Optimizing data acquisition for MLDS: when is it valid to take a short-cut?

Guillermo Aguilar, Jakob Grünwald and Marianne Maertens

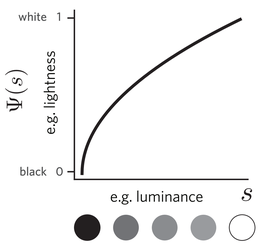
Computational Psychology, Technische Universität Berlin, Germany

Contact
guillermo.aguilar@mail.tu-berlin.de

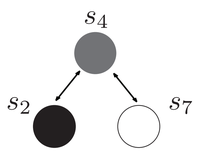
Deutsche Forschungsgemeinschaft
German Research Foundation

Background

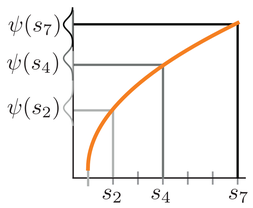
- Maximum Likelihood Difference Scaling (MLDS) is a robust method to estimate perceptual scales [1, 2]



Method of triads



"Which difference is bigger?"
(s_2, s_4) or (s_4, s_7)



Decision

$$\Delta = [\psi(s_7) - \psi(s_4)] - [\psi(s_4) - \psi(s_2)]$$

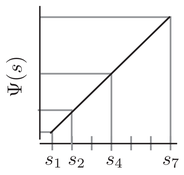
if $\Delta > 0 \rightarrow R = 1$
otherwise $R = 0$

- Requires large amount of trials
- e.g. 10 stimulus levels: 120 triads

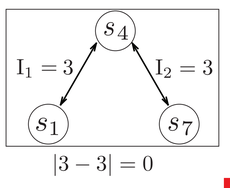
Scaling is informative but time consuming

Proposal to increase efficiency

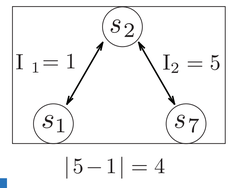
- Shooner & Mullen (2022) proposed to use a subset of all triads
- Prior knowledge of scales: adjacent stimuli are perceptually equidistant
- Select triads with specific difference between pairs of stimuli:



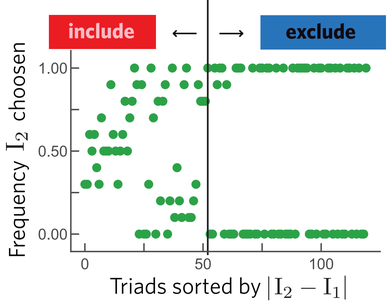
selection rule
$|I_2 - I_1| \leq 1$



$|3 - 3| = 0$



$|5 - 1| = 4$

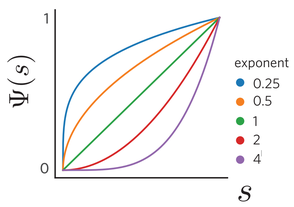


- With k=10 stimulus levels: 52 instead of 120 triads!

⟶ Will the suggested subsampling scheme allow to recover other (non-linear) scales?

Method: Simulating perceptual scales for subset of trials

1. Ground truth scale



2. Design

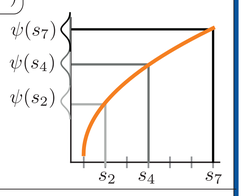
triad
s_2, s_4, s_7
s_1, s_5, s_7
...

3. Simulated observer

$\psi(s) \sim N(\Psi(s), \sigma^2)$

Internal response

s_2, s_4, s_7
$\sigma = 0.05$



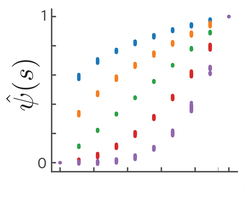
Decision

$$\Delta = [\psi(s_7) - \psi(s_4)] - [\psi(s_4) - \psi(s_2)]$$

if $\Delta > 0 \rightarrow R = 1$, otherwise $R = 0$

compare

5. Estimated scale



MLDS routine

4. Simulated data

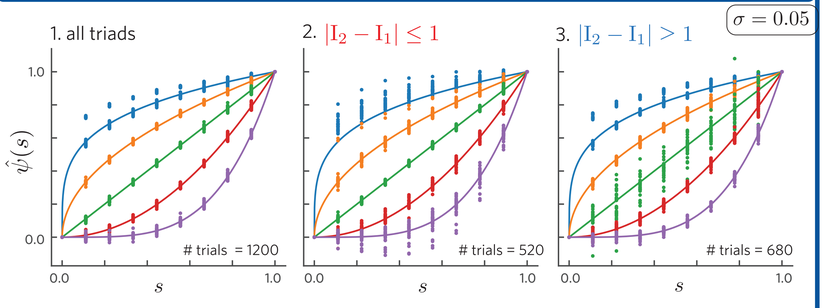
triad	R
s_2, s_4, s_7	0
s_1, s_5, s_7	1
...	

x 1000 times

- For different ground truth scales
- 3 sampling schemes: all triads, $|I_2 - I_1| \leq 1$, $|I_2 - I_1| > 1$

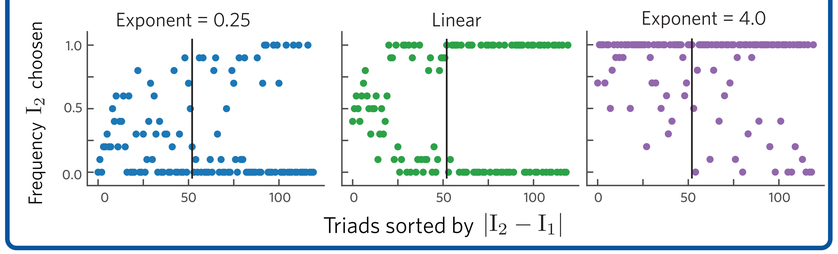
Results: Simulated scales for one noise level

$\sigma = 0.05$



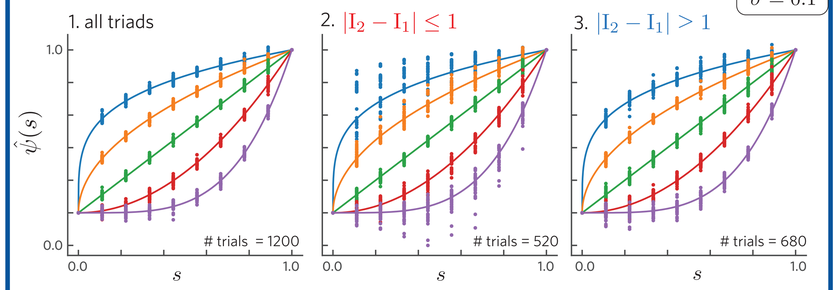
Subsampling recovers linear scales

Subsampling might introduce bias for non-linear scales

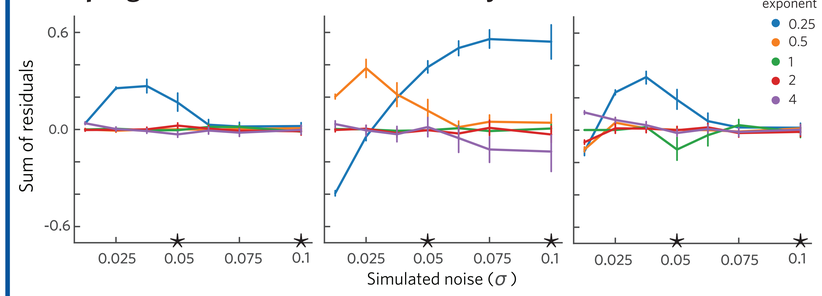


Results: Effect of subsampling for different noise levels

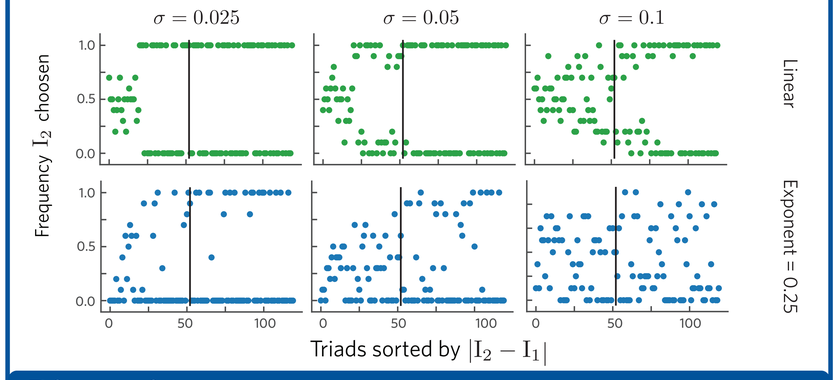
$\sigma = 0.1$



Sampling scheme and noise affect validity of scale estimation



Sampled subset must not include only difficult or only easy trials



Discussion

- Subsampling can increase efficiency of data collection with prior knowledge of scale shape
- Without prior knowledge estimated scales might misestimate true internal scales
- Visualizing the frequency of responses for sorted triads helps to sample the right regime

Acknowledgements

This work was supported by the German Research Foundation (DFG MA5127/5-1 to Marianne Maertens).

References

[1] Maloney & Yang (2003)
[2] Knoblauch & Maloney (2012)
[3] Shooner & Mullen (2022)