \begin{bmatrix} x_\nu \\ \dot{x}_\nu \end{bmatrix} \\ \mathbf{y}_f &= \begin{bmatrix} x_n \\ \dot{x}_n \\ f_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -k_e & -b_e \end{bmatrix} \begin{bmatrix} x_n \\ \dot{x}_n \end{bmatrix} \end{aligned} \quad (37)$$

where  $k_e$  and  $b_e$  are zero when not in contact with the environment. For this to be LTI, we assume the environment is unmoving and its impedance is constant. An advantage of this model is that we do not have to consider switching contact, which is handled implicitly by the novice.

The operator model is given by

$$m_o \ddot{x}_o + b_o \dot{x}_o = f_h + f_o. \quad (38)$$

The operator changes their applied force depending on the haptic device force, to track a desired trajectory. In particular, to choose a representative response, we can represent the system in state space and use servo control with pole placement, choosing

$$f_o = -K_o \mathbf{x}_o + k_a \int (x_o^* - x_o) dt \quad (39)$$

where  $\mathbf{x}_o = [x_o, \dot{x}_o]^\top$ , and  $x_o^*$  is the operator's desired motion. The state feedback gains are  $K_o = [k_{o1} \ k_{o2}]$ . These can be tuned to obtain a response that resembles the recorded operator data. We ignore the operator's desired velocity for simplicity as it introduces a zero eigenvalue and the operator is primarily interested in position. Setting  $x_a = \int (x_o^* - x_o) dt$ , we find the following model:

$$\begin{aligned} \dot{\mathbf{x}}_O &= \begin{bmatrix} \dot{x}_o \\ \dot{x}_a \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{k_{o1}}{m_o} & \frac{-b_o - k_{o2}}{m_o} & \frac{k_a}{m_o} \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_o \\ x_a \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} x_o^* + \begin{bmatrix} 0 \\ \frac{1}{m_o} \\ 0 \end{bmatrix} f_h \end{aligned}$$

$$\mathbf{y}_o = \begin{bmatrix} x_o \\ \dot{x}_o \\ f_o \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k_{o1} & -k_{o2} & k_a \end{bmatrix} \begin{bmatrix} \mathbf{x}_o \\ x_a \end{bmatrix} \quad (40)$$

where the gain on  $f_h$  is denoted  $E_o$ .

We now combine this with (9). Let  $\mathbf{x}_\nu = [x_\nu \ \dot{x}_\nu]^\top$  be the state of the virtual tool. The state equations of the operator, novice, and virtual tool are then

$$\begin{aligned} \dot{\mathbf{x}}_O &= \mathbf{A}_o \mathbf{x}_O + \mathbf{B}_o \mathbf{x}_o^* + \mathbf{E}_o (\mathbf{K}_{no} \mathbf{C}_f \mathbf{x}_n - \mathbf{K}_{oo} \mathbf{C}_o \mathbf{x}_O) \\ &= (\mathbf{A}_o - \mathbf{K}_{oo} \mathbf{C}_o) \mathbf{x}_O + \mathbf{E}_o \mathbf{K}_{no} \mathbf{C}_f \mathbf{x}_n + \mathbf{B}_o \mathbf{x}_o^* \end{aligned} \quad (41)$$

$$\dot{\mathbf{x}}_f = \mathbf{A}_f \mathbf{x}_n + \mathbf{B}_{f1} \mathbf{x}_\nu + \mathbf{B}_{f2} (\mathbf{K}_{on} \mathbf{C}_o \mathbf{x}_E - \mathbf{K}_{nn} \mathbf{C}_f \mathbf{x}_f) \quad (42)$$

$$\dot{\mathbf{x}}_\nu = \mathbf{K}_{on} \mathbf{C}_o \mathbf{x}_O - \mathbf{K}_{nn} \mathbf{C}_f \mathbf{x}_n \quad (43)$$

where  $\mathbf{B}_{f1}$  and  $\mathbf{B}_{f2}$  are the two columns of  $\mathbf{B}_f$ . We now define a combined state  $\mathbf{x} = [\mathbf{x}_O^\top \ \mathbf{x}_n^\top \ \mathbf{x}_\nu]^\top$ . This allows us to state the system as one equation

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} \mathbf{A}_o - \mathbf{E}_o \mathbf{K}_{oo} \mathbf{C}_o & \mathbf{E}_o \mathbf{K}_{no} \mathbf{C}_f & \mathbf{0}_{3 \times 1} \\ \mathbf{B}_{f2} \mathbf{K}_{on} \mathbf{C}_o & \mathbf{A}_f - \mathbf{B}_{f2} \mathbf{K}_{nn} \mathbf{C}_f & \mathbf{B}_{f1} \\ \mathbf{K}_{on} \mathbf{C}_o & -\mathbf{K}_{nn} \mathbf{C}_f & 0 \end{bmatrix} \mathbf{x} \\ &\quad + \begin{bmatrix} 1 \\ \mathbf{0}_{5 \times 1} \end{bmatrix} \mathbf{x}_o^* \\ \mathbf{y} &= \begin{bmatrix} \mathbf{C}_o & \mathbf{0}_{3 \times 2} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{C}_f & \mathbf{0}_{3 \times 1} \end{bmatrix} \mathbf{x}. \end{aligned} \quad (44)$$

The dimensions of the matrices are  $\mathbf{A} \in \mathbb{R}^{6 \times 6}$ ,  $\mathbf{B} \in \mathbb{R}^{6 \times 1}$ , and  $\mathbf{C} \in \mathbb{R}^{6 \times 6}$ . The  $\mathbf{A}$  matrix can be used to determine stability.

This analysis has only considered the case of time delay  $\tau = 0$ . For nonzero delay, we must use (11) instead of (9). The system model with time delays is thus

$$\begin{aligned} \dot{\mathbf{x}}_O &= \mathbf{A}_o \mathbf{x}_O + \mathbf{B}_o \mathbf{x}_o^* + \mathbf{E}_o e^{-s\tau} f_h \\ \dot{\mathbf{x}}_f &= \mathbf{A}_f \mathbf{x}_n + \mathbf{B}_{f1} \mathbf{x}_\nu + \mathbf{B}_{f2} \dot{\mathbf{x}}_v \\ \dot{\mathbf{x}}_\nu &= \mathbf{K}_{on} \mathbf{C}_o e^{-s\tau} \mathbf{x}_O - \mathbf{K}_{nn} \mathbf{C}_f \mathbf{x}_n. \end{aligned}$$

Combining this with (11) and (12), and with a slight abuse of notation to include time delays in the matrices, we obtain

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} \mathbf{A}_o - \mathbf{E}_o \mathbf{K}_{oo} \mathbf{C}_o e^{-s\tau} & \mathbf{E}_o \mathbf{K}_{no} \mathbf{C}_f e^{-s2\tau} & \mathbf{0}_{3 \times 1} \\ \mathbf{B}_{f2} \mathbf{K}_{on} \mathbf{C}_o e^{-s\tau} & \mathbf{A}_f - \mathbf{B}_{f2} \mathbf{K}_{nn} \mathbf{C}_f & \mathbf{B}_{f1} \\ \mathbf{K}_{on} \mathbf{C}_o e^{-s\tau} & -\mathbf{K}_{nn} \mathbf{C}_f & 0 \end{bmatrix} \mathbf{x} \\ &\quad + \begin{bmatrix} 1 \\ \mathbf{0}_{5 \times 1} \end{bmatrix} \mathbf{x}_o^* \\ \mathbf{y} &= \begin{bmatrix} \mathbf{C}_o & \mathbf{0}_{3 \times 2} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{C}_f & \mathbf{0}_{3 \times 1} \end{bmatrix} \mathbf{x}. \end{aligned} \quad (45)$$

This can also be written in the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0 \mathbf{x}(t) + \sum_{n=1}^2 \mathbf{A}_n \mathbf{x}(t - n\tau) + \mathbf{B} \mathbf{x}_o^*(t)$$

$$\begin{aligned} \mathbf{A}_0 &= \begin{bmatrix} \mathbf{A}_o & \mathbf{0}_{3 \times 2} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{2 \times 3} & \mathbf{A}_f - \mathbf{B}_{f2} \mathbf{K}_{nn} \mathbf{C}_f & \mathbf{B}_{f1} \\ \mathbf{0}_{1 \times 3} & -\mathbf{K}_{nn} \mathbf{C}_f & 0 \end{bmatrix} \\ \mathbf{A}_1 &= \begin{bmatrix} -\mathbf{E}_o \mathbf{K}_{oo} \mathbf{C}_o & \mathbf{0}_{3 \times 2} & \mathbf{0}_{3 \times 1} \\ \mathbf{B}_{f2} \mathbf{K}_{on} \mathbf{C}_o & \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 1} \\ \mathbf{K}_{on} \mathbf{C}_o & \mathbf{0}_{1 \times 2} & 0 \end{bmatrix} \\ \mathbf{A}_2 &= \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{E}_o \mathbf{K}_{no} \mathbf{C}_f & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 2} & 0 \end{bmatrix}. \end{aligned} \quad (46)$$

According to [40] and [41], the system is asymptotically stable independent of time delays if

$$\mu(\mathbf{A}_0) + \sum_{n=1}^2 \|\mathbf{A}_n\|_2 < 0 \quad (47)$$

where  $\mu(\mathbf{A}) = \frac{1}{2} \lambda_{\max}(\mathbf{A}^\top + \mathbf{A})$  and  $\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^\top \mathbf{A})}$ , where  $\lambda_{\max}$  is the maximum eigenvalue. This can be used to test specific known gains and parameters, but is too conservative for most applications as it guarantees stability for all time delays. In HT, it is more important to achieve transparency at reasonable delays since stability does not pose a safety hazard.

## REFERENCES

- [1] D. Black, Y. O. Yazdi, A. H. H. Hosseiniabadi, and S. Salcudean, "Human teleoperation—A haptically enabled mixed reality system for teleultrasound," *Human–Comput. Interact.*, vol. 39, pp. 1–24, 2023.
- [2] D. Black, M. Nogami, and S. Salcudean, "Mixed reality human teleoperation with device-agnostic remote ultrasound: Communication and user interaction," *Comput. Graph.*, vol. 118, pp. 184–193, 2024.
- [3] D. Black and S. Salcudean, "Human-as-a-robot performance in mixed reality teleultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 18, pp. 1–8, 2023.
- [4] D. G. Black, D. Andjelic, and S. E. Salcudean, "Evaluation of communication and human response latency for (human) teleoperation," *IEEE Trans. Med. Robot. Bionics*, vol. 6, no. 1, pp. 53–63, Feb. 2024.
- [5] D. Black and S. Salcudean, "Robust object pose tracking for augmented reality guidance and teleoperation," *IEEE Trans. Instrum. Meas.*, vol. 73, 2024, Art. no. 9509815.
- [6] D. Black, A. H. H. Hosseiniabadi, N. R. Pradnyawira, M. Pol, M. Nogami, and T. Salcudean, "Towards differential magnetic force sensing for ultrasound teleoperation," in *Proc. IEEE World Haptics Conf.*, 2023, pp. 333–339.
- [7] D. Black, N. R. Pradnyawira, M. Nogami, A. H. H. Hosseiniabadi, and S. Salcudean, "Low-profile 6-axis differential magnetic force/torque sensing," *IEEE Trans. Med. Robot. Bionics*, vol. 6, no. 3, pp. 992–1003, Aug. 2024.
- [8] N. Smith-Guerin, L. Al Bassit, G. Poisson, C. Delgorgue, P. Arbeille, and P. Vieyres, "Clinical validation of a mobile patient-expert tele-echography system using ISDN lines," in *Proc. 4th Int. IEEE EMBS Special Topic Conf. Inf. Technol. Appl. Biomed.*, 2003, pp. 23–26.
- [9] A. Saracino et al., "Haptic feedback in the da Vinci research kit (DVRK): A user study based on grasping, palpation, and incision tasks," *Int. J. Med. Robot. Comput.-Assist. Surg.*, vol. 15, no. 4, 2019, Art. no. e1999.
- [10] O. A. Van der Meijden and M. P. Schijven, "The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: A current review," *Surg. Endoscopy*, vol. 23, pp. 1180–1190, 2009.
- [11] C. C. Alleblas, M. P. Vleugels, S. F. Coppus, and T. E. Nieboer, "The effects of laparoscopic graspers with enhanced haptic feedback on applied forces: A randomized comparison with conventional graspers," *Surg. Endoscopy*, vol. 31, pp. 5411–5417, 2017.

- [12] G. Tholey, J. P. Desai, and A. E. Castellanos, "Force feedback plays a significant role in minimally invasive surgery: Results and analysis," *Ann. Surg.*, vol. 241, no. 1, pp. 102–109, 2005.
- [13] N. Diolaiti and C. Melchiorri, "Teleoperation of a mobile robot through haptic feedback," in *Proc. IEEE Int. Workshop Haptic Virtual Environ. Their Appl.*, 2002, pp. 67–72.
- [14] T. M. Lam, H. W. Boschloo, M. Mulder, and M. M. Van Paassen, "Artificial force field for haptic feedback in UAV teleoperation," *IEEE Trans. Syst., Man, Cybern., A, Syst. Humans*, vol. 39, no. 6, pp. 1316–1330, Nov. 2009.
- [15] A. Bolopion and S. Régnier, "A review of haptic feedback teleoperation systems for micromanipulation and microassembly," *IEEE Trans. Automat. Sci. Eng.*, vol. 10, no. 3, pp. 496–502, Jul. 2013.
- [16] D. A. Lawrence, "Stability and transparency in bilateral teleoperation," *IEEE Trans. Robot. Automat.*, vol. 9, no. 5, pp. 624–637, Oct. 1993.
- [17] G. Burdea and J. Zhuang, "Dexterous telerobotics with force feedback—An overview Part 2: Control and implementation," *Robotica*, vol. 9, no. 3, pp. 291–298, 1991.
- [18] J. Kim, P. H. Chang, and H.-S. Park, "Transparent teleoperation using two-channel control architectures," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 1953–1960.
- [19] K. Hashtrudi-Zaad and S. E. Salcudean, "Analysis of control architectures for teleoperation systems with impedance/admittance master and slave manipulators," *Int. J. Robot. Res.*, vol. 20, no. 6, pp. 419–445, 2001.
- [20] S. E. Salcudean, K. Hashtrudi-Zaad, S. Tafazoli, S. P. DiMaio, and C. Reboulet, "Bilateral matched impedance teleoperation with application to excavator control," *IEEE Control Syst. Mag.*, vol. 19, no. 6, pp. 29–37, Dec. 1999.
- [21] H.-K. Lee and M. J. Chung, "Adaptive controller of a master-slave system for transparent teleoperation," *J. Robot. Syst.*, vol. 15, no. 8, pp. 465–475, 1998.
- [22] K. Hashtrudi-Zaad and S. E. Salcudean, "Transparency in time-delayed systems and the effect of local force feedback for transparent teleoperation," *IEEE Trans. Robot. Automat.*, vol. 18, no. 1, pp. 108–114, Feb. 2002.
- [23] T. B. Sheridan, "Space teleoperation through time delay: Review and prognosis," *IEEE Trans. Robot. Automat.*, vol. 9, no. 5, pp. 592–606, Oct. 1993.
- [24] E. Nuño, L. Basañez, and R. Ortega, "Passivity-based control for bilateral teleoperation: A tutorial," *Automatica*, vol. 47, no. 3, pp. 485–495, 2011.
- [25] G. Niemeyer and J.-J. Slotine, "Stable adaptive teleoperation," *IEEE J. Ocean. Eng.*, vol. 16, no. 1, pp. 152–162, Jan. 1991.
- [26] A. Azimnejad, M. Tavakoli, R. V. Patel, and M. Moallem, "Transparent time-delayed bilateral teleoperation using wave variables," *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 3, pp. 548–555, May 2008.
- [27] A. C. Smith and K. Hashtrudi-Zaad, "Adaptive teleoperation using neural network-based predictive control," in *Proc. IEEE Conf. Control Appl.*, 2005, pp. 1269–1274.
- [28] A. C. Smith and K. Hashtrudi-Zaad, "Smith predictor type control architectures for time delayed teleoperation," *Int. J. Robot. Res.*, vol. 25, no. 8, pp. 797–818, 2006.
- [29] S. Soroushpour and A. Shahdi, "Model predictive control for transparent teleoperation under communication time delay," *IEEE Trans. Robot.*, vol. 22, no. 6, pp. 1131–1145, Dec. 2006.
- [30] X. Xu, B. Cizmeci, C. Schuwerk, and E. Steinbach, "Model-mediated teleoperation: Toward stable and transparent teleoperation systems," *IEEE Access*, vol. 4, pp. 425–449, 2016.
- [31] W.-H. Zhu and S. E. Salcudean, "Stability guaranteed teleoperation: An adaptive motion/force control approach," *IEEE Trans. Autom. Control*, vol. 45, no. 11, pp. 1951–1969, Nov. 2000.
- [32] A. Haddadi and K. Hashtrudi-Zaad, "Robust stability of teleoperation systems with time delay: A new approach," *IEEE Trans. Haptics*, vol. 6, no. 2, pp. 229–241, Apr./Jun. 2013.
- [33] B. Hannaford and J.-H. Ryu, "Time-domain passivity control of haptic interfaces," *IEEE Trans. Robot. Automat.*, vol. 18, no. 1, pp. 1–10, Feb. 2002.
- [34] J.-H. Ryu, C. Preusche, B. Hannaford, and G. Hirzinger, "Time domain passivity control with reference energy following," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 5, pp. 737–742, Sep. 2005.
- [35] D. Black and S. Salcudean, "Linearity, time invariance, and passivity of a novice person in human teleoperation," 2025, *arXiv:2504.11653*.
- [36] N. Hogan, "Stable execution of contact tasks using impedance control," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 4, 1987, pp. 1047–1054.
- [37] P. Mitra and G. Niemeyer, "Model-mediated telemanipulation," *Int. J. Robot. Res.*, vol. 27, no. 2, pp. 253–262, 2008.
- [38] R. J. Anderson and M. W. Spong, "Asymptotic stability for force reflecting teleoperators with time delay," *Int. J. Robot. Res.*, vol. 11, no. 2, pp. 135–149, 1992.
- [39] J. Lawrence, J. Bernal, and C. Witzgall, "A purely algebraic justification of the Kabsch-Umeayama algorithm," *J. Res. Natl. Inst. Standards Technol.*, vol. 124, pp. 1–6, 2019.
- [40] T. Mori and M. Kuwahara, "A way to stabilize linear systems with delay," *IFAC Proc. Vols.*, vol. 14, no. 2, pp. 165–170, 1981.
- [41] E. Cheres, Z. Palmor, and S. Gutman, "Quantitative measures of robustness for systems including delayed perturbations," *IEEE Trans. Autom. Control*, vol. 34, no. 11, pp. 1203–1204, Nov. 1989.



**David G. Black** (Member, IEEE) was born in Mainz, Germany. He received the B.A.Sc. degree in engineering physics and the Ph.D. degree in electrical and computer engineering from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2021 and 2025, respectively.

He has worked as a Robotics Engineer with A&K Robotics, Vancouver and ImFusion GmbH, Munich, Germany, and as a Research Student with the BC Cancer Research Centre. From 2018 to 2019, he was a Systems Engineer in advanced development with Carl Zeiss Meditec AG, Oberkochen, Germany, and has been a Consultant and Collaborator since 2019. He is currently a Postdoctoral Fellow with the UBC.



**Septimiu E. Salcudean** (Life Fellow, IEEE) was born in Cluj, Romania. He received the B.Eng. (Hons.) and M.Eng. degrees in electrical engineering from McGill University, Montreal, QC, Canada, in 1979 and 1981, respectively, and the Ph.D. degree in electrical engineering from the University of California, Berkeley, CA, USA, in 1986.

He was a Research Staff Member with the IBM T.J. Watson Research Center from 1986 to 1989. He then joined the University of British Columbia (UBC), Vancouver, BC, and is currently a Professor with the Department of Electrical and Computer Engineering, where he holds the C.A. Laszlo Chair in Biomedical Engineering and a Canada Research Chair. He has courtesy appointments with the UBC School of Biomedical Engineering and the Vancouver Prostate Centre, Vancouver.

Dr. Salcudean has been a Coorganizer for the Haptics Symposium, a Technical Editor and a Senior Editor for IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, and on the program committees of the International Conference on Robotics and Automation, International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), and International Conference on Information Processing in Computer-Assisted Intervention (IPCAI) Conferences. He is currently on the Steering Committee of the IPCAI conference and on the Editorial Board for the *International Journal of Robotics Research*. He is a Fellow of MICCAI and the Canadian Academy of Engineering.