

# 1 Stochastic characterization of small-scale algorithms for human sensory 2 processing

3 Peter Neri<sup>a)</sup>

4 *Aberdeen Medical School, Institute of Medical Sciences, Aberdeen, Scotland AB25 2ZD, United Kingdom*

5 (Received 2 August 2010; accepted 13 November 2010; published online xx xx xxxx)

AQ:  
#1

6 Human sensory processing can be viewed as a functional  $\mathcal{H}$  mapping a stimulus vector  $\mathbf{s}$  into a  
7 decisional variable  $r$ . We currently have no direct access to  $r$ ; rather, the human makes a decision  
8 based on  $r$  in order to drive subsequent behavior. It is this (typically binary) decision that we can  
9 measure. For example, there may be two external stimuli  $\mathbf{s}^{[0]}$  and  $\mathbf{s}^{[1]}$ , mapped onto  $r^{[0]}$  and  $r^{[1]}$  by  
10 the sensory apparatus  $\mathcal{H}$ ; the human chooses the stimulus associated with largest  $r$ . This kind of  
11 decisional transduction poses a major challenge for an accurate characterization of  $\mathcal{H}$ . In this  
12 article, we explore a specific approach based on a behavioral variant of reverse correlation techni-  
13 ques, where the input  $\mathbf{s}$  contains a target signal corrupted by a controlled noisy perturbation. The  
14 presence of the target signal poses an additional challenge because it distorts the otherwise unbiased  
15 nature of the noise source. We consider issues arising from both the decisional transducer and the  
16 target signal, their impact on system identification, and ways to handle them effectively for system  
17 characterizations that extend to second-order functional approximations with associated small-scale  
18 cascade models. © 2010 American Institute of Physics. [doi:10.1063/1.3524305]

19  
20 During the past 50 years, reverse correlation has become  
21 the elective methodology for the characterization of sen-  
22 sory neurons. Starting with the 1970s, similar tools have  
23 been developed for probing the sensory processes medi-  
24 ating vision/audition in humans. In both single-neuron  
25 electrophysiology and sensory psychophysics, the domi-  
26 nant model has been one where a linear filter is followed  
27 by a static nonlinearity (e.g., spike generation in single  
28 neurons, behavioral choice in human observers). For  
29 white-noise inputs and provided this simple model is ad-  
30 equate, results from reverse correlation experiments are  
31 relatively easy to interpret and the linear kernel is an  
32 appropriate descriptor for the process of interest. How-  
33 ever, many processes operating within both neurons and  
34 observers are not adequately captured by the linear-  
35 nonlinear cascade model; in these instances, the linear  
36 kernel characterization must be augmented by additional  
37 nonlinear kernels. Methods for characterizing nonlinear  
38 kernels have been developed for application with single  
39 neurons, but the extension to human observers is not  
40 trivial due to significant differences between the two sys-  
41 tems relating to the characteristics of the sensory input to  
42 the system, as well as the nature of the available output  
43 from the system. We tackle this problem at both theoret-  
44 ical and experimental levels, and show how some of the  
45 distortions in the kernel estimation procedure that are  
46 idiosyncratic to the human sensory process can be turned  
47 to the experimenter's advantage and exploited to gain  
48 additional information on the process at hand.

## I. INTRODUCTION

### A. Animal sensors as signal detection devices

50

51 The most important goal of sensory processing is to  
52 drive adequate behavior: the animal must process informa-  
53 tion about the environment in order to take action that will  
54 maximize its survival. Natural selection has produced so-  
55 phisticated sensory systems, which have been conceptualized  
56 in the form of signal detection devices during the past 50  
57 years;<sup>17,18,20</sup> the application of signal detection theory (SDT)  
58 to animal sensory processing still represents the most robust  
59 approach for interpreting quantitative measurements of this  
60 phenomenon.<sup>38</sup> We consequently adopt SDT here without  
61 questioning its applicability, while at the same time recog-  
62 nizing that it may not provide an adequate description for  
63 every aspect of animal sensory processing.

64 Implicit to SDT is the decision variable assumption:<sup>64</sup>  
65 the animal's decision comes in the form of a choice among a  
66 set of sensory stimuli, where the choice is based on one  
67 figure of merit for each stimulus. Regardless of the dimen-  
68 sionality of the incoming stimulation, each stimulus is there-  
69 fore mapped to a single scalar value; this value is meant to  
70 reflect how likely the corresponding stimulus is to contain  
71 the signal of interest. The animal then chooses the stimulus  
72 associated with largest estimated likelihood. The estimated  
73 likelihood is the actual likelihood if the animal operates like  
74 a Bayesian device;<sup>20</sup> if not, then it is simply some figure of  
75 merit constrained by what the animal's neural hardware is  
76 capable of doing.

### B. Output distortion: The decisional transducer

77

78 There is a fundamental distinction between the operator  
79 that maps input vectors into scalars on the one hand, and the  
80 operator that converts scalars into decisions on the other. The  
81 former (which we call  $\mathcal{H}$ ) is a functional that encompasses

<sup>a)</sup>Electronic mail: peter.neri@abdn.ac.uk.

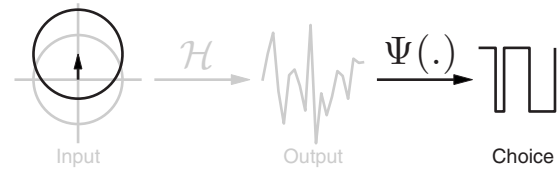


FIG. 1. The sensory process as conceptualized here. Gray items delineate the system of interest and the ideal setting for accessing and characterizing it: a controlled unbiased input with elliptical symmetry (Gaussian noise source) is mapped to a scalar output value by the functional  $\mathcal{H}$  which captures the relevant aspects of the sensory apparatus. Our goal is to characterize  $\mathcal{H}$ . When applied to human sensory processing, this framework must be modified in two substantial ways, highlighted by the black items: (1) the input must contain a target signal, thus shifting and biasing the input distribution on the left (Sec. I C); (2) the scalar output must be converted into a (typically binary) choice by the decisional transducer  $\Psi$  (Sec. I B). The challenge is to characterize  $\mathcal{H}$  despite the gross distortions introduced by these two transformations.

82 all processing layers believed to reflect general and physi-  
 83 ologically relevant properties of how visual processing oper-  
 84 ates, while the latter (which we call  $\Psi$ ) is dependent upon  
 85 the specific question that is asked of observers and the form  
 86 in which the response is acquired. Our primary interest is in  
 87 characterizing  $\mathcal{H}$ , not  $\Psi$ —indeed the desired goal is to char-  
 88 acterize  $\mathcal{H}$  regardless of the specific  $\Psi$  that is instantiated by  
 89 the chosen experimental protocol.

90 To provide a concrete example, we imagine that our in-  
 91 terest is in how the human visual system processes the direc-  
 92 tion of moving objects. We select some stimulus specifics,  
 93 say dots moving at a certain speed. One member of our labo-  
 94 ratory runs experiment 1 in which the observer sees two such  
 95 stimuli on each trial, and must choose which one moves  
 96 better (where a precise definition of “better” is irrelevant for  
 97 the current discussion). Another member runs experiment 2  
 98 in which the observer sees four such stimuli on every trial,  
 99 and must choose one out of four. A third member of the  
 100 laboratory runs experiment 3 in which the observer only sees  
 101 one stimulus on every trial and must judge how well it  
 102 moves on a rating scale from 1 to 10. The three experiment-  
 103 ers are measuring outputs of different kinds: binary in ex-  
 104 periment 1, ternary in experiment 2, on a 10-point rating  
 105 scale in experiment 3. These different outputs are necessarily  
 106 associated with different decisional rules on the part of the  
 107 observer, yet our hope is that the final characterization of the  
 108 system will be one and the same: the neural circuit used by  
 109 human observers to process moving stimuli. Not simply be-  
 110 cause we wish the ensuing characterization to be representa-  
 111 tive of the experiments carried out in the laboratory, but  
 112 more importantly because we wish it to be representative of  
 113 visual processing in the natural environment where humans  
 114 must cope with a wide range of different tasks. We therefore  
 115 split the description of the system into a general-purpose  
 116 operator ( $\mathcal{H}$ ) followed by an *ad hoc* decisional transducer  
 117 ( $\Psi$ ). We want  $\mathcal{H}$ ; the question is whether we can bypass  $\Psi$   
 118 get at it, which is one of the main concerns of this article.

### 119 C. Input distortion: The target signal

120 The other main concern relates to the unavoidable pres-  
 121 ence of a target signal. As we have detailed above, our ability  
 122 to measure behavior relies on querying human participants  
 123 and recording their response. We therefore need to pose a  
 124 question, and it must be well defined in order to allow mean-  
 125 ingful treatment of the subject. Within the context of SDT,  
 126 the question involves selecting a stimulus that contains a  
 127 certain signal, requiring us to define a signal-to-be-detected.  
 128 This apparently innocuous requirement has fundamental im-  
 129 plications for reverse correlation methodologies like the one  
 130 described here. These approaches rely on sensory inputs con-  
 131 sisting of noisy processes with unbiased characteristics, such  
 132 as Gaussian white noise.<sup>40</sup> The introduction of a consistent  
 133 signal source, in the form of a sensory target for the human  
 134 to detect, represents a significant departure from these re-  
 135 quirements.

136 It is legitimate to ask whether one could not simply do  
 137 away with the target, and ask human participants to press  
 138 buttons while observing pure noise (e.g., Ref. 19). This ap-

proach (which amounts to an extreme version of previous  
 strategies where the analysis is restricted to target-absent tri-  
 als in yes-no tasks<sup>1,4,52,75</sup>) presents insurmountable problems.  
 First, it is very difficult to prompt a stable behavioral strategy  
 on the part of the observer because, among other reasons, no  
 meaningful correct/incorrect feedback can be provided (there  
 is no correct or incorrect answer when there is no signal).  
 Second, it is impossible to know whether the observer is  
 performing the task at all: he/she may be pressing buttons  
 randomly without any way for the experimenter to know,  
 because there is no objective manner of establishing whether  
 the system is detecting anything (there is nothing to be de-  
 tected). Third, worst and most important of all, it is impos-  
 sible to know whether the observer is performing a different  
 task from the one specified by the experimenter: observers  
 may perform some task and generate some form of measur-  
 able outcomes, but the experimenter cannot be certain as to  
 what task that is. These and related issues exclude the pos-  
 sibility that robust results may be obtained in the total ab-  
 sence of a target signal. Furthermore, there are situations in  
 which adding a target can actually increase the ability of the  
 experimenter to characterize a certain class of mechanisms  
 (see Sec. VI A). The impact of the target signal on a reliable  
 characterization of  $\mathcal{H}$  is the second main concern of this  
 article. The two topics discussed above, introduction of a  
 target signal and decisional transduction, are highlighted and  
 summarized in Fig. 1. As emphasized in the figure, the  
 former affects what goes into the system (input), the latter  
 affects what comes out of the system (output).

### D. Structure of this article

We start by fleshing out the conceptual framework out-  
 lined above, i.e., define the input,  $\mathcal{H}$ ,  $\Psi$ , etc. This is done in  
 Sec. II where we approximate  $\mathcal{H}$  using a Volterra functional  
 expansion (Sec. II B). This approximation recasts  $\mathcal{H}$  in terms  
 of system kernels; within Sec. II we therefore consider one  
 specific way of estimating these kernels based on  
 cross-correlation<sup>47,68</sup> (Sec. II C). Only a brief overview is  
 given in Sec. II, to be followed by more detailed treatments  
 in Secs. III–V. Section III introduces some basic results re-  
 lating to the handling of  $\Psi$ , i.e., how it can be bypassed to

179 access the  $\mathcal{H}$  kernels. We then consider issues relating to  
 180 first-order kernel estimation in Sec. IV and second-order ker-  
 181 nel estimation in Sec. V. Many of the difficulties that arise in  
 182 connection with kernel estimation considered in these sec-  
 183 tions relate to the presence of the target signal (as already  
 184 mentioned in Sec. I C). Finally, in Sec. VI we consider how  
 185 kernel estimation can be used to infer/inform the structure of  
 186 simple algorithms that attempt to simulate human sensory  
 187 processing; we do this via a combination of both theoretical  
 188 and experimental materials. Because the connection between  
 189 kernel structure and potential underlying cascade model is  
 190 currently best understood only for a standard set of simple  
 191 models,<sup>16,27,41,86</sup> we show how well-established nonlinear  
 192 models in the vision literature can be approximated using  
 193 these simple cascades (Secs. VI A and VI B) and then con-  
 194 sider relevant examples from real data (Secs. VI C–VI F). Of  
 195 the four data sets which we evaluate, two are easily inter-  
 196 preted using the tools described in this article (Secs. VI C  
 197 and VI D) while the remaining two present significant chal-  
 198 lenges for the proposed approach (Secs. VI E and VI F).

## 199 E. Notation

200 In treating these topics we adopt the following conven-  
 201 tions: vectors/matrices are indicated in bold letters, e.g.,  $\mathbf{H}$ ;  
 202 occasionally we refer to individual elements (e.g.,  $k$ ) via  
 203  $\mathbf{H}(x_k)$  (we do not adopt  $\mathbf{H}_k$  or  $H_k$  to avoid additional sub-  
 204 scripting and to accommodate the  $:$  notation detailed in the  
 205 next line). When we index using  $:$  we take the entire corre-  
 206 sponding vector dimension, e.g.,  $\mathbf{H}_2(:, x_k)$  is a 1D vector con-  
 207 sisting of the  $m$  elements  $\mathbf{H}_2(x_j, x_k)$  for  $j$  from 1 to  $m$  for a  
 208 fixed  $k$ . We denote the inner product by  $\langle \cdot, \cdot \rangle$  and the outer  
 209 product by  $\otimes$  (we do not adopt  $\mathbf{x}^T \mathbf{y}$  and  $\mathbf{xy}^T$  to avoid addi-  
 210 tional superscripting). When matrices appear as arguments of  
 211  $\langle \cdot, \cdot \rangle$ , the Frobenius inner product is assumed [i.e., if  $\mathbf{A}$  and  $\mathbf{B}$   
 212 are two matrices,  $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{AB}^T)$ ]. Using this notation,  
 213  $\langle \mathbf{H}_2(:, x_k), \mathbf{H}_1 \rangle$  is the inner product between 1D vector  
 214  $\mathbf{H}_2(:, x_k)$  and 1D vector  $\mathbf{H}_1$ .  $\langle \cdot \rangle_i$  stands for (ensemble) aver-  
 215 age over an infinite number of trials  $m$ , i.e.,  $\lim_{m \rightarrow \infty} 1/m \sum_m$   
 216 (also  $\langle \mathbf{x} \rangle_i = \langle \mathbf{x}, 1/m \rangle$  if  $x_i$  are elements of  $\mathbf{x}$ ). The following  
 217 common symbols are adopted:  $*$  for convolution,  $\star$  for cross-  
 218 correlation, and  $\circ$  for Hadamard (entrywise) product. We use  
 219  $\triangleq$  for matrix assignments  $\mathbf{A} \triangleq \mathbf{B}$  to indicate  $\mathbf{A} = (\mathbf{B} + \mathbf{B}^T)/2$   
 220 (assign symmetry by definition to  $\mathbf{A}$  whenever not explicit in  
 221  $\mathbf{B}$ ). We adopt the digital signal processing notation for  
 222 Kronecker  $\delta: \delta[x] = 1$  for  $x = 0$ ,  $\delta[x] = 0$  for  $x \neq 0$ , where  $x$  is  
 223 integer.

## 224 II. GENERAL FRAMEWORK

### 225 A. Input, task, and decisional transducer

226 Our focus is on the two alternative forced choice (AFC)  
 227 designs (detailed below), arguably the most robust and  
 228 best understood experimental protocol in sensory  
 229 psychophysics.<sup>20</sup> We occasionally consider the implications  
 230 of response bias (Sec. III B 2) for which the 2AFC design  
 231 approximates a yes-no protocol. This is undesirable for a  
 232 number of reasons,<sup>20</sup> particularly in relation to nonlinear ker-  
 233 nel estimation.<sup>47,50</sup>

In the 2AFC design, two stimuli are presented on every  
 trial: stimulus  $\mathbf{s}^{[0]}$  contains the nontarget  $\mathbf{t}^{[0]}$ , while stimulus  
 $\mathbf{s}^{[1]}$  contains the target  $\mathbf{t}^{[1]}$ . Noise  $\mathbf{n}$  is added to both so that,  
 on trial  $i$ ,  $\mathbf{s}_i^{[q]} = \mathbf{n}_i^{[q]} + \mathbf{t}^{[q]}$  ( $q=0$  for nontarget, 1 for target). The  
 statistics governing every element of  $\mathbf{n}$  is the same across  
 stimuli and trials and is a Gaussian noise source. Some re-  
 sults (Sec. III A) are applicable to any circularly symmetric  
 noise distribution (see also Ref. 61), but we restrict the spe-  
 cifics of the derivations here to Gaussian noise for which  
 $\langle \mathbf{n}_i \rangle_i = 0$ ,  $\langle \mathbf{n}_i(x_j) \mathbf{n}_i(x_k) \rangle_i = 0$  for  $j \neq k$  and  $= \sigma_N^2 = \mathbf{W}$  for  $j = k$ ,  
 $\langle \mathbf{n}_i(x_1) \cdots \mathbf{n}_i(x_{2M+1}) \rangle_i = 0$  for  $M = 0, 1, 2, \dots$ ,  $\langle \mathbf{n}_i(x_1) \cdots \mathbf{n}_i(x_{2M}) \rangle_i$   
 $= \Sigma \Pi \langle \mathbf{n}_i(x_j) \mathbf{n}_i(x_k) \rangle_i$ , where  $\Sigma \Pi$  stands for summation over all  
 distinct ways of partitioning the  $2M$  random variables into  
 products of averages of pairs.<sup>68</sup>

In line with SDT, each stimulus  $\mathbf{s}_i^{[q]}$  maps to a scalar  
 response  $r_i^{[q]}$ . This is the transformation we are specifically  
 interested in, i.e., the functional  $\mathcal{H}: \mathbf{s} \rightarrow r$ . However we do  
 not have access to  $r$ , but rather to a binary response  $z_i$  (0 for  
 incorrect and 1 for correct detection), which is generated  
 based on a comparison between  $r_i^{[1]}$  and  $r_i^{[0]}$ . Following SDT,  
 we make the very reasonable assumption that the probability  
 of making a correct response  $p(z_i = 1)$  on trial  $i$  (which we  
 abbreviate as  $p_i$ ) is  $p_i = \Psi(r_i^{[1]} - r_i^{[0]}) = \Psi(r_i)$ . We can make a  
 number of statements about  $\Psi: \Psi(0) = 0.5$  (the percentage of  
 correct responses should be at chance if the output of the  
 system is the same for target and nontarget),  $\lim_{x \rightarrow -\infty} \Psi(x)$   
 $= 0$ ,  $\lim_{x \rightarrow \infty} \Psi(x) = 1$ , and  $\Psi$  is monotonically increasing.<sup>14</sup> In  
 this article, we use a polynomial approximation for  $\Psi$ , which  
 is technically adequate only if  $\Psi$  is analytic. However, with a  
 few exceptions our conclusions are not based on accurate  
 approximations of  $\Psi$ , so they can be at least qualitatively  
 extended to nonanalytic transducers too. In practice,  $\Psi$  al-  
 most certainly conforms to a well-behaved sigmoid curve so  
 that our polynomial approximation poses no concern. In  
 principle, the only candidate exception would be the unit  
 step function [ $u(x) = 0$  for  $x < 0$  and  $= 1$  for  $x > 0$ , not analytic  
 at 0], but it is not even approximately realizable in human  
 sensory processing due to the presence of a very significant  
 amount of internal noise.<sup>50</sup> No deterministic transducer can  
 therefore be operating in the human observer, requiring that  
 $\Psi$  be specified statistically to reflect the characteristics of the  
 internal noise source. We can capture  $\Psi$  adequately using a  
 relatively shallow Weibull function, cumulative Gaussian  
 function, or hyperbolic tangent (any of these choices would  
 be appropriate here). When we approximate  $\Psi$  near its oper-  
 ating point  $\bar{r} = \langle r_i \rangle_i$  using

$$\hat{\Psi}_{\bar{r}}(r_i) \approx \sum_{d=0}^{\bar{d}} \Psi^{(d)}(r_i - \bar{r})^d, \quad (1)$$

we do not necessarily use  $\Psi^{(d)}$  to indicate local derivatives as  
 in the Taylor series, but the coefficients that minimize

$$\int \mathcal{N}(r - \bar{r}) (\Psi(r) - \hat{\Psi}_{\bar{r}}(r))^2 dr \quad (2)$$

subject to the inequality constraint

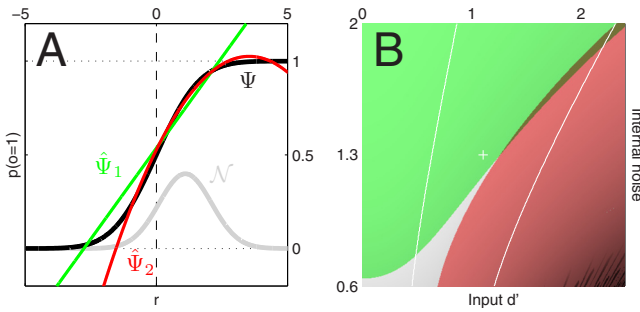


FIG. 2. (Color online) Realistic characterization of  $\Psi$ . (a) Gray trace shows a plausible distribution for  $r$ , the differential output generated by  $\mathcal{H}$ , assumed normal ( $\mathcal{N}$ ) in line with SDT. Black trace presents a plausible characteristic for  $\Psi$  in the form of a cumulative Gaussian function with a standard deviation of 1.3; this corresponds to the most basic 2AFC decisional rule (choose stimulus with largest associated output) corrupted by a veridical internal noise source (Ref. 49). Green trace shows first-order approximation  $\hat{\Psi}_1$  to  $\Psi$ , red trace shows second-order approximation [see Eq. (1)]. (b) Explained variance for first-order approximation  $\hat{\Psi}_1$  within the range spanned by  $\mathcal{N}$  as a function of input  $d'$  [before the addition of internal noise (Ref. 49)] and internal noise. Green region indicates acceptable range for  $\hat{\Psi}_1$ , defined as explained variance  $>90\%$ . Red region indicates range for which second-order approximation  $\hat{\Psi}_2$  is markedly superior to  $\hat{\Psi}_1$ , defined as explained variance for the former exceeding the latter by 10%. White lines delimit the typical region for humans corresponding to a 2AFC correct performance between 65% and 85%.

AQ:  
#18

$$\frac{\partial}{\partial r} \hat{\Psi}_{\tilde{d}}(\bar{r} + r, \theta) > 0, r \in (-2, 2) \quad (3)$$

for a specific choice of approximation order  $\tilde{d}$ , parametrization  $\theta$ , and mean response  $\bar{r}$ , where  $r$  is in units of its expected standard deviation across trials and  $\mathcal{N}(x)$  is the normal distribution. The minimizer in Eq. (2) is weighted by  $\mathcal{N}$  to target the operating range for  $r$ , which we assume to be normally distributed following the standard practice in SDT;<sup>20</sup> Equation (2) is therefore  $\langle (\Psi(r_i) - \hat{\Psi}_{\tilde{d}}(r_i))^2 \rangle_i$  written as expectancy. The requirement in Eq. (3) enforces positive monotonicity on  $\Psi$  within the range of interest ( $\pm 2$  standard deviations around  $\bar{r}$ ), which is theoretically necessary to enable meaningful treatments of the mapping from stimulus to percept.<sup>14</sup>

Figure 2 uses the approximation in Eqs. (1)–(3) to demonstrate that given current experimental knowledge relating to  $\Psi$ , approximations of low-order (sometimes as low as linear) are often adequate for our purposes. Figure 2(a) shows plausible examples of what  $\Psi$  may look like and where  $\mathcal{N}(r - \bar{r})$  would have to be located with relation to it in order to generate 75% correct responses, which is a typical choice for threshold performance;<sup>20,46</sup>  $\Psi$  is modeled as a cumulative Gaussian function centered at 0 with standard deviation  $\theta_1$ . Corresponding first- and second-order  $\Psi$  approximations are shown in green and red, respectively. Figure 2(b) plots the explained variance for different values of internal noise ( $\theta_1$ ) and input  $d'$  (Ref. 49) ( $\bar{r}$ ); the specific choice shown in (a) corresponds to the white cross in (b). The green region indicates acceptable first-order approximations; it can be seen that this region extends to most of the “physiological” range (defined here as the range of 65%–85% correct performance and bracketed by the white con-

tours). However, there is a substantial portion of this range for which a second-order approximation is preferable (red region). Throughout this study we therefore consider both first-order and second-order approximations to  $\Psi$ , depending on which one is necessary to demonstrate specific results.

## B. Volterra expansion of $\mathcal{H}$

We approximate  $\mathcal{H}$  using an adapted Volterra expansion,<sup>40,50,68</sup>

$$r^{[q]} = \sum_d \langle \mathbf{H}_d, \Theta_d(\mathbf{s}^{[q]}) \rangle, \quad (4)$$

where  $\Theta_d(j_1, \dots, j_d, \mathbf{s}) = \Pi_d \mathbf{s}(x_{j_d})$  is the  $d$ th degree monomial matrix of  $\mathbf{s}$  (same feature mapping used by polynomial classifiers in machine learning<sup>15,70</sup>). For example,  $\Theta_1 = \mathbf{s}$  and  $\Theta_2 = \mathbf{s} \otimes \mathbf{s}$ .  $\mathbf{H}_d$  therefore enjoys symmetry by definition. Equation (4) can be viewed as a generalization of the Taylor series to several variables.<sup>41</sup> From an experimental viewpoint, our interest is restricted to second-order approximations because higher-order kernel characterizations are impractical (previous work has shown that reliable characterizations of  $\mathbf{H}_2$  require at least 6–10k trials per human observer,<sup>50</sup> already approaching the upper limit for feasible projects on a sizeable number of participants). We occasionally approximate to third-order [e.g., Eq. (20)] to study the impact of third-order system kernels on lower-order kernel estimation.

Readers familiar with Volterra expansion will recognize that Eq. (4) should be expanded using convolution;<sup>40,68</sup> please refer to Sec. III B 3 for details on why the psychophysical variant can be equivalently expanded using inner products. Different from the output of the convolution-based expansion, which is ordinarily expressed as a time-varying function,<sup>40</sup> the output of Eq. (4) is a scalar (the decision variable): it generates a static snapshot from the underlying dynamic behavior of the system.<sup>36</sup> The lack of temporal dynamics for the output does not preclude characterization of the temporal dynamics associated with the sensory process that analyzes the input (see Fig. 9 for relevant examples); however, it represents an important limitation on the experimenter’s ability to constrain the resulting characterization.<sup>35</sup> The issue is one of resolving power: because we only acquire a binary response from the human participant for every temporal segment of the stimulus that we present on every trial (and we assume that the process is static from trial to trial), our ability to resolve the temporal scale of the output is poorer than our ability to control the temporal scale of the input.

Alternative experimental protocols have been attempted in which a response is acquired from the participant at any time during a long stimulus presentation,<sup>11,66,72</sup> but these methods have not afforded higher temporal resolution of the resulting system characterization (in some cases the resolution seemed lower, see Ref. 50 for a discussion of this issue in relation to a comparison between Refs. 11 and 53); even if somewhat higher resolution is obtained by prompting fast responses, it is then difficult to know whether the observed effects are due to perceptual or motor preparation.<sup>73</sup> The limitation in resolving the decisional output at high temporal

316  
317  
318  
319  
320  
321  
322  
323

324

AQ:

#2

326

327

328

329

330

331

332

333

334

AQ:

#3

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371 resolution derives from at least two factors: (1) the intrinsic  
372 timescale of the read-out process supporting behavioral de-  
373 cisions (the “psychological moment”) is expected to lie in  
374 the  $\sim 200$  ms range;<sup>85</sup> (2) the motor act by which humans  
375 convey a perceptual decision is inherently noisy on the range  
376 of 200–500 ms (e.g., Ref. 72). Because of these and poten-  
377 tially other factors, there is an intrinsic limit to the temporal  
378 resolution at which behavioral decisions can be accessed ex-  
379 perimentally; this limit is likely to exceed the temporal scale  
380 at which most sensory detectors operate. Nevertheless, the  
381 temporal dynamics of the perceptual process can be esti-  
382 mated at higher resolution by modulating the temporal prop-  
383 erties of the input on a finer scale; we discuss a relevant  
384 example in Sec. VI E (see also Ref. 50).

385 The above-detailed limitations become less relevant if  
386 we transition from low-level sensory processing on a scale of  
387  $\sim 20$ – $200$  ms to higher-level cognitive processing on a scale  
388 of  $\sim 2$ – $20$  s (generically referred to as learning). The non-  
389 linear dynamics (sometimes chaotic<sup>23</sup>) of this class of phe-  
390 nomena has been characterized using a variety of methods,<sup>22</sup>  
391 including nonlinear kernel estimation:<sup>23</sup> examples come from  
392 a wide range of psychological functions such as manual  
393 tracking,<sup>24</sup> handwriting dynamics,<sup>23</sup> associative memory,<sup>45</sup>  
394 visual illusions,<sup>21,78,79</sup> decision making,<sup>28</sup> prediction of event  
395 sequences,<sup>23,43,56,74</sup> and even cognitive development.<sup>81</sup> In  
396 these experiments, the output is sampled at a temporal reso-  
397 lution comparable to the modulation applied to the input, but  
398 the associated temporal scale is typically two orders of mag-  
399 nitude greater than the relevant scale for sensory signal de-  
400 tection. Furthermore, it is unclear whether the critical param-  
401 eter for the dynamics of these phenomena is time *per se* or  
402 rather the occurrence of a cognitive event: first-order kernels  
403 for priming (a form of short-term memory) decay as a func-  
404 tion of the number of sensory events,<sup>37</sup> to some extent inde-  
405 pendent of event duration, similar to other memory  
406 phenomena.<sup>45</sup> These cognitive processes are no doubt inter-  
407 esting but they fall outside the focus of interest for the  
408 present article, which is primarily on phenomena that ap-  
409 proach the maximum affordable temporal resolution at which  
410 behavior can be accessed for empirical measurement. In Sec.  
411 VI F we consider an application to visual adaptation, a pro-  
412 cess that belongs to the class of dynamic phenomena which  
413 operate on a significantly longer timescale.

#### 414 C. Kernel estimation via cross-correlation

415 Following the established methods,<sup>3</sup> first-order kernel  
416 estimates can be expressed as

$$417 \quad \hat{\mathbf{H}}_1 = \sum_{q,z} (2\delta[q-z] - 1) \hat{\mathbf{H}}_1^{[q,z]},$$

$$418 \quad \hat{\mathbf{H}}_1^{[q,1]} = \frac{1}{\bar{p}} \langle p_i \mathbf{n}_i^{[q]} \rangle_i, \quad (5)$$

$$419 \quad \hat{\mathbf{H}}_1^{[q,0]} = \frac{1}{(1-\bar{p})} \langle (1-p_i) \mathbf{n}_i^{[q]} \rangle_i,$$

420 where  $\sum_{q,z} = \sum_{q=0}^1 \sum_{z=0}^1$  and  $\bar{p} = \langle p_i \rangle_i$ . Expressed this way,  $\hat{\mathbf{H}}_1^{[q,z]}$   
421 is the expected value of  $\mathbf{n}^{[q]}$  when restricted to trials on

which the observer responded  $z$  (see Ref. 3). Second-order  
kernel estimates can be similarly expressed as<sup>47,50</sup>

$$422 \quad \hat{\mathbf{H}}_2 = \sum_{q,z} (2\delta[q-z] - 1) (\hat{\mathbf{H}}_2^{[q,z]} - \hat{\mathbf{H}}_1^{[q,z]} \otimes \hat{\mathbf{H}}_1^{[q,z]}), \quad (6)$$

423 where

$$424 \quad \hat{\mathbf{H}}_2^{[q,1]} = \frac{1}{\bar{p}} \langle p_i \mathbf{n}_i^{[q]} \otimes \mathbf{n}_i^{[q]} \rangle_i, \quad (7)$$

$$425 \quad \hat{\mathbf{H}}_2^{[q,0]} = \frac{1}{1-\bar{p}} \langle (1-p_i) \mathbf{n}_i^{[q]} \otimes \mathbf{n}_i^{[q]} \rangle_i. \quad (7)$$

426 It is clear from the subtraction term in Eq. (6) (involving  $\hat{\mathbf{H}}_1$ )  
427 that  $\hat{\mathbf{H}}_2$  represents a differential covariance matrix. In prin-  
428 ciple, there appears to be no obvious reason why one should  
429 not simply compute the differential second-order moment  
430  $\sum_{q,z} (2\delta[q-z] - 1) \hat{\mathbf{H}}_2^{[q,z]}$ ; this would seem a more natural  
431 choice to keep in line with established practice.<sup>40</sup> We show  
432 later (Sec. V B) that covariance is a more robust estimator of  
433  $\mathbf{H}_2$  under certain conditions, thus motivating the correction  
434 term in Eq. (6). We also introduce the following notation for  
435 estimates obtained from target versus nontarget stimuli, re-  
436 spectively:

$$437 \quad \hat{\mathbf{H}}_1^{[1]} = \sum_z (2\delta[z-1] - 1) \hat{\mathbf{H}}_1^{[1,z]}, \quad (8)$$

$$438 \quad \hat{\mathbf{H}}_1^{[0]} = \sum_z (2\delta[z] - 1) \hat{\mathbf{H}}_1^{[0,z]}.$$

439 Clearly,  $\hat{\mathbf{H}}_1 = \hat{\mathbf{H}}_1^{[1]} + \hat{\mathbf{H}}_1^{[0]}$ . We define  $\hat{\mathbf{H}}_2^{[1]}$  and  $\hat{\mathbf{H}}_2^{[0]}$  similarly,  
440 i.e., using Eq. (6) restricted to  $q=1$  or  $q=0$  so that  $\hat{\mathbf{H}}_2$   
441  $= \hat{\mathbf{H}}_2^{[1]} + \hat{\mathbf{H}}_2^{[0]}$ .

442 The following expressions are immediately derived from  
443 those listed above:

$$444 \quad \hat{\mathbf{H}}_1^{[q,0]} = -\frac{\bar{p}}{1-\bar{p}} \hat{\mathbf{H}}_1^{[q,1]}, \quad (8)$$

$$445 \quad \hat{\mathbf{H}}_2^{[q,0]} = \frac{\mathbf{I}w}{1-\bar{p}} - \frac{\bar{p}}{1-\bar{p}} \hat{\mathbf{H}}_2^{[q,1]}, \quad (9)$$

$$446 \quad \hat{\mathbf{H}}_1^{[q,0]} \otimes \hat{\mathbf{H}}_1^{[q,0]} = -\frac{\bar{p}^2}{(1-\bar{p})^2} \hat{\mathbf{H}}_1^{[q,1]} \otimes \hat{\mathbf{H}}_1^{[q,1]}, \quad (9)$$

447 where  $\mathbf{I}$  is the identity matrix. From Eq. (8) we expect first-  
448 order subclass estimates from correct and incorrect trials to  
449 be scaled versions of each other, where the scaling factor is  
450 determined by the relative percentage of the two trial types.  
451 This result is known.<sup>3</sup> Equation (9) shows that a similar scal-  
452 ing factor applies to second-order subclass estimates, al-  
453 though with an additional diagonal term. From these equa-  
454 tions, we can readily derive the following useful expressions:

$$455 \quad \hat{\mathbf{H}}_1 = \frac{1}{1-\bar{p}} (\hat{\mathbf{H}}_1^{[1,1]} - \hat{\mathbf{H}}_1^{[0,1]}), \quad (10)$$

$$\sum_{q,z} (2\delta[q-z]-1)\hat{\mathbf{H}}_2^{[q,z]} = \frac{1}{1-\bar{p}}(\hat{\mathbf{H}}_2^{[1,1]} - \hat{\mathbf{H}}_2^{[0,1]}), \quad (11)$$

$$\sum_{q,z} (2\delta[q-z]-1)\hat{\mathbf{H}}_1^{[q,z]} \otimes \hat{\mathbf{H}}_1^{[q,z]} = \frac{1-2\bar{p}}{(1-\bar{p})^2}(\hat{\mathbf{H}}_1^{[1,1]} \otimes \hat{\mathbf{H}}_1^{[1,1]} - \hat{\mathbf{H}}_1^{[0,1]} \otimes \hat{\mathbf{H}}_1^{[0,1]}), \quad (12)$$

where Eq. (12) refers to the covariance correction term in Eq. (6). It will be relevant for further treatment of this term (Sec. V B) that the factor  $1-2\bar{p}/(1-\bar{p})^2$  is negative because  $\bar{p} > \frac{1}{2}$ .

#### D. Experimental cross-check

We need to check that the framework detailed above works well in real applications, because it forms the basis for all subsequent analytical treatments of behavioral kernel estimation presented here. To appreciate the importance of providing a cross-check, consider Eq. (5): this is different from what is actually done when computing  $\hat{\mathbf{H}}_1^{[q,1]}$  in the laboratory. Suppose we are computing  $\hat{\mathbf{H}}_1^{[1,1]}$ , in the laboratory, we collect say  $10k$  trials on a given observer, and select the subset of those trials on which the observer responded correctly ( $z=1$ ). We then select the subset of noise fields on which we presented the target ( $q=1$ ), and average only these noise fields. This is not a probabilistic procedure; if we were to write it down, it would look more like

$$\hat{\mathbf{H}}_1^{[q,z]} = \langle \mathbf{n}_{i_k}^{[q]} \rangle_{i_k}, i_k \in A_z, \quad (13)$$

where  $A_z$  is the indexing subset for trials corresponding to response  $z$  given a specific finite collection of trials out of all possible trials otherwise indexed by  $i$ . In Eq. (5) the average is taken over this infinite set of all possible trials, weighted by  $p_i$ . We cannot actually compute Eq. (5), not even approximately for a finite number of trials, because we do not have direct access to a description of  $p_i$  (i.e.,  $\Psi$ ). We therefore need to check that we are on solid grounds when abstracting from Eq. (13) to Eq. (5) to ensure that our analytical results can be transferred back to the laboratory.

We can attempt a cross-check of this kind by verifying the empirical applicability of Eqs. (8) and (9), which are directly derived from Eq. (5), to data sets from our laboratory (detailed in Refs. 50 and 51). These equations involve kernel estimates for incorrect ( $\hat{\mathbf{H}}_d^{[q,0]}$ ) and correct ( $\hat{\mathbf{H}}_d^{[q,1]}$ ) trials, which we can measure from data using Eq. (13), and the percentage of correct responses  $\bar{p}$ , which we can easily estimate for a given experiment. We write Eq. (8) as  $\hat{\mathbf{H}}_1^{[q,0]} = k\hat{\mathbf{H}}_1^{[q,1]}$ , and compute the scaling factor  $k$  either by linear regression between  $\hat{\mathbf{H}}_1^{[q,0]}$  and  $\hat{\mathbf{H}}_1^{[q,1]}$ , or from  $-\bar{p}/1-\bar{p}$  as prescribed by Eq. (8). If this equation is correct, we expect the two estimates to match. Figure 3(a) plots the scaling factor computed using method 1 on the abscissa versus method 2 on the ordinate. Figure 3(b) plots the same for  $\hat{\mathbf{H}}_2$  after accounting for the additional term  $\mathbf{I}\mathbf{w}/1-\bar{p}$  in Eq. (9). For both first-order [panel (a)] and second-order [panel (b)] kernel estimates, we have that the measured scaling factor on the abscissa falls between the expected value on the ordinate

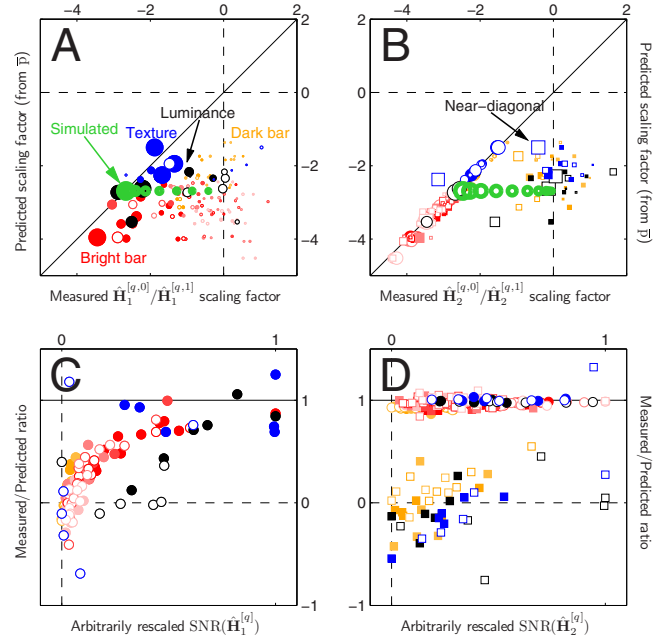


FIG. 3. (Color online) Experimental cross-check for Eqs. (8) and (9). Abscissa plots the linear regression coefficient (scaling factor) for kernel estimates from correct vs incorrect trials for first-order kernels in (a) and second-order kernels in (b). Ordinate plots the corresponding predictions computed from the percentage of correct responses (estimate for  $\bar{p}$ ) using Eqs. (8) and (9). Solid symbols refer to target estimates ( $q=1$ ), open symbols refer to nontarget ( $q=0$ ); for these experiments,  $\mathbf{t}^{[0]}=0$ . Black and blue symbols show data for the luminance and texture experiments detailed in Ref. 50; red and yellow symbols show data for the spatial uncertainty experiments involving bright and dark-bar detection detailed in Ref. 51. Color saturation for the latter data set reflects spatial uncertainty (more saturated, less uncertain). Symbol size scales with kernel SNR [Eq. (14)]. Green symbols refer to the average of 100 iterations of a simulated L(N) model; increasing symbol sizes (correlations) were obtained by increasing the number of simulated trials per iteration from 100 to 102 400 in logarithmic steps. In (b), circles refer to coefficients computed from the entire second-order kernel, squares only to near-diagonal region (see Sec. II D). (c) and (d) replot the data from (a) and (b) with SNR on the abscissa and the ratio between the two values in (a) and (b) on the ordinate.

(when points fall on the solid unity line) and 0 (when points fall on the vertical dashed line). Does the latter trend reflect a failure of Eqs. (8) and (9)? Not necessarily, because a value of 0 on the abscissa is expected for noisy kernel measurements (see green symbols showing simulated results where less trials were used as proxy for less reliable measurements). In particular, according to this interpretation we expect that scaling factors near 0 would correspond to lower kernel signal-to-noise ratios defined as

$$\text{SNR}(\hat{\mathbf{H}}_d) = \frac{dN\langle \hat{\mathbf{H}}_d^2, 1/m \rangle}{\mathbf{w}^d}, \quad (14)$$

where  $d$  can be 1 or 2,  $m$  is the full dimensionality of  $\hat{\mathbf{H}}_d$  (i.e., the inner product in the numerator is simply the mean square), and  $N$  is the number of noise fields that went into computing  $\hat{\mathbf{H}}_d$ . Equation (14) extends the standard definitions of SNR (e.g., Ref. 46) to second-order kernels. Symbol size scales with SNR in Figs. 3(a) and 3(b); particularly in the case of first-order kernels (a), it is clear that smaller symbols are closer to the vertical dashed line, while bigger ones are closer to the solid unity line (in conformity with the expected

AQ:  
#19

509  
510  
511  
512  
513  
514  
515  
516  
517

518

519  
520  
521  
522  
523AQ:  
524 #5  
525  
526  
527

528 tation detailed above). This trend can be exposed more  
529 readily by replotting the data as done in (c), where the ordi-  
530 nate plots the ratio between the two quantities in (a) and the  
531 abscissa plots the arbitrarily rescaled SNR. Equation (8) pre-  
532 dicted that points should fall on the horizontal solid line; panel  
533 (c) shows that they asymptote to this value as SNR grows  
534 [all data sets show a significant positive correlation  $p$   
535  $< 0.01$  with the exception of the texture data set (blue),  
536 which is only marginally significant at  $p=0.06$ ].

537 Results for second-order kernels [panels (b) and (d)] re-  
538 quire a slightly more involved analysis. When the entire  $\hat{\mathbf{H}}_2$   
539 kernel is considered (circles), Eq. (9) holds for all data sets  
540 and kernel subtypes [ $\hat{\mathbf{H}}_2^{[1]}$  (solid) and  $\hat{\mathbf{H}}_2^{[0]}$  (open)]. There is  
541 also no clear dependence on SNR ( $p > 0.05$  for all data sets).  
542 This analysis treats diagonal and nondiagonal regions of the  
543 kernel in the same way, but there are good reasons for in-  
544 specting them separately: (1) the diagonal region corre-  
545 sponds to differential variance, rather than covariance; (2)  
546 Eq. (9) contains a term that specifically affects the diagonal  
547 alone; (3) one of the simplest cascade models used in neuro-  
548 science applications, the Hammerstein model<sup>26</sup> (see Sec.  
549 IV B), predicts second-order modulations only within the di-  
550 agonal region.<sup>26,41,86</sup> These issues are highlighted by the  
551 square symbols in Fig. 3(b), which refer to the kernel region  
552 immediately adjacent to the diagonal (corresponding to 1's in  
553 the matrix  $\mathbf{M}(x_\nu, x_\xi) \triangleq \delta[\nu - \xi - 1]$ ). For this region, the mea-  
554 sured scaling factor is often close to 0 (several square sym-  
555 bols cluster around vertical dashed line in b) and there is a  
556 dependence on SNR similar to that observed for first-order  
557 kernels (square symbols in d), although only for dark-bar  
558 (yellow) and texture (blue) data sets ( $p < 0.005$ ). The bright-  
559 bar data set (red symbols) is particularly interesting, in that  
560 Eq. (9) holds closely for both diagonal and near-diagonal  
561 regions (all red symbols fall on solid lines in b and d) and  
562 there is no detectable dependence on SNR ( $p > 0.05$ ).

563 To summarize the results from Fig. 3, in general (i.e.,  
564 across the data sets presented here) Eqs. (8) and (9) apply  
565 well. Most departures are at least in part attributable to limi-  
566 tations imposed by measurement noise, which curtails our  
567 ability to verify the applicability of the predictions.

### 568 III. THE DECISIONAL BOTTLENECK AND ITS 569 IMPLICATIONS

#### 570 A. Benchmark result: Bypassing $\Psi$ for Wiener 571 systems

AQ572 #6 In this section, we show that for the simple LN cascade  
573 ( $\mathbf{H}_d=0$  for  $d > 1$  and  $N$  referring to  $\Psi$ ) we have  $\hat{\mathbf{H}}_1 \propto \mathbf{H}_1$  and  
574  $\hat{\mathbf{H}}_2=0$ . The former result is well-established (generally re-  
575 ferred to as Busgang theorem<sup>12</sup>), the latter is specific to the  
576 context explored here and differs from that obtained in stan-  
577 dard applications of Lee–Schetzen cross-correlation for  
578 which  $\hat{\mathbf{H}}_2 \propto \mathbf{H}_1 \otimes \mathbf{H}_1$ .<sup>41,86</sup> For the system of interest  $r_i - \bar{r}$   
579  $= \langle \mathbf{H}_1, \mathbf{n}_i^{[1]} - \mathbf{n}_i^{[0]} \rangle$  [from Eq. (4)]. We substitute it into Eq. (1)  
580 and insert the resulting  $\Psi$  approximation into Eq. (5),

$$\hat{\mathbf{H}}_1^{[1,1]}(x_\nu) = \frac{1}{\bar{p}} \left\langle \sum_{d=1}^{\infty} \Psi^{(d)} \left[ \sum_j \mathbf{H}_1(x_j) (\mathbf{n}_i^{[1]}(x_j) - \mathbf{n}_i^{[0]}(x_j)) \right]^d \right. \\ \left. \times \mathbf{n}_i^{[1]}(x_\nu) \right\rangle_i, \quad (15)$$

where the dot product term within  $\langle \rangle$  has been expanded into  
583 a sum for clarity. For each value of  $d$ , the term under  $\sum_d$   
584 consists of a sum of terms of the form  
585

$$\left[ \prod_{k=1}^a \mathbf{H}_1(x_{y_k}) \mathbf{n}_i^{[1]}(x_{y_k}) \prod_{k=1}^b -\mathbf{H}_1(x_{v_k}) \mathbf{n}_i^{[0]}(x_{v_k}) \right] \mathbf{n}_i^{[1]}(x_\nu), \quad (16)$$

where  $y$  and  $v$  are indexing sets. These terms mostly vanish,  
587 e.g., whenever  $a$  is even and/or  $b$  is odd. When they do not  
588 vanish, they can be written as  $k\mathbf{H}_1$ , where  $k > 0$  because  $b$   
589 must be even. Equation (15) can therefore be rewritten com-  
590 pactly as  
591

$$\hat{\mathbf{H}}_1^{[1,1]}(x_\nu) = \omega_1 \mathbf{H}_1(x_\nu), \quad (17)$$

where  $\omega_1$  does not depend on  $\nu$  (note that we are implicitly  
593 exploiting the circular symmetry of the input noise source  
594 here; more generally, this result depends on elliptical  
595 symmetry<sup>61</sup>).  $\hat{\mathbf{H}}_1^{[0,1]}$  follows a similar expression except  $\mathbf{n}_i^{[1]}$   
596 and  $-\mathbf{n}_i^{[0]}$  are swapped, leading to a sign inversion,  
597

$$\hat{\mathbf{H}}_1^{[0,1]}(x_\nu) = -\omega_1 \mathbf{H}_1(x_\nu). \quad (18)$$

We then have from Eq. (10) that  $\hat{\mathbf{H}}_1 \propto \mathbf{H}_1$ .  
599

Following a similar procedure for  $\hat{\mathbf{H}}_2$ , we have  
600

$$\hat{\mathbf{H}}_2^{[1,1]}(x_\nu, x_\xi) = \frac{1}{\bar{p}} \left\langle \sum_{d=1}^{\infty} \Psi^{(d)} \left[ \sum_j \mathbf{H}_1(x_j) (\mathbf{n}_i^{[1]}(x_j) - \mathbf{n}_i^{[0]}(x_j)) \right]^d \right. \\ \left. \times \mathbf{n}_i^{[1]}(x_\nu) \mathbf{n}_i^{[1]}(x_\xi) \right\rangle_i, \quad (19)$$

where the terms summed under  $\sum_d$  are of the form  
603

$$\left[ \prod_{k=1}^a \mathbf{H}_1(x_{y_k}) \mathbf{n}_i^{[1]}(x_{y_k}) \prod_{k=1}^b -\mathbf{H}_1(x_{v_k}) \mathbf{n}_i^{[0]}(x_{v_k}) \right] \mathbf{n}_i^{[1]}(x_\nu) \mathbf{n}_i^{[1]}(x_\xi). \quad (20)$$

Using the same logic adopted for  $\hat{\mathbf{H}}_1^{[1,1]}$ , we obtain  
605

$$\hat{\mathbf{H}}_2^{[1,1]} = \hat{\mathbf{H}}_2^{[0,1]} = \omega_2 \mathbf{H}_1 \otimes \mathbf{H}_1, \quad (21)$$

where there is no sign inversion between  $\hat{\mathbf{H}}_2^{[1,1]}$  and  $\hat{\mathbf{H}}_2^{[0,1]}$   
607 because both  $a$  and  $b$  must be even for terms not to vanish in  
608 Eq. (16). Substituting into Eqs. (11) and (12) and then Eq. (6)  
609 leads to  $\hat{\mathbf{H}}_2=0$ . This result applies for any polynomial expansion  
610 of  $\Psi$  and here it is referred to as the “benchmark” result.  
611

Later in this article, we demonstrate that kernel subclass  
612 estimates show strong target-related modulations, a result  
613 that has been experimentally observed on numerous occasions  
614 for first-order kernels.<sup>1,4,50,52,75,82</sup> This observation  
615 leads one to question whether it may not be advisable to rely  
616 exclusively on target-absent estimates  $\hat{\mathbf{H}}_1^{[0]}$  and  $\hat{\mathbf{H}}_1^{[0]}$ . This  
617 option may be available (depending on whether  $\mathbf{t}^{[0]}=0$  and  
618

619 the amount of data is sufficient to afford curtailing it) for  $\hat{\mathbf{H}}_1$   
 620 because it does not violate the benchmark result; however, it  
 621 is not viable for  $\hat{\mathbf{H}}_2$  because in this case  $\hat{\mathbf{H}}_2^{[0]} \neq 0$ , even  
 622 though  $\mathbf{H}_2=0$  (thus violating the benchmark result), as ap-  
 623 parent from the above demonstration [Eq. (17)] that  $\hat{\mathbf{H}}_2=0$   
 624 results from  $\hat{\mathbf{H}}_2^{[1,1]}$  and  $\hat{\mathbf{H}}_2^{[0,1]}$  canceling each other out in Eq.  
 625 (11). If  $\hat{\mathbf{H}}_2$  is computed by omitting either one (as in  $\hat{\mathbf{H}}_2^{[0]}$ ),  
 626 the benchmark result does not hold any longer.  $\hat{\mathbf{H}}_2^{[0]}$  is there-  
 627 fore inadequate for the purposes examined here because it is  
 628 largely affected by the decisional nonlinearity which (as dis-  
 629 cussed in Sec. I B) we are trying to circumvent in order to  
 630 characterize  $\mathcal{H}$ .

### 631 B. Handling peridecisional operators (those applied 632 right before choice)

#### 633 1. Assimilation of nonlinearities into $\Psi$

634 Given the framework  $p_i = \Psi(\mathcal{H}(s_i^{[1]}) - \mathcal{H}(s_i^{[0]}))$  (Sec. II A)  
 635 and associated benchmark result (Sec. III A), there are con-  
 636 ditions under which  $\mathcal{H}$  can be augmented by applying an  
 637 operator  $\Phi$  to its output  $r_i$  without altering our analysis in  
 638 any significant way. The theoretical interest in this question  
 639 stems from a number of considerations (see below), but  
 640 chiefly from the possibility that peridecisional (i.e., immedi-  
 641 ately preceding the psychophysical choice) static nonlinearities  
 642 may be incorporated into  $\Psi$  so as to extend the applica-  
 643 bility of the benchmark result to  $\mathcal{H}$  systems that are not  
 644 strictly linear. To provide an example of how this analysis  
 645 may be applied, we start with a noiseless system, i.e.,  $\Psi(x)$   
 646  $= u(x)$  (deterministic choice transducer—see Sec. II A), and  
 647 the application of a plausible static nonlinearity  $\Phi$ ; by plau-  
 648 sible we mean that it is strictly monotonic and continuous  
 649 within the operating range of  $r$  (see also Ref. 14). We can  
 650 then easily incorporate  $\Phi$  into  $\Psi$ :  $\Psi(\Phi(\mathcal{H}(s_i^{[1]}))$   
 651  $- \Phi(\mathcal{H}(s_i^{[0]}))) = \Psi(\mathcal{H}(s_i^{[1]}) - \mathcal{H}(s_i^{[0]}))$ . It is easy to see why this  
 652 simple result, also known as Birdsall's theorem,<sup>33,80</sup> applies  
 653 in the noiseless case:  $\Phi(y) - \Phi(x)$  preserves the sign of  $y - x$ ,  
 654 which is the only determinant of choice under a noiseless  $\Psi$ .  
 655 We then have that the benchmark result applies even though  
 656 the system was not written as linear- $\Psi$  because its output is  
 657 identical to the linear- $\Psi$  version.

658 The extension to the noisy case is not trivial, unless  $\Phi$  is  
 659 applied after the internal noise source in which case Bird-  
 660 sall's theorem still applies. If applied before, our goal is  
 661  $\Psi(\Phi(\mathcal{H}(s_i^{[1]})) - \Phi(\mathcal{H}(s_i^{[0]}))) = \tilde{\Psi}(\mathcal{H}(s_i^{[1]}) - \mathcal{H}(s_i^{[0]}))$ , where  $\tilde{\Psi}$   
 662 hopefully retains the same overall characteristics as  $\Psi$   
 663 [ $\tilde{\Psi}(0) = 0.5$ ,  $\lim_{x \rightarrow -\infty} \tilde{\Psi}(x) = 0$ ,  $\lim_{x \rightarrow \infty} \tilde{\Psi}(x) = 1$ , monotoni-  
 664 cally increasing, see Sec. II A]. This is rarely the case;  
 665 consider for example  $\Phi(x) = e^x$ . The map  $(r^{[1]} - r^{[0]})$   
 666  $\rightarrow (e^{r^{[1]}} - e^{r^{[0]}})$  is not injective, so it is not possible to specify  
 667 a unique  $\tilde{\Psi}$  on  $r^{[1]} - r^{[0]}$ ; this, in turn, means that the bench-  
 668 mark result is not applicable and  $\hat{\mathbf{H}}_2 \neq 0$  for sizeable internal  
 669 noise.

#### 670 2. A notable example: Response bias

671 The critical requirement specified in Sec. III B 1 is that a  
 672 unique  $\tilde{\Psi}$  can be identified. We now consider its implications  
 673 for what is perhaps the most significant source of artifactual

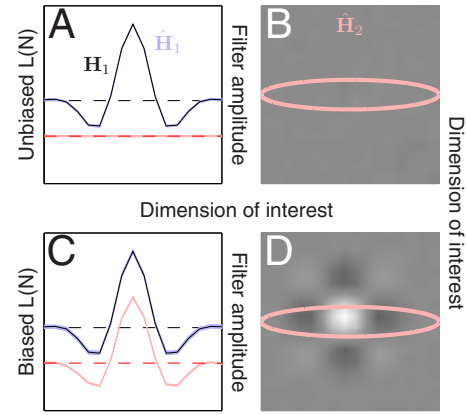


FIG. 4. (Color online) Response bias violates the benchmark result. The latter is demonstrated in (a) and (b): for a simulated system output  $\langle \mathbf{H}_1, s \rangle$  subjected to the decisional transducer  $\Psi$  [modeled using the unit step function  $u(x)$ ], we obtain  $\hat{\mathbf{H}}_1 \propto \mathbf{H}_1$  (compare blue shading, which shows  $\pm 1$  standard deviation on  $\hat{\mathbf{H}}_1$  for 100 simulations of 100k trials each, to black trace for  $\mathbf{H}_1$ ) and  $\hat{\mathbf{H}}_2 = 0$  [panel (b)]. Black/blue traces in (a) are plotted to the zero line indicated by the black dashed horizontal line, while light-red shading [corresponding to  $\hat{\mathbf{H}}_2$  slice indicated by oval in (b)] is plotted to the zero line indicated by the red dashed horizontal line. In the presence of response bias [panels (c) and (d)], modeled as the application of a static nonlinearity ( $e^x$ ) to the output from one of the two stimuli (randomly choosing between  $s^{[0]}$  and  $s^{[1]}$  on each trial), the benchmark result no longer holds for  $\hat{\mathbf{H}}_2$ , which is now  $\neq 0$  (d) and more specifically  $\propto \mathbf{H}_1 \otimes \mathbf{H}_1$  as expected from theoretical considerations (see Sec. III B 2). Filter amplitude has been arbitrarily rescaled for each curve/surface, but using the same scaling factor for a given curve/surface between the top and bottom rows, so as to allow direct comparison between biased and unbiased kernels.

AQ:  
#20

674 results in human psychophysics: response bias (see Ref. 69  
 675 for a recent example). Response bias can introduce nonlinear  
 676 distortions, potentially leading to  $\hat{\mathbf{H}}_2 \neq 0$  even though  $\mathbf{H}_2$   
 677  $= 0$ . We can model response bias as the application of a non-  
 678 linearity to the output from one of the two intervals before  
 679 both outputs are combined and submitted to  $\Psi$  for choice  
 680 behavior. This procedure warps the  $L(N)$  layer into a mixed  
 681  $L\tilde{N}(N)$  plus  $L(N)$  cascade, where  $\tilde{N}$  cannot be assimilated  
 682 into  $(N)$  using the result in Sec. III B 1 (see below). For this  
 683 combined cascade (which stands for  $\mathcal{H}$ +bias rather than  $\mathcal{H}$   
 684 alone), we know that the effective  $\mathbf{H}_2 \neq 0$  and therefore  $\hat{\mathbf{H}}_2$   
 685  $\neq 0$ .

686 To demonstrate this result, it is sufficient to rewrite the  
 687 decision variable as  $r_{i_1} = \Phi(r_{i_1}^{[1]}) - r_{i_1}^{[0]}$  on half the trials  $i_1$   
 688 (when  $t^{[1]}$  is presented in the distorted interval) and  $r_{i_2} = r_{i_2}^{[1]}$   
 689  $- \Phi(r_{i_2}^{[0]})$  on the remaining half. We cannot assimilate  $\Phi$  into  
 690  $\Psi$  because the map  $(r^{[1]} - r^{[0]}) \rightarrow (\Phi(r^{[1]}) - r^{[0]})$  is not injec-  
 691 tive, i.e., it is not possible (in general) to rewrite the process  
 692 using only one  $\tilde{\Psi}$  (Sec. III B 1). If we then compute  $\hat{\mathbf{H}}_2$  from  
 693 the  $i_1$  trials, it will expose the  $\mathbf{H}_2$  kernel from the Wiener  
 694 system  $\Phi(r)$ , which is of the form  $\mathbf{H}_1 \otimes \mathbf{H}_1$  [see Ref. 41 and  
 695 Eq. (17)]. The same structure is exposed when computing  
 696 from the  $i_2$  trials, thus biasing  $\hat{\mathbf{H}}_2$ . This result is easily dem-  
 697 onstrated via simulation, as shown in Fig. 4; notice the clear  
 698  $\mathbf{H}_1 \otimes \mathbf{H}_1$  structure for  $\hat{\mathbf{H}}_2$  [panel (d)].

699 It is important to emphasize that the result detailed in the  
 700 previous paragraph applies regardless of the presence/

701 absence of internal noise [i.e., even if  $\Psi(x)=u(x)$ ], because  
 702 the map  $(r^{[1]}-r^{[0]}) \rightarrow (\Phi(r^{[1]})-r^{[0]})$  will in general not be  
 703 homomorphic with respect to the order relation, i.e.,  $y > x$   
 704 does not necessarily mean that  $y > \Phi(x)$  [ $\Phi(x)=e^x$  provides a  
 705 useful example], leading to a different choice from  $r^{[1]}-r^{[0]}$   
 706 depending on the specific values of  $r^{[1]}$  and  $r^{[0]}$ . In this sense,  
 707 response bias implements a more pervasive distortion than  
 708 the symmetrically applied nonlinearity discussed in Sec.  
 709 III B 1. The simulated results in Fig. 4 refer to a noiseless  
 710 system.

711 Response bias is a well-known issue in psychophysical  
 712 experiments where only one stimulus is presented on every  
 713 trial (either  $s^{[0]}$  or  $s^{[1]}$  randomly), and the observer is re-  
 714 quired to report on which one he/she perceived.<sup>20</sup> This class  
 715 of experimental protocols, known as yes-no, should be  
 716 avoided whenever possible. AFC protocols, when imple-  
 717 mented effectively, eliminate bias. This goal is achieved well  
 718 in the case of spatial AFC designs, where, for example,  $s^{[0]}$   
 719 or  $s^{[1]}$  appear simultaneously on opposite sides of the screen  
 720 and the observer is asked to indicate which side contains  $s^{[1]}$ .  
 721 There are situations, however, when spatial AFC designs are  
 722 not viable (e.g., auditory experiments); in this case, the  
 723 choice is between first and second intervals, for which some  
 724 bias may be expected.<sup>89</sup> A powerful method for reducing bias  
 725 in these instances is to provide the participant with trial-by-  
 726 trial feedback (correct/incorrect), which tends to make the  
 727 observer/listener converge toward a near-optimal (bias-free)  
 728 strategy.

### 729 3. Formulation of end linearity as inner product

730 When the last linear stage in  $\mathcal{H}$  is linear, it can be written  
 731 as an inner product regardless of how it is formulated.<sup>50</sup> This  
 732 result is a direct consequence of the decision variable as-  
 733 sumption (Sec. II A):  $\mathcal{H}$  must return a single scalar value as  
 734 argument to  $\Psi$ , whatever the dimensionality of the vector  
 735 returned by the previous stage. If we call this vector  $\mathbf{r}$ , and  
 736 we map it to scalar  $r$  via a linear transformation, we can  
 737 always write  $\mathbf{r} \rightarrow r$  as  $\langle \mathbf{g}, \mathbf{r} \rangle$  for some appropriate choice of  $\mathbf{g}$ .  
 738 Any preceding linear stage is similarly assimilated into a dot  
 739 product, which allows us to rewrite the Volterra expansion in  
 740 Eq. (4) using  $\langle \cdot, \cdot \rangle$  rather than  $*$  (as originally formulated by  
 741 Volterra<sup>41,68</sup>). To see this explicitly for a second-order Volt-  
 742 erra operator, first we can write its output using  
 743 convolution,<sup>40</sup>

$$744 \quad \mathbf{r} = \check{\mathbf{H}}_1 * \mathbf{s} + \check{\mathbf{H}}_2 * (\mathbf{s} \otimes \mathbf{s}). \quad (18)$$

745 We must now apply  $\mathbf{r} \rightarrow r$  so that we can submit the output to  
 746  $\Psi$  for choice behavior (see above). When we express  $\mathcal{H}$  this  
 747 way, we can always rewrite its output either using Eq. (18) or  
 748 equivalently Eq. (4),

$$749 \quad \langle \mathbf{g}, \check{\mathbf{H}}_1 * \mathbf{s} + \check{\mathbf{H}}_2 * (\mathbf{s} \otimes \mathbf{s}) \rangle = \langle \mathbf{H}_1, \mathbf{s} \rangle + \langle \mathbf{H}_2, (\mathbf{s} \otimes \mathbf{s}) \rangle, \quad (19)$$

750 where  $\mathbf{H}_d = \check{\mathbf{H}}_d * \mathbf{g}$  [for  $d=2$  this is  $\mathbf{H}_2(x_\nu, x_\xi)$   
 751  $= \sum_j \check{\mathbf{H}}_2(x_{j-\nu}, x_{j-\xi}) \mathbf{g}(x_j)$ ] and  $*$  has been replaced by  $\langle \cdot \rangle$ . Be-  
 752 cause in general we have no direct access to  $\mathbf{g}$ , the two for-  
 753 mulations are indistinguishable; our access is restricted to  $\mathbf{H}_d$   
 754 in the form of  $\hat{\mathbf{H}}_d$  (with associated limitations, e.g.,  
 755 Sec. IV A).

## 756 C. The -(N) notation

757 Taken together, the two simple results detailed above  
 758 (Secs. III A and III B 1) allow us to bypass the decisional  
 759 transducer  $\Psi$  and access  $\mathcal{H}$  using established tools from non-  
 760 linear systems analysis. First we focus on the most basic  
 761 system: linear ( $\mathbf{H}_d=0$  for  $d>1$ ). If we were reading off  $\mathcal{H}$   
 762 directly, we would expect  $\hat{\mathbf{H}}_2=0$  (as well as  $\hat{\mathbf{H}}_1 \propto \mathbf{H}_1$ ). Due to  
 763 the presence of  $\Psi$ , we are effectively monitoring the output  
 764 of a Wiener linear-nonlinear cascade where the static nonlin-  
 765 earity is introduced by  $\Psi$ . As mentioned earlier, the standard  
 766 applications of nonlinear kernel analysis return  $\hat{\mathbf{H}}_2 \propto \mathbf{H}_1$   
 767  $\otimes \mathbf{H}_1$  for this cascade system.<sup>41</sup> Our goal is to develop an  
 768 estimation strategy for  $\mathbf{H}_2$  that bypasses  $\Psi$ , i.e., one in which  
 769 we can treat the linear+ $\Psi$  system as simply linear and obtain  
 770  $\hat{\mathbf{H}}_2=0$ . This benchmark result is achieved by Eq. (6) (see  
 771 Sec. III A). Later in the article (Sec. V), we show that Eq. (6)  
 772 also approximates  $\mathbf{H}_2$  when  $\neq 0$ , although distortions may be  
 773 present. From Sec. III B 1 we also have that under restricted  
 774 conditions, the application of any number of plausible  $\Phi$  late  
 775 transducers does not impact our discussion, thus extending  
 776 the generality of the benchmark result. We indicate the by-  
 777 passing of  $\Psi$  by bracketing the end-nonlinearity in cascade  
 778 model formulations of  $\mathcal{H}$ : for example, if we assume a Ko-  
 779 renberg LNL model (a cascade consisting of a linear filter, a  
 780 static nonlinearity, and an additional linear filter, see Sec.  
 781 IV C) for  $\mathcal{H}$ , we indicate this model as LNL(N) where the  
 782 end decisional nonlinearity is bracketed to indicate that it  
 783 does not affect our treatment of the subject. This approach  
 784 allows us to apply the known results for LNL cascade  
 785 models<sup>40,41,86</sup> directly to  $\mathcal{H}$  (Ref. 50) (see, for example,  
 786 Sec. VI D).

## 787 IV. FIRST-ORDER KERNEL ESTIMATE

### 788 A. Nonlinear distortions

789 For  $\mathcal{H}$  linear, we know that  $\hat{\mathbf{H}}_1 \propto \mathbf{H}_1$  (Sec. III A). For a  
 790 nonlinear system, the main distortion on  $\hat{\mathbf{H}}_1$  is introduced by  
 791 the presence of the target signal, regardless of  $\Psi$ ; we can  
 792 therefore demonstrate this result analytically using a simple  
 793 linear approximation to  $\Psi$ . We also ignore kernels of order  
 794 higher than 3 ( $\mathbf{H}_d=0$  for  $d>3$ ). Given these approximations,  
 795 we have

$$796 \quad \hat{\mathbf{H}}_1^{[q,1]}(x_\nu) = \frac{\mathbf{w}}{\bar{p}} \Psi^{(1)}(2q-1) [\mathbf{H}_1 + 2\langle \mathbf{H}_2(x_\nu, \cdot), \mathbf{t}^{[q]} \rangle$$

$$797 \quad + 3\mathbf{w} \langle \mathbf{H}_3(x_\nu, \cdot, \cdot), \mathbf{I} \rangle + 3\langle \mathbf{H}_3(x_\nu, \cdot, \cdot), \mathbf{t}^{[q]} \rangle$$

$$798 \quad \otimes \mathbf{t}^{[q]}]. \quad (20)$$

799 Readers familiar with Wiener orthogonalization will recog-  
 800 nize that the term  $3\mathbf{w} \langle \mathbf{H}_3(x_\nu, \cdot, \cdot), \mathbf{I} \rangle$  can be avoided by rede-  
 801 fining  $\hat{\mathbf{H}}_1$  as a Wiener (rather than Volterra) kernel estimate.<sup>40</sup>  
 802 This distinction is not interesting in the present context  
 803 mainly for two reasons. First, in most applications we as-  
 804 sume  $\mathbf{H}_d=0$  for  $d>2$ ,<sup>47,50</sup> in this case, there is no distinction  
 805 between Volterra and Wiener first- and second-order  
 806 kernels.<sup>40</sup> Second, even if we approximate to third-order as  
 807 in Eq. (20), only the first-order kernel differs between

808 Volterra and Wiener representations. The difference, repre-  
 809 sented by the term highlighted above, is negligible compared  
 810 to the distortions introduced by terms containing  $\mathbf{t}$ . The dis-  
 811 tinction between target-present and target-absent kernels is  
 812 far more relevant to  $\hat{\mathbf{H}}_1$ , so here we focus on this aspect of  
 813 the estimation procedure rather than adopting a Wiener ker-  
 814 nel representation. This representation is equivalent to the  
 815 Volterra representation<sup>40,86</sup> and amounts to little more than a  
 816 change in notation for the purposes examined here.  
 817 Further distortions can be demonstrated by approximat-  
 818 ing  $\Psi$  to second-order (we set  $\mathbf{H}_3=0$  to simplify calcula-  
 819 tions),

$$820 \quad \hat{\mathbf{H}}_1^{[q,1]}(x_\nu) = \frac{\mathbf{w}}{\bar{p}} \Psi^{(1)}(2q-1)[\mathbf{H}_1(x_\nu) + 2\langle \mathbf{H}_2(x_\nu, :), \mathbf{t}^{[q]} \rangle]$$

$$821 \quad + \frac{4\mathbf{w}^2}{\bar{p}} \Psi^{(2)}(2\langle \mathbf{H}_2, \mathbf{H}_2(x_\nu, :) \otimes \mathbf{t}^{[q]} \rangle$$

$$822 \quad + \langle \mathbf{H}_2(x_\nu, :), \mathbf{H}_1 \rangle). \quad (21)$$

823 The above expression also demonstrates that even if we rely  
 824 on target-absent estimates alone for an experiment where  
 825  $\mathbf{t}^{[0]}=0$  (e.g., Ref. 50), we obtain

$$826 \quad \hat{\mathbf{H}}_1^{[0,1]}(x_\nu) = -\frac{\mathbf{w}}{\bar{p}} \Psi^{(1)} \mathbf{H}_1(x_\nu) + \frac{4\mathbf{w}^2}{\bar{p}} \Psi^{(2)} \langle \mathbf{H}_2(x_\nu, :), \mathbf{H}_1 \rangle,$$

827 exposing a residual departure from  $\hat{\mathbf{H}}_1^{[0]} \propto \mathbf{H}_1$ .

### 828 B. Special case: Hammerstein NL(N) cascade

829 The output of a Hammerstein cascade is defined as  
 830  $\mathbf{f} * \Phi(\mathbf{s})$ :<sup>26</sup> the underlying block-model consists of a front-end  
 831 static nonlinearity followed by a linear filter (dynamic linear-  
 832 ity). As explained in Sec. III B 1, the  $*$  operation can be  
 833 replaced by an inner product because  $\mathcal{H}$  must return a scalar  
 834 decision variable to obtain  $\langle \mathbf{f}, \Phi(\mathbf{s}) \rangle$ . For this cascade, we  
 835 have  $\mathbf{H}_2(x_\nu, x_\xi) = \delta[\nu - \xi] (\Phi^{(2)}/\Phi^{(1)}) \mathbf{H}_1(x_\nu)$  and  $\mathbf{H}_3(x_\nu, x_\xi, x_\zeta)$   
 836  $= \delta[\nu - \xi] \delta[\nu - \zeta] (\Phi^{(3)}/\Phi^{(1)}) \mathbf{H}_1(x_\nu)$ , where  $\Phi^{(d)}$  is the  $d$ th co-  
 837 efficient in the Taylor expansion of  $\Phi$  (see, for example, Ref.  
 838 86). We consider the case  $\mathbf{t}^{[1]} = \rho \delta[\tau]$  and  $\mathbf{t}^{[0]}=0$ . Equation  
 839 (20) then reduces to

$$840 \quad \hat{\mathbf{H}}_1^{[1,1]}(x_\nu) = \frac{\mathbf{w}}{\bar{p}} \Psi^{(1)} \left( \mathbf{H}_1(x_\nu) (1 + 3\mathbf{w}\Phi^{(3)}) \right.$$

$$841 \quad \left. + \frac{\rho}{\Phi^{(1)}} \mathbf{H}_1(x_\nu) \delta[\nu - \tau_0] (\Phi^{(2)} + 3\rho\Phi^{(3)}) \right),$$

$$842 \quad \hat{\mathbf{H}}_1^{[1,1]}(x_\nu) = (\omega_3 + \rho\omega_4 \delta[\nu - \tau_0]) \mathbf{H}_1(x_\nu), \quad (23)$$

843 where  $\omega_3$  and  $\omega_4$  do not depend on  $\nu$ . Equation (23) shows  
 844 that the  $\hat{\mathbf{H}}_1^{[1,1]}$  estimate is a signal-distorted image of  $\mathbf{H}_1$ .  
 845 From Eq. (20) we also have  $\hat{\mathbf{H}}_1^{[0,1]} \propto \mathbf{H}_1$ .

### 846 C. Special case: Korenberg LNL(N) cascade

847 The output of a Korenberg (sometimes termed  
 848 “sandwich”) cascade is defined as  $\mathbf{g} * \Phi(\mathbf{f} * \mathbf{s})$ :<sup>31</sup> the underly-  
 849 ing block-model consists of a front-end linear filter, a static

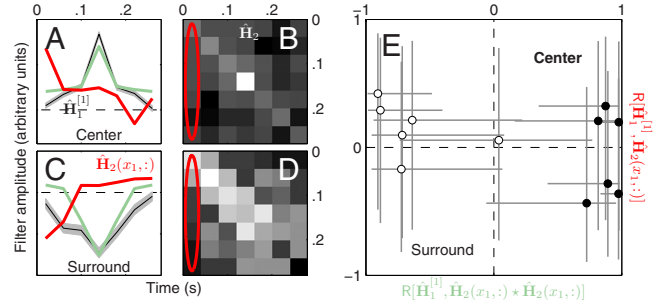


FIG. 5. (Color online) Target-present first-order estimates ( $\hat{\mathbf{H}}_1^{[1]}$ ) for Korenberg systems [black traces in panels (a) and (c)] do not resemble the underlying front-end filter  $\mathbf{f}$  but  $\mathbf{f} * \mathbf{f}$ . Data set from Ref. 50; observers were required to perform an orientation discrimination task for a stimulus divided into center and surround. (b) and (d) show the corresponding second-order kernel estimates. If we assume a Korenberg cascade  $\langle \mathbf{g}, \Phi(\mathbf{f} * \mathbf{s}) \rangle$  (see Sec. IV C),  $\mathbf{f}$  can be estimated via  $\hat{\mathbf{H}}_2(x_1, :)$  indicated by red oval and replotted in (a) and (c). Black traces show target-present first-order kernel estimates ( $\hat{\mathbf{H}}_1^{[1]}$ ), green traces show the autocorrelation of the red traces. (e) plots the correlation coefficient between red and black traces on the ordinate vs the correlation coefficient between green and black traces on the abscissa for each subject separately (different data points) and for both center (solid) and surround (open). Error bars show  $\pm 1$  SEM. Traces in (a) and (c) have been independently and arbitrarily rescaled, including sign inversion for red and green in (e).

nonlinearity, and an additional linear stage (the nonlinearity is therefore “sandwiched” in between the two linear stages). We replace the end  $*$  with inner product  $\mathcal{H}$  (Sec. III B 1), and for  $\langle \mathbf{g}, \Phi(\mathbf{f} * \mathbf{s}) \rangle$  we have

$$850 \quad \mathbf{H}_d(x_{\nu_1}, \dots, x_{\nu_d}) = \Phi^{(d)} \sum_k \mathbf{g}(x_k) \mathbf{f}(x_k - x_{\nu_1}) \cdots \mathbf{f}(x_k - x_{\nu_d}). \quad (24) \quad 854$$

We consider the special case  $\mathbf{g}=1$  and  $\mathbf{t}^{[1]} = \rho \delta$  ( $\mathbf{t}^{[0]}=0$ ),

$$855 \quad \hat{\mathbf{H}}_1^{[1,1]} = \frac{\mathbf{w}}{\bar{p}} \Psi^{(1)} (\Phi^{(1)} \langle \mathbf{f}, 1 \rangle + 2\rho \Phi^{(2)} \mathbf{f} * \mathbf{f} + 3\Phi^{(3)} \times (\rho^2 \mathbf{f} * \mathbf{f}^2 + \mathbf{w} \langle \mathbf{f} * \mathbf{f}^2, 1 \rangle). \quad (25) \quad 856 \quad 857$$

The inner products do not depend on  $\nu$ , so we can rewrite this compactly as  $\mathbf{f} * \mathbf{f} + \rho \omega_5 \mathbf{f} * \mathbf{f}^2$ . At low SNR (small  $\rho$ ), we therefore expect  $\hat{\mathbf{H}}_1^{[1]} \propto \mathbf{f} * \mathbf{f}$  (see also Ref. 51).

We can attempt to verify whether this analytical result is born out by data for measurements of human texture processing; we previously established that the relevant data set is consistent with a Korenberg cascade<sup>50</sup> (see Sec. VID for details). If we assume that this was the underlying structure in the human observers, we can estimate  $\mathbf{f}$  via  $\hat{\mathbf{H}}_2(x_1, :)$  (see Refs. 41 and 86 for details relating to this well-established result) shown in red in Fig. 5(a);  $\hat{\mathbf{H}}_1^{[1]}$  is shown in black. We expect that the autocorrelation of the red trace, shown in green, is proportional to the black trace; this prediction seems approximately correct for both stimulus center [panel (a)] and surround [panel (c)]. To quantify the relationship across observers, panel (e) plots the correlation coefficients for each observer separately; the abscissa plots the correlation between  $\hat{\mathbf{H}}_1^{[1]}$  and the green trace, while the ordinate plots the correlation between  $\hat{\mathbf{H}}_1^{[1]}$  and the red trace. If  $\hat{\mathbf{H}}_1^{[1]}$  returned an estimate of the underlying impulse response  $\mathbf{f}$ ,

878 we would expect symbols to fall either below or above the  
 879 horizontal dashed line (0 correlation on ordinate); instead,  
 880 the ordinate values are no different from 0 (t-test across ob-  
 881 servers,  $p=0.68$  for center and  $p=0.13$  for surround). From  
 882 Eq. (25) we expect  $\hat{\mathbf{H}}_1^{[1]}$  to return an estimate of the autocor-  
 883 relation of  $\mathbf{f}$ , not  $\mathbf{f}$  itself; indeed the abscissa values are sig-  
 884 nificantly different from 0 (points lie away from vertical  
 885 dashed line at  $p < 10^{-5}$  for center and  $p < 0.01$  for surround).  
 886 This result provides indirect confirmation of the applicability  
 887 of Eq. (25) to real data.

## 888 V. SECOND-ORDER KERNEL ESTIMATE

### 889 A. Basic expressions

890 By combining Eqs. (1), (4), and (7) and adopting the  
 891 same procedure used for  $\hat{\mathbf{H}}_1$ , we have the corresponding ex-  
 892 pressions for  $\hat{\mathbf{H}}_2$ ,

$$893 \quad \hat{\mathbf{H}}_2^{[q,1]} = \frac{\mathbf{w}\mathbf{I}}{\bar{p}}\Psi^{(0)} + \frac{2\mathbf{w}^2}{\bar{p}}\Psi^{(1)}(2q-1) \\ 894 \quad \times \left( \mathbf{H}_2 + 3 \sum_j \mathbf{H}_3(x_j, :, :) \mathbf{t}^{[q]}(x_j) \right) \quad (26)$$

895 for a third-order approximation of  $\mathcal{H}$  and first-order approxi-  
 896 mation of  $\Psi$ , and

$$897 \quad \hat{\mathbf{H}}_2^{[q,1]} = \frac{\mathbf{w}\mathbf{I}}{\bar{p}}\Psi^{(0)} + \frac{2\mathbf{w}^2}{\bar{p}}\Psi^{(1)}(2q-1)\mathbf{H}_2 \\ 898 \quad + \frac{2\mathbf{w}^2}{\bar{p}}\Psi^{(2)}(4\mathbf{A}^{[q]} + \mathbf{B} + \mathbf{C}) \quad (27)$$

899 for a second-order approximation of  $\mathcal{H}$  and second-order ap-  
 900 proximation of  $\Psi$ , where

$$901 \quad \mathbf{A}^{[q]} \triangleq \mathbf{H}_1 \otimes \sum_j \mathbf{H}_2(x_j, :) \mathbf{t}^{[q]}(x_j) + \sum_{j,k} (\mathbf{H}_2(x_j, :) \\ 902 \quad \otimes \mathbf{H}_2(x_k, :)) \mathbf{t}^{[q]}(x_j) \mathbf{t}^{[q]}(x_k),$$

$$903 \quad \mathbf{B} = \mathbf{H}_1 \otimes \mathbf{H}_1 + 2\mathbf{w} \sum_j (\mathbf{H}_2(:, j) \otimes \mathbf{H}_2(:, j)),$$

$$904 \quad \mathbf{C} = \mathbf{I} \left[ \langle \mathbf{H}_1^2, 1 \rangle + 2 \sum_j \mathbf{H}_1(x_j) \langle \mathbf{H}_2(j, :), \mathbf{t}^{[1]} + \mathbf{t}^{[0]} \rangle + 2\mathbf{w} \langle \mathbf{H}_2^2, 1 \rangle \right. \\ 905 \quad \left. + \sum_j 2 \langle (\mathbf{H}_2(:, x_j) \otimes \mathbf{H}_2(:, x_j)), \mathbf{t}^{[1]} \otimes \mathbf{t}^{[1]} + \mathbf{t}^{[0]} \otimes \mathbf{t}^{[0]} \rangle \right]. \quad (28)$$

906 These three terms have been kept separate to emphasize that  
 907  $\mathbf{A}^{[q]}$  depends on  $q$  (i.e., it differs for target and nontarget  
 908 subestimates), whereas  $\mathbf{B}$  and  $\mathbf{C}$  do not [i.e., they cancel out  
 909 in Eq. (11)], and that  $\mathbf{C} \propto \mathbf{I}$  (the quantity within  $\square$  is a scalar).

### 910 B. Why covariance and not second-order 911 moment?

912 If we use a second-order moment estimate  $\tilde{\mathbf{H}}_2$  for  $\mathbf{H}_2$ , we  
 913 have from Eq. (27),

$$\tilde{\mathbf{H}}_2 = \sum_{q,z} (2\delta[q-z] - 1) \hat{\mathbf{H}}_2^{[q,z]} \quad 914 \\ = \frac{1}{1-\bar{p}} \left( \frac{4\mathbf{w}^2}{\bar{p}} \Psi^{(1)} \mathbf{H}_2 + \frac{8\mathbf{w}^2}{\bar{p}} \Psi^{(2)} \mathbf{A} \right), \quad 915$$

where  $\mathbf{A} = \mathbf{A}^{[1]} - \mathbf{A}^{[0]}$  is the distorting term, i.e., causing de- 916  
 917 parture from the desired relationship  $\hat{\mathbf{H}}_2 \propto \mathbf{H}_2$ . We rewrite  
 918 this term below in a slightly different format from Eq. (28) to  
 919 emphasize the structure induced by  $\mathbf{t}$ ,

$$920 \quad \mathbf{A}(x_\nu, x_\xi) \triangleq \mathbf{H}_1(x_\nu) \langle \mathbf{H}_2(x_\xi, :), \mathbf{t}^{[1]} - \mathbf{t}^{[0]} \rangle + \langle \mathbf{H}_2(x_\nu, :), \\ 921 \quad \otimes \mathbf{H}_2(x_\xi, :), \mathbf{t}^{[1]} \otimes \mathbf{t}^{[1]} - \mathbf{t}^{[0]} \otimes \mathbf{t}^{[0]} \rangle.$$

This term is proportional to the covariance subtraction term 922  
 923 [from Eq. (21)],

$$\sum_{q,z} (2\delta[q-z] - 1) \hat{\mathbf{H}}_1^{[q,z]} \otimes \hat{\mathbf{H}}_1^{[q,z]} = \left( \frac{2\mathbf{w}\Psi^{(1)}}{\bar{p}(1-\bar{p})} \right)^2 (1-2\bar{p})\mathbf{A}. \quad 924$$

When we estimate  $\mathbf{H}_2$  using covariance [Eq. (6)], we there- 925  
 926 fore obtain

$$\hat{\mathbf{H}}_2 = \frac{4\mathbf{w}^2}{\bar{p}(1-\bar{p})} \left( \Psi^{(1)} \mathbf{H}_2 + \mathbf{A} \left[ 2\Psi^{(2)} - (\Psi^{(1)})^2 \frac{1-2\bar{p}}{\bar{p}(1-\bar{p})} \right] \right). \quad 927$$

Our goal is  $\hat{\mathbf{H}}_2 \propto \mathbf{H}_2$  so we want the expression within  $\square$  to 928  
 929 vanish. If we adopt a Taylor expansion for  $\Psi$ , we can rewrite  
 930 this condition as

$$\ddot{\Psi} = \dot{\Psi}^2 \frac{1-2\bar{p}}{\bar{p}(1-\bar{p})}. \quad 931$$

At threshold ( $\bar{p} \sim \frac{3}{4}$ ) this translates into the requirement 932  
 933  $\ddot{\Psi}/\dot{\Psi}^2 \sim -\frac{8}{3}$ . For a realistic internal noise source and thresh-  
 934 old performance [Fig. 2(a)], we have  $\ddot{\Psi}/\dot{\Psi}^2 \sim -3$  at  $\bar{r}$ , which  
 935 is within 15% of the above target value. Therefore, by com-  
 936 puting  $\hat{\mathbf{H}}_2$  in the form of a differential covariance matrix, we  
 937 may compensate for 85% of the distortions introduced into  
 938 the second-order moment estimate.

We demonstrate this result in Fig. 6 for simulated Ham- 939  
 940 merstein (top row) and Korenberg (bottom row) cascades  
 941 (Secs. IV B and IV C). First-order kernels [black traces in (a)  
 942 and (e)] are best estimated via  $\hat{\mathbf{H}}_1^{[0]}$  (gray traces); when  
 943 target-present trials are included, the resulting estimates  $\hat{\mathbf{H}}_1$   
 944 (blue traces) are distorted, as expected from Eqs. (23)–(25).  
 945 Expected second-order kernels are shown in (b) and (f);  
 946 these were not simulated, but computed as predicted in the  
 947 absence of  $\Psi$  (see Secs. IV B and IV C). The last two col-  
 948 umns show estimates obtained from differential second-order  
 949 moments (third column) and covariance (fourth column). As  
 950 further emphasized by replotted diagonals in (a) and (e),  
 951 (see captions) second-order moment estimates  $\tilde{\mathbf{H}}_2$  present  
 952 clear distortions near target position (negative dip in green  
 953 trace, compare with red trace). These distortions are largely  
 954 eliminated when computing differential covariance  $\hat{\mathbf{H}}_2$  (light  
 955 red), as we predict from the equations detailed above. Notice,  
 956 however, that some distortions persist, in particular the nega-  
 957 tive flanks outside the diagonal in (d) (indicated by an ar-  
 958 row). For a Hammerstein cascade  $\mathbf{H}_2(x_\nu, x_\xi) \propto \delta[\nu - \xi] \mathbf{H}_1(x_\nu)$  959

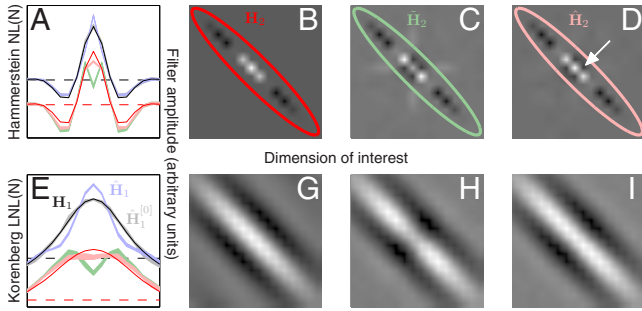


FIG. 6. (Color online) Second-order kernel estimates from differential second-order moments ( $\hat{H}_2$ ) vs covariance ( $\hat{H}_2$ ). Top row shows the results for a simulated Hammerstein NL(N) cascade ( $\langle \mathbf{f}, e^s \rangle$ ) (Sec. IV B), bottom row for Korenberg LNL(N) ( $\langle e^{f^*s}, \mathbf{g} \rangle$ ) (Sec. IV C), where  $\mathbf{g}$  is a Gaussian envelope with standard deviation equal to  $\approx \frac{1}{2}$  the period of the  $\mathbf{f}$  carrier. (a) Black trace shows  $\mathbf{f}$  (which is  $\propto \mathbf{H}_1$ ), blue trace shows the first-order kernel estimate, and gray trace shows the target-absent first-order kernel estimate [ $\mathbf{t}^{[0]}$  was set to 0 for these simulations, and  $\mathbf{t}^{[1]} = \rho \delta$  (peak at center of dimension of interest, zero elsewhere) with  $\rho$  selected to yield 75% correct responses]. Red/green traces plot the diagonals of kernels shown in (b) and (d). Black trace in (e) shows  $\mathbf{H}_1$  (not  $\mathbf{f}$ ) for Korenberg, i.e.,  $\mathbf{f} * \mathbf{g}$ . [(b) and (g)] Expected second-order kernel structure if  $\Psi$  is bypassed successfully (see Secs. IV B and IV C). [(c) and (h)] Second-order kernel estimates using differential second-order moments. [(d) and (i)] Second-order kernel estimates using differential covariance. Plotting conventions similar to Fig. 4.

AQ:  
#22

959 (see Refs. 26, 41, and 86 and Sec. IV B): no modulation is  
960 present outside the diagonal in  $\hat{H}_2$  [see panel (b)], a charac-  
961 teristic that  $\hat{H}_2$  fails to estimate correctly.

## 962 VI. SMALL-SCALE ALGORITHMS

963 Cascade models like those considered in Secs. IV B and  
964 IV C consist of a few processing components assembled into  
965 a relatively simple circuit: they are small-scale networks. A  
966 significant advantage of considering sensory processing at  
967 the level of small-scale circuitry is that it affords scalability  
968 across systems: Korenberg cascades can be successfully ex-  
969 ploited to describe processes within the fly brain<sup>27,48</sup> as well  
970 as the human brain,<sup>50</sup> despite obvious macroscopic anatomical  
971 differences between these two neural structures. Although  
972 their behavior can be unexpectedly complex and gives rise to  
973 somewhat unexpected results (e.g., Ref. 9), it is legitimate to  
974 ask whether they can provide a reasonable account of a phe-  
975 nomenon like perceptual processing, which originates from a  
976 large-scale neuronal assembly (the brain). A related concern  
977 can be raised in connection with the Volterra cascade [Eq.  
978 (4)]: given that it is empirically feasible to obtain kernel  
979 characterizations only up to second-order (Sec. II B), is it  
980 reasonable to expect that this approach may yield a useful  
981 description of complex processes like human vision?

982 There is really no way to answer this question outside  
983 the laboratory: the ultimate and most relevant test of any  
984 approach for the characterization of human sensory process-  
985 ing is whether the approach is able to provide an accurate  
986 reconstruction of the sensory process as conveyed to us via  
987 empirical measurements. Lacking extensive experimental  
988 evidence, there is no reason to prefer other approaches over  
989 the one outlined above; in fact, the knowledge accumulated  
990 so far suggests that this may be the most fruitful framework  
991 for understanding sensory processing within limited con-

texts, such as relatively simple detection tasks. Starting with  
pioneering work in the 1960s by a number of investigators  
(e.g., Barlow,<sup>5</sup> Mountcastle<sup>44</sup>), quantitative studies of sensory  
processing in humans have adopted the working hypothesis known as the lower envelope principle (LEP). According to this notion, humans presented with a specific sensory task monitor the output of only a restricted subset of their available neural circuitry; this subset consists of the neural components that are most sensitive for the task at hand.<sup>62</sup> Sensitivity is defined within the context of SDT (Ref. 20) and is therefore a well-characterized concept.

If we were to provide a concrete example of what the LEP translates to, we could imagine an observer who is presented with moving dots and asked to detect drift in a specified direction. According to the LEP, the observer bases his/her choice not on the output of his/her entire brain from occipital to frontal cortex, but only on the output from visual cortex. Further, the LEP postulates that the observer relies not on the entirety of visual cortex, but only on the area within visual cortex that is most responsive to motion; current knowledge would indicate MT+ as being the relevant area.<sup>8</sup> Even further, the LEP would restrict the relevant signals only to the subset of MT+ neurons with spatial receptive fields located within the region of visual space spanned by the stimulus, not on neurons responsive to other regions. And finally, the LEP may be interpreted to indicate that of these neurons only those responsive to the target direction would be monitored by the observer.<sup>59</sup> Those neurons and the associated circuitry would represent the bulk of  $\mathcal{H}$ .

The view expressed above is no doubt simplistic and opens to a number of criticisms: to mention one, there is not even consensus over the notion that motion processing may be restricted to the MT region.<sup>34</sup> However, there is convincing experimental evidence that it may be applicable at least within the context of specific sensory stimuli, tasks, and the experimenter's ability to measure the quantities of interest.<sup>62</sup> The most compelling evidence in this respect dates back to the classic work by Newsome and collaborators in the 1990s,<sup>59</sup> who demonstrated that the behavioral choice expressed by primates in response to a visual motion task could be biased in a systematic and predictable manner by applying electrical stimulation to a restricted portion of visual cortex. For example, if the animal is asked to indicate the direction of a moving stimulus by pressing button 1 for leftward motion and button 2 for rightward motion, the proportion of button 1 presses can be increased by electrically stimulating the subregion of MT cortex containing neurons responsive to leftward motion.<sup>58</sup> Subsequent work has shown that similar results can be obtained for more complex stimuli and tasks such as face recognition.<sup>2</sup> These studies do not provide conclusive evidence for the LEP in its most extreme formulation, but they strongly suggest that weak variants may be applicable,<sup>62</sup> for example, versions of the LEP in which it is hypothesized that the subset of neural resources monitored by the observer may only roughly correspond to the sensitive neurons with the inclusion of a number of irrelevant components, and furthermore that the extent of the envelope may depend on attention.<sup>57</sup>

1050 Even if one were to accept the LEP in some form, it is  
 1051 not clear that the relevant neural circuitry would be ad-  
 1052 equately captured by small-scale cascade models: again this  
 1053 may seem oversimplistic. However, the question of interest  
 1054 here is not whether the relevant circuitry actually conforms  
 1055 to such descriptors, but rather whether its output in the form  
 1056 of a behavioral choice is consistent with such descriptors  
 1057 within the precision afforded by the empirical measurements.  
 1058 When the problem is viewed from the standpoint of an algo-  
 1059 rithmic solution, the latter specification is critical: there is an  
 1060 intrinsic limit to how detailed any characterization of human  
 1061 sensory processing can be for detection/discrimination at  
 1062 threshold; beyond this limit further algorithmic distinctions  
 1063 are irrelevant. Currently, this limit is determined primarily by  
 1064 the number of single-trial choices that can be feasibly re-  
 1065 corded in the laboratory from a given participant, as well as  
 1066 the presence of a sizeable noise source that is intrinsic to the  
 1067 participant and therefore not under direct experimental  
 1068 control.<sup>49</sup> Both factors are integral to the process of record-  
 1069 ing behavior regardless of available technology, so it is un-  
 1070 likely that their impact will be reduced in the future (al-  
 1071 though it is possible to envisage, for example, that drugs may  
 1072 be developed to specifically reduce internal sensory noise). It  
 1073 is possible to obtain indirect estimates of internal noise in  
 1074 individual participants,<sup>10</sup> and use these to establish an upper  
 1075 limit on how well different models can be expected to predict  
 1076 human choices on individual trials.<sup>55</sup> It then becomes an em-  
 1077 pirical question whether small-scale algorithms can approach  
 1078 this limit or not. In previous work,<sup>50,51</sup> we have shown that  
 1079 they often do.

1080 Below we consider two examples of models that have  
 1081 been successful in a vast number of applications in percep-  
 1082 tion science, with the goal of demonstrating how they can be  
 1083 approximated via Korenberg cascades (already popular in  
 1084 neuroscience<sup>76,88</sup>) and subsequently analyzed using some of  
 1085 the tools described in this article. We use general formula-  
 1086 tions to emphasize their algorithmic potential for the solution  
 1087 of simple processing problems on a small scale. In both  
 1088 cases, a nonstatic nonlinearity is approximated by a combi-  
 1089 nation of static nonlinearities and linear operators,<sup>30,60</sup> the  
 1090 advantage of the latter being that it forms the backbone of  
 1091 well-studied cascade systems for which kernel interpretation  
 1092 is relatively straightforward (Secs. IV B and IV C). Figure 7  
 1093 summarizes these approximations; below we refer to them as  
 1094 “toy” examples to emphasize that they are model-based, not  
 1095 data based. We then consider four “real” examples from data.  
 1096 The first two examples (Secs. VI C and VI D) consist of  
 1097 kernel measurements which could be interpreted based on  
 1098 the kind of approximation presented in Fig. 7; the first ex-  
 1099 ample is for a Hammerstein cascade, the second one for a  
 1100 Korenberg cascade.

1101 The third and fourth examples are in a sense counterex-  
 1102 amples, in that the approach outlined in Fig. 7 did not offer a  
 1103 satisfactory interpretation for the data. In the third example  
 1104 (Sec. VI E), we consider a data set for which there was no  
 1105 straightforward, simple approximation of the kind presented  
 1106 in Fig. 7; in this case, we resorted to a dynamically nonlinear  
 1107 model [Fig. 9(f)] that was able to capture the data satisfac-  
 1108 torily while at the same time retaining physiological plausi-

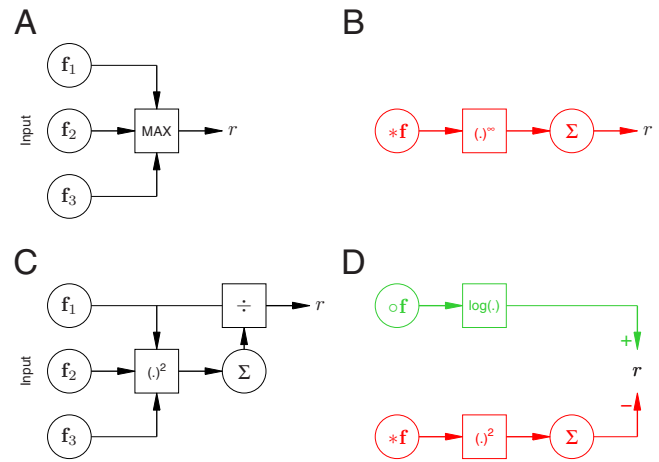


FIG. 7. (Color online) Approximation of MAX uncertainty (a) and normalization (c) models using Wiener (green) and Korenberg (red) cascades (Refs. 26 and 31). (a) is approximated in (b); (c) in (d). All symbols correspond to those used in the main text; see in particular Secs. VI A and VI B.

AQ:  
#23

1109 bility. This example in particular demonstrates how the utili-  
 1110 ty of kernel characterization is not restricted to small-scale  
 1111 static approximations of dynamic systems; the information it  
 1112 provides can be effectively used to constrain plausible dyn-  
 1113 amic models. The fourth example (Sec. VI F) makes a related  
 1114 point but from a different perspective: it shows how dynamic  
 1115 nonlinearities that occur on a relatively long timescale (per-  
 1116 ceptual adaptation in this case) can be approached by treat-  
 1117 ing each dynamic state as static on the much shorter timescale  
 1118 of the experimental characterization, and by inducing and  
 1119 characterizing different dynamic states separately one at a  
 1120 time. In the specific case of stereo motion adaptation consid-  
 1121 ered in Sec. VI F, we used reverse correlation to estimate  
 1122  $\hat{H}_1$  kernels after prolonged viewing of an adapting stimu-  
 1123 lus containing moving elements at specific 3D positions, and  
 1124 compared them to kernels obtained in the absence of adapta-  
 1125 tion. The adaptor caused the system to shift onto a new per-  
 1126 ceptual state; we were able to measure specific signatures  
 1127 of this shift at the level of the  $\hat{H}_1$  kernels. Although these  
 1128 measurements in isolation are not able to capture the highly  
 1129 nonlinear and nonstatic nature of perceptual adaptation, they  
 1130 can be combined to generate an informative picture of the  
 1131 overall phenomenon at a greater level of detail than afforded  
 1132 by measurements that do not rely on kernel estimation (see  
 1133 Ref. 54 for the specific rationale behind this statement).  
 1134

AQ:  
#11

### A. Toy example 1: MAX uncertainty model

1135 The MAX uncertainty model<sup>65</sup> [Fig. 7(a)] can be ap-  
 1136 proximated using a Korenberg cascade [Fig. 7(b)] by ex-  
 1137 pressing the max operation as an  $\ell^\infty$  norm.<sup>51</sup> For this model,  
 1138 we have  
 1139

$$\mathbf{H}_d(x_{v_1}, \dots, x_{v_d}) = \Phi^{(d)} \sum_k \mathbf{g}(x_k)^d \mathbf{f}(x_k - x_{v_1}) \cdots \mathbf{f}(x_k - x_{v_d}),$$

(29) 1140

1141 where  $\mathbf{g}$  is raised to the  $d$  power because the uncertainty  
 1142 model is written as  $\langle \Phi(\mathbf{g} \circ (\mathbf{f} * \mathbf{s}), 1) \rangle$ , i.e., the sandwiched

AQ:  
#12

1143 nonlinearity is applied after weighting by  $\mathbf{g}$  (see Ref. 51); see  
 1144 Sec. IV C and compare Eq. (29) with Eq. (24).  
 1145 If we restrict our analysis to the special case considered  
 1146 in Sec. IV C, i.e.,  $\mathbf{g}=1$  and  $\mathbf{t}^{[1]}=\rho\delta(\mathbf{t}^{[0]}=0)$ , we have (at low  
 1147 SNR)  $\hat{\mathbf{H}}_1^{[1]} \propto \mathbf{f} \star \mathbf{f}$  (as detailed in Sec. IV C). From Eq. (20) we  
 1148 also have  $\hat{\mathbf{H}}_1^{[0]} = \omega$ , i.e., the target-absent first-order kernel is  
 1149 essentially featureless. More generally, it is  $\propto \mathbf{g} \star \mathbf{f}$  and there-  
 1150 fore reflects the extent of the uncertainty window, while  $\hat{\mathbf{H}}_1^{[1]}$   
 1151 provides an indirect image of the front-end filter  $\mathbf{f}$  (see Sec.  
 1152 IV C for a tentative experimental validation of this result and  
 1153 Ref. 83 for a discussion of related notions). The uncertainty  
 1154 model therefore provides an interesting example of how both  
 1155 target-present and target-absent first-order estimates can de-  
 1156 liver useful information in a complementary fashion.

### 1157 B. Toy example 2: Divisive normalization

1158 The other model we consider here is the normalization  
 1159 model commonly used to implement contrast gain control<sup>25</sup>  
 1160 [Fig. 7(c)] with output,

$$1161 \quad r = \frac{\langle \mathbf{f}, \mathbf{s} \rangle}{k + \langle (\mathbf{f} \star \mathbf{s})^2, 1 \rangle}. \quad (30)$$

1162 If internal noise is relatively small, we can log this expres-  
 1163 sion without affecting our discussion (Sec. III B 1),

$$1164 \quad r = \log(\langle \mathbf{f}, \mathbf{s} \rangle) - \log\left(1 + \frac{\langle (\mathbf{f} \star \mathbf{s})^2, 1 \rangle}{k}\right) + \log(k). \quad (31)$$

1165 We have assumed  $r > 0$  in Eq. (30) (necessary to log it); if it  
 1166 is not, we can approximate it as such by adding a large  
 1167 constant (which has no effect on choice), logging and using  
 1168  $\log(1+x) \sim x$  for  $x \ll 1$ . We can ignore the last additive term  
 1169 in Eq. (31) (it does not affect the final psychophysical re-  
 1170 sponse). If  $k$  is large compared to the normalizing term (its  
 1171 numerator), we can approximate the above expression to

$$1172 \quad r = \log(\langle \mathbf{f}, \mathbf{s} \rangle) - \frac{1}{k} \langle (\mathbf{f} \star \mathbf{s})^2, 1 \rangle.$$

1173 This is now the combined output of a Wiener LN system  
 1174  $\log(\langle \mathbf{f}, \mathbf{s} \rangle)$ , where  $N$  is the log operation, and a Korenberg  
 1175 LNL system  $-1/k \langle (\mathbf{f} \star \mathbf{s})^2, 1 \rangle$ , where  $N$  is the squaring opera-  
 1176 tion [Fig. 7(d)]. We emphasize that several assumptions (e.g.,  
 1177 small internal noise) and approximations were involved in  
 1178 this derivation, therefore its utility is primarily qualitative.

### 1179 C. Real example 1: Spatial uncertainty approximated 1180 by Hammerstein cascade

1181 This section reproduces results from a more extensive  
 1182 publication<sup>51</sup> on the topic of spatial uncertainty. Figure 8(a)  
 1183 shows a set of  $\hat{\mathbf{H}}_1$  kernels for detecting a luminance bar  
 1184 appearing within a specified spatial range; the range could be  
 1185 very narrow (no spatial uncertainty), coded by the blue sym-  
 1186 bols, or very wide (large spatial uncertainty), coded by the  
 1187 red symbols. Cyan and magenta colors refer to intermediate  
 1188 uncertainty levels. As expected, the kernels spread over a  
 1189 progressively wider spatial range with increasing uncertainty  
 1190 [blue kernel peaks at target location (center of plot), red ker-  
 1191 nel spreads across x axis]. Figures 8(b)–8(e) show corre-

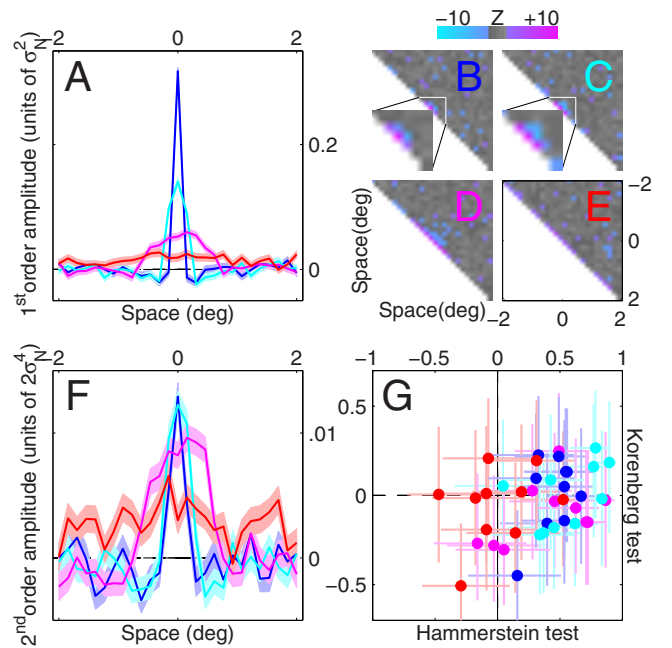


FIG. 8. (Color online) First-order and second-order kernels for visual detection under spatial uncertainty (Ref. 51) (data from >110K trials). (a) Aggregate  $\hat{\mathbf{H}}_1$  kernels for four different uncertainty levels; uncertainty increases from blue to cyan, magenta, and red. [(b)–(e)] Aggregate second-order kernels for four different uncertainty levels, color-coded for  $|Z| > 2$  (red for positive, blue for negative). For the two smaller uncertainty levels [(b) and (c)], the central regions of the kernels are magnified for ease of inspection. (f) Aggregate  $\hat{\mathbf{H}}_2$  diagonals. (g) Correlation between  $\hat{\mathbf{H}}_1$  kernel and  $\hat{\mathbf{H}}_2$  diagonal (Hammerstein test) is plotted on the x axis vs correlation between  $\hat{\mathbf{H}}_1$  and  $\sum_k \hat{\mathbf{H}}_2(:, k)$  on the y axis for each observer separately. The former correlation tests for the applicability of a Hammerstein NL(N) cascade (Sec. IV B), the latter for the applicability of a Korenberg LNL(N) cascade (Sec. IV C); if the phenomenon of interest is well approximated by either cascade, the corresponding correlation is expected to differ from 0 (see Sec. VI C for details on these two tests). Error bars and shading show  $\pm 1$  SEM.

AQ:  
#24AQ:  
#25

sponding  $\hat{\mathbf{H}}_2$  kernels; because we observed modulations primarily within the diagonal region, we replot  $\hat{\mathbf{H}}_2$  diagonals in Fig. 8(f). It is apparent that  $\hat{\mathbf{H}}_1$  kernels [Fig. 8(a)] and  $\hat{\mathbf{H}}_2$  diagonals [Fig. 8(f)] share similar characteristics, a potential signature of the Hammerstein cascade (see below).

As a first check of whether the general strategy outlined in Fig. 7 was applicable to this data set, we carried out two standard tests in nonlinear system identification based on the following two results:<sup>41,86</sup> for a Hammerstein cascade  $\mathbf{H}_1 \propto \text{diag}(\mathbf{H}_2)$ ; for a Korenberg cascade  $\mathbf{H}_1 \propto \sum_k \mathbf{H}_2(:, k)$ . We therefore expect a positive correlation between the first-order kernel and the diagonal of the second-order kernel for a Hammerstein cascade; failing that, we test for a positive correlation between the first-order kernel and the second-order marginal to check for compatibility with a Korenberg cascade. This strategy is a standard practice in neuroscience applications (e.g., Refs. 16 and 27) and it essentially relies on the logic of Fig. 7, whereby the problem is simplified by approximating the biological system using a well-understood model with tractable nonlinear operators.

Figure 8(g) shows the result of both tests, Hammerstein on the x axis versus Korenberg on the y axis. With the only

1214 exception of the largest uncertainty level (red), all other ker-  
 1215 nels show significant correlations between  $\hat{H}_1$  and  $\text{diag}(\hat{H}_2)$   
 1216 (blue, cyan, and magenta data points fall to the right of the  
 1217 vertical dashed line at  $p < 0.01$ ), supporting the applicability  
 1218 of the Hammerstein model. Correlations between  $\hat{H}_1$  and  
 1219  $\sum_k \hat{H}_2(:, k)$  (Korenberg test) were no different from 0 (at  $p$   
 1220  $> 0.05$ ) for any of the four uncertainty levels (data points fall  
 1221 on the horizontal dashed line), indicating that modulations  
 1222 outside the diagonal region contained primarily noise (which  
 1223 eliminated the diagonal correlation). From this analysis, we  
 1224 conclude that the observed kernel structure is consistent with  
 1225 a Hammerstein model (Sec. IV B), with no clear evidence  
 1226 that this model needs further elaboration into a Korenberg  
 1227 cascade.<sup>51</sup>

#### 1228 D. Real example 2: Texture processing approximated 1229 by Korenberg cascade

1230 We have already used this data set in Fig. 5 when dis-  
 1231 cussing  $\hat{H}_1^{[1]}$  estimates and their relation to the system front-  
 1232 end filter (Sec. IV C). Figure 9(a) uses plotting conventions  
 1233 similar to Figs. 5(a) and 5(c), except both center and sur-  
 1234 round data are plotted in the same panel (black and blue  
 1235 traces, respectively) and red is used for the  $\hat{H}_2$  diagonal as  
 1236 opposed to  $\hat{H}_2(x_1, :)$  (yellow refers to data for the surround).  
 1237 As in the previous example,  $\hat{H}_1$  kernels and  $\hat{H}_2$  diagonals  
 1238 share similar characteristics; however, notice that there is a  
 1239 sign inversion for the surround kernels (compare blue trace  
 1240 with yellow trace), which is not observed for the center ker-  
 1241 nels (compare black with red). This sign inversion is not  
 1242 accommodated by a Hammerstein cascade with a single  
 1243 static nonlinearity (see Sec. IV B and Ref. 50 for details).  
 1244 Figure 9(b) demonstrates the sign inversion across observers:  
 1245 correlations between  $\hat{H}_1$  and  $\text{diag}(\hat{H}_2)$  [plotted on the x axis  
 1246 as in Fig. 8(g)] are significantly positive for center data and  
 1247 significantly negative for surround data ( $p < 0.01$ ). A similar  
 1248 result is obtained for the Korenberg test (y axis), but in this  
 1249 case the underlying cascade is able to account for the sign  
 1250 inversion by applying a linear filter before the static nonlin-  
 1251 earity with opposite sign for center and surround.<sup>50</sup> From this  
 1252 analysis, we conclude that the observed kernel structure is  
 1253 consistent with a Korenberg LNL(L) cascade; this result is in  
 1254 line with the existing models of human texture processing,  
 1255 which are mostly of the LNL type,<sup>7,32</sup> also termed filter-  
 1256 rectify-filter models.<sup>88</sup>

#### 1257 E. Real example 3: Nonlinear dynamics on neuronal 1258 timescale

1259 In a previous publication,<sup>50</sup> we documented the temporal  
 1260 dynamics of a simple visual process involving the detection  
 1261 of a luminance increment within a circular region surrounded  
 1262 by a large annulus. Figure 9(c) [plotted to the same conven-  
 1263 tions used for Fig. 9(a)] shows  $\hat{H}_1$  (cold colors) and  $\hat{H}_2$  di-  
 1264 agonal (warm colors) for both center (black/red) and sur-  
 1265 round (blue/yellow). As expected from the antagonistic  
 1266 nature of center-surround interactions in human vision,<sup>77</sup> the  
 1267 surround  $\hat{H}_1$  takes opposite polarity with respect to the center

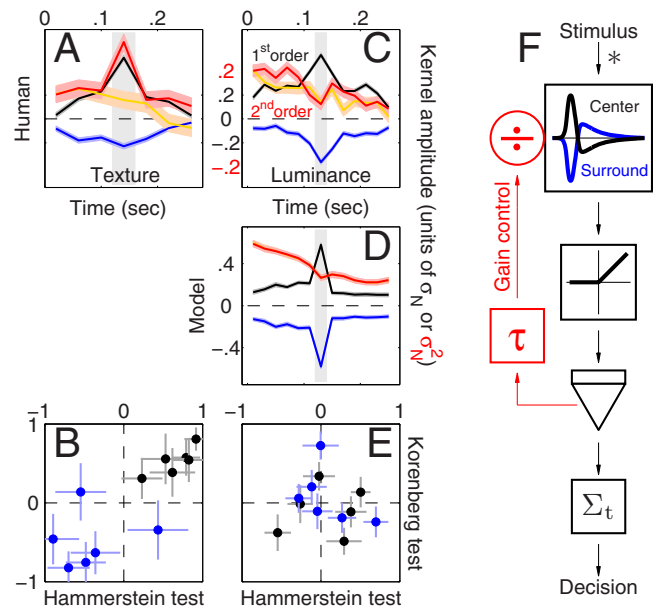


FIG. 9. (Color online) Temporal dynamics of texture (a) and luminance (c) processing (Ref. 50) (data from  $>56K$  trials). (a) shows aggregate (across observers) first-order (black/red) and second-order (red/yellow) kernels for both center and surround of an annular stimulus defined by texture modulations (see Ref. 50 for details). Temporal dynamics was captured over a range of  $\sim \frac{1}{4}$  s (see x axis). (c) is plotted to the same conventions as (a) but shows data for a stimulus defined by luminance modulations. (d) Luminance kernels generated by the model in (f) (traces for center and surround second-order kernels are not visible separately because of overlapping). (b) and (e) plot the results from the Hammerstein test on the x axis vs the Korenberg test on the y axis (see caption to Fig. 8 and Sec. VI C) for texture and luminance, respectively. (f) shows the luminance model structure (constructed around knowledge from retinal circuitry): input stimulus is convolved (\*) with center-surround receptive field (black and blue smooth traces) using first-order filter from primate retinal ganglion cells (Ref. 6). The output from this stage is half-wave rectified and fed onto an accumulator (inverted triangular symbol) implemented by differential equation (see Ref. 50 for details), then integrated across time ( $\Sigma_t$ ) to yield scalar decision variable further converted into psychophysical decision. Time-varying signal from accumulator stage is carried back into circuit with temporal delay  $\tau=60$  ms to control gain of linear stage via divisive inhibition (refer to Sec. VI E for more details). Gain-control loop is highlighted in red. The model was used to simulate a psychophysical experiment adopting same parameters selected for human observers. Shaded regions in (a) and (c) show  $\pm 1$  SEM, in (d)  $\pm 1$  SD across repeated Monte Carlo simulations. Gray rectangular regions indicate the time of occurrence of the target signal.

$\hat{H}_1$  (compare blue and black traces). The  $\hat{H}_2$  diagonal is essentially identical for center and surround (compare red and yellow traces) and its time-course differs significantly from the  $\hat{H}_1$  profile, a feature that proved challenging for subsequent modeling efforts (see below).

As with the data sets discussed above, we carried out the two tests that were able to account for those previous results [see Figs. 8(g) and 9(b)]. However, as shown in Fig. 9(e), both tests failed: correlation coefficients were not statistically different from 0 (at  $p > 0.05$ ) for either model (Hammerstein/Korenberg) or stimulus region (center/surround), as evident from the clustering of data points around 0 (intersection of dashed lines). This lack of correlation is primarily a consequence of the different temporal dynamics we consistently observed between first- and second-order kernels [these differences were not observed for texture processing, see Fig. 9(a)].

1285 Failing the two basic tests detailed above [Fig. 9(e)],  
 1286 there is no obvious way to proceed in terms of attempting an  
 1287 approximation using a simple block-model of the kind out-  
 1288 line in Fig. 7. This is not to say that such an approximation is  
 1289 impossible: a sufficient number of Korenberg cascades can  
 1290 approximate virtually any nonlinear system of physiological  
 1291 interest via variants of the Wiener–Bose model<sup>30,60,86</sup> and  
 1292 methods have been developed to identify the relevant struc-  
 1293 ture for specific applications (e.g., Refs. 13 and 63). How-  
 1294 ever, it may then become more parsimonious to explore sen-  
 1295 sible and simpler models containing a plausible dynamic  
 1296 nonlinearity. Figure 9(f) shows an example from this class of  
 1297 models; its time-varying response  $\mathbf{r}$  is determined by the  
 1298 following differential equation:

$$1299 \quad \frac{d\mathbf{r}}{dt} = \left[ \frac{\mathbf{r}(t)}{b\mathbf{r}(t-\tau) + 1} \right]_+,$$

1300 where  $\mathbf{r} = \mathbf{f} * \mathbf{s}_{\text{ctr}} - \mathbf{f} * \mathbf{s}_{\text{sur}}$ ,  $\mathbf{f}$  is the front-end impulse response  
 1301 of a primate retinal ganglion cell [shown by the black trace  
 1302 in Fig. 9(f)],  $\mathbf{s}_{\text{ctr}}$  is the center stimulus,  $\mathbf{s}_{\text{sur}}$  the surround  
 1303 stimulus,  $b = \frac{1}{5}$ ,  $[\ ]_+$  is half-wave rectification, and  $\tau = 60$  ms.  
 1304 For the purpose of generating a decision variable, we simply  
 1305 summed  $\mathbf{r}$  across time ( $r = (\mathbf{r}, 1)$ ). The critical component is  
 1306 the delayed gain-control feedback loop [red in Fig. 9(f)],  
 1307 which is able to generate dynamics similar to those observed  
 1308 experimentally [Fig. 9(e)]. The choice of a servo-mechanism  
 1309 was prompted primarily by the damped profile of the  $\hat{\mathbf{H}}_2$   
 1310 diagonal [red trace in Fig. 9(c)], which is common in non-  
 1311 linear feedback systems,<sup>39,87</sup> and by the knowledge that gain  
 1312 control is a well-characterized phenomenon in both  
 1313 vertebrate<sup>67,71</sup> and invertebrate<sup>84</sup> retinas. Interestingly, our  
 1314 earlier work on the dynamics of directional processing in  
 1315 humans<sup>53</sup> required the introduction of a 90 ms temporal de-  
 1316 lay in the divisive feedback loop of the motion detector  
 1317 model, a figure comparable with the 60 ms used here.

#### 1318 F. Real example 4: Nonlinear dynamics on cognitive 1319 timescale

1320 Visual adaptation is a complex phenomenon that oper-  
 1321 ates on a variety of timescales,<sup>29</sup> typically much longer than  
 1322 those we considered in Secs. VI D and VI E, and provides a  
 1323 clear example of nonstatic behavior. If we look at a motion-  
 1324 less object at a given time A, it will look motionless; if we  
 1325 then observe a moving pattern for a prolonged length of time  
 1326 ( $\sim 1$  min) and subsequently look back to the same motion-  
 1327 less object at a given time B, it will appear to move in a  
 1328 direction opposite to that of the moving pattern we had seen  
 1329 previously. This motion after-effect is robust and can be eas-  
 1330 ily experienced even in natural vision.<sup>42</sup> Because the visual  
 1331 input is the same at time A as it is at time B, and yet our  
 1332 visual system delivers an entirely different percept, the un-  
 1333 derlying phenomenon is clearly dynamic. However, the rel-  
 1334 evant dynamics is slow (order of several seconds) compared  
 1335 to the dynamics we have been interested in so far (Sec.  
 1336 VI E); it is therefore reasonable to expect that we may be  
 1337 able to capture the effect of adaptation on kernel structure by

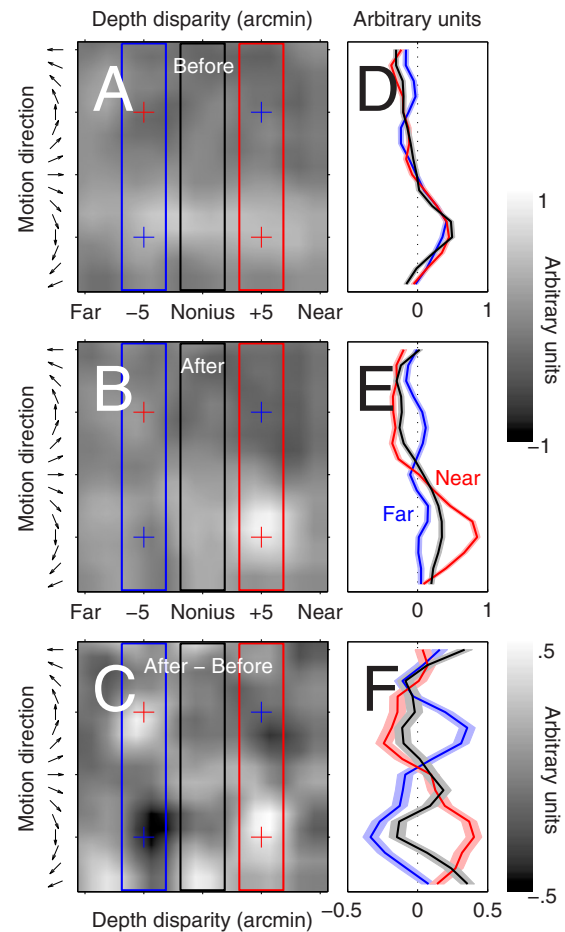


FIG. 10. (Color online) Effect of visual adaptation on first-order kernels for processing motion in depth (Ref. 54) (data from  $\sim 20$ K trials). [(a) and (b)] Smoothed (using Gaussian filter with SD equal to 1 surface pixel)  $\hat{\mathbf{H}}_1$  kernels (average across observers) obtained before (a) and after (b) adaptation; (c) the differential filter [(b) minus (a)]. The stimulus consisted of moving elements presented at different depths; motion direction is plotted on the y axis, depth (as delivered via binocular disparity) on the x axis. Red crosses indicate signal locations for the target, blue crosses for the nontarget [see Ref. 54 (particularly Fig. 1) for details on the stimulus and experimental protocol]. (a) and (b) are plotted to the gray scale shown by the top right legend; (c) is plotted to the scale shown by the bottom right legend. [(d)–(f)] Slices along direction of motion averaged across the disparity range indicated by the corresponding dashed rectangular regions in (a)–(c) (red for near, blue for far, black for nonius fixation, i.e., the depth level corresponding to where the eyes fixate). Shaded regions show  $\pm 1$  SEM. Units are arbitrary because averaging was performed after observer-by-observer normalization (Ref. 54).

AQ:  
#27

probing the system with brief stimuli under different states of  
 adaptation, a strategy that has been successfully exploited in  
 physiology (e.g., Ref. 16).

Figure 10(a) shows an  $\hat{\mathbf{H}}_1$  kernel defined across the dimen-  
 sions of depth (stereoscopic disparity on the x axis) and motion  
 direction (y axis); the reader is referred to Ref. 54 for details  
 on the stimulus and experimental design. The target signal was  
 delivered to the locations indicated by the red crosses; some kernel  
 modulations colocalize with the signal, although only to a coarse  
 extent [see Fig. 10(c) for cross-sections]. Figure 10(b) shows  
 the same measurement performed after prolonged adaptation, and  
 Fig. 10(c) shows the difference between Figs. 10(b) and 10(a).  
 It is clear that adaptation induced specific changes within the  
 kernels; these

changes were meaningfully related to the target signal structure [Fig. 10(f)]. The characterization afforded by this approach is of course very limited: we only obtain two snapshots of the adaptation process, not a detailed time-course of its dynamics. However, each snapshot is remarkably detailed; this level of detail allows the experimenter to draw a number of significant conclusions about the underlying mechanism (see Ref. 54 for an extended treatment of this topic).

## VII. CONCLUSIONS

We have examined a methodology for characterizing second-order functional approximations to a general-purpose sensory apparatus  $\mathcal{H}$  embedded within realistic experimental constraints. From a general theoretical viewpoint, the two main constraints are those summarized in Fig. 1: a distortion of the input caused by the target signal  $\mathbf{t}$  and a distortion of the output caused by the decisional transducer  $\Psi$ . Our goal is to estimate a first-order kernel  $\hat{\mathbf{H}}_1$  and a second-order kernel  $\hat{\mathbf{H}}_2$  that characterize  $\mathcal{H}$  in the sense of a Volterra approximation [Eq. (4)]. These estimates can then be used to build and constrain processing algorithms operating at small scales (Sec. VI), i.e., via a few components connected in relatively simple ways.

For the first-order kernel estimate  $\hat{\mathbf{H}}_1$ , there are important differences between estimates obtained from noise fields containing the target ( $\hat{\mathbf{H}}_1^{[1]}$ ) and those not containing the target ( $\hat{\mathbf{H}}_1^{[0]}$ ), whenever this distinction is applicable (i.e.,  $\mathbf{t}^{[0]} = 0$ ). In general,  $\hat{\mathbf{H}}_1^{[0]}$  provides a less distorted estimate of  $\hat{\mathbf{H}}_1$  [although still potentially distorted by higher-order kernels, see Eq. (22)]. There are instances, however, when  $\hat{\mathbf{H}}_1^{[0]}$  is not particularly informative, e.g., uncertainty models (see Sec. IV C). In these instances, the front-end filter  $\mathbf{f}$  may be estimated more effectively using  $\hat{\mathbf{H}}_1^{[1]}$ , with the caveat that this estimator reflects  $\mathbf{f} \star \mathbf{f}$  not  $\mathbf{f}$  (see Fig. 5 and Ref. 51).

For the second-order kernel estimate  $\hat{\mathbf{H}}_2$ , the distinction between target-present and target-absent subestimates is not relevant, because these do not allow reliable measurements of  $\mathbf{H}_2$  when considered separately. We have shown in Sec. III A that it is possible to bypass  $\Psi$  for a benchmark linear system ( $\mathbf{H}_d = 0$  for  $d > 1$ ) to obtain  $\hat{\mathbf{H}}_2 = 0$ . This result ensures that any modulation within  $\hat{\mathbf{H}}_2$  must reflect departures from linearity, and not simply artifacts introduced by  $\Psi$ ; any estimator for  $\mathbf{H}_2$  must satisfy this benchmark result.  $\hat{\mathbf{H}}_2^{[1]}$  and  $\hat{\mathbf{H}}_2^{[0]}$  do not, making them unsuitable for the application of interest. We have further demonstrated that the second-order estimator based on differential covariance  $\hat{\mathbf{H}}_2$  is more robust than that based on differential second-order moments  $\tilde{\mathbf{H}}_2$  (see Sec. V B). A more effective estimation strategy for  $\mathbf{H}_2$  is therefore achieved via differential covariance, as expressed in Eq. (6).

A major issue, which we have chosen to ignore in this article, is whether it is at all possible to reduce the treatment of animal sensory processing to a functional of the kind exemplified by  $\mathcal{H}$  here. This is of course a critical issue for the applicability of the tools we have considered so far. Whether

the framework adopted here is applicable or not will surely depend on the specific aspect of sensory processing that is being considered, on the animal system which implements it, and on the kinds of stimuli and measurements that are being carried out; ultimately, it is an empirical question that can only be answered effectively by transferring the theoretical concepts explored here to the laboratory.

## ACKNOWLEDGMENTS

This work was supported by the Royal Society (University Research Fellowship) and the Medical Research Council (New Investigator Research Grant).

- <sup>1</sup>Abbey, C. K. and Eckstein, M. P., "Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments," *J. Vision* **2**, 66–78 (2002).
- <sup>2</sup>Afraz, S. R., Kiani, R., and Esteky, H., "Microstimulation of inferotemporal cortex influences face categorization," *Nature (London)* **442**, 692–695 (2006).
- <sup>3</sup>Ahumada, A. J., "Classification image weights and internal noise level estimation," *J. Vision* **2**, 121–131 (2002).
- <sup>4</sup>Ahumada, A. J., Marken, R., and Sandusky, A., "Time and frequency analyses of auditory signal detection," *J. Opt. Soc. Am. A Opt. Image Sci. Vis* **57**, 385–390 (1975).
- <sup>5</sup>Barlow, H. B., Levick, W. R., and Yoon, M., "Responses to single quanta of light in retinal ganglion cells of the cat," *Vision Res.* **11**, 87–101 (1971).
- <sup>6</sup>Benardete, E. A. and Kaplan, E., "The dynamics of primate M retinal ganglion cells," *Visual Neurosci.* **16**, 355–368 (1999).
- <sup>7</sup>Bergen, J. R. and Landy, M. S., "Computational models of visual texture segregation," in *Computational Models of Visual Processing*, edited by M. Landy and J. Movshon (MIT, Cambridge, MA, 1991), pp. 253–271.
- <sup>8</sup>Born, R. T. and Bradley, D. C., "Structure and function of visual area MT," *Annu. Rev. Neurosci.* **28**, 157–189 (2005).
- <sup>9</sup>Borst, A., Flanagan, V. L., and Sompolinsky, H., "Adaptation without parameter change: Dynamic gain control in motion detection," *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6172–6176 (2005).
- <sup>10</sup>Burgess, A. E. and Colborne, B., "Visual signal detection. IV. Observer inconsistency," *J. Opt. Soc. Am. A* **5**, 617–627 (1988).
- <sup>11</sup>Busse, L., Katzner, S., Tillmann, C., and Treue, S., "Effects of attention on perceptual direction tuning curves in the human visual system," *J. Vision* **8**, 1–13 (2008).
- <sup>12</sup>Bussgang, J. J., "Cross-correlation functions of amplitude-distorted Gaussian signals," MIT Research Laboratory Electricity Technical Report No. 216, August 1952.
- <sup>13</sup>Citron, M. C. and Marmarelis, V. Z., "Applications of minimum-order Wiener modeling to retinal ganglion cell spatiotemporal dynamics," *Biol. Cybern.* **57**, 241–247 (1987).
- <sup>14</sup>Falmagne, J.-C., *Elements of Psychophysical Theory* (Oxford University Press, New York, 1985).
- <sup>15</sup>Franz, M. O. and Schölkopf, B., "A unifying view of Wiener and Volterra theory and polynomial kernel regression," *Neural Comput.* **18**, 3097–3118 (2006).
- <sup>16</sup>French, A. S., Korenberg, M. J., Jarvilehto, M., Kouvalainen, E., Juusola, M., and Weckstrom, M., "The dynamic nonlinear behavior of fly photoreceptors evoked by a wide range of light intensities," *Biophys. J.* **65**, 832–839 (1993).
- <sup>17</sup>Geisler, W. S., "Ideal observer theory in psychophysics and physiology," *Phys. Scr.* **39**, 153–160 (1989).
- <sup>18</sup>Gold, J. I. and Shadlen, M. N., "The neural basis of decision making," *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
- <sup>19</sup>Gosselin, F. A. and Schyns, P. G., "Superstitious perceptions reveal properties of internal representations," *Psychol. Sci.* **14**, 505–509 (2003).
- <sup>20</sup>Green, D. M. and Swets, J. A., *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
- <sup>21</sup>Gregson, R. A. and Britton, L. A., "The size-weight illusion in 2-D nonlinear psychophysics," *Percept. Psychophys.* **48**, 343–356 (1990).
- <sup>22</sup>Guastello, S. J., "Nonlinear dynamics in psychology," *Discrete Dyn. Nat. Soc.* **6**, 11–29 (2001).
- <sup>23</sup>Heath, R. A., *Nonlinear Dynamics: Techniques and Applications in Psychology* (Lawrence Erlbaum Associates, New Jersey, 2000).

- 1476<sup>24</sup>Heath, R. A. and Fulham, R., "Applications of system identification and adaptive filtering techniques in human information processing," in *Cognition, Information Processing and Motivation*, edited by G. d'Ydewalle (Elsevier Science, Amsterdam, 1985), pp. 117–147.
- 1477
- 1478
- 1479
- 1480<sup>25</sup>Heeger, D. J., Simoncelli, E. P., and Movshon, J. A., "Computational models of cortical visual processing," *Proc. Natl. Acad. Sci. U.S.A.* **93**, 623–627 (1996).
- 1481
- 1482
- 1483<sup>26</sup>Hunter, I. W. and Korenberg, M. J., "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models," *Biol. Cybern.* **55**, 135–144 (1986).
- 1484
- 1485
- 1486<sup>27</sup>James, A. C., "Nonlinear operator network models of processing in the fly lamina," in *Nonlinear Vision*, edited by R. B. Pinter and B. Nabet (CRC, Boca Raton, FL, 1992), pp. 39–73.
- 1487
- 1488
- 1489<sup>28</sup>Kelly, A., Heathcote, A., Heath, R., and Longstaff, M., "Response-time dynamics: Evidence for linear and low-dimensional nonlinear structure in human choice sequences," *Q. J. Exp. Psychol. A* **54**, 805–840 (2001).
- 1490
- 1491
- 1492<sup>29</sup>Kohn, A., "Visual adaptation: Physiology, mechanisms, and functional benefits," *J. Neurophysiol.* **97**, 3155–3164 (2007).
- 1493
- 1494<sup>30</sup>Korenberg, M. J., "Parallel cascade identification and kernel estimation for nonlinear systems," *Ann. Biomed. Eng.* **19**, 429–455 (1991).
- 1495
- 1496<sup>31</sup>Korenberg, M. J. and Hunter, I. W., "The identification of nonlinear biological systems: LNL cascade models," *Biol. Cybern.* **55**, 125–134 (1986).
- 1497
- 1498
- 1499<sup>32</sup>Landy, M. S. and Graham, N., "Visual perception of texture," in *The Visual Neurosciences*, edited by L. M. Chalupa and J. S. Werner (MIT, Cambridge, MA, 2004), pp. 1106–1118.
- 1500
- 1501
- 1502<sup>33</sup>Lasley, D. J. and Cohn, T. E., "Why luminance discrimination may be better than detection," *Vision Res.* **21**, 273–278 (1981).
- 1503
- AQ504<sup>34</sup>Lennie, P., "Single units and visual cortical organization," *Perception* **27**, 889–935 (1998).
- #1505
- 1506<sup>35</sup>Luce, R. D., "Four tensions concerning mathematical modeling in psychology," *Annu. Rev. Psychol.* **46**, 1–27 (1995).
- 1507
- 1508<sup>36</sup>Luce, R. D., *Response Times: Their Role in Inferring Elementary Mental Organization* (Oxford University Press, New York, 1986).
- 1509
- 1510<sup>37</sup>Maljkovic, V. and Martini, P., "Implicit short-term memory and event frequency effects in visual search," *Vision Res.* **45**, 2831–2846 (2005).
- 1511
- 1512<sup>38</sup>Maloney, L. T. and Zhang, H., "Decision-theoretic models of visual perception and action," *Vision Res.* **50**, 2362–2374 (2010).
- 1513
- 1514<sup>39</sup>Marmarelis, P. Z., "Identification of nonlinear biological systems using Laguerre expansions of kernels," *Ann. Biomed. Eng.* **21**, 573–589 (1993).
- 1515
- 1516<sup>40</sup>Marmarelis, P. Z. and Marmarelis, V. Z., *Analysis of Physiological Systems: The White-Noise Approach* (Plenum, New York, 1978).
- 1517
- 1518<sup>41</sup>Marmarelis, V. Z., *Nonlinear Dynamic Modeling of Physiological Systems* (Wiley IEEE, Piscataway, NJ, 2004).
- 1519
- 1520<sup>42</sup>Mather, G., Pavan, A., Campana, G., and Casco, C., "The motion after-effect reloaded," *Trends Cogn. Sci.* **12**, 481–487, 2008.
- 1521
- 1522<sup>43</sup>Metzger, M. A. and Theisz, M. F., "Forecast: Program to obtain forecasts from subjects for successive values of chaotic time series," *Behav. Res. Methods Instrum. Comput.* **26**, 387–394 (1994).
- 1523
- 1524<sup>44</sup>Mountcastle, V. B., LaMotte, R. H., and Carli, G., "Detection thresholds for stimuli in humans and monkeys: Comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand," *J. Neurophysiol.* **35**, 122–136 (1972).
- 1525
- 1526<sup>45</sup>Murdoch, B. B., *Human Memory: Theory and Data* (Lawrence Erlbaum Associates, MD, 1974).
- 1527
- 1528
- 1529<sup>46</sup>Murray, R. F., Bennett, P. J., and Sekuler, A. B., "Optimal methods for calculating classification images: Weighted sums," *J. Vision* **2**, 79–104 (2002).
- 1530
- 1531
- 1532
- 1533
- 1534<sup>47</sup>Neri, P., "Estimation of nonlinear psychophysical kernels," *J. Vision* **4**, 82–91 (2004).
- 1535
- 1536<sup>48</sup>Neri, P., "Fast-scale adaptive changes of directional tuning in fly tangential cells are explained by a static nonlinearity," *J. Exp. Biol.* **210**, 3199–3208 (2007).
- 1537
- 1538
- 1539<sup>49</sup>Neri, P., "How inherently noisy is human sensory processing?," *Psychon. Bull. Rev.* (in press).
- 1540
- AQ541<sup>50</sup>Neri, P., "Nonlinear characterization of a simple process in human vision," *J. Vision* **9**, 1–29 (2009).
- #1542
- 1543
- 1544<sup>51</sup>Neri, P., "Visual detection under uncertainty operates via an early static, not late dynamic, nonlinearity," *Frontiers in Computational Neuroscience* (to be published).
- AQ545
- #1546<sup>52</sup>Neri, P. and Heeger, D. J., "Spatiotemporal mechanisms for detecting and identifying image features in human vision," *Nat. Neurosci.* **5**, 812–816 (2002).
- 1547
- 1548
- 1549<sup>53</sup>Neri, P. and Levi, D., "Temporal dynamics of directional selectivity in human vision," *J. Vision* **8**, 1–11 (2008).
- 1550
- 1551<sup>54</sup>Neri, P. and Levi, D. M., "Evidence for joint encoding of motion and disparity in human visual perception," *J. Neurophysiol.* **100**, 3117–3133 (2008).
- 1552
- 1553
- 1554<sup>55</sup>Neri, P. and Levi, D. M., "Receptive versus perceptive fields from the reverse-correlation viewpoint," *Vision Res.* **46**, 2465–2474 (2006).
- 1555
- 1556<sup>56</sup>Neuringer, A. and Voss, C., "Approximating chaotic behavior," *Psychol. Sci.* **4**, 113–119 (1993).
- 1557
- 1558<sup>57</sup>Newsome, W. T., "Visual attention: Spotlights, highlights and visual awareness," *Curr. Biol.* **6**, 357–360 (1996).
- 1559
- 1560<sup>58</sup>Newsome, W. T., Britten, K. H., and Movshon, J. A., "Neuronal correlates of a perceptual decision," *Nature (London)* **341**, 52–54 (1989).
- 1561
- 1562<sup>59</sup>Newsome, W. T., Shadlen, M. N., Zohary, E., and Britten, K. H., "Visual motion: Linking neuronal activity to psychophysical performance," in *The Cognitive Neurosciences*, edited by M. S. Gazzaniga (MIT, Cambridge, 1995), pp. 401–414.
- 1563
- 1564
- 1565<sup>60</sup>Palm, G., "On representation and approximation of nonlinear systems. Part II: Discrete time," *Biol. Cybern.* **34**, 49–52 (1979).
- 1566
- 1567<sup>61</sup>Paninski, L., "Convergence properties of three spike-triggered analysis techniques," *Network* **14**, 437–464 (2003).
- 1568
- 1569<sup>62</sup>Parker, A. J. and Newsome, W. T., "Sense and the single neuron: Probing the physiology of perception," *Annu. Rev. Neurosci.* **21**, 227–277 (1998).
- 1570
- 1571<sup>63</sup>Paulin, M. G., "A method for constructing data-based models of spiking neurons using a dynamic linear-static nonlinear cascade," *Biol. Cybern.* **69**, 67–76 (1993).
- 1572
- 1573
- 1574<sup>64</sup>Pelli, D. G., "Noise in the visual system may be early," in *Computational Models of Visual Processing*, edited by M. Landy and A. J. Movshon (MIT, Cambridge, 1991), pp. 147–152.
- 1575
- 1576
- 1577<sup>65</sup>Pelli, D. G., "Uncertainty explains many aspects of visual contrast detection and discrimination," *J. Opt. Soc. Am. A* **2**, 1508–1532 (1985).
- 1578
- 1579
- 1580<sup>66</sup>Ringach, D. L., "Tuning of orientation detectors in human vision," *Vision Res.* **38**, 963–972 (1998).
- 1581
- 1582<sup>67</sup>Sakai, H. M., Wang, J. L., and Naka, K., "Contrast gain control in the lower vertebrate retinas," *J. Gen. Physiol.* **105**, 815–835 (1995).
- 1583
- 1584<sup>68</sup>Schetzen, M., *The Volterra and Wiener Theories of Nonlinear Systems* (Wiley, New York, 1980).
- 1585
- 1586<sup>69</sup>Schneider, K. A. and Komlos, M., "Attention biases decisions but does not alter appearance," *J. Vision* **8**, 1–10 (2008).
- 1587
- 1588<sup>70</sup>Schölkopf, B. and Smola, A. J., *Learning with Kernels* (MIT, Cambridge, MA, 2002).
- 1589
- 1590<sup>71</sup>Shapley, R. and Enroth-Cugell, C., "Visual adaptation and retinal gain controls," *Prog. Retinal Res.* **3**, 263–346 (1984).
- 1591
- 1592<sup>72</sup>Shub, D. E. and Richards, V. M., "Psychophysical spectro-temporal receptive fields in an auditory task," *Hear. Res.* **251**, 1–9 (2009).
- 1593
- 1594<sup>73</sup>Simpson, W. A., Braun, W. J., Barga, C., and Newman, A. J., "Identification of the eye-brain-hand system with point processes: A new approach to simple reaction time," *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 1675–1690 (2000).
- 1595
- 1596
- 1597<sup>74</sup>Smithson, M., "Judgment under chaos," *Org. Behav. Hum. Decis. Process* **69**, 58–66 (1997).
- 1598
- 1599
- 1600<sup>75</sup>Solomon, J. A., "Noise reveals visual mechanisms of detection and discrimination," *J. Vision* **2**, 105–120 (2002).
- 1601
- 1602<sup>76</sup>Spekreijse, H. and Oosting, H., "Linearizing: A method for analysing and synthesizing nonlinear systems," *Kybernetik* **7**, 22–31 (1970).
- 1603
- 1604<sup>77</sup>Spillmann, L., "Foveal perceptive fields in the human visual system measured with simultaneous contrast in grids and bars," *Pfluegers Arch.* **326**, 281–299 (1971).
- 1605
- 1606
- 1607<sup>78</sup>Stewart, I. N. and Peregoy, P. L., "Catastrophe theory modeling in psychology," *Psychol. Bull.* **94**, 336–362 (1983).
- 1608
- 1609<sup>79</sup>Ta'edd, L. K., Ta'eed, O., and Wright, J. E., "Determinants involved in the perception of the Necker cube: An application of catastrophe theory," *Behav. Sci.* **33**, 97–115 (1988).
- 1610
- 1611
- 1612<sup>80</sup>Tanner, W. P., "Physiological implications of psychophysical data," *Ann. N.Y. Acad. Sci.* **89**, 752–765 (1961).
- 1613
- 1614<sup>81</sup>Thelen, E. and Smith, L. B., *H Dynamic Systems Approach to the Development of Cognition and Action* (MIT, Cambridge, MA, 1995).
- 1615
- 1616<sup>82</sup>Thomas, J. P. and Knoblauch, K., "Frequency and phase contributions to the detection of temporal luminance modulation," *J. Opt. Soc. Am. A Opt. Image Sci. Vis* **22**, 2257–2261 (2005).
- 1617
- 1618
- 1619<sup>83</sup>Tjan, B. S. and Nandy, A. S., "Classification images with uncertainty," *J. Vision* **6**, 387–413 (2006).
- 1620
- 1621<sup>84</sup>van Hateren, J. H. and Snippe, H. P., "Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells," *Vision Res.* **41**, 1851–1865 (2001).
- 1622
- 1623
- 1624<sup>85</sup>von der Malsburg, C., "The what and why of binding: The modeler's perspective," *Neuron* **24**, 95–104 (1999).
- 1625

- 1626** <sup>86</sup>Westwick, D. T. and Kearney, R. E., *Identification of Nonlinear Physiological Systems* (Wiley IEEE, Piscataway, NJ, 2003). **1631**
- 1627** <sup>86</sup>Westwick, D. T. and Kearney, R. E., *Identification of Nonlinear Physiological Systems* (Wiley IEEE, Piscataway, NJ, 2003). **1632**
- 1628** <sup>87</sup>Wiener, N., *Cybernetics or Control and Communication in the Animal and the Machine* (Wiley, New York, 1948). **1633**
- 1629** <sup>87</sup>Wiener, N., *Cybernetics or Control and Communication in the Animal and the Machine* (Wiley, New York, 1948). **1634**
- 1630** <sup>88</sup>Wilson, H. R., “Non-Fourier cortical processes in texture, form, and motion perception,” in *Cerebral Cortex*, edited by P. S. Ulinski, E. G. Jones, and A. Peters (Plenum, New York, 1999), Vol. 14, pp. 445–477. **1635**
- <sup>89</sup>Yeshurun, Y., Carrasco, M., and Maloney, L. T., “Bias and sensitivity in two-interval forced choice procedures: Tests of the difference model,” *Vision Res.* **48**, 1837–1851 (2008).

**AUTHOR QUERIES — 010091CHA**

- #1 Au: Please verify insertions in affiliation.
- #2 Au: Please check change of “1-D” to “1D” and define if possible.
- #3 Au: Please verify italic “k” associated with the word “trials” throughout.
- #4 Au: Please verify edits in the sentence “first-order kernels for priming ...” to make sure your meaning was preserved.
- #5 Au: Please define SNR if possible.
- #6 Au: Please define LN if possible.
- #7 Au: Please verify edits in the sentence “To see this explicitly for a second-order ...” to make sure your meaning was preserved.
- #8 Au: Please define LNL if possible.
- #9 Au: Please verify use of single left parens in Eq. (25).
- #10 Au: Please define MT if possible.
- #11 Au: Please spell out 3D if possible.
- #12 Au: Please verify use of single left parens in math notation that immediately follows.
- #13 Au: Please verify insertion of volume no. in Ref. 5.
- #14 Au: All references must come in alphabetical order. Please check renumbering of Refs. 35 and 36; Refs. 49 and 50; Refs. 54 and 55; and Refs. 64 and 65.
- #15 Au: Please update Ref. 49.
- #16 Au: Please supply the full journal title, the CODEN, and/or ISSN for the journal in Ref. 51.
- #17 Au: Please update Ref. 51.
- #18 AU: Please reword text (or caption of Fig. 2) without color words, as readers of print will see black and white figures.
- #19 AU: Please reword text (or caption of Fig. 3) without color words, as readers of print will see black and white figures.
- #20 AU: Please reword text (or caption of Fig. 4) without color words, as readers of print will see black and white figures.
- #21 AU: Please reword text (or caption of Fig. 5) without color words, as readers of print will see black and white figures.
- #22 AU: Please reword text (or caption of Fig. 6) without color words, as readers of print will see black and white figures.
- #23 AU: Please reword text (or caption of Fig. 7) without color words, as readers of print will see black and white figures.
- #24 AU: Please reword text (or caption of Fig. 8) without color words, as readers of print will see black and white figures.
- #25 Au: Please verify insertion of “cascade” in the sentence “if the phenomenon of interest is well...” in Fig. 8 caption.
- #26 AU: Please reword text (or caption of Fig. 9) without color words, as readers of print will see black and white figures.
- #27 AU: Please reword text (or caption of Fig. 10) without color words, as readers of print will see black and white figures.