

Introduction

As fMRI measurements are complex, expensive, and often target low-amplitude effects, the statistical optimization of stimulus sequences for specific experimental targets plays an integral role in the experiment design process. We provide an extremely flexible framework to perform such optimizations using the genetic algorithm (GA) introduced by Wager and Nichols [6] and improved by Kao et al. [4]. Our Python implementation improves on many aspects of existing solutions, including higher parameterization, API availability, and detailed documentation. Because of its modular structure, our implementation is highly adaptable to specific use cases in both human and animal fMRI. Nevertheless, we also provide an easy to use default algorithm, with widely applicable and properly cited presets. The implementation allows full model (design matrix construction and covariance matrix computation) as well as fitness measure specification.

Mathematical Foundation

GLM:

$Y = X\beta + \epsilon$,
where X is the design matrix, Y a vector of random variables and ϵ is noise

Best linear unbiased estimator (BLUE):

$$\hat{\beta} = (X'X)^{-1}X'y$$

Assumptions:

- $\mathbb{E}[\epsilon_i] = 0$
- $\text{Var}(\epsilon_i) = \sigma^2$
- $\text{Cov}(\epsilon_i, \epsilon_j) = 0, \forall i \neq j$

The last assumption is generally not fulfilled for fMRI data unless it is whitened.

How to Measure the Fitness of a Sequence?

Idea: Minimize variance of parameter estimates $\hat{\beta}$

Covariance matrix:

$$\text{Cov}(\hat{\beta})_{ij} := \mathbb{E}[(\hat{\beta}_i - \beta_i)(\hat{\beta}_j - \beta_j)] \quad \text{Cov}(\hat{\beta}) = (X'X)^{-1}X' \text{Cov}(\epsilon)X(X'X)^{-1} = X^+ \text{Cov}(\epsilon)(X^+)'$$

Examples for fitness measures: $\text{tr}(\text{Cov}(\hat{\beta}))$, $\det(\text{Cov}(\hat{\beta}))$

For details see Chapter 15 of [5].

Synthetic Data

In order to validate the results of the algorithm synthetic data was used. The data is based on a 1500s long resting state scan of a mouse brain. Activation is estimated and added to the time course of every voxel in the region of interest. The simplest estimate for the activation is the convolution of the haemodynamic response function (HRF) and the stimulus sequence. More complex estimations take non-linearity effects into account using Volterra series [1].

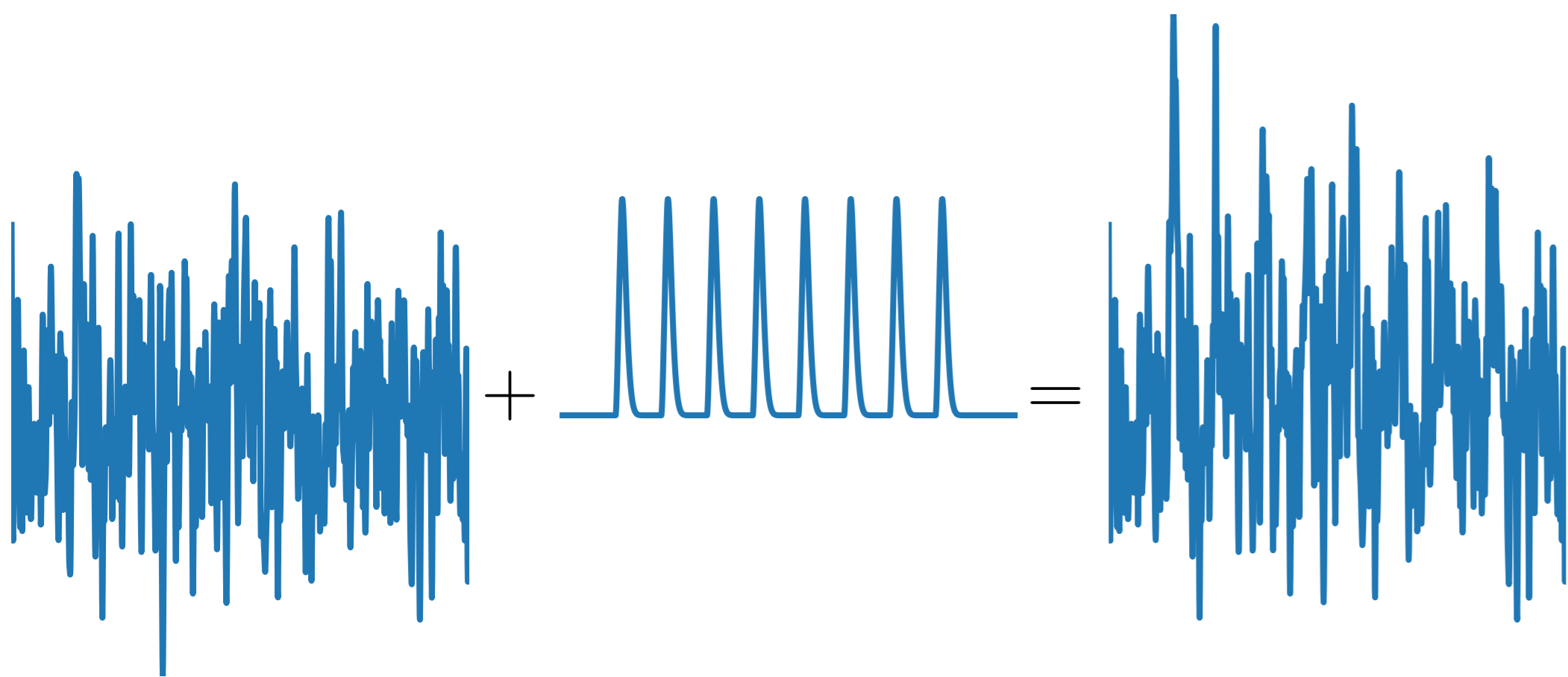


Figure 1: Computation of synthetic data for the time course of one voxel. Left: Time course of one voxel (noise) exit Middle: Assumed activation (convolution of sequence and HRF) exit Right: Sum of noise and activation (synthetic time course).

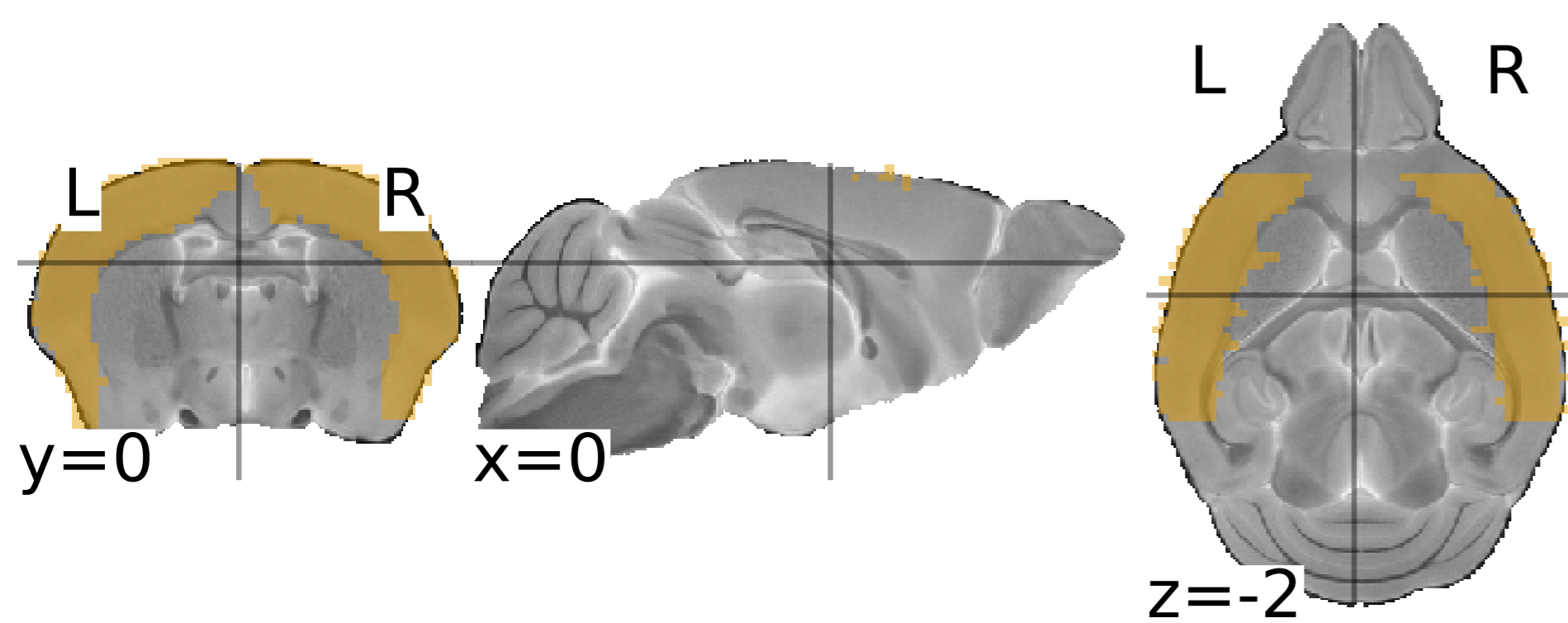


Figure 2: Region of interest used for the synthetic data. It is the cortex of a mouse brain.

Validation of fitness measure

Theoretically, the fitness measures $\text{tr}(\text{Cov}(\hat{\beta}))$ and $\det(\text{Cov}(\hat{\beta}))$ should directly correspond to the sensitivity index $d' := Z(\text{true positive rate}) - Z(\text{false positive rate})$, where Z is the inverse of the cumulative Gaussian distribution. However, there are a lot of parameters (noise whitening, non-linearity, HRF-shape) that have to be taken into account. Figure 3 shows that the fitness measures predicts the sensitivity index calculated from the synthetic data for the parameters chosen. The analysis was done using SAMRI [3] a Python wrapper for Nipype specifically suited for small animal data.

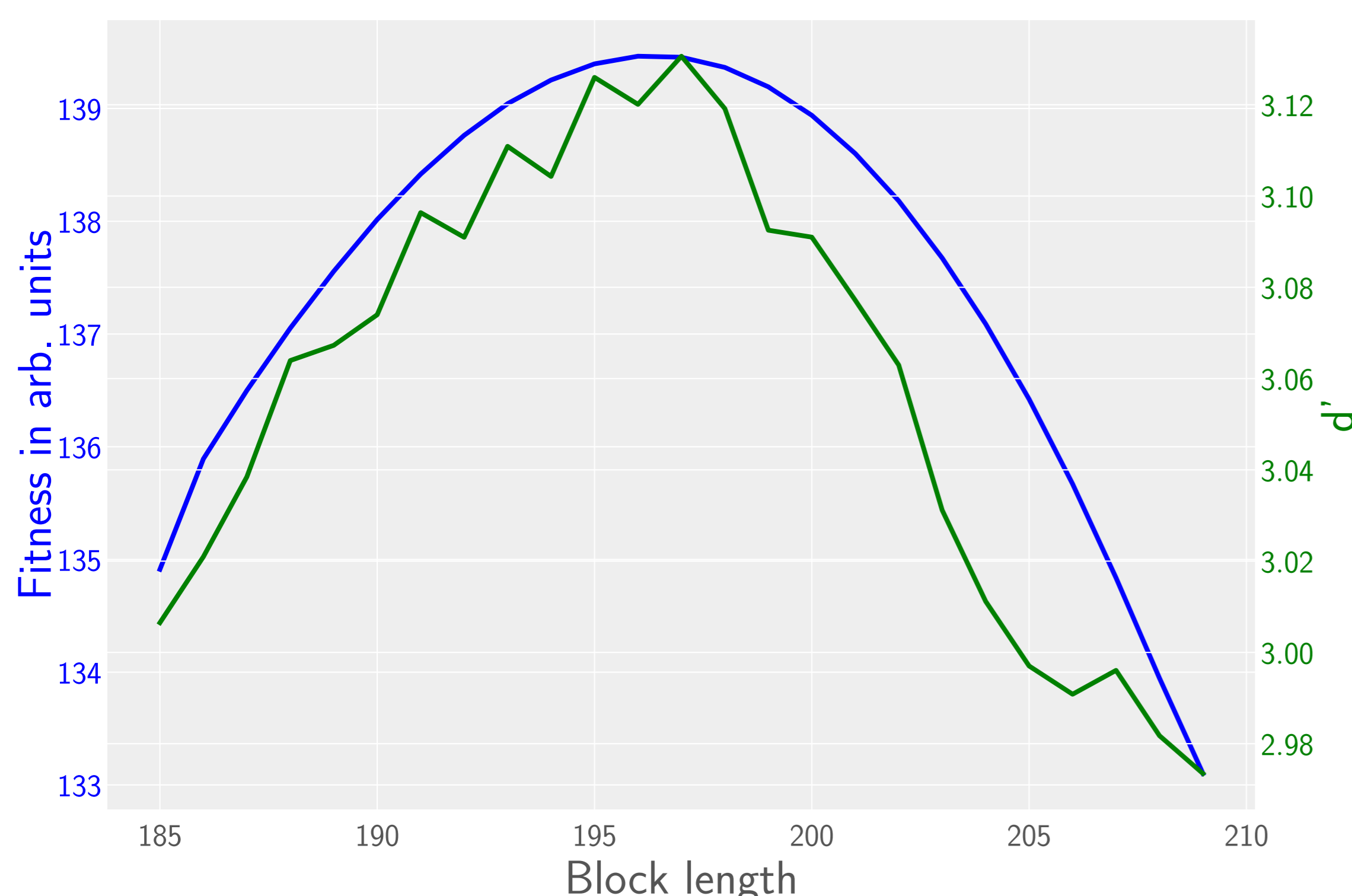


Figure 3: Correlation of theoretical fitness measure and sensitivity index of synthetic data.

Results

It is documented that block designs tend to be optimal for contrast detection (c.f. Chapter 8 of [5]). In accordance with that our algorithm found that a sequence with 196s blocks is optimal given the custom HRF used in Figure 1. However, for some stimulus types such long blocks can result in habituation — or even tissue damage — and therefore should be avoided. We restricted our search to blocks of length 5s to 40s and rest intervals up to 200s. Figure 5 shows that the best fitness is achieved with 40s stimulus blocks and 91s rest in between. A plot of the optimal sequence can be found in Figure 6. The t-values in the region of interest are significantly higher for the optimized sequence compared to original sequence, i.e. the ability to detect the signal is higher.

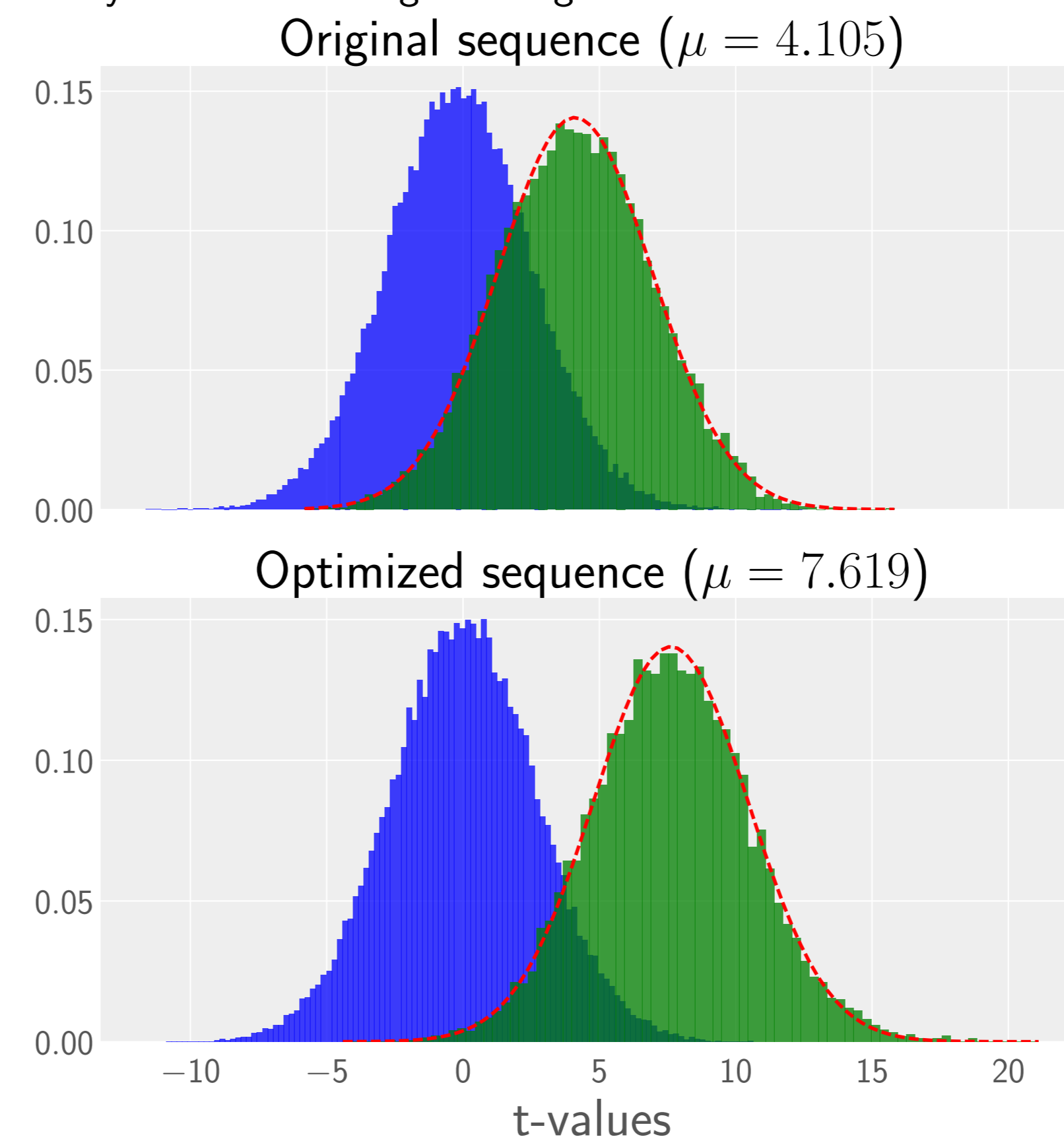


Figure 4: Comparison of the t-value distributions of synthetic data generated using the original sequence and the optimized sequence. The green histogram are the values in the region of interest and blue are the remaining voxels.

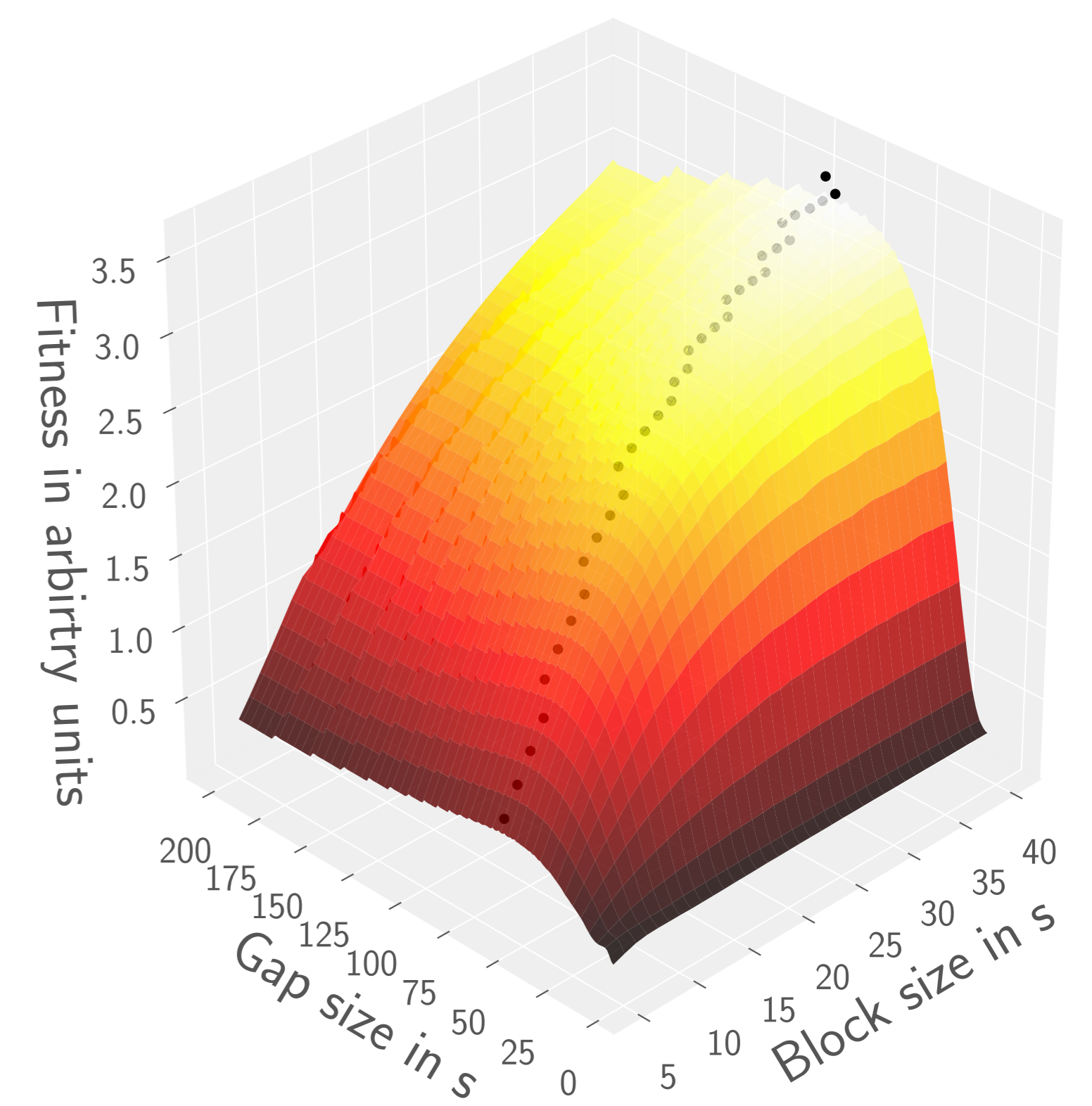


Figure 5: Fitness for different stimulus duration and different resting periods (gaps) in between. The black dots show the optimal resting interval for a given stimulus block length.

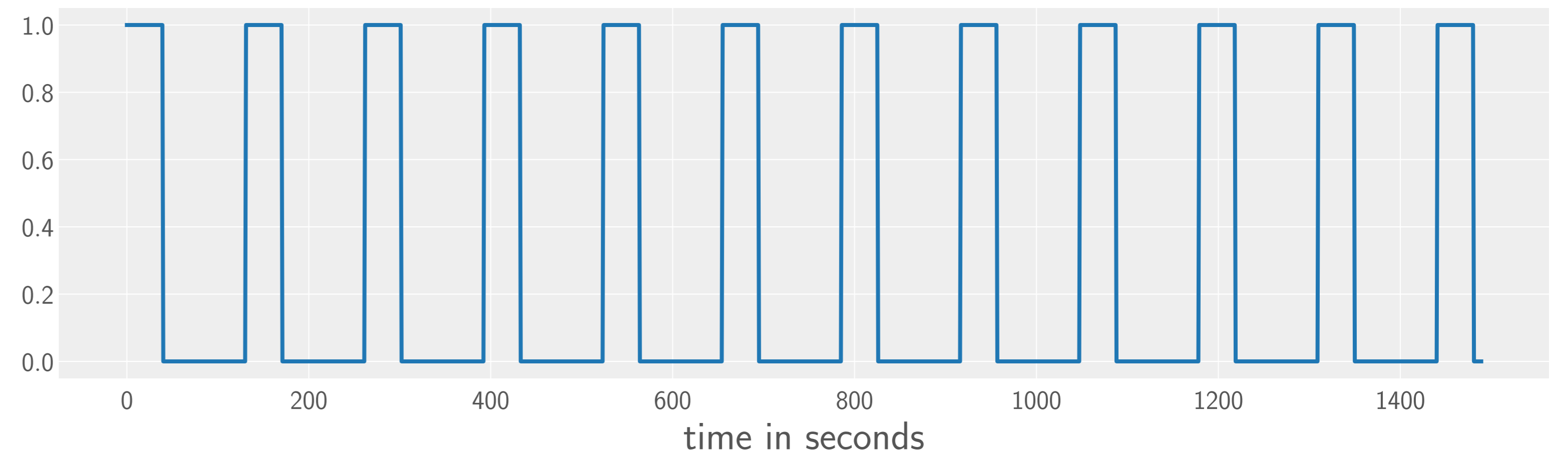


Figure 6: Optimized block style sequence.

Outlook

- Improve GA initial population and choice of immigrants to allow independent stimulus block and resting period optimization.
- Confirm results with wet work experiments in mice and/or rats with optogenetic and electric stimulation.
- Improve handling of non-linearity → Find better estimate of Volterra kernels.
- Reduce dependency on HRF.

Contrast Optimized Stimulus player (COSplay)

COSplay is a MicroPython/Pyboard based delivery system for BIDS-formatted [2] sequences — including those generated with COSgen. The system receives a TTL pulse trigger signal from the MRI machine and then delivers a sequence of stimuli. It uses TTL pulses or short circuiting to steer stimulation devices such as lasers for optogenetic stimulation or electric stimulation boxes for sensory stimulation.

- In total 6 different stimulation devices can be connected simultaneously.
- Two of the output channels can deliver analogue pulses of varying amplitude.
- The stimuli are delivered with an accuracy up to $1\mu\text{s}$.
- Furthermore we developed a small program that can display the status of the system on the computer of the MRI machine and store delivered sequences with the scan data.

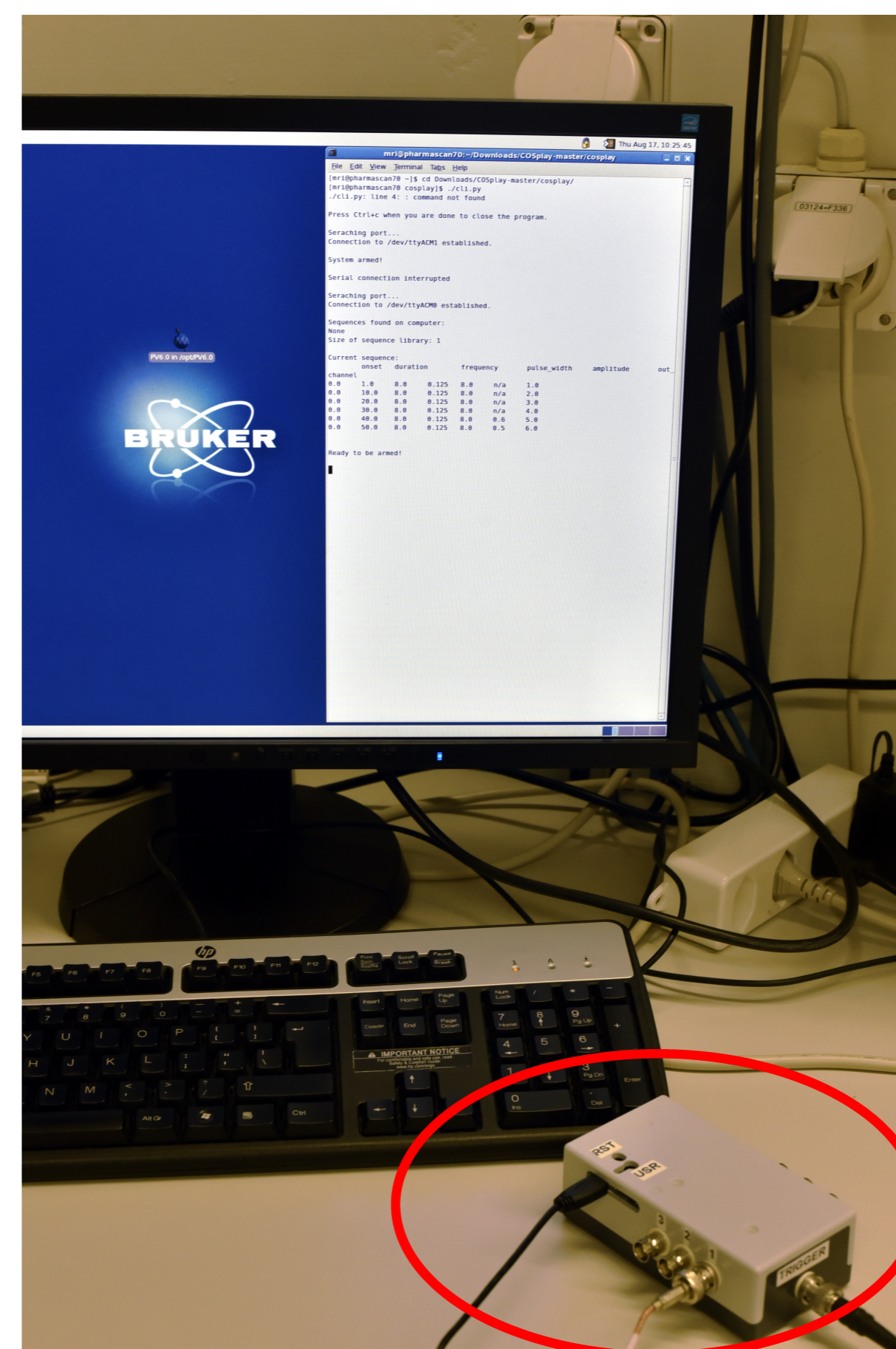


Figure 7: Python based system for delivery of stimuli.

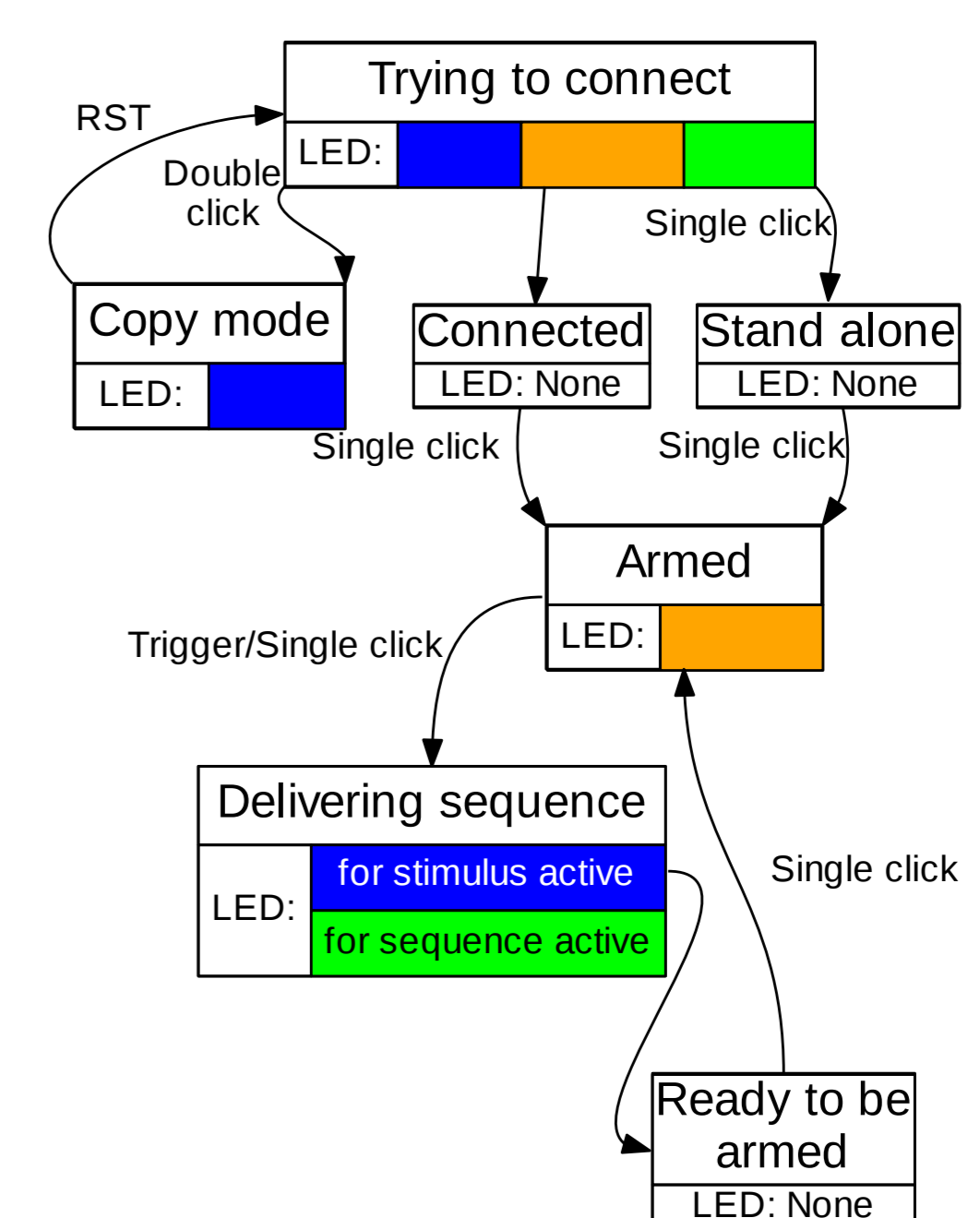


Figure 8: Workflow.

References

- [1] K. J. Friston, O. Josephs, G. Rees, and R. Turner. Nonlinear event-related responses in fmri. *Magnetic Resonance in Medicine*, 39(1):41–52, 1998.
- [2] K. J. Gorgolewski et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3:160044 EP –, Jun 2016. Article.
- [3] H.-I. Ioanas, M. Marks, and M. Rudin. Samri - small animal magnetic resonance imaging via python. *GitHub*, 2017.
- [4] M. H. Kao, A. Mandal, N. Lazar, and J. Stufken. Multi-objective optimal experimental designs for event-related fMRI studies. *Neuroimage*, 44(3):849–856, Feb 2009.
- [5] W. Penny, K. Friston, J. Ashburner, S. Kiebel, and T. Nichols. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier, 2007.
- [6] T. D. Wager and T. E. Nichols. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *Neuroimage*, 18(2):293–309, Feb 2003.