Velocity Estimation Algorithms for Audio-Haptic Simulations Involving Stick-Slip

Stephen Sinclair, Marcelo M. Wanderley, *Member, IEEE*, and Vincent Hayward, *Fellow, IEEE*

Abstract—With real-time models of friction that take velocity as input, accuracy depends in great part on adequately estimating velocity from position measurements. This process can be sensitive to noise, especially at high sampling rates. In audio-haptic acoustic simulations, often characterized by friction-induced, relaxation-type stick-slip oscillations, this gives a gritty, dry haptic feel and a raspy, unnatural sound. Numerous techniques have been proposed, but each depend on tuning parameters so that they may offer a good trade-off between delay and noise rejection. In an effort to compare fairly, each of thirteen methods considered in the present study was automatically optimized and evaluated; finally a subset of these were compared subjectively. Results suggest that no one method is ideal for all gain levels, though the best general performance was found by using a sliding-mode differentiator as input to a Kalman integrator. An additional conclusion is that estimators do not approach the quality available in physical velocity transduction, and therefore such sensors should be considered in haptic device design.

Index Terms—Haptics, friction, velocity estimation

◆

1 INTRODUCTION

COMPLEX oscillations arising from non-linear mechanics are commonly found in the real world, leading to high-frequency behaviour. For instance, a dry finger tip sliding on an otherwise smooth surface generates wideband noise [1]. In the case of musical instruments, non-linear behaviour is fundamental to their operation, and must be taken into account during simulation; such simulations may be used to train skills, create music and sounds, study their physics, or study human sensorimotor behaviour [2]. Wideband, audible properties of such phenomena imply high simulation rates, viz. 20 to 40 kHz, exacerbating any noise issues due to sampling and differentiation.

In simulation, audio and haptic feedback signals may be generated synchronously for interactive applications; in such a case, noise issues are not only annoying, but may affect the haptic properties of the simulated material (e.g., modifying perception of hardness [3]), or even result in low-quality audio synthesis which defeats the purpose of the simulation.

For simulating the bowed string, at least two parameters, velocity and pressure, must be accounted for to control the friction-induced vibrations between a bow and a string [4]. In implementing such a simulation, we encountered the challenge that friction-driven dynamics depend on velocity, but force feedback devices almost invariably employ displacement sensors, necessitating differentiation. This led to

an unacceptable level of noise, destroying realism of both the sound and feel of the simulation.

One solution, employed in past efforts, is to include an inertial element in the mechanical model, which acts as a filter to attenuate the noise portion of the signal [4]. However, filtering in general adds delay and decreases the capabilities of the simulation, for instance, by limiting the stiffnesses that can be represented. Noise removal thus comes with a trade-off, precluding certain types of simulations. Like in most engineering systems, it is greatly preferable to eliminate noise at its source rather than attempt to filter it out.

In this work, a variety of velocity estimation methods are compared with velocity sensing in terms of this trade-off. A stochastic multi-objective optimisation was used to tune parameters in order to eliminate manual tuning and human partiality. We emphasize that although the performance/quality trade-off is well-known in the haptics community, evaluation is often in the context of simple simulations such as virtual walls, where damped interaction occurs only briefly during transients, and noisiness at the threshold may be mostly ignored, or even considered to contribute serendipity to a wall's feeling of being "crisp" and hard, or vice versa contribute to softness [3]. The unilateral virtual spring is a useful test because it forms the basis of a variety of force feedback rendering methods. However, in the present study we consider that friction-driven dynamics represents a class of interaction with particular challenges for haptic rendering: large bandwidth response, velocity dependence, and a fundamental connection to sound synthesis; this latter point forces consideration for high frequency simulation. The bowed string is an example, but these requirements also apply to scratching textures, or stick-slip action on sticky surfaces.

- S. Sinclair and V. Hayward are with Institut des Systèmes Intelligents et de Robotique, UPMC University Paris 06, Paris 75005, France.
  E-mail: stephen.sinclair@isir.upmc.fr, vincent.hayward@upmc.fr.
- M.M. Wanderley is with the Input Devices and Music Interaction Laboratory (IDMIL) at McGill University, Montréal, QC H3A 0G4, Canada,and the Centre for Interdisciplinary Research in Music Media Technology (CIRMMT) , Montréal, QC H3A 1E3, Canada.
  E-mail: marcelo.wanderley@mcgil.ca.

2 LIMITATIONS TO VELOCITY ESTIMATION FROM A POSITION SIGNAL

Indirect velocity acquisition implies choices in numerical estimation methods and sensing apparatus. Individual

Fig. 1. Bowed string digital waveguide from [5], enhanced with a string-on-bow force output to provide haptic feedback. The *Device* block represents the impedance haptic device, which outputs estimated velocity and takes a force command as input.

sensor signals can be processed in a variety of ways, and multiple sensors and estimates can also be combined.

Estimation from a lower derivative necessarily adds some delay; position is delayed by one time step relative to the force command signal, and so velocity, since it must take into account previous position samples, is delayed by at minimum two time steps [6]. More generally, for discrete position-controlled systems it is necessary to consider the Courant-Friedrichs-Lewy condition, which says that for explicit finite difference schemes, velocity may only be known within a quantum defined by time and space resolution:

$$\Delta > v_C T, \ v_C < \frac{\Delta}{T}, \tag{1}$$

where Δ is the spatial resolution, T is the temporal resolution, and v_C is the critical velocity of one quantum Δ/T. Thus, sampling faster may improve velocity resolution, but only if position resolution is sufficient [7], leading to demanding hardware specifications.

The effect of signal delay principally impacts impedance range [8]. Reduced delay has been shown to improve the impedance range of a damped virtual wall in several cases, including two-sliding observer [9] and adaptive windowing [10] estimation methods. Sensing higher derivatives may allow for less delay, since more information can be inferred about the same instant in time.

3   VELOCITY-COUPLED BOWED STRING

In sound synthesis, a well-recognized method for bowed string physical modeling called the *digital waveguide* was proposed by Smith and is based on D'Alembert's traveling wave solution to the wave equation [5]. In this formulation, traveling waves in a 1-DOF mechanical system are represented by a delay line loop, with linear losses aggregated into a filter, and propagation time modeled using pure delay. By placing a non-linear element correctly, a variety of instruments, namely the bowed string and single-reed bore (e.g., clarinet), can be simulated very efficiently in real time.

In previous work, the possibility of bypassing the velocity estimation problem by modifying Smith's formulation to leverage *position-dependent* friction methods has been explored [11], [12]. Although such tricks were shown to work, limitations were apparent [12], particularly for extreme playing gestures—force/velocity combinations outside the normal range of playing. In the current work, a physically-correct coupling to the digital waveguide, based on velocity, was therefore used to increase realism.

As Smith's formulation is intended as an audio synthesis model, it samples the string velocity for listening purposes, but does not include a string-on-bow force output needed for haptic feedback. Our proposed coupling includes a *Device* block representing a one-sample delay, having a velocity output and a force input, shown in Fig. 1. Force on the device (the "bow friction force," F_b) is calculated from the string velocity v_s and the bow-string impedance R_b. This implies that the velocity of the device must be available, which is the subject of the current work.

This coupling is based on the observation [13],

$$F_b = -R_b(v_\Delta)v_\Delta, \tag{2}$$

where v_Δ is the velocity difference between bow and string. It can be shown [14] that this solution corresponds with the *implicit junction* coupling method for digital waveguides [15]. Note that R_b is itself a function of velocity, therefore the impedance is time-varying. It must be selected to correspond with ρ, the bow junction transmission coefficient, which is also a velocity-dependent relation.

4   EQUIPMENT

4.1   Haptic Device

We used the Ergon_X *Transducteur Gestuelle Rétroactif* (TGR), designed at ACROE/ICA, INPG, Grenoble [16]. It is a modular device composed of several linear sensor/actuator pairs which can be coupled in a variety of ways. For the current work, we used the device in a side-ways, 1-DOF configuration, as seen in Fig. 2.

Fig. 2. The TGR device in a horizontal 1-DOF configuration, coupled to the Trans-Tek Series 100 LVT tachometer. The PCB Piezotronics accelerometer, not visible here, was adhered to the backside of the handle using wax recommended for the purpose.

A dedicated signal processor board provides a variable-frequency internal interrupt clock, and features 16-bit analog input and output. The motors can support impulsive linear forces up to 200 N, or 60 N sustained.

For our application, the digital waveguide simulation can run at 35 kHz, although lower rates were used during testing in order to execute multiple estimators simultaneously.

4.2 Sensors
The TGR device features built-in LVDT displacement sensors for each motor. In addition, a tachometer and an accelerometer were attached to the end effector. The tachometer was the Series 100 linear velocity transducer (LVT), model 0112-0000, from Trans-Tek. It has a flat frequency response up to 500 Hz. The moving magnet has a mass of 15 g. We measured a noise amplitude of 0.115 mm/s after analog-to-digital conversion when at rest.

The accelerometer was the model 352C22 from PCB Piezotronics. It weighs 0.5 grams, has a measurement range of $\pm 4{,}900$ m/s^2 (± 500 gravities), and a frequency range up to 10 kHz. We measured a constant RMS noise level of 0.22 m/s^2 (0.021 gravities) from this sensor when at rest.

For the LVDT sensors, we measured an RMS amplitude of 15 μm in the sampled analog signal when at rest. The total displacement range is 2.2 cm.

All sensors were verified to give a Gaussian normal distribution when at rest. A photo of the device coupled to the velocity sensor can be seen in Fig. 2.

5 DIFFERENTIATORS

This section describes several methods for determining velocity based on a sampled position signal. Although the following collection cannot be said to be complete, several estimator and observer approaches are included that have been previously proposed in the haptics literature.

5.1 Backward Difference
Defining x_k as the signal $x(t)$ at time $t = k\tau$, where τ is the sample period, the differential $\frac{dx}{dt}$ as can be approximated,

$$\hat{v}_k = \frac{x_k - x_{k-z}}{z\tau}, \qquad (3)$$

where $\hat{v}_k$ is the velocity estimate for sample k, and $z \geq 1$ is some number of time steps. Precision can thus be had at the expense of time.

Since this defines velocity as the change in position over time, it is effectively a poor average over the sample period [17]. One solution is to use averaging or some other low-pass filter on the $\hat{v}_k$ estimate [8]. In this work we employ a second-order Butterworth filter as a comparison case.

5.2 Least Squares Fit
The averaging approach of backward difference is based on an assumption of a constant slope of a linear fit over the window. It follows that a higher-order fit may provide an improved approximation by taking the derivative at the most recent sample time.

A linear least squares estimator can be expressed as a set of finite impulse response (FIR) filter coefficients [17]. The differential of an N-order polynomial is,

$$\frac{d\hat{x}_k}{dt} = c_1 + 2c_2 t_k + \cdots + N c_N t_k^{N-1}, \qquad (4)$$

$$\hat{x} = Ac. \qquad (5)$$

Therefore, A is matrix of size $N \times M$, representing a linear combination of the last M samples, and the sum of squares of the error can be minimized by,

$$c = (A^T A)^{-1} A^T x = A^\dagger x. \qquad (6)$$

A vector $\dot{q} = [0 \ 1 \ 2M \ 3M^2 \ \cdots \ NM^{N-1}]$ can be used to take the derivative with respect to time,

$$\frac{d\hat{x}}{dt} = \dot{q}^T A^\dagger x = \dot{h}^T x = \dot{v}_k, \qquad (7)$$

where $\dot{h}^T$ represents a linear combination of previous samples, i.e. the desired FIR coefficients [17].

5.3 Adaptive Windowing
The first-order adaptive windowing filter (FOAW) iteratively compares a noise estimate against a given expected margin to dynamically select the smallest acceptable window size for a linear fit at each sample [10]. In pseudocode,

for **i** in **1..N**,
  $\mathbf{w_i} = [x_{k-i} \cdots x_k]$
  $\mathbf{L_i} = \texttt{line\_fit}(\mathbf{w_i})$
  if $\max(|\mathbf{w_i}\text{-}\mathbf{L_i}|) < \mathbf{e}$:
    continue
  else
  return $\texttt{slope}(\mathbf{L_{i-1}})$

for some a priori estimate of expected signal noise **e**. The *best-fit* method, where the `line_fit` routine is a least squares linear fit, was used [10].

5.4 Levant's Differentiator
The two-sliding method as a differentiator was proposed by Levant [18] as a "robust" and "exact" differentiator, hence we refer to it as Levant's differentiator, following Chawda et al. [9]. r-sliding mode, or higher-order sliding modes (HOSM), are a model-free method to maintain a constraint up to its rth derivative [19]. Therefore two-sliding mode control uses a two-stage process to impose finite-time convergence for constraint $\sigma = \dot{\sigma} = 0$, where, for an observer w on signal x,

$$\sigma = w - x(t), \qquad (8)$$

$$\dot{\sigma} = u - \dot{x}(t), \qquad (9)$$

where u is an observer on the differential. If $\dot{w} = u$, then $u(t)$ can be taken as an estimate of $\dot{x}(t)$.

Levant [18] gives the control laws for the differentiator as,

$$\dot{w} = u, \qquad (10)$$

$$u = u_1 - \lambda \, |\, w - x(t)\,|^{1/2} \, \text{sgn}\,(w - x(t)), \qquad (11)$$

$$\dot{u}_1 = -\alpha \,\mathrm{sgn}\,(w - x(t)), \tag{12}$$

where $\alpha,\ \lambda > 0$, and u_1 is an additional observer state.

Here, α is some proportion of C, and λ some proportion of $\sqrt{C}$, where $C > 0$ is the Lipschitz constant. This guarantees local differentiability of $\dot{x}(t)$ if its derivative stays under a limit, i.e. if the following inequality holds:

$$|\dot{x}(t_k) - \dot{x}(t_{k-1})| \le C|t_k - t_{k-1}|. \tag{13}$$

Recommended choices [18] are,

$$\alpha = 1.1C, \tag{14}$$

$$\lambda = \sqrt{C}, \tag{15}$$

which we use in our implementation [9].

It is noted that a significant advantage for this technique is that performance increases with sampling rate, as opposed to backward-difference techniques which worsen [9]. However, switching noise can be detrimental and it is recommended to follow with a low-pass filter [18], potentially harming this advantage. This method was therefore tested with and without this post-filter, again employing a second-order Butterworth.

5.5   Sensor Fusion

Two methods to combine position and accelerometer signals were evaluated. The *complementary filter* [20] is a pair of filters of complementary frequency bands applied to different sensor signals and summed to arrive at a complete spectrum. Low- and high-frequency filters are respectively applied to the differentiated position signal, to remove high frequency noise, and the integrated accelerometer signal, to remove low-frequency bias.

Another approach is to make use of statistical maximization to estimate the most likely true signal value at a given time. The Kalman filter is such a technique, which at each step updates a prediction based on a system model, and then corrects the prediction using an optimal combination of weighted error judgements based on the expected covariance of each input with the measurement error and model uncertainty [21].

5.6   Kalman Filter

The haptic device is modeled as a mass driven by an unknown external force. Using notation from Bishop and Welch [21], a linear system model of a driven mass is,

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}, \tag{16}$$

$$P_k^- = AP_{k-1}A^T + Q, \tag{17}$$

where A and B are the linear model coefficients, $\hat{x}_k^-$ and $\hat{x}_k$ are the predicted and estimated states, u_k is the command signal, P_k^- and P_k are the predicted and estimated error covariance, and Q is the provided process covariance.

The state prediction $\hat{x}_k^-$ is updated according to a process model, and this prediction is corrected by measurements according to reliability expressed by covariance R. The measurement reliability along with the predicted error covariance

determines an optimal gain K on the residual [21]. We do not model the human input, and therefore set $B = 0$.

We have $\hat{x}_k$ and $\hat{x}_k^-$ as three-vectors of the form $[x\ \dot{x}\ \ddot{x}]^T$ and A is a 3×3 matrix. Since we have a discrete system, A must express an update for $\hat{x}_k^-$ over time τ, where time $t = k\tau$, and similarly Q must describe the discrete propagation of noise through the model. From Bar-Shalom et al. [20, p. 274], for a third-order system,

$$x_{k+1} = Fx_k + \Gamma v_k,$$

where v_k is the process noise, then,

$$F = \begin{bmatrix} 1 & \tau & \frac{1}{2}\tau^2 \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{bmatrix} \qquad \Gamma = \begin{bmatrix} \frac{1}{2}\tau^2 \\ \tau \\ 1 \end{bmatrix}. \tag{18}$$

For (16) and (17) we set $A = F$, and,

$$Q = \Gamma\sigma^2\Gamma^T = \begin{bmatrix} \frac{1}{4}\tau^4 & \frac{1}{2}\tau^3 & \frac{1}{2}\tau^2 \\ \frac{1}{2}\tau^3 & \tau^2 & \tau \\ \frac{1}{2}\tau^2 & \tau & 1 \end{bmatrix}\sigma^2, \tag{19}$$

where σ is the power spectral density of the process noise.

The Kalman filter has been applied to velocity estimation for applications in speed control: Kim et al. [22] applied this method to the two estimates available from timer- and event-driven encoder count measurement; Bélanger et al. [23] used a time-varying Kalman filter to update state estimates at intervals driven by both the encoder and clock events.

5.7   Hybrid Solutions

Although the Kalman estimator should perform optimally, our accelerometer measurement does not adhere to its bias-free assumptions. Correction of measurement bias at the expense of noise should therefore be of interest.

Two so-called *hybrid* methods were constructed, using noisy position-based acceleration estimates. Double-application of a differentiator was added to high-passed acceleration measurements before feeding the mixed data to the Kalman update. The FOAW and Levant estimators were both tried.

We therefore define four hybrid conditions as:

KALLEV:	$K(x, L(L(x)))$
KALLEVACC:	$K(x, L(L(x)) + H(\ddot{x}))$
KALFOAW:	$K(x, F(F(x)))$
KALFOAWACC:	$K(x, F(F(x)) + H(\ddot{x})),$

where x is the position measurement, $\ddot{x}$ is the acceleration measurement, K is a Kalman filter with position and acceleration as input, L is the Levant differentiator, F is the FOAW differentiator, and H is a 20 Hz high-pass filter to remove accelerometer DC error. No matching low-pass filter was applied to the position signal; rather, the Kalman update is intended to deal with the zero-centered position noise.

This method leverages on the one hand the improved noise and delay performance of the non-linear differentiators, and on the other hand, the optimal estimation properties of the Kalman filter. To verify that accelerometer data is used as intended, hybrid methods were evaluated with and without adding accelerometer measurements, as defined above.

6 OPTIMISATION AND NUMERICAL EVALUATION

Several example signals were recorded while interacting with the bowed string model at increasing friction feedback gain. The LVT tachometer was used to drive the velocity input, and therefore this signal was used as the comparison case for evaluating the estimators.

Using an appropriate error metric, estimators were numerically tuned and compared according to noise/delay criteria, described below.

6.1 Signal Recordings

The TGR was mounted in the 1-DOF horizontal orientation as described in Section 4.1. Fifteen recordings of bowing gestures were made at increasing levels of friction force gain. Position, velocity, and acceleration were recorded for 7 seconds per recording at 5 kHz.

Since the system was 1-DOF, friction gain was controlled with reference to $F_{\max}$, the maximum available friction force, rather than discussing a "friction coefficient," which would depend on normal force. In the case of the bowed string model, this refers to,

$$F_{\max} = \max F_b = \max \mu R_b(v_\Delta) \cdot v_\Delta. \quad (20)$$

Thus $F_{\max}$ implicitly determines a gain μ. The 15 recordings spanned values of $F_{\max}$ from 0 to 12.1 N in an approximately logarithmic distribution: 0, 0.3, 0.5, 0.8, 1.1, 1.4, 1.6, 1.9, 2.2, 3.3, 4.4, 5.5, 7.7, 9.9, 12.1 N.

6.2 An Error Metric for Signal Comparison

Many works (e.g. [9], [17], [24], [25]) make use of root mean square error (RMSE) to evaluate signal differences. This is calculated for a digital signal as,

$$\text{RMSE}(x_1, x_2) = \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} (x_1[k\tau] - x_2[k\tau])^2}, \quad (21)$$

where τ is the sample period, and operator $y[t]$ is time-quantized access into signal y at time step $\lfloor t/\tau \rfloor$.

In practice we found that using RMSE to compare the performance of velocity estimators often produced an unfair advantage to noisy-but-fast estimators, leading to unsatisfactory choices. Since delay leads to large error during transients, the bowed string signal, which consists of a series of transients, is particularly sensitive to this issue.

Considering noise and delay as two distinct sources of error that are confounded by RMSE, an approach is to distinguish them by estimating delay and removing it before calculating the error due to noise. Then, the two objectives can be weighted appropriately during optimisation.

Delay was determined by peak cross-correlation between the estimate and the measured velocity. This delay was then removed by time shifting, and the RMSE between the base signal and the delay-corrected signal was measured, providing the noise measurement.

This procedure can be summarized as,

$$D_r^e(\mathbf{x}) = \underset{k}{\arg\max}\ (v_r \star \tilde{v}_r)[k], \quad (22)$$

Fig. 3. Error evaluation and parameter optimization process.

$$E_r^e(\mathbf{x}) = \text{RMSE}(v_r, \text{SHIFT}[D_r^e(\mathbf{x})]) \quad (23)$$

$$= \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} (v_r[k\tau] - \tilde{v}_r[\tau(k + D_r^e(\mathbf{x}))])^2}, \quad (24)$$

$$v_r = \bar{y}_r', \quad (25)$$

$$\tilde{v}_r = e(\mathbf{x}, \bar{y}_r, \bar{y}_r''), \quad (26)$$

where D is the relative global delay in samples between two signals detected via a maximum cross-correlation analysis corrected by a time-shift operator SHIFT. The reference v_r is the measured LVT signal for recording r, and $\tilde{v}_r$ is the estimated velocity signal determined by some estimator $(e, \mathbf{x})$ applied to position and acceleration measurements r, where e is the estimator algorithm and $\mathbf{x}$ is the argument vector to the estimator. Sensor measurements for position (LVDT), velocity (LVT), and acceleration (accelerometer) for recording r are denoted $\bar{y}_r, \bar{y}_r', \bar{y}_r''$, respectively.

This method was applied to artificially delayed and filtered examples based on the recordings and verified that sufficient levels of accuracy in E and D were achieved [14]. Although a filter imposes frequency-dependent phase delay, we consider D to be an acceptable scalar measure of "global" signal delay, in the sense that the time-location of impulsive events are preserved after correcting for it. In the work that follows, E and D for each recording are seen as a set of objectives to be minimized.

6.3 Estimator Tuning and Comparison

The above error metric was used to determine an optimally tuned parameter set for each estimator based on our data recordings, minimizing both the delay and the delay-corrected error. A stochastic global optimisation approach was used in combination with an *objective sum* strategy for combining these two measures [26]. A block diagram describing the parameter optimization process can be found in Fig. 3.

The objective sum is a special case of the *weighted sum* method for global objective F_g,

$$F_{\mathrm{g}}(\mathbf{x}) = \sum_{i=0}^{j-1} \frac{w_i(F_i(\mathbf{x}) + o_i)}{s_i}, \qquad (27)$$

where all weights $w_i = 1$. Here, $j = 2n$ for n recordings since we have two objectives per recording, and o_i and s_i are normalization offsets and scaling factors per objective. Other choices of weights would allow expression of preference for some objectives over others, but we wish to balance our criteria evenly. Objectives F_i are E_r and D_r for recording r,

$$F_i(\mathbf{x}) = [E_0(\mathbf{x}) \; D_0(\mathbf{x}) \; \cdots \; E_{n-1}(\mathbf{x}) \; D_{n-1}(\mathbf{x})]. \qquad (28)$$

An adaptive normalization procedure, defined below, ensures that these objectives are comparable. An optimiser for each estimator e then finds the parameter set $\mathbf{x}_e$ that minimizes the global objective F_{g}^e for that estimator.

6.3.1 Normalization

In order for $w_i = 1$ to be effectively true, the objectives F_i must be of comparable magnitude. Not only are error and delay of different units, but between recordings we can expect the estimates of these values to vary; signals recorded with higher friction gain have stronger transients and therefore are more sensitive to delay. The scaling and offset were therefore adjusted adaptively according to the range defined by the best and worst results from a previous iteration.

6.3.2 Procedure

On iteration k, we compute E_r^e and D_r^e, the error and delay for estimator e for recording r, and compute normalized objectives F_i^e,

$$F_{2r}^e(k, K) = \frac{E_r^e\big(\mathbf{x}_e(k)\big) - \min_e E_r^e\big(\mathbf{x}_e(K)\big)}{\max_e E_r^e\big(\mathbf{x}_e(K)\big) - \min_e E_r^e\big(\mathbf{x}_e(K)\big)} \qquad (29)$$

$$= \frac{E_r^e\big(\mathbf{x}_e(k)\big) + O_{2r}(K)}{S_{2r}(K)}, \qquad (30)$$

$$F_{2r+1}^e(k, K) = \frac{D_r^e\big(\mathbf{x}_e(k)\big) - \min_e D_r^e\big(\mathbf{x}_e(K)\big)}{\max_e D_r^e\big(\mathbf{x}_e(K)\big) - \min_e D_r^e\big(\mathbf{x}_e(K)\big)} \qquad (31)$$

$$= \frac{D_r^e(\mathbf{x}_e(k)) + O_{2r+1}(K)}{S_{2r+1}(K)}, \qquad (32)$$

where (S_i, O_i) are scalings and offsets for error and delay for each recording across all estimators. This normalizes error and delay functions for n recordings, $0 \le r < n - 1$, into the range $[0, 1]$ at iteration k, based on the bounds of a different iteration K. The global objective $F_{\mathrm{g}}^e = \sum_{i=0}^{2n-1} F_i^e$ was minimized for each estimator, and used to produce a ranking.

Though the scale changes dynamically, it is of course necessary to keep it the same when comparing across iterations. Comparisons therefore have the form $F_{\mathrm{g}}^e(k, b_k) < F_{\mathrm{g}}^e(k - 1, b_k)$, where b_k is the iteration of the best-so-far F_{g}^e at iteration k. Stated algorithmically, the steps are as follows:

TABLE 1
Final Parameters for Each Estimator Based on Global Optimisation across All Recordings

Estimator	Parameters	
COMP1	LPF = 317 Hz	HPF = 317 Hz
COMP2	LPF = 256 Hz	HPF = 2,254 Hz
FOAW	Size = 17	Noise = 1.001×10^{-4} m/s
KALPOS	$Q = 0.191$	
KALMAN	$Q = 108$	$R_a = 6.28 \times 10^4$
KALLEV	$Q = 6.38 \times 10^4$	$R_a = 195$
KALLEVACC	$Q = 2,954$	$R_a = 64.6$
KALFOAW	$Q = 77.7$	$R_a = 8.06$
KALFOAWACC	$Q = 88.5$	$R_a = 3.85$
LEASTSQ	$N = 2$	$M = 33$
LEVANT	$C = 19.7$	
LEVANTLP	$C = 3.6$	LPF = 293 Hz
LOWPASS	LPF = 256 Hz	

Position covariance $R_p = 2 \times 10^{-10}$ was set a priori, based on measurement. Accelerometer input was pre-filtered with HPF = 20 Hz.

1) At iteration $k := 0$, for each estimator,
2) Select parameter set $\mathbf{x}_e(0)$.
3) Compute E_r^e and D_r^e for all estimators on all recordings.
4) Let $b_0 = 0$, and memorize all $S_i(b_0)$, $O_i(b_0)$.
5) For every subsequent iteration $k := k + 1$,
6) Select parameter set $\mathbf{x}_e(k)$.
7) Compute E_r^e and D_r^e for all estimators on all recordings.
8) If $F_{\mathrm{g}}^e(k, b_k) < F_{\mathrm{g}}^e(k - 1, b_k)$, let $b_{k+1} = k$ and memorize all $S_i(k)$, $O_i(k)$. Otherwise, let $b_{k+1} = b_k$.

Comparisons are therefore based on the same scaling, but the scale is adjusted as estimator performance improves. At each iteration the parameters of a selected estimator are perturbed, and retained only if the global objective shows an improvement. We halt the procedure if no improvements are found in 1,000 new parameter sets for any estimator.

6.3.3 Parameter Selection

Since objectives F_i^e in the space of $\mathbf{x}_e$ are in general non-convex and noisy for most e, we avoided gradient-based search strategies. Instead, *pure adaptive search* [27] was employed, a model-free stochastic sampling approach. Random sampling of the region around the best-so-far discovered point with increasing density allows to "zoom in" on the best parameters. Though slow, this method ensures a thorough, unbiased sampling of the objective function without involving tuning of optimiser hyperparameters.

7 NUMERICAL RESULTS

Results of the optimisation give a parameter set and a normalized scalar for each estimator, from which we can calculate a global ranking as well as evaluate individual performance for each recording. Labels in Table 2 will be used throughout the remainder of this section for discussion of optimisation results.

7.1 Parameter Selection

Final parameter selection based on the global criterion is given in Table 1. Observations on parameter choices follow:

TABLE 2
Complete List of Estimators Tested, with Description and a List of Parameters to Be Tuned

Estimator	Description	Parameters
COMP1	Sum of a second-order Butterworth low- and high-pass complementary filter pair, configured with identical cut-off frequencies.	LPF = HPF
COMP2	Sum of a second-order Butterworth low- and high-pass complementary filter pair, with independently-optimised cut-off frequencies.	LPF, HPF
FOAW	First-order adaptive windowing filter, *best-fit* method.	Max. window size, noise margin
KALPOS	Second-order Kalman filter with a double-integrator process model, pos. measurement, pos. covariance set to $R_p = 2 \times 10^{-10}$.	Process covariance Q.
KALMAN	Third-order Kalman filter with a double-integrator process model, pos. and accel. measurement, pos. covariance set constant at $R_p = 2 \times 10^{-10}$.	Process covariance Q, accel. covariance R_a.
KALLEV	Like KALMAN, with accel. measurement replaced with double-differentiated pos. using LEVANT.	Process covariance Q, accel. covariance R_a.
KALLEVACC	Like KALMAN, with accel. measurement added to double-differentiated pos. using LEVANT, $C = 100$.	Process covariance Q, accel. covariance R_a.
KALFOAW	Like KALMAN, with accel. measurement replaced with double-differentiated pos. using FOAW of size 12, noise margins 10^{-4} m and 0.5 m/s.	Process covariance Q, accel. covariance R_a.
KALFOAWACC	Like KALMAN, with accel. measurement added to double-differentiated pos. using FOAW of size 12, noise margins 10^{-4} m and 0.5 m/s.	Process covariance Q, accel. covariance R_a.
LEASTSQ	Least squares polynomial fit expressed as a set of FIR filter coefficients.	Poly. order N, window size M.
LEVANT	Levant's differentiator, a two-sliding observer driven to follow pos.	Max. accel. C
LEVANTLP	The LEVANT estimator followed by a second-order low-pass Buttworth filter.	Max. accel. C, LPF cut-off freq.
LOWPASS	Second-order Butterworth LPF.	LPF cut-off freq.

FOAW The LVDT position noise was measured to have a maximum error of 6.4×10^{-5} m. The optimal error margin was selected as 1.001×10^{-4}, about twice larger than expected, but within the correct order of magnitude.

LEVANT The Lipschitz bound on acceleration, C, poses a problem for optimisation across recordings, since the maximum expected acceleration changes with friction gain, which reaches about 200 m/s^2 for the larger gain settings. A worst-case approach would be used in manual tuning, and a value of $C > 100$ or more should be expected, but the optimiser instead allows some temporary drift as $C \approx 20$ is exceeded, which we regard as erroneous behaviour. As demonstrated in Fig. 4, in LEVANTLP the filtering effectively "covers up" these divergences, as well as problems with large switching noise, allowing smaller C to be selected.

KALMAN While it is possible to set the measurement covariance parameters R based on measured noise amplitude, the process covariance Q is more difficult to tune and it is common to determine it using an automatic optimisation procedure [21]. Since modifying both R_p and R_a can lead to multiple equivalent solutions, position covariance was held constant at $R_p = \sigma_p^2 = 2 \times 10^{-10}$, where $\sigma_p = 6.4 \times 10^{-5}$ m/s was the measured noise amplitude.

R_a selected by the optimiser was not of the same order of magnitude as measured $\sigma_a^2 = 0.048$ ($\sigma_a = 0.22$ m/s^2), perhaps reflecting problems due to low-frequency error, since Kalman assumes a zero-centered noise source. This seems confirmed by noticing that R_a is lower in the hybrid cases (KALLEV, KALFOAW), and there is additionally a small decrease in R_a when accelerometer data is included (KALLEVACC, KALFOAWACC). In general,

selected Q shows the opposite trend, compensating for distrusted R_a and vice-versa.

7.2 Performance Grouping

Error and delay estimates varied across recordings, but good consistency was found within recordings across estimators. Results were therefore normalized by expressing them in terms of percentage of the LOWPASS results.
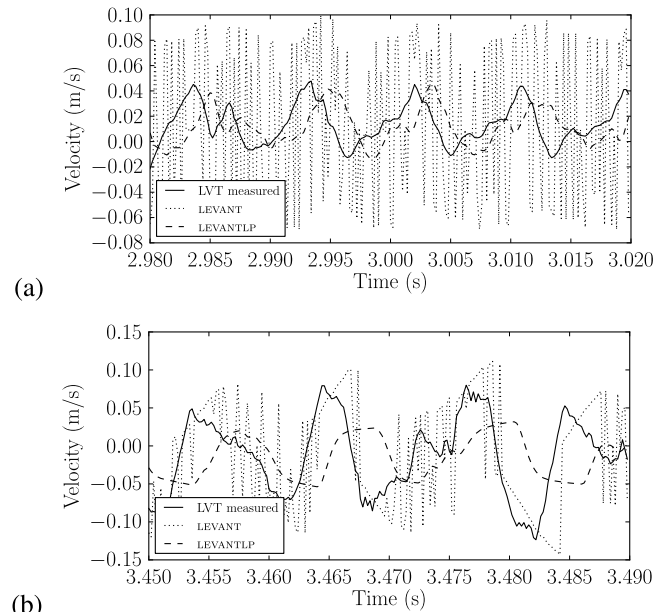

(a)

(b)

Fig. 4. Examples of Levant's differentiatior behaviour. For both graphs, $F_{\max} = 7.67$ N, LEVANT $C = 19.7$, LEVANTLP $C = 3.6$, LPF = 293 Hz. (a) Switching noise exceeds slip amplitude. (b) Lipschitz bound is exceeded for short intervals, e.g. from $t = 3.480$ to 3.487 s. In both cases, the low-pass post-filtering effectively removes switching noise and covers up temporary divergence, allowing for a smaller choice of C. The trade-off is an increase in delay and a reduction of the sharpness of the peaks.
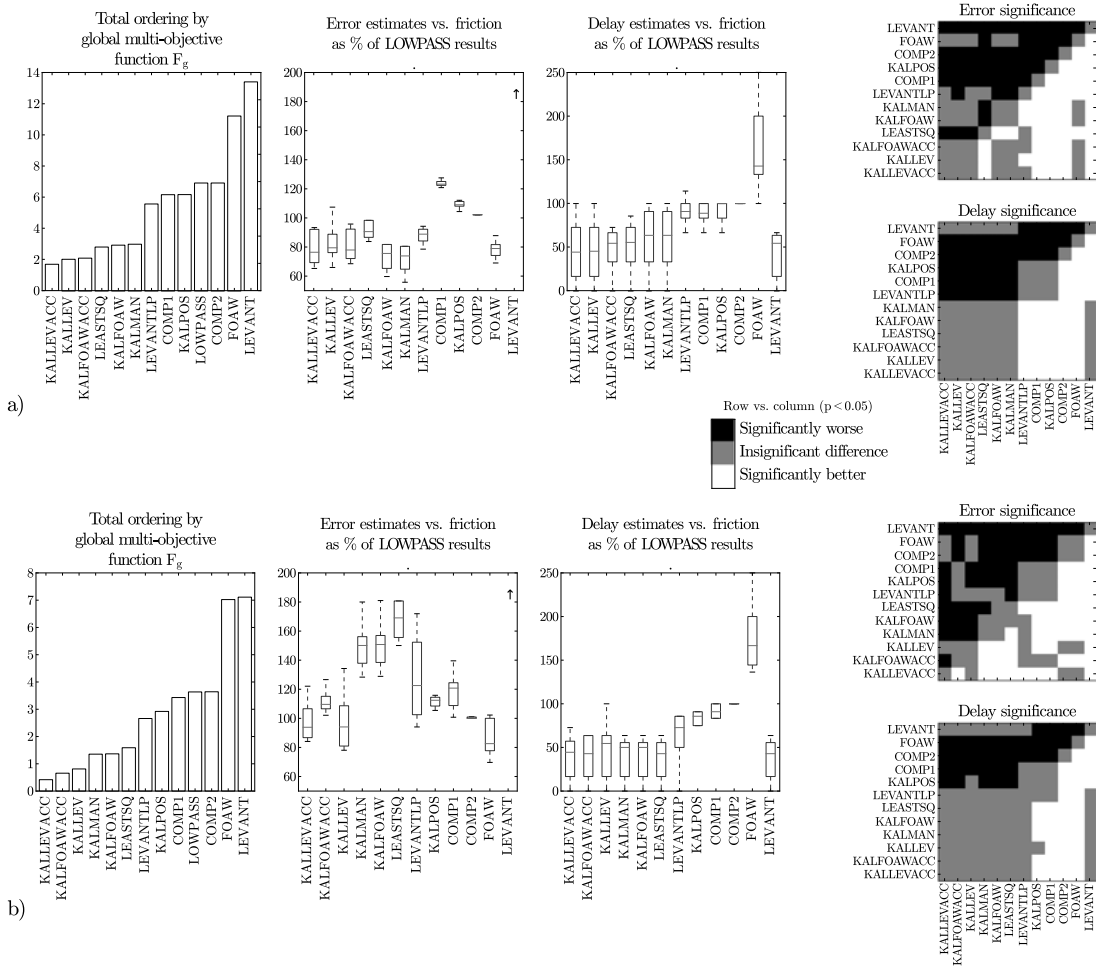
Fig. 5. Global objective performance, left, with error and delay evaluation, middle and right, for each recording using optimised parameter sets for each estimator, non-parametric box plot form. Results are given in left-to-right order of best (lowest) to worst (highest). Significance was determined using the Wilcoxon rank-sums test with threshold of $p < 0.05$. Note that LEVANT error results are present above the graph limits, with a median of 560 percent. a) 13 recordings with $F_{max} = 0$ to 7.7 N were included. b) High-gain case: seven recordings with $F_{max} = 2.2$ to 12.1 N were included.

An exception was the two highest-gain recordings, $F_{max} = 9.9$ and 12.1 N, where estimator performance was inconsistent with other gain settings, e.g. rankings differed. For this reason they were excluded from the analysis as outliers, but optimisation was performed a second time including only the higher-gain half of the data set from 2.2 up to 12.1 N.

Results for these two groups are seen in Figs. 5a and 5b respectively. A Wilcoxon rank-sums test [28] was used to determine significance between pairs, with significance determined by probability threshold of two distributions being the same of $p < 0.05$. It can be seen that the multi-objective optimisation has correctly balanced the two criteria: the top performers on the global criterion also performed best for both error and delay.

7.3 Validation

A separate data set was recorded with the same friction values and very similar gestures as our test set, and kept aside for validation. Using the parameter sets from Table 1, we evaluated error and delay for the second data set.

Taking the difference between results at matching friction levels, the average percentage of absolute error difference was 10.3 percent, with a standard deviation of 7.76 percent.

The average delay difference was 2.60 samples with a standard deviation of 2.02 samples. This represents a large percentage difference (75.6 percent), but it is comparable to delay differences across recordings in the original dataset, reflected in the large variance seen in the delay results of Fig. 5. In other words, the delay estimator can vary a fair amount between recordings, even for the same friction gain. Therefore we ascribe delay differences primarily to imperfect delay estimation rather than to differences between data sets.

An additional observation is that differences in mean delay between data sets were not significant across estimators, but there was a correlation between the mean delay absolute difference and standard deviation, (Pearson's $r = 0.72$, $p < 0.05$.) Small mean delays received smaller deviation in delay, meaning that time-accurate estimators remained accurate for the validation data set. We conclude therefore that the performance of estimators on data outside the training set is acceptable, and over-fitting did not occur.

7.4 Discussion

Fig. 5a shows that for all but the highest gain settings, most approaches performed as well as or better than LOWPASS on both error and delay criteria, in some cases with a median improvement of about 30 percent on error, and 50 percent on delay.

The third-order Kalman estimators were the best performers on both criteria, although differences between them were generally not significant. The performance of the hybrid approaches—the Kalman filter combined with nonlinear estimators—indicate some possible trends toward improvement in median delay, but variance is too high to draw conclusions.

The second-order least squares fit (LEASTSQ) turns out to be a very effective estimator in terms of delay. Although its error performance is not as good as other methods, its delay performance achieved a very competitive global ranking.

In comparison, the FOAW estimator had very good error performance, but suffered badly from delay. This was surprising because the FOAW algorithm is designed specifically to improve on delay and accuracy during transients—if we consider that our data consists of continual small transients due to stick-slip, we expected the window to be small on average. For FOAW to be useful here, short windows should trigger at roughly the period of the waveform, but this was not the case. Instead, the FOAW window was at maximum for over 98 percent of time steps: the adaptive filter was thus not reacting to stick-slip transients. We conclude that our data is pathological for this algorithm, indicating a distinction between the current case and more typical virtual wall scenarios.

Overall, no significant differences could be found for the best six estimators on both error and delay criteria simultaneously. Therefore we can conclude that the KALMAN estimator provides the best global results, since it is not significantly different than LEASTSQ in delay, and beats it by a large margin in error performance. No significant improvements could be found by including measured acceleration, nor by means of using the FOAW or Levant methods to improve measured acceleration via the hybrid estimators.

The high-gain results, Fig. 5b, show one statistically significant benefit of using the hybrid Kalman methods—the KALFOAW estimator does no better than KALMAN, but KALFOAWACC is significantly better on error. However, the KALLEV and KALLEVACC are the best performers and are not different, therefore advantages of including accelerometer information remain inconclusive. Notably, the Levant-Kalman estimators are the only ones to beat LOWPASS on both error and delay criteria in the high-gain case.

8 SUBJECTIVE EVALUATION

In the previous section, some estimation and measurement methods were shown to reduce signal noise while introducing less delay than a low-pass filter. According to teleoperator theory, lower delay should lead to an improvement in the impedance range of the display, while to the benefit of audio-haptic interaction, it should minimize amplitude of perceived noise in the velocity signal that stimulates the acoustic model.

To verify these predictions, the subjective noise performance was rated by human operators; secondly, participants were asked to determine the friction impedance range.[1]

1. This study was performed with the approval of McGill University's Research Ethics Board, REB #105-0908.

Fig. 6. A participant interacting with the device in experimental conditions. The left hand is manipulating a knob while the right hand interacts with the haptic device. The device is seen in its horizontal configuration, coupled to the velocity transducer.

The perceptual rating task was necessary because we found, informally, that measurements of stationary noise did not well-reflect the relative perceived noisiness of the estimators. This may be due to differing spectral signatures of estimator noise [14]. We opted therefore to judge perceived quality directly using human participants.

8.1 Methodology

The study consisted of two parts, a rating task and an adjustment task. These were performed consecutively, with the rating task performed first, and the adjustment task second.

There were nine participants recruited from the university lab between the ages of 25 and 35, where seven were male and two were female. All self-reported to have normal hearing and normal cutaneous feeling. Of these, four participants played an instrument, though only one played a stringed instrument, the violin. None played an instrument professionally.

In the rating task, participants rated the "noisiness" of eight estimators on a normalized scale, freely switching between them to compare. Participants were not instructed how to interpret the word "noisiness," however they were allowed to browse the conditions and discover the noisier ones themselves, making the concept evident. They were encouraged to consider both the sound and feel in their judgement. They were asked to determine the "best" and then the "worst" estimator, establishing the highest and lowest knob positions respectively, and then to rate the remaining estimators within this subjective range. All participants finished this task in less than 5 minutes.

For the adjustment task, participants were given control over a single variable, the friction gain, during each trial. The objective was to determine the maximum gain at which the model behaviour "breaks down" and no longer performs as a string simulation, which we identify with the point of marginal stability. This is explained further below.

8.2 Apparatus

The same apparatus as used for our data recordings, the Ergon_X TGR device in a horizontal configuration, was employed. A photo of a participant performing the experiment is presented in Fig. 6. For both tasks, participants
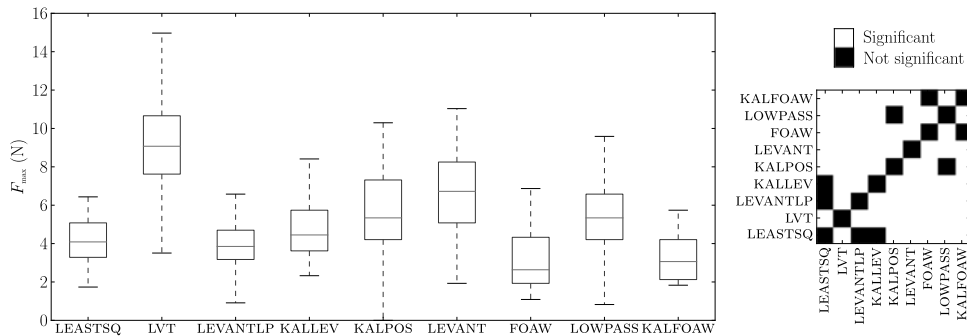
Fig. 7. Maximum friction force F_{max} for bowed string interaction. Significance is determined by Wilcoxon's rank-sums test, with $p < 0.05$ indicating confidence that distributions are not the same. We see that LEVANT performs best after the tachometer, agreeing with the hypothesis that delay is a significant factor for the impedance range, however the next best estimators are KALPOS and LOWPASS, which did not perform exceptionally well for delay. Therefore we must assume there are additional factors influencing performance at high gain.

listened to the string model using headphones.[2] Participants were seated and grasped the handle with their right hand, while their left hand controlled a set of eight MIDI knobs.[3]

In the first part of the experiment, moving a knob resulted in the associated condition being selected, while also adjusting the rating for that condition.

For the second part, participants were given control over a single MIDI knob controlling F_{max} in order to perform the adjustment task. Each trial started with $F_{max} = 0$, and was increased logarithmically by turning to the right until discovering where unexpected behaviour began, fine-tuning it by compensating hysteresis if necessary. Six samples per condition per participant were taken.

8.3 Stimulus

For the rating task, participants interacted with the model running at 5 kHz by moving the handle along its single free axis, and listened to the sound of the string as well as any background noise produced by the selected estimator. Friction forces were exerted with a constant $F_{max} = 1$ N.

For the adjustment task, we characterise the impedance range for the estimator as the maximum displayable friction gain, F_{max} such that the display remains stable. Participants were told to find the margin of stability, apparent due to device oscillation and clear distortion in the sound.

Participants were encouraged to use the breakdown of the string behaviour as a cue to differentiate stable and unstable conditions. Specifically, they were told to "find the point immediately before the model no longer feels or sounds like a bowed string." A few training trials were used to demonstrate.

8.4 Conditions

The estimators used in this experiment were the subset from Table 2 that depend only on position: these were LOWPASS, LEASTSQ, FOAW, LEVANT, LEVANTLP, KALPOS, KALLEV, KALFOAW, as well as LVT, a direct usage of the tachometer signal. Accelerometer-based methods were not included.

Estimators were executed concurrently on the Toro hardware to allow quick switching between conditions. Under this computational load, the system executed at a maximum

of 5 kHz, with the exception of KALFOAW, which was only able to run at 4 kHz.[4] For this reason, KALFOAW was left out of the rating task, but included in the impedance task.

9   RESULTS

The friction gain adjustment results are given in Fig. 7, and the rating task is analysed in Fig. 8. A scatter plot comparison is available in Fig. 9.

9.1   Maximum Impedance Judgement

We notice firstly, in Fig. 7, that the tachometer, condition LVT, features nearly twice the impedance range compared to any estimator. Secondly, although it was very noisy, LEVANT, the non-filtered Levant differentiator, performed second best, agreeing with the hypothesis that delay, independent of noise, is an important factor for the range of stability. Confirming this, the FOAW estimator performed worst, and was also a weak performer on the delay criterion in Fig. 5.

However, the next-best impedance ranges after LEVANT belong to KALPOS and LOWPASS, which did not perform exceptionally well on the delay criterion in Fig. 5. This is unexpected, since it does not correspond with numerical delay estimates, and it therefore suggests that there are other factors affecting behaviour at high gain.

We note that variance in Fig. 7 is mostly due to inter-subject differences, while intra-subject variance was on average 61 percent lower. Participants may have used different criteria to judge marginal stability, but distribution differences were nonetheless statistically significant, therefore normalization was not necessary. Furthermore, individual rankings were strongly correlated to the global ranking, (Spearman's $\rho > 0.9$ for all participants.) We additionally note that variance may represent a real, ambiguous range in which a higher likelihood of instability is gradually approached.

9.2   Subjective Noisiness Ratings

The subjective ratings, Fig. 8, are more easily interpreted. The best and worst performers for noisiness, corresponding to LVT and LEVANT, were selected unanimously. The

---

4. Prior to seeing the computational load implied by concurrent execution, this study was to run at a rate of 20 kHz, to allow a more natural-sounding bowed string; at 5 kHz, the sound is degraded, but recognizable as a string.
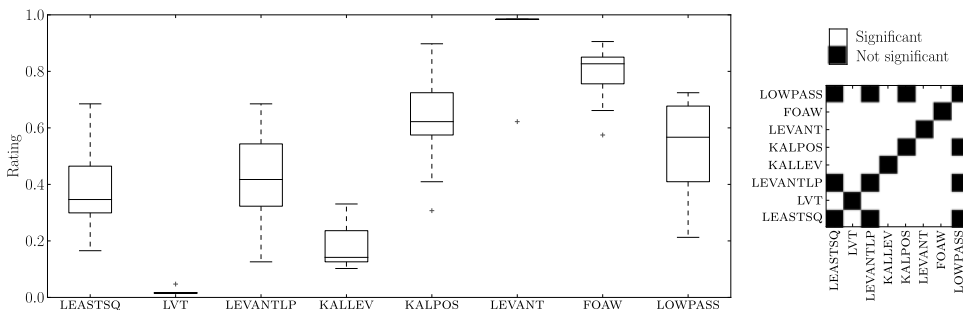
Fig. 8. Subjective ratings of noisiness for several estimators. Significance is determined by Wilcoxon's rank-sums test, with $p < 0.05$ indicating confidence that distributions are not the same. The small deviations of LVT and LEVANT are due to instructions that the noisiest and least noisy condition be maximized and minimized within the range—the choices of best and worst were unanimous. The KALFOAW estimator was not included in these ratings.

preference for KALLEV is indicated clearly, which confirms its performance on the noise criterion as predicted by Fig. 5.

10 DISCUSSION

These results show exceptionally good performance for direct measurement using the tachometer, and also show that some estimators are indeed preferable. KALLEV seems to allow better noise rejection than LOWPASS while featuring similar impedance range. Finally, it is clear that a 250 Hz low-pass filter actually ranks fairly well in practice.

We note that rankings derived from noisiness rating medians did not correlate strongly with numerical delay-corrected error medians, nor did similar median rankings from max. impedance ratings correlate strongly with rankings based on numerically-judged delay, (Spearman's $\rho = 0.61$ and $\rho = 0.43$, respectively.) However, this may be due simply to lack of significance between the best-rated conditions in Fig. 5.

11 CONCLUSION

In this article we addressed the question of noise/delay trade-off in velocity estimation, as applied to audio-haptic interaction with an acoustic bowed string model.

From our selection of differentiators, the numerical results indicate that for friction-coupled audio-haptic interaction, velocity is best estimated using a third-order Kalman
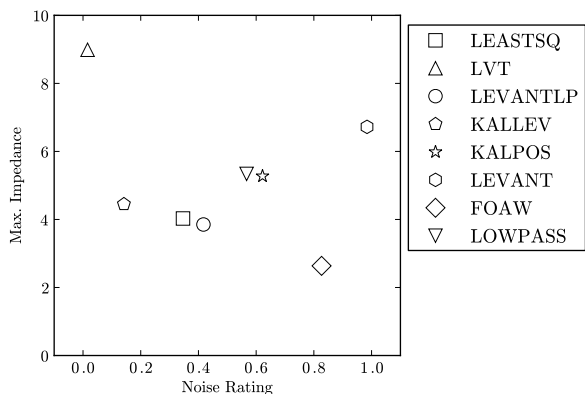

Fig. 9. Scatter plot of median subjective results from Figs. 7 and 8. Estimators closer to the top-left corner are better. Results are reflective of the estimators with the tuned parameter set. Each point is one sample (the "best" according to our criteria) of a contour in this space as estimator parameters are adjusted.

approach; a non-linear two-sliding observer-based differentiator driving a statistically optimal linear observer to remove switching noise gave the best performance on our objective sum. We may assume that a more detailed process model taking device characteristics into account might bring further improvements.

The inclusion of measured accelerometer data only contributed significantly in the high-gain case, and only provided mild improvements. This indicates either that accelerometer data could not provide significantly better information, or more likely that in the high-frequency region where accelerometer data can contribute, the small transient stick-slip peaks are too small relative to noise to be correctly measured by our criteria. These high-frequency transients are important to audio perception of the bowed string acoustics, therefore it is possible that better measures are needed to properly optimise for acoustic models. For instance, an objective function based on models of human perception could be relevant for future work.

It is interesting to note that methods such as FOAW that have previously shown good results for virtual wall models do not perform particularly well in this scenario, similarly due to the small transients. This implies that some estimators may be more appropriate to specific classes of interaction.

Our choice of the *objective sum* method is not the only possible optimisation strategy. Since estimators were allowed to optimise on either axis, there is no consistency between estimators on any one measure, leading to difficulties in subjective comparison. For example, one cannot listen to each optimised estimator and judge its noise qualities, since another parameter selection for the same estimator may reduce noise at the expense of delay. An alternative could be to use a unilateral constraint on one variable while optimising the other.

We performed a two-axis evaluation with participants to determine how numerical results corresponded with subjective quality. Unexpectedly, numerical delay estimates did not well-predict the ranking of estimators for impedance performance, nor did error estimates well-predict noisiness ratings. This further supports the need for better predictive models of the practical effects of noise and delay on human perception and machine performance in the context of wideband signals.

The study indicated that no estimator could approach the velocity transducer in terms of perceived noise and attainable impedance, therefore such sensing hardware should be considered for integration into force feedback devices.

ACKNOWLEDGMENTS

REFERENCES

[1] M. Wiertlewski, C. Hudin, and V. Hayward, "On the 1/f noise and non-integer harmonic decay of the interaction of a finger sliding on flat and sinusoidal surfaces," in *Proc. IEEE World Haptics Conf.*, 2011, pp. 25–30.

[2] J.-L. Florens, A. Razafindrakoto, A. Luciani, and C. Cadoz, "Optimized real-time simulation of objects for musical synthesis and animated image synthesis," in *Proc. Int. Comput. Music Conf.*, 1986, pp. 65–70.

[3] Y. Visell, B. L. Giordano, G. Millet, and J. R. Cooperstock, "Vibration influences haptic perception of surface compliance during walking," *PLoS one*, vol. 6, no. 3, p. e17697, 2011.

[4] J.-L. Florens, "Expressive bowing on a virtual string instrument," in *Proc. 5th Int. Gesture Workshop Gesture-Based Communication in Human-Computer Interaction*, Apr. 2003, vol. 2915, pp. 487–496.

[5] J. O. Smith, "Efficient simulation of the reed-bore and bow-string mechanisms," in *Proc. Int. Comput. Music Conf.*, 1986, pp. 275–280.

[6] V. Hayward and K. MacLean, "Do it yourself haptics: Part I," *Robot. Autom. Mag.*, vol. 14, no. 4, pp. 88–104, 2007.

[7] G. Campion and V. Hayward, "Fundamental limits in the rendering of virtual haptic textures," in *Proc 1st Joint Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2005, pp. 263–270.

[8] J. E. Colgate and J. M. Brown, "Factors affecting the z-width of a haptic display," in *Proc. IEEE Conf. Robot. Autom.*, 1994, pp. 3205–3210.

[9] V. Chawda, O. Celik, and M. K. O'Malley, "Application of Levant's differentiator for velocity estimation and increased Z-width in haptic interfaces," presented at the World Haptics, Istanbul, Turkey, Jun. 2011,

[10] F. Janabi-Sharifi, V. Hayward, and C.-S. J. Chen, "Discrete-time adaptive windowing for velocity estimation," *IEEE Trans. Cont. Sys. Technol.*, vol. 8, no. 6, pp. 1003–1009, Nov. 2000,

[11] S. Sinclair, G. Scavone, and M. M. Wanderley, "Audio-haptic interaction with the digital waveguide bowed string," in *Proc. Int. Comput. Music Conf.*, Montreal, QC, Canada, 2009, pp. 275–278.

[12] S. Sinclair, M. M. Wanderley, V. Hayward, and G. Scavone, "Noise-free haptic interaction with a bowed-string acoustic model," presented at the World Haptics Conf., Istanbul, Turkey, Jul. 2011,

[13] J. O. Smith. (2010). *Physical audio signal processing: For virtual musical instruments and audio effects*. W3K Publishing [Online]. Available: http://ccrma.stanford.edu//jos/pasp/

[14] S. Sinclair, "Velocity-driven audio-haptic interaction with real-time digital acoustic models," Ph.D. dissertation, Schulich School of Music, McGill Univ., Montreal, QC, Canada, 2012.

[15] E. Berdahl, G. Niemeyer, and J. Smith, III, "Using haptic devices to interface directly with digital waveguide-based musical instruments," in *Proc. 9th Int. Conf. New Interfaces Musical Expression*, 2009, pp. 183–186.

[16] J. Florens, A. Luciani, C. Cadoz, and N. Castagné, "ERGOS: Multi-degrees of freedom and versatile force-feedback panoply," in *Proc. Eurohaptics*, 2004, pp. 356–360.

[17] R. H. Brown, S. C. Schneider, and M. G. Mulligan, "Analysis of algorithms for velocity estimation from discrete position versus time data," *IEEE Trans. Ind. Electron.*, vol. 39, no. 1, pp. 11–19, Feb. 1992.

[18] A. Levant, "Robust exact differentiation via sliding mode technique," *Automatica*, vol. 34, no. 3, pp. 379–384, 1998.

[19] A. Levant, "Principles of 2-sliding mode design," *Automatica*, vol. 43, no. 4, pp. 576–586, 2007.

[20] Y. Bar-Shalom, T. Kirubarajan, and X.-R. Li, *Estimation with Applications to Tracking and Navigation*. New York, NY, USA: Wiley, 2002.

[21] G. Bishop and G. Welch, "An introduction to the Kalman filter," in *Proc. SIGGRAPH 2001 Course Notes*, 2001, pp. 1–82.

[22] H. Kim, J. Choi, and S. Sul, "Accurate position control for AC servo motor using novel speed estimator," in *Proc. Int. Conf. Ind. Electron., Control Instrum.*, 1995, vol. 1, pp. 627–632.

[23] P. Bélanger, P. Dobrovolny, A. Helmy, and X. Zhang, "Estimation of angular velocity and acceleration from shaft-encoder measurements," *Int. J. Robot. Res.*, vol. 17, no. 11, pp. 1225–1233, 1998.

[24] E. Kilic, O. Baser, M. Dolen, and E. I. Konukseven, "An enhanced adaptive windowing technique for velocity and acceleration estimation using incremental position encoders," in *Proc. Int. Conf. Sig. Electron. Sys.*, Gliwice, Poland, Sep. 2010, pp. 61–64.

[25] S. P. Chan, "Velocity estimation for robot manipulators using neural network," *J. Intell. Robot. Syst.*, vol. 23, no. 2, pp. 147–163, 1998.

[26] R. Marler and J. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidisciplinary Optim.*, vol. 26, no. 6, pp. 369–395, 2004.

[27] N. Patel, R. Smith, and Z. Zabinsky, "Pure adaptive search in monte carlo optimization," *Math. Program.*, vol. 43, no. 1, pp. 317–328, 1989.

[28] R. L. Ott and M. Longnecker, *An Introduction to Statistical Methods and Data Analysis*, 6th ed. Belmont, CA, USA: Brooks/Cole Cengage Learning, 2010.

Stephen Sinclair received the PhD degree in music technology from McGill University, Montreal, QC, Canada, in 2012. He is currently a postdoctoral fellow at the Université Pierre et Marie Curie (Paris VI), Paris, France, in the Institut des Systèmes Intelligents et de Robotique. His current research interests include haptic rendering systems, sensory integration, human-machine interaction, signal processing and control, and robotics.

Marcelo M. Wanderley (M'XX) received the PhD degree from the Université Pierre et Marie Curie (Paris VI), Paris, France, in acoustics, signal processing, and computer science applied to music. He is currently a William Dawson scholar and an associate professor of music technology at the Schulich School of Music, McGill University, Montreal, QC, Canada, where he directs the Input Devices and Music Interaction Laboratory and the Centre for Interdisciplinary Research in Music Media and Technology. He was the chair of the 2003 International Conference on New Interfaces for Musical Expression and coauthored the textbook *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard (A-R Editions)*. His current research interests include gestural control of sound, and input device design and evaluation. He is a member of the IEEE.

Vincent Hayward (F'08) received the Dr-Ing. degree from the University of Paris XI, Paris, France, in 1981. He was a postdoctoral fellow and then as a visiting assistant professor at Purdue University, in 1982, and joined CNRS, Paris, France, as Chargé de Recherches in 1983. In 1987, he joined the Department of Electrical and Computer Engineering at McGill University, Montreal, QC, Canada, as an assistant, associate and then full professor in 2006. He was the director of the McGill Center for Intelligent Machines from 2001 to 2004 and held the "Chaire internationale d'haptique" at the Université Pierre et Marie Curie (UPMC), Paris, France, from 2008 to 2010. He is currently a professor at UPMC. His current research interests include haptic device design, haptic perception, and robotics. He is a fellow of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.