<u>SMAUG Toolbox: Starting Guide</u>

Introduction

Welcome to the Single Molecule Analysis Unified Graphical (SMAUG) Toolbox! This software package contains the code for Segment Clustering, a hypothesis generation tool for single molecule distance/conductance breaking traces described in Bamberger et al. 2020 [link]. However, as the name suggests, the SMAUG toolbox contains additional graphical analysis tools for investigating datasets of breaking traces, all incorporated into a single unified framework. These tools range from the simple and common (e.g. 1D and 2D histograms) to more advanced statistical techniques developed in the literature (e.g. conductance correlation histograms). Capabilities for other clustering approaches not described in Bamberger et al. 2020 are also included.

Software Requirements

The SMAUG toolbox has been fully tested using MATLAB R2019b, but will likely work with older version of MATLAB as well. Please contact us if you run into any forward- or backward-compatibility issues (monti@email.arizona.edu). In addition to a general download of MATLAB, the following common MATLAB add-ons should be downloaded and installed: Curve Fitting Toolbox, Statistics and Machine Learning Toolbox, Mapping Toolbox, and the Parallel Computing Toolbox (only needed if parallelization is going to be used to speed up clustering).

Navigating the Documentation

Here is a high-level guide for how to explore and use this software package:

1. Before doing anything else, run the "RUN_ME" function in the top-level SMAUG directory. This function adds all SMAUG sub-folders to your MATLAB path so that any function can be run from anywhere inside the package. This will need to be run each time you re-open MATLAB, unless you save your search path (see https://www.mathworks.com/help/matlab/matlab_env/what-is-the-matlab-search-path.html for additional details).
2. The following tutorials can be completed in any order to learn about different aspects of the SMAUG toolbox:
   - All SMAUG analysis functions require raw data to be stored in the same "Trace Structure" format. Use the live script "How_To_Format_Input_Data.mlx" to learn how to put your data into this format. Once you've done this, you'll be able to easily employ all SMAUG tools!
   - If you just want to quickly start using Segment Clustering as described in Bamberger et al. 2020, the live script "Quick_Introduction_To_Clustering.mlx" is

where to start. The first section quickly shows how to use the main Segment Clustering tools by applying them to an example dataset, and later sections show examples of more advanced clustering functionality.

- If you want to explore the breadth of analysis tools that the SMAUG toolbox has to offer, use the live script "Quick_Introduction_To_Other_Analysis_Tools.mlx". This guide goes through a series of quick examples using an example dataset to demonstrate why SMAUG is truly a "unified" graphical analysis toolbox.

3. For a more complete list of the different tools available in this package and their capabilities, see "Detailed_Function_Guide.pdf".
4. For detailed information on any particular function in this package, see the comments at the very start of its ".m" file, which will include a brief description of what the function does and what each of its inputs and outputs are.

Fair Use and Making Additions

The SMAUG toolbox is publicly available free of charge under a Creative Commons Attribution-NonCommercial 4.0 International License (to view a copy of this license, visit http://creativecommons.org/licenses/by-nc/4.0/).   This essentially means that anyone can make use of any part of this software package for any non-commercial use, as long as that use is attributed to the SMAUG toolbox (see below). Users are actively encouraged not only to use existing SMAUG tools for their own research, but also to make modifications and add new tools and capabilities to the package. Any additions that seem generally useful to the single molecule transport community will, with the consent of the adding users, be incorporated with attribution into future releases of the SMAUG toolbox.

Getting Help

If you run into apparent software bugs or encounter difficulties while trying to use this package, please contact us at monti@email.arizona.edu. We want people to use this code, so we are very motivated to fix any problems or help with any confusion!

How to Cite

If you end up using any of this code for your own research, then in addition to citing Bamberger et al. 2020 (if applicable), please cite the software package itself. Recommended language is something like:

Data analysis in this work was carried out in part using the SMAUG Toolbox maintained by LabMonti^{TM}, which is publicly available at https://github.com/LabMonti/SMAUG-Toolbox.