Short Paper

Addendum to "Dimensional Reduction for 6D Vibrotactile Display"

Ruisi Zhang, *Student Member, IEEE*, Trevor J. Schwehr, and Jake J. Abbott, *Senior Member, IEEE*

***Abstract*—In our recent article, we found a general quadratic weighting function to predict how a six-dimensional (6D) vibrotactile stimulus rendered at the haptic interaction point (HIP) of a kinesthetic haptic interface, with the stylus held in a precision pen-hold grasp, is mapped to an equivalent one-dimensional (1D) stimulus that is normalized by the detection threshold. However, in that work we did not constrain the weighting function to be positive semidefinite, and as a result, the model will not generalize well to all future 6D inputs. In this addendum, we reconsider the original data set, incorporating the positive-semidefinite constraint. We find that as few as four independent parameters are required to describe the coupling, and the model has one fewer coupling term than originally proposed. We also describe the process of fitting a positive-semidefinite function to a new stylus at some specific excitation frequency.**

Index Terms—High-frequency vibrations, pen-hold grasp, tool-mediated vibrotactile perception, untethered magnetic haptic interface.

I. Introduction

In our recent article [1], we explored how a six-dimensional (6D) vibrotactile stimulus rendered at the haptic interaction point (HIP) of a kinesthetic haptic interface, with the stylus held in a precision pen-hold grasp, can be mapped to a haptically equivalent 1D stimulus that is normalized by the detection threshold (i.e., a 6D stimulus at the detection threshold maps to a unit-magnitude 1D stimulus). This knowledge will enable maximizing the efficiency of vibrotactile display for any given haptic interface. In [1], we gathered a large human-subjects data set in which we determined detection thresholds for 45 configurations representing distinct combinations of three orthogonal forces and three orthogonal torques rendered at the HIP, at a single frequency of 108 Hz (corresponding to the peak sensitivity in an earlier study). Using this data set, we found a general quadratic weighting function of the form

$$\tilde{P} = V^\top W V \tag{1}$$

to predict the 1D normalized stimulus (i.e., the detection threshold is indicated by $\tilde{P} = 1$) for a given 6D vibrotactile stimulus, where $V = [f_x\ f_y\ f_z\ \tau_x\ \tau_y\ \tau_z]^\top$ is the signed magnitudes of the 6D vibrotactile stimulus at a given frequency (using the coordinate system defined in Fig. 1), and W is a 6×6 symmetric weighting matrix whose elements serve a dual purpose of normalizing the stimulus values and describing the coupling between the six orthogonal components at the HIP. For brevity, we refer the reader to [1] for a critical review of related prior work, and details of the experiment with 12 human subjects that was used to gather the data set.

We have subsequently come to the realization that the weighting matrix should be constrained to be positive semidefinite. Otherwise, it will be possible to give a new 6D vibrotactile stimulus for which the model predicts a 1D normalized stimulus that is negative, which is nonsensical. That is, the model will not generalize sufficiently well to new data. In this addendum, we impose the positive-semidefinite constraint on our weighting matrix, using the same data set from [1]. Including this constraint increases the complexity of solving for W, so we were not able to apply the same methodology that we used in [1] to solve for W while avoiding fitting an overly complex model to the data. Instead, we apply methodologies from machine learning [2] to avoid over-fitting while solving for W, which we believe to be superior to our previous methodology. In the end, our conclusions about the structure of W are slightly different than our conclusions in [1], and we enumerate those differences.

II. Solving for a Positive-Semidefinite W

A positive-semidefinite matrix W can be decomposed using the Cholesky decomposition, $W = LL^\top$, where L is a lower-triangular matrix. A 6×6 positive-semidefinite W matrix has only 21 independent parameters due to its symmetry, and that fact is explicitly encoded in the 21 elements in its associated L matrix:

$$L = \begin{bmatrix} L_{11} & 0 & 0 & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 & 0 & 0 \\ L_{31} & L_{32} & L_{33} & 0 & 0 & 0 \\ L_{41} & L_{42} & L_{43} & L_{44} & 0 & 0 \\ L_{51} & L_{52} & L_{53} & L_{54} & L_{55} & 0 \\ L_{61} & L_{62} & L_{63} & L_{64} & L_{65} & L_{66} \end{bmatrix} \tag{2}$$

We can also pack these parameters into a 21×1 array as

$$\Lambda = [L_{11} \cdots L_{61}\, L_{22} \cdots L_{62}\ \cdots\ \cdots\ \cdots L_{66}]^\top \tag{3}$$

For a data set comprising N 6D best-estimate-threshold (BET) stimuli V_j, we can use the gradient-descent method in Algorithm 1 to solve for the L that minimizes the sum-of-squares error (SSE), which assumes a $\log_{10}$ mapping (based on Stevens' power law [3]) from skin-absorbed power to stimulus intensity:

$$\mathrm{SSE} = \sum_N \left(\log_{10} \tilde{P}\right)^2, \tag{4}$$

Manuscript received July 1, 2020; revised September 23, 2020 and November 3, 2020; accepted November 29, 2020. Date of publication December 8, 2020; date of current version June 16, 2021. This work was supported by the National Science Foundation under Grant 1423273. This article was recommended for publication by Associate Editor Dr. Matteo Bianchi and Editor-in-Chief Lynette Jones upon evaluation of the reviewers' comments. *(Corresponding author: Ruisi Zhang.)*

The authors are with the Department of Mechanical Engineering and the Robotics Center, University of Utah, Salt Lake City, UT 84112 USA (e-mail: ruisi.zhang@utah.edu; trevor.schwehr@utah.edu; jake.abbott@utah.edu).

Digital Object Identifier 10.1109/TOH.2020.3043095

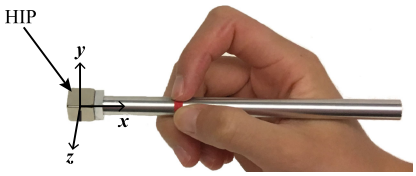

Fig. 1. Stylus of the untethered magnetic haptic interface used in the original study [1], held in a precision pen-hold grasp. Forces and torques are applied at the haptic interaction point (HIP). The coordinate system used defines x along the stylus, y pointing upward, and z pointing toward the subject when the stylus is held in front of the subject as shown. This coordinate system should be viewed as constant in the frame of the hand, not the stylus.

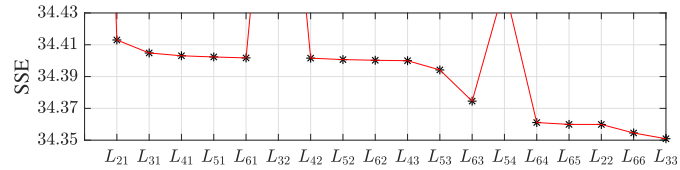

Fig. 2. Process for eliminating independent elements of L, from left to right. From L_{21} to L_{65}, each entry L_{ij} shows the average validation SSE after that element is set equal to zero. The entries L_{22} and L_{66} show the average validation SSE after those elements are constrained as $L_{22}^2 = L_{32}^2 + L_{33}^2$ and $L_{66}^2 = L_{54}^2 + L_{55}^2$, respectively. Note that the average validation SSE values for the original 21-parameter model and the models that eliminated L_{32} and L_{54} are off the chart.

A. Eliminating Independent Parameters in L

We would like to avoid over-fitting a model to our data set because an over-fit model will not generalize well to future data. To this end, we begin with the model with 21 independent elements as in (2), and attempt to eliminate each of the off-diagonal elements (i.e., set them equal to zero) one by one, starting from L_{21}, moving down the first column, and then repeating column by column until reaching L_{65}; this strategy is motivated by the structure of the W matrix associated with L (see Appendix III). For each of these element tests, we compare the SSE of the model with the term eliminated versus the SSE from the previous best model. We use a 12-fold cross-validation method in which we use the $N = 495$ training data from 11 of the subjects to fit a model, and we leave one subject out as our validation data set from which the SSE is calculated, and repeat until each of the 12 subjects has been used as the validation set [2]. A given element is eliminated if the average SSE of the 12 validation sets decreases when the element has been eliminated, indicating that the new reduced L matrix is more robust at predicting data that was not included in the training set.

When using Algorithm 1, we establish the initial condition L_0 by explicitly solving for a diagonal L to give just the diagonal terms in W (which are found from the first six of our 45 configurations). We use a tolerance $\epsilon = 1.0\text{e}{-10}$, a step size $\gamma = 0.1$, a finite difference $\delta = 1.0\text{e}{-10}$, and a maximum number of iterations of $i_{\max} = 1.0\text{e}6$.

Figure 2 shows the results for the average validation SSE as we eliminate independent parameters in L. The resulting L matrix has just eight parameters: L_{11}, L_{22}, L_{32}, L_{33}, L_{44}, L_{54}, L_{55}, and L_{66}.

Next, using the reduced L matrix above, we use the same 12-fold cross-validation method to evaluate if the model can be further simplified with two additional constraints as suggested by our study in [4]: $W_{22} = W_{33}$ and $W_{55} = W_{66}$. These constraints are each premised on the respective symmetry in the stylus (see Section II-C for more on this). These constraints are converted into constraints on the remaining terms of L as $L_{22}^2 = L_{32}^2 + L_{33}^2$ and $L_{66}^2 = L_{54}^2 + L_{55}^2$, respectively (see Appendix III). As indicated in Fig. 2, both of these constraints lead to a further reduction in the average validation SSE.

Finally, we observed that L_{33} was many orders of magnitude smaller than the other remaining terms, so we hypothesized that it could also be eliminated. As indicated in Fig. 2, this leads to a further reduction in the average validation SSE.

B. Final Model

The resulting final L matrix has five independent parameters—L_{11}, L_{32}, L_{44}, L_{54}, and L_{55}—as well as the constraints $L_{22} = L_{32}$ (i.e., $W_{22} = W_{33}$) and $L_{66} = \sqrt{L_{54}^2 + L_{55}^2}$ (i.e., $W_{55} = W_{66}$). After fitting this five-parameter L to the full 12-subject data set ($N = 540$), our final positive-semidefinite $W = LL^\top$ is:

$$W = \begin{bmatrix} 2.6\text{e}6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.3\text{e}6 & 3.3\text{e}6 & 0 & 0 & 0 \\ 0 & 3.3\text{e}6 & 3.3\text{e}6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8.1\text{e}10 & 6.6\text{e}8 & 0 \\ 0 & 0 & 0 & 6.6\text{e}8 & 5.2\text{e}8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5.2\text{e}8 \end{bmatrix} \quad (5)$$

where the upper-left 3×3 block has units N^{-2} and the lower-right 3×3 block has units $\text{N}^{-2} \cdot \text{m}^{-2}$. The final weighting matrix has five independent weights: one weight for each of the the f_x and τ_x modes, one common weight for the f_y and f_z modes as well as their coupling term, one common weight for the τ_y and τ_z modes, and a coupling term between τ_x and τ_y.

Using this final five-parameter weighting matrix, we compute the predicted normalized skin-absorbed power for the full data set comprising 45 configurations and 12 subjects; the results are shown in Fig. 3. We report the predicted normalized skin-absorbed power in a $\log_{10}$ scale, since our cutaneous system perceives vibrotactile stimuli in this scale based on Stevens' power law [3]. Note that $\log_{10}(\tilde{P}) > 0$ predicts a stimulus being above the detection threshold, and $\log_{10}(\tilde{P}) < 0$ predicts a stimulus being below the detection threshold, whereas the actual data represents BET values. Qualitatively, the model does a good job of capturing the thresholds obtained across configurations, with the vast majority of data grouped near $\log_{10}(\tilde{P}) = 0$; intersubject variance remains, which is unavoidable for any one-size-fits-all model.

C. Discussion

Recall that the numerical values in (5) are for our specific stylus, excited at 108 Hz. For a different stylus and/or different frequency, the values in the L matrix can be determined using the same procedure described here, with $W = LL^\top$.

It is our conjecture that a simpler fitting procedure may be possible with only knowledge of the diagonal elements of W, now that we are equipped with the form of the L and W matrices. It is possible to experimentally determine W_{11}, W_{22} and/or W_{33}, W_{44}, and W_{55} and/or W_{66}. It is our conjecture that setting $W_{23} = W_{32} = \sqrt{W_{22}W_{33}}$ and $W_{45} = W_{54} = 0.102\sqrt{W_{44}W_{55}}$ will maintain the same relative coupling relationships between f_y and f_z and between τ_x and τ_y, respectively. In this case, as few as four experiments would be required to characterize a new stylus and/or new frequency.

In [4], we experimentally determined the detection threshold values across a wide range of frequencies (20–250 Hz), for the same stylus used here, in each of six orthogonal modes independently. That is, we know the values for each of the diagonal

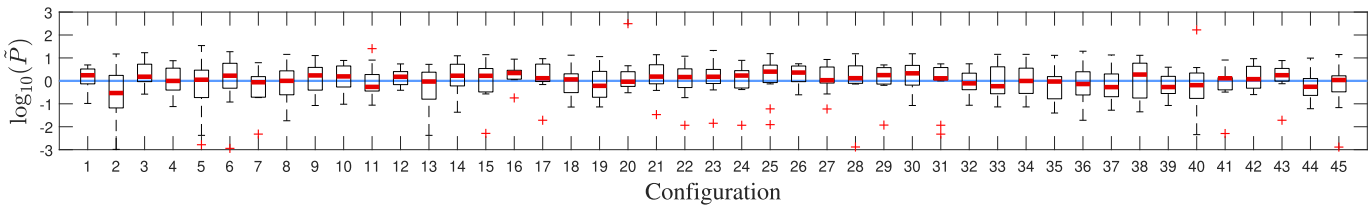

Fig. 3. Box-whisker plot showing estimated normalized skin-absorbed power in a $\log_{10}$ scale using our proposed six-parameter W, for all 45 configurations (at 108 Hz) and 12 subjects tested. A perfect model would result in an estimated value of 0 (indicated with a blue line).

Algorithm 1: Solve for L using gradient-descent method

Input: Training data set V with N training data, initial condition L_0, step size γ, tolerance ϵ, maximum number of iterations $i_{\max}$, finite-difference value δ, set of constraints to impose on L

Procedure L
  $L \leftarrow L_0$
  for $j \leftarrow 1$ to N **do**
    $G_0(j) \leftarrow \log_{10}(V_j^\top L L^\top V_j)$ ▷ $\tilde{P}_j = V_j^\top L L^\top V_j$
  end for
  $\mathrm{SSE}_0 \leftarrow G_0^\top G_0$
  $\Lambda_0 \leftarrow [L_{11} \cdots L_{61}\, L_{22} \cdots L_{62} \cdots \cdots \cdots L_{66}]^\top$
  for $i \leftarrow 1$ to $i_{\max}$ **do**
    for $j \leftarrow 1$ to N ▷ build Jacobian $\partial G / \partial \Lambda$
      for $k \leftarrow 1$ to 21 **do**
        $\Lambda_f \leftarrow \Lambda_{i-1}$
        $\Lambda_b \leftarrow \Lambda_{i-1}$
        $\Lambda_f(k) \leftarrow \Lambda_f(k) + \delta$ ▷ forward difference
        Construct L_f using Λ_f and impose L constraints
        $\Lambda_b(k) \leftarrow \Lambda_b(k) - \delta$ ▷ backward difference
        Construct L_b using Λ_b and impose L constraints
        $J_{jk} \leftarrow \frac{\log_{10}(V_j^\top L_f L_f^\top V_j) - \log_{10}(V_j^\top L_b L_b^\top V_j)}{2\delta}$
      end for
  end for
  $\Lambda_i \leftarrow \Lambda_{i-1} - \gamma J^\top G_{i-1}$ ▷ gradient descent
  Construct L using Λ_i and impose L constraints
  for $j \leftarrow 1$ to N **do**
    $G_i(j) \leftarrow \log_{10}(V_j^\top L L^\top V_j)$
  end for
  $\mathrm{SSE}_i \leftarrow G_i^\top G_i$
  if $\|\mathrm{SSE}_i - \mathrm{SSE}_{i-1}\| < \epsilon$ **then**
    return L ▷ end if tolerance reached
  end if
  end for
  Return L ▷ end if max iterations reached
end procedure

elements of W at each frequency tested. Using the simple fitting procedure described above, one could approximate the complete W matrix at each frequency tested for the stylus used in this paper and [4]. One could also interpolate values at frequencies not tested.

Many styluses are approximately radially symmetric, as ours was in this study, but it should be possible to apply the modeling in this work to an asymmetric stylus. We would not expect the $W_{22} = W_{33}$ and $W_{55} = W_{66}$ constraints to hold in general for an asymmetric stylus. The equivalent constraints on L would thus be removed, requiring up to two addition independent parameters. It is our conjecture that the simplified fitting procedure described above would still apply.

III. Comparison to Previous Results

For comparison, the final weighting matrix proposed in [1] is

$$W = \begin{bmatrix} 1.5\text{e}7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.2\text{e}7 & 4.8\text{e}7 & 0 & 0 & 0 \\ 0 & 4.8\text{e}7 & 2.2\text{e}7 & -1.2\text{e}9 & 0 & 0 \\ 0 & 0 & -1.2\text{e}9 & 2.0\text{e}10 & 2.0\text{e}9 & 0 \\ 0 & 0 & 0 & 2.0\text{e}9 & 7.8\text{e}7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7.8\text{e}7 \end{bmatrix} \tag{6}$$

with the same units as (5). Note: there is typo in exponent values in [1] due to an error in unit conversion, which has been corrected in (6). There is a coupling term between f_z and τ_x, which we do not see in (5). In a sense, that coupling term was over-fitting the data, since W was not positive semidefinite. There are other differences between the two studies that result in changes in the numerical values of W. Our study in [1] eliminated off-diagonal elements of W based on their statistical significance with a (somewhat arbitrary) significance of $\alpha = 0.1$, which our current method does not require. In fitting the final model to the 12-subject data set, our study in [1] used only the median BET of each configuration as the training set ($N = 45$), whereas our current method uses the BETs for the entire data set ($N = 540$). Finally, our study in [1] found an optimal model by minimizing error in a linear scale, whereas our current method minimizes error in a $\log_{10}$ scale, which better matches human sensory perception. If we compare the results of Fig. 3 to the analogous results in [1], we find that the results are qualitatively very similar. That is, although we expect the model with the positive-semidefinite constraint to result in a higher overall error, the difference appears negligible.

Appendix A
Cholesky Decomposition $W = LL^\top$

The symmetric W matrix resulting from the L matrix in (2) is:

$$W_{11} = L_{11}^2 \tag{7}$$
$$W_{12} = W_{21} = L_{11} L_{21} \tag{8}$$
$$W_{13} = W_{31} = L_{11} L_{31} \tag{9}$$
$$W_{14} = W_{41} = L_{11} L_{41} \tag{10}$$
$$W_{15} = W_{51} = L_{11} L_{51} \tag{11}$$
$$W_{16} = W_{61} = L_{11} L_{61} \tag{12}$$
$$W_{22} = L_{21}^2 + L_{22}^2 \tag{13}$$
$$W_{23} = W_{32} = L_{21} L_{31} + L_{22} L_{32} \tag{14}$$

$$W_{24} = W_{42} = L_{21}L_{41} + L_{22}L_{42} \tag{15}$$
$$W_{25} = W_{52} = L_{21}L_{51} + L_{22}L_{52} \tag{16}$$
$$W_{26} = W_{62} = L_{21}L_{61} + L_{22}L_{62} \tag{17}$$
$$W_{33} = L_{31}^2 + L_{32}^2 + L_{33}^2 \tag{18}$$
$$W_{34} = W_{43} = L_{31}L_{41} + L_{32}L_{42} + L_{33}L_{43} \tag{19}$$
$$W_{35} = W_{53} = L_{31}L_{51} + L_{32}L_{52} + L_{33}L_{53} \tag{20}$$
$$W_{36} = W_{63} = L_{31}L_{61} + L_{32}L_{62} + L_{33}L_{63} \tag{21}$$
$$W_{44} = L_{41}^2 + L_{42}^2 + L_{43}^2 + L_{44}^2 \tag{22}$$
$$W_{45} = W_{54} = L_{41}L_{51} + L_{42}L_{52} + L_{43}L_{53} + L_{44}L_{54} \tag{23}$$
$$W_{46} = W_{64} = L_{41}L_{61} + L_{42}L_{62} + L_{43}L_{63} + L_{44}L_{64} \tag{24}$$
$$W_{55} = L_{51}^2 + L_{52}^2 + L_{53}^2 + L_{54}^2 + L_{55}^2 \tag{25}$$
$$W_{56} = W_{65} = L_{51}L_{61} + L_{52}L_{62} + L_{53}L_{63} + L_{54}L_{64} + L_{55}L_{65} \tag{26}$$
$$W_{66} = L_{61}^2 + L_{62}^2 + L_{63}^2 + L_{64}^2 + L_{65}^2 + L_{66}^2 \tag{27}$$

REFERENCES

[1] R. Zhang, T. J. Schwehr, and J. J. Abbott, "Dimensional reduction for 6D vibrotactile display," *IEEE Trans. Haptics*, vol. 13, no. 1, pp. 102–108, Jan.–Mar. 2020.

[2] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 569–575, Mar. 2010.

[3] S. S. Stevens, "On the psychophysical law," *Psychol. Rev.*, vol. 64, no. 3, pp. 153–181, 1957.

[4] R. Zhang, A. J. Boyles, and J. J. Abbott, "Six principal modes of vibrotactile display via stylus," in *Proc. IEEE Haptics Symp.*, 2018, pp. 313–318.