Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 46 (2006) 259–266
Vision Research
www.elsevier.com/locate/visres

Singularities in the inverse modeling of 2AFC contrast discrimination data

Mikhail Katkov, Misha Tsodyks, Dov Sagi *

Department of Neurobiology/Brain Research, The Weizmann Institute of Science, Rehovot 76100, Israel

Received 16 January 2005; received in revised form 16 September 2005

Abstract

Analytical calculations show that two-alternative force-choice data are not always suitable for specifying the parameters of the underlying discrimination model. Experimentally, we show here that in the case of contrast discrimination in humans, a variety of models spanning a large range of parameters can explain the data within an experimental error. Monte-Carlo simulations indicate that the number of trials in psychophysical experiments is not the limiting factor in estimating the parameters in contrast discrimination. These results can therefore explain the contradictory conclusions made by different groups about the relationship between the response to contrast and the noise amplitude.

Keywords: Contrast discrimination; 2AFC; Modeling; Psychophysics; Signal detection theory; Noise

1. Introduction

Human performance in psychophysical tasks, where stimuli with different intensities have to be discriminated, depends on the magnitude of the internal response in the brain, and on its trial-by-trial variability (Green & Swets, 1966). The specification of these components is important for understanding the characteristics of the internal modules that process the presented stimuli. However, these components cannot be measured independently in a psychophysical experiment. Recent attempts to resolve this issue in the case of contrast discrimination led to conflicting results, indicating that the internal noise magnitude is either increasing with contrast (Kontsevich, Chen, & Tyler, 2002a; Lu & Dosher, 1999) or constant (Gorea & Sagi, 2001), or approximately constant (Foley & Legge, 1981). Here, we show a limitation of the two-alternative force-choice (2AFC) method in separating signal and noise in contrast discrimination and provide an account for the various results reported in the literature.

In the 2AFC procedure, an observer reports which one of two stimuli with different intensities, presented sequentially in a single trial, contains the target (e.g., the stimulus with the higher intensity). Signal detection theory (SDT) assumes that each stimulus evokes a one-dimensional internal response that varies across trials, and that the observer's performance depends on the distribution of internal responses. Two simplifying assumptions are usually added in practice: (1) the distribution of the internal responses is Gaussian, and (2) the decision is made by comparing the internal responses for the two stimuli. The percentage of correct discriminations under these assumptions is given by Green and Swets (1966) and Thurstone (1927):

$$P_{s_1,s_2} = \Phi\left(\frac{R_{s_2} - R_{s_1}}{\sqrt{\sigma_{s_1}^2 + \sigma_{s_2}^2}}\right), \qquad (1)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt,$$

where $\Phi(x)$ is a cumulative distribution function for normal distribution, R_{s_1}, R_{s_2}, are the mean internal responses, σ_{s_1}, σ_{s_2} are trial-by-trial standard deviations of the internal

* Corresponding author. Tel.: +972 8 934 3747; fax: +972 8 934 4131.
E-mail address: Dov.Sagi@Weizmann.ac.il (D. Sagi).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2005.09.022

responses, and P_{s_1,s_2} is the probability of a correct discrimination between stimulus s_1 and s_2. The values of P_{s_1,s_2} are measured in the experiment, whereas the R's and σ's are the parameters of the model. Eq. (1) has four unknown variables for one pair of stimulus intensities, leading to an ambiguous solution for the R's and σ's. By increasing the number of stimuli and pairs, it is possible to define an overdetermined set of equations, and thus, to estimate the parameters of the model (McNicol, 1972; Torgerson, 1958).

A mathematical analysis of the above SDT equations shows two types of problems when attempting to recover model parameters (i.e., the internal response and noise values). First, there is the possibility of multiple solutions for the same set of data (Iverson, 1987), as for example when solving a quadratic equation. Second, since the data have only limited precision and constitute only an estimate of the actual performances, very different solutions may correspond to very similar performances (Katkov, Tsodyks, & Sagi, submitted). Thus, two data sets with small differences, which are within experimental errors, may result from very different values of parameters. The models that give rise to such data are termed singular models. Not all sets of performances correspond to singular models. For non-singular data, small finite deviations in the measurements produce small finite changes in the model's parameters. These non-singular models are expected to exhibit a higher stability in the presence of experimental errors, i.e., the parameters of such models should be relatively easy to estimate. The mathematical analysis of Katkov et al. (submitted) identified four families of singular models, related to the above SDT equations, among which is the family that includes models with constant noise. The term family is defined here as a set of models that explain the data, within the experimental error. Thus, the size of a family depends on the number of trials. When a family includes a singular model, all models that are not distinguishable from the singular model are also included, and the range of parameters explaining the data expands substantially, making a reliable estimation of the "true" model almost impossible.

As will be shown here by way of simulations, 2AFC performances that are consistent with the constant noise model are also in agreement with non-constant noise models when the measurement error is taken into account. This ambiguity in the solutions can be resolved by applying additional constraints on the underlying model, as is usually the practice, by assuming a specific functional dependency of noise on the response. An earlier attempt to construct an unconstrained contrast discrimination model was made by Foley and Legge (1981), who used low-contrast grating stimuli at around contrast detection threshold. Here, we analyzed contrast discrimination data covering a large range of contrasts, without putting any constraints on the underlying model. Specifically, we checked whether the contrast discrimination data can be described by a constant noise model. We also checked the effects of a limited number of trials on the parameter estimation. To this end, we estimated the parameters of the model that describes the experimental data, and checked the sensitivity of the parameters to small variations in performance. To check the effects of finite sampling, we performed Monte-Carlo simulations of 2AFC experiments in the SDT framework for different types of models, including models describing the measured data.

2. Methods

2.1. Psychophysics

The temporal 2AFC paradigm was used in the psychophysical experiment (Fig. 1A). The stimuli consisted of a

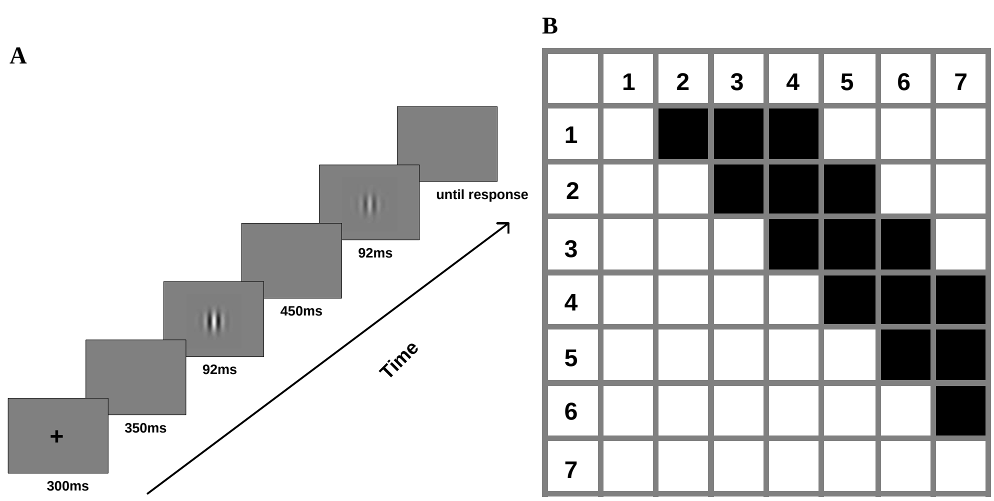

Fig. 1. Experimental scheme. (A) Schematic representation of the presented stimuli. The following sequence of frames was presented to the observers: fixation point (300 ms), blank screen (350 ms), stimulus (92 ms), blank screen (450 ms), stimulus (92 ms), and blank screen until the response from the observer was obtained. (B) Contrast pairs used in the experiment. The row and column titles represent the index of the stimulus. The dark squares represent pairs of stimuli (defined by the row and column indexes) used in the comparisons.

single Gabor Patch (GP) added at the center of the screen to an otherwise uniform gray display with a 36 cd/m^2 mean luminance. A GP is the product of a cosine grating and a Gaussian envelope. The spatial frequency of the carrier grating was 9.5 cpd and the width of the envelope was equal to the period of the grating. The contrast of the GPs (defined here as the GP amplitude divided by the screen's mean luminance) was varied. Two stimuli with different contrasts were presented to the observer during a trial and the observer had to report which of the two stimuli appeared to have a higher contrast. The stimuli were presented on a Mitsubishi Diamond Pro 900u computer monitor (19 in.).

The experiments were performed with each observer covering the high (less than 61.5%) and middle (slightly above detection threshold) contrast ranges. A total of seven contrast levels were chosen in each range. Four neighboring contrasts were used in pair-wise discriminations for each GP contrast (schematically represented in Fig. 1B). All pairs chosen for comparison were checked so that they had neither too many nor a too small number of errors (60–90% of correct responses) in a short session. If some of the performances were too high or too low, then the set of contrasts was corrected, and the test was repeated.

Three observers (17–27 years old) participated in the experiments and all of them had normal or corrected-to-normal vision. Approximately 600 pair-wise discriminations were performed for each pair by each observer.

2.2. Monte-Carlo simulations

The performances of a hypothetical observer were modeled using Monte-Carlo simulations. For this purpose, a number of stimuli and a set of pairs for comparison were fixed. Then, a hypothetical observer was assigned a mean internal response and noise amplitudes for each stimulus ("true" model), which were kept *constant* for all trials during the simulations. The mean internal response as a function of the stimulus index was of two types: (1) a linear function of the index and (2) a function generated by sorting uniformly distributed random numbers in increasing order. Noise amplitudes as a function of the stimulus index were constant or sorted uniformly distributed random numbers either in descending or ascending order. Additionally, Monte-Carlo simulations were performed during the sensitivity test (explained later) for the parameters of models derived from the experimental data. For the purpose of normalization, the smallest and highest mean internal responses were set to 0 and 1, respectively. Adding a common constant to all R values and scaling all the parameters by the same value does not change the model's performance (Eq. (1)).

The simulation procedure consisted of modeling 2AFC trials. A fixed number of trials for each pair was chosen. During each trial, an internal response for each stimulus in the pair was computed using a generator of normally distributed random numbers. The mean and the standard deviation of the generator were set to the mean internal response and noise amplitude for a corresponding stimulus. The trial was marked as correct when the differences between the internal responses had the same sign as the differences between the mean internal responses assigned to the stimuli. At the end of the simulation, the rate of correct responses was computed for each stimulus pair.

2.3. Parameter estimation

To estimate the parameters of the model (R's and σ's in Eq. (1)), we computed an observer model that minimizes a cost function based on the differences between the measured and modeled performances. We used the following cost function:

$$E(x) = \sum_{(i,j)} \left[\Phi\left(\frac{R_j - R_i}{\sqrt{\sigma_i^2 + \sigma_j^2}} \right) - P_{i,j} \right]^2, \tag{2}$$

where x is the combined vector of the mean of the internal responses (R) and noise amplitudes (σ), and $P_{i,j}$ is the correct discrimination rate measured for the pair (i,j). The summation is performed over all pairs chosen for the pair-wise discrimination. In the constant noise model, all σ_i values were forced to be a single value ($\sigma_i \equiv \sigma$) for estimating the parameters.

To determine the sensitivity of the estimated parameters to experimental errors, one should check the deviation of the estimated parameters as a result of small changes in performance. In the singular case we expect large changes in the parameters. To this end, we modified each $P_{i,j}$ obtained in the simulation or measured in the experiment by adding a normally distributed random value with a zero mean and a standard deviation equal to the standard error computed from simulated or measured data, correspondingly. Then, parameters were estimated for the modified performances. This procedure was repeated several times. The sensitivity of the parameters was also tested by repeating the simulation–estimation procedure several times, starting with the same initial conditions. The results were compatible in both cases.

The compatibility of the data with a constant noise model can be checked by the χ^2 goodness of fit presented in Kontsevich et al. (2002a). The estimations of the probability of correct responses in the experiment after n trials have a binomial distribution over different realizations, with a mean P (predicted by the model, see Eq. (1)) and a variance $P(P-1)/n$. A single pair χ^2 can be computed as

$$\chi^2 = \frac{(P - \frac{m}{n})^2}{P(1-P)/n}, \tag{3}$$

where m is the number of correct responses in n trials.

The overall χ^2 error is the sum over all pairs. The number of degrees of freedom is the number of the experimental comparisons minus the number of estimated independent parameters. In the case of a constant noise model the number of independent variables is the number of stimuli minus one, since one variable can be used for choosing the initial point in the internal response axis and does not change the performances.

Each observer performed experiments in two ranges of contrasts. The recovered constant noise models were used to combine the results obtained in different contrast ranges into one model. The models for different ranges were first scaled to have the same noise amplitudes and then a constant was added to all values of the mean internal responses for each model, producing the best fit (in the least squares sense) for the following equation with fitting parameters γ, α_1, α_2:

$$R(c) = \alpha_1 c^{\gamma} + \alpha_2. \quad (4)$$

Eq. (4) represents the high-contrast limit of the gain control model (Legge & Foley, 1980).

3. Results

3.1. Psychophysics

The parameters of the model best fitted to the discrimination data and their sensitivity to the experimental error are presented in Fig. 2 for three observers (see details in Section 2). It can be seen that the noise amplitudes, obtained for the slightly perturbed performances, have a large variability. There are two possible explanations for the parameter variability: the limited number of trials, and the presence of singularity (see Section 1). In the latter case, increasing the number of trials does not change variability much. In order to check which explanation is valid here, we tested if any singular model can describe the data. To this end, we fixed all noise amplitudes to be a single value during the fitting and found the best constant noise model (a singular model) that describes the data. We found that there are no appreciable differences in the performances predicted by the two models fitted to the data (the constant noise model and the unconstrained one). It is also reflected in the goodness of fit test. The goodness of fit test, when applied to the performances predicted by the constant noise model best describing the data, showed a significance level approaching 1.0 in five experiments and was around 0.6 in the sixth one, meaning that the data are very well described by the constant noise models. Thus, the variability of the model parameters is mostly explained by its singularity and not by the finite sampling. Moreover, the obtained noise amplitudes for modified performances were either decreasing or increasing functions of the contrast, similar to the simulation results obtained for the constant noise models (see Section 3.2 below, Figs. 4B and 5).

Three observers performed the experiments, each in two ranges of contrast. The parameters of the constant noise models obtained in the two experiments, performed by each observer, were merged and are presented in Fig. 3. The noise amplitude of these plots is equal to one for all stimuli and the mean internal response for the smallest contrast is set to zero for the purpose of normalization. Note

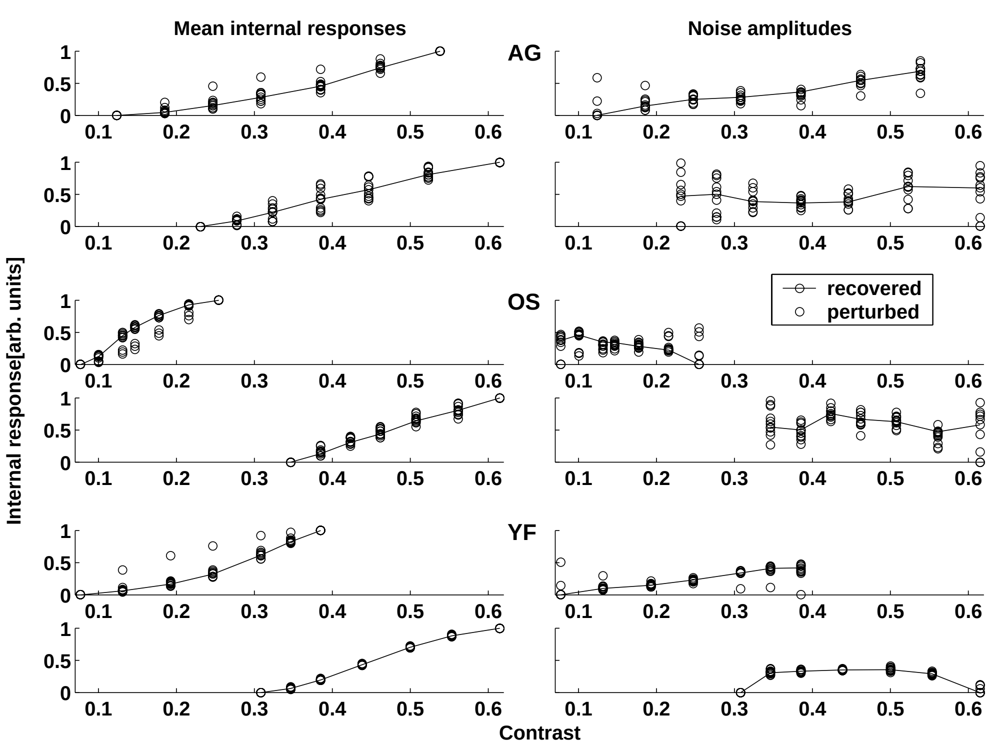

Fig. 2. Stability of the models obtained for the experimental data. The mean internal response and noise are presented on the plots in the left and right columns, respectively. Each row represents different experiments (observer and contrast range). The plots for one observer are grouped together. The solid lines denote the recovered model and the open circles denote the parameters best describing the perturbed performances. A large variability can be seen in all cases, except the case represented by the last row.

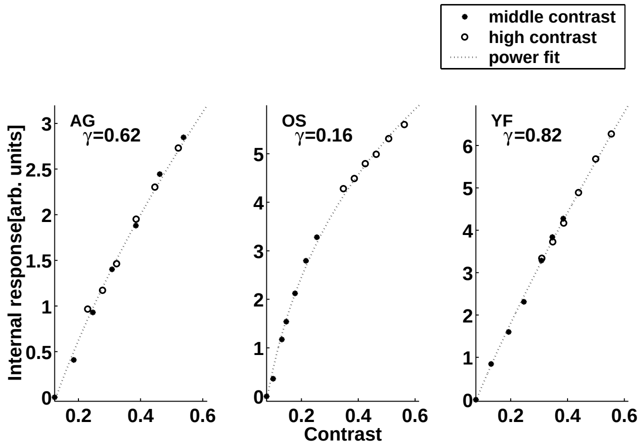

Fig. 3. Combined mean internal response functions are presented for the three observers who participated in the experiment (different plots). The noise amplitude for all plots is set to one. The dots represent the mean internal response values for the range of middle contrast, open circles—for the range of high contrast, and the dotted curves represent the power fit (Eq. (4)).

that the fitted parameters in the overlapping regions of contrast, obtained in independent experiments, are consistent. There is a good fit of the parameters to the power function for all observers, as suggested in the literature for the gain control model (see Section 2 for details). However, since the constant noise model is a singular one, and a singularity implies that a large family of models can explain the same data, the 2AFC method does not allow a unique specification of model parameters, as seen in Figs. 2 and 5.

3.2. Monte-Carlo simulations

The estimated model parameters are extremely sensitive to small changes in performance when the true model approaches a singular one (Katkov et al., submitted). However, even for non-singular models, there is a deviation of the estimated model from the true one. To study the effects of finite sampling, we performed a Monte-Carlo simulation of psychophysical experiments using the SDT model previously described (see details in Section 2). Based on the simulated performances, we estimated the parameters of the model. To this end, we minimized the cost function, which is defined as the sum of the squared differences between performances predicted by the model and the simulated ones. The comparisons between the true parameters and the ones estimated by the model are illustrated in Fig. 4.

It can be seen that the mean internal responses and noise amplitudes were recovered reliably in all cases except where the true model is a singular one (Fig. 4B). In this case, the shape of the recovered noise amplitude is very different from the true one. In the following text, we focused on the constant noise singular models, since the contrast discrimination data are compatible with it (see Section 3.1).

The variability of the model parameters estimated in different runs (simulated performances followed by the estimation of the model parameters using the same "true"

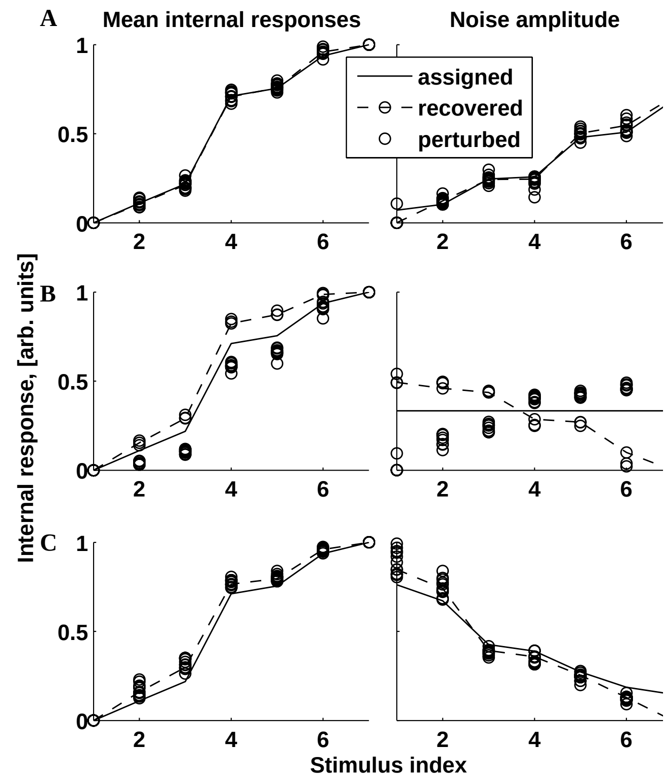

Fig. 4. Stability of the recovered mean response values and noise amplitudes (simulation). The mean internal responses and noise amplitudes are presented on the plots in the left and right columns, respectively, for three noise dependencies on stimuli: (A) increasing, (B) constant, and (C) decreasing. The solid lines represent the values chosen before the simulation ("true" model), and the dashed lines represent the recovered ones. The open circles denote the parameters best describing the perturbed performances. The models were recovered from simulated performances that were obtained with 2000 trials per pair.

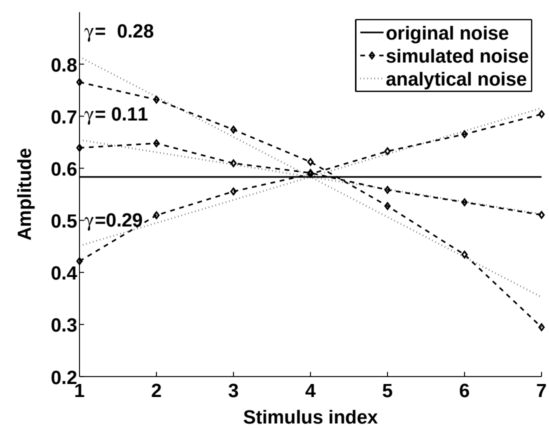

Fig. 5. Examples of noise amplitudes computed analytically and estimated from simulated performances. The picture shows a comparison of noise amplitudes estimated from the simulated data and ones obtained by analytical computation for different values of parameter γ (Eq. (5)). The "true" model is a constant noise model here. The performance for all models presented here is almost identical.

model) exhibits behavior expected from the analytical calculations: the constant noise true model (one of the singular models) shows a large variability of estimated parameters, whereas the other models exhibit much less variability. Moreover, the shape of different solutions can also be predicted from the analytical calculations. For example, for the (singular) constant noise model, the corresponding family of models can be parametrized in the following way (Katkov et al., submitted):

$$R'_i = \frac{(1+\gamma R_i)R_i}{1+\gamma}$$
$$\sigma'_i = \frac{(1+2\gamma R_i)\sigma}{1+\gamma}. \quad (5)$$

Here, R_i is the mean internal response of the original singular model, σ represents the noise in the constant noise true model, R'_i and σ'_i are the parameters of a model from the family parametrized by γ, which result in similar performances (formally $\gamma \in [-\frac{1}{2}, +\infty]$, and $\gamma = 0$ corresponds to the original constant noise model). Fig. 5 compares analytical models obtained for several γ values (Eq. (5)) and those obtained in the simulation. It can be seen that every model obtained in the simulations can be closely approximated by one of the models in the family described by Eq. (5). Thus, we can also conclude that for the limited number of trials used in the psychophysical experiments, the sensitivity of the model parameters to the perturbation of the performances in the singular case is higher than that defined by the finite sampling in the non-singular case. Thus, if the data are compatible with a singular model, any model from the corresponding family will describe the data well. Moreover, a family of a singular model spans a larger space of parameters than models defined by finite sampling.

Next, we investigated whether statistical tests can be used to test whether a computed model belongs to a singular family or not. More specifically, whether the range of models that are compatible with a given singular model is dependent on the sampling size. The smaller the sampling size is, the higher is the probability that a specific set of performances are compatible with some singular set of performances. We tested this possibility on the simulated data, applying a χ^2 goodness of fit test (see details in Section 2) to the constant noise model best describing the estimated performances (in the least squares sense). When the true model had constant noise (a singular model) and the reconstructed parameters were very different (Fig. 4B), the significance level p was approximately one. For non-singular models, where the model parameters are reliably reconstructed, the test showed $p < 10^{-15}$, thus rejecting the constant noise model. Therefore, the statistical tests can indeed be used to specify the type of the recovered model (singular/non-singular).

The simulation results show that predictions of the analytical calculations hold for a realistic number of trials. Specifically, there is a qualitative difference in the sensitivity of the estimated parameters regarding the experimental errors between the two classes of models. The simulation results also show that the least square estimator can reliably reconstruct the parameters of the non-singular model, and that the χ^2 goodness of fit test can be used to determine which type of the model (singular or not) is more appropriate to describe the data.

4. Discussion

Our analytical calculations show that a singularity exists in the parameterization of the performances using the 2AFC procedure (Katkov et al., submitted). This singularity leads to an ambiguity in the estimation of the parameters for some types of models. For example, for the model with noise amplitudes independent of the stimulus intensities, there is a family of other models with large differences in parameters and close performances. Note that when performance differences between two models are smaller than the standard deviation of performances obtained in the experiment, the models are practically indistinguishable. In the singular case the span of parameters, bounded by the standard deviation of performances, decreases with the number of trials very slowly, remaining wide even for an unrealistically large number of trials. On the other hand, Monte-Carlo simulations show that the parameters of the models outside singular families can be reliably estimated by minimizing the sum of the squared differences between the predicted and measured performances (the least squares estimator). Whether the model is from the family of singular models or not can be established using the χ^2 goodness of fit test.

An analysis of the SDT model shows that it is practically impossible to distinguish between models of the same family. The finite number of trials make the situation even more dramatic. For example, models with internal responses having a Poisson distribution with a mean above 5–10, which can be approximated by the Gaussian distribution having a variance equal to the mean, produce performances for pair-wise discrimination that are almost identical to a certain Gaussian constant noise model. Thus, within the number of trials used in the psychophysical experiments, it will be almost impossible to discriminate between these two types of models in the 2AFC experiments. Nevertheless, ambiguity can be removed if some constraint is applied to the model parameters. For example, an a priori assumption that the noise is constant yields a unique model. An a priori assumption of Poisson noise leads to another unique model. Thus, assumptions used in different versions of the Thurstonian scaling method (like any other assumptions) may produce a unique solution (as in Foley & Legge (1981)), which can be unique, however, only under the assumptions used in the method.

The experimental results obtained in this work, in the ranges of middle and high contrasts, are well described by the constant noise model. This is consistent with some of the results presented in the literature (Foley & Legge, 1981; Gorea & Sagi, 2001). Nevertheless, the present analysis shows that the true model for human observers is not necessarily a constant noise model, since this model is a singular one, and for every constant noise model there are many other models belonging to one family (see Eq. (5)) that can describe the data as well (Katkov et al., submitted). This finding can explain the variety of results obtained by different groups.

Several authors modeled contrast discrimination using a noise function that increases with internal response. Lu and Dosher (1999) used an equivalent noise paradigm employing external noise. The experimental results were modeled using a processing scheme that included both multiplicative and additive noise, as well as transduction nonlinearities. The model used is mathematically equivalent to a theory of contrast gain control (Legge & Foley, 1980) with constant noise (Lu & Dosher, 1998). Thus, it is mute in the context of the present discussion regarding the dependency of noise on the response magnitude. Kontsevich et al. (2002a) used an SDT model, assuming that the mean internal response and the noise amplitudes are power functions of contrast. The 2AFC paradigm was used in that work and thus, a singularity problem may appear. However, the power-function assumption used by Kontsevich et al. (2002a) constrained the model parameters and could therefore hide singularities. In Fig. 6 we present some simulations that illustrate this point. We assume here that the true model is of the Legge and Foley (1980) type, as shown in Fig. 6A, and that internal noise is *independent* of the response (a singular model). A simulation of the contrast discrimination task was performed (see Section 2) with contrast pairs consisting of three base contrasts, each paired with 20 contrasts equally spaced between zero and one (200 trials per measurement). Based on the obtained performances, we estimated the parameters of the model presented in Kontsevich et al. (2002a):

$$R(c) = c^p, \quad \sigma(c) = kR(c)^q.$$

The estimated parameters obtained from repeating the simulation/estimation procedure are presented in Fig. 6C. The obtained models exhibit a behavior very different from that of the original true model. To estimate the confidence interval of the recovered parameters, we chose one of the models and performed several simulation/estimation procedures. The variability of the estimated models is presented in Fig. 6D. The range of estimated parameters and χ^2 are close to that presented in Kontsevich et al. (2002a). As can be seen, the variability of the constrained parameters and models is not wide. In contrast, when no constraints are applied to the model parameters, their variability is much larger (Fig. 6E), as expected for a singular model. Thus, possibly the conclusion of Kontsevich et al. (2002a) could be affected by the constraint imposed on the model.

A new approach for analyzing noise in detection tasks was recently developed by Gorea and Sagi (2001). It was

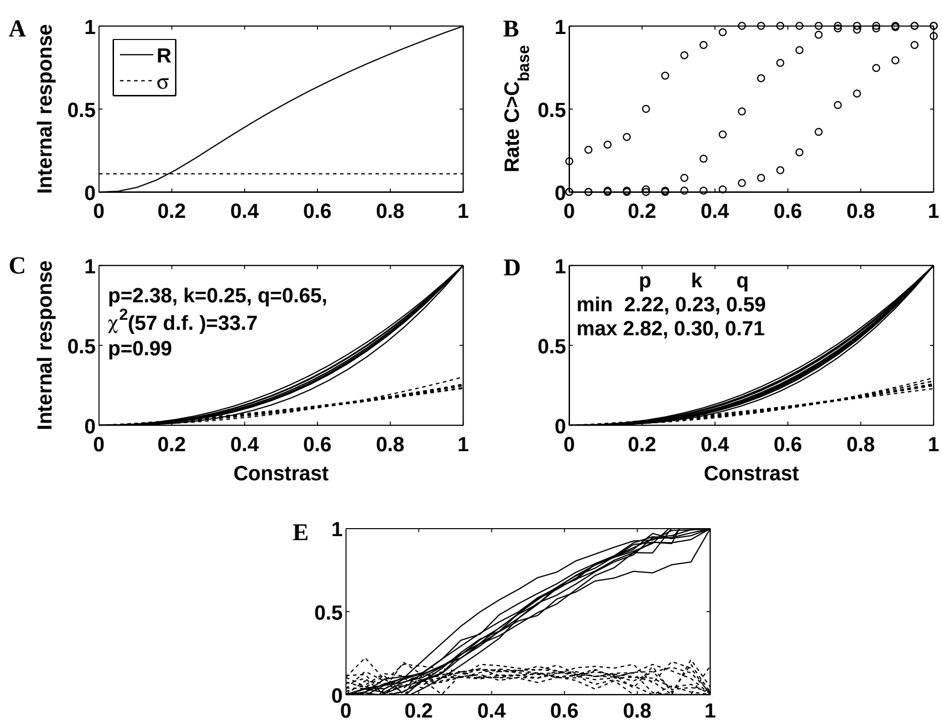

Fig. 6. Power fit of the contrast gain control model. Power fits to data simulated from the Legge and Foley (1980) model with constant noise. (A) The transducer function used: $R \sim c^{2.6}/(c^{2.0} + 0.1)$. (B) Psychometric curves obtained from simulated performances of the model using 200 trials for each contrast pair. Psychometric curves are shown for three base contrasts: 0.25, 0.5, 0.75. (C) Best-fitted power function from ten simulations, showing a very small variability of predictions. The average parameters for the fit are presented in the inset (see details in the text). (D) The variability of estimated parameters for one of the models presented in plot (C). (E) The variability of estimated parameters if no constraints are applied. All the models were obtained from fitting simulated performances (200 trials). Each curve represents the estimated model from separately simulated performances. The variability of estimated models gives an estimate of the confidence intervals. Minimal and maximal values of estimated parameters are presented in the inset.

suggested that observers use a fixed internal reference (R_{ref}) when making decisions regarding the presence or absence of equally probable targets that are presented within the same behavioral context (Gorea & Sagi, 2000; Sagi & Gorea, 2004). This reference can be estimated from the false-alarm rates associated with the different targets. False-alarm rates in turn depend on the noise amplitude for non-target stimuli (σ) and on the internal reference (according to SDT, $Z_{FA} = (R_{ref} - \bar{R}_{no\ target})/\sigma$, where $\bar{R}_{no\ target}$ is the mean internal response in the absence of a target). Thus, in the presence of a fixed reference, the task associated with the larger σ is predicted to produce a higher false-alarm rate, with Z_{FA} inversely proportional to its corresponding σ (see the discussion in Kontsevich, Chen, Verghese, & Tyler (2002b) & Gorea & Sagi (2002)). This prediction was recently confirmed in experiments using external noise, where the paired targets had different external-noise amplitudes (with amplitude ratios up to 3), thus effectively producing different internal noise levels (Sagi & Gorea, 2004). To examine the dependency of internal noise on contrast, Gorea and Sagi (2001) paired targets with different base-contrasts (non-target stimuli) and found equal false-alarm rates, as predicted by equal noise amplitudes. These results are in agreement with the results of the present study by showing constant noise as a function of contrast. Since the Gorea and Sagi (2001) method does not use the 2AFC procedure, the corresponding parameterization is not constrained by the singularities reported here.

Despite the fact that we do not know the exact model parameters, the excellent data fit suggests that the basic assumptions underlying the model (Gaussian distribution of internal responses and the decision made by comparing internal responses) are reasonable. Note that when the noise amplitude is independent of the stimulus intensity, the decision strategy based on comparing the responses used in our model is equivalent to the decision based on a likelihood ratio and hence is optimal (Green & Swets, 1966), i.e., our results indicate that human performance on the contrast discrimination task approximates that of an ideal observer. Also note that the compatibility of the data with the models, where the noise amplitudes are independent of the stimulus in the contrast discrimination task, allows for characterization (still ambiguous) of the transducer function by other techniques such as the threshold versus contrast curve technique.

5. Conclusion

In the present work, we analyzed a model of human performance in the 2AFC contrast discrimination task within the SDT framework. We found that the model parameters describing the discrimination data exhibit a singularity. This singularity does not allow estimating the parameters of the model with reasonable precision. Nevertheless, in the Monte-Carlo simulations, we found that if the true model does not belong to any family of singular models, the least squares estimator can be used to determine the parameters of the model with good precision, for an experimentally reasonable number of trials. Since the sensitivity of the parameters to changes in performances is much higher in the singular case than in the non-singular one, even for a small number of trials, it is possible to use the goodness of fit test to separate these two types of models. Any prior assumption made regarding the model parameters can hide singularities and as a result, may lead to a unique model. The resulting model, however, is unique only within the context of the assumptions made.

Acknowledgments

The authors thank Annat Judelman for helpful discussions. This research was supported by a grant from the Pearl M. Levine Trust, Irving B. Harris Foundation, by the Israel Academy of Science, and by the Bi-national US–Israel Science Foundation.

References

Foley, J. M., & Legge, G. E. (1981). Contrast detection and near threshold discrimination in human vision. *Vision Research, 21*, 1041–1053.

Gorea, A., & Sagi, D. (2000). Failure to handle more than one internal representation in visual detection tasks. *Proceedings of the National Academy of Sciences of the United States of America, 97*(22), 12380–12384.

Gorea, A., & Sagi, D. (2001). Disentangling signal from noise in visual contrast discrimination. *Nature Neuroscience, 4*(11), 1146–1150.

Gorea, A., & Sagi, D. (2002). Reply to 'The unique criterion constraint: a false alarm?'. *Nature Neuroscience, 5*(8), 707–708.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Iverson, G. J. (1987). Thurstonian psychophysics: Case III. *Journal of Mathematical Psychology, 31*, 219–247.

Katkov, M., Tsodyks, M., & Sagi, D. (submitted). Analysis of a two-alternative force-choice signal detection theory model. *Journal of Mathematical Psychology*.

Kontsevich, L. L., Chen, C. C., & Tyler, C. W. (2002a). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research, 42*(14), 1771–1784.

Kontsevich, L. L., Chen, C. C., Verghese, P., & Tyler, C. W. (2002b). The unique criterion constraint: a false alarm? *Nature Neuroscience, 5*(8), 707.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70*(12), 1458–1471.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*(9), 1183–1198.

Lu, Z. L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A, 16*, 764–778.

McNicol, D. (1972). *A primer of signal detection theory*. London: George Allen & Unwin Ltd., Sydney: Australian Publishing Company.

Sagi, D., & Gorea, A. (2004). The unique decision criterion: constant internal response or false-alarm rate? *Journal of Vision, 4*, 454.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychophysical Review, 34*, 273–287.

Torgerson, W. S. (1958). *Theory and method of scaling*. New York: Wiley.