Sandia National Laboratories
Unclassified Unlimited Release (UUR)
SAND2019-8293D

STCHbook: An AWSM Automated Analysis GUI

Srinath A. Nandakumar[1,2], Anthony H. McDaniel[2]
[1] Purdue University, West Lafayette, IN [2] Sandia National Laboratories, Livermore, CA

Motivations

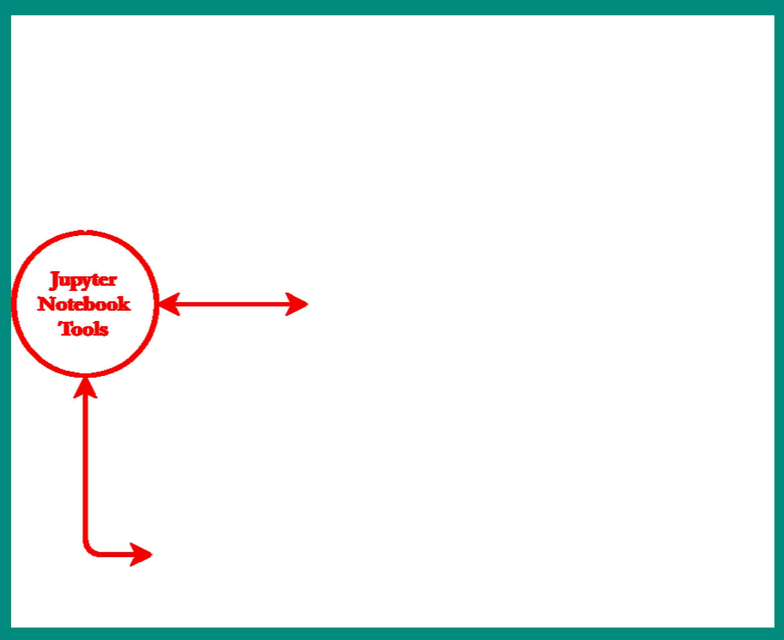


Figure 1: HydroGEN EMN Data Schematic

HydroGEN is one of seven Energy Materials Networks (EMN) funded by the US Department of Energy (DOE). Specifically, HydroGEN is a consortium for research on "advanced water splitting materials" (AWSM) for renewable and sustainable hydrogen generation[1]. EMNs facilitate access and collaboration between various laboratories, academia, and industry. This requires a robust infrastructure for data management and sharing, centered around a CKAN database as shown in the schematic on the left. Laboratories, such as Sandia, operate "research nodes" that offer various capabilities to generate useful data. "Data Users" are consortium partners who will extend upon the initial data to develop implementations of novel technologies. Intuitive tools built on the Jupyter Notebook platform are proposed to largely automate the preliminary data analysis and to better facilitate data sharing. These tools are tested on data from one of Sandia's material characterization nodes that generates chromatograms from mass spectra.

Figure 2: HydroGEN EMN Participants[1]

User Interface (GUI)

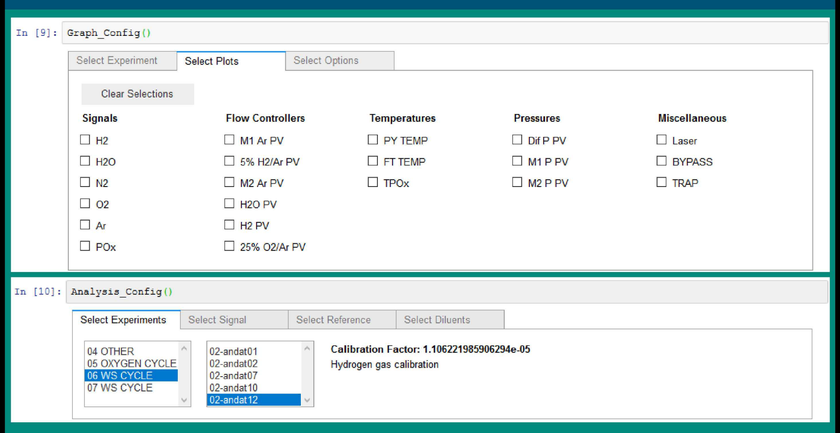


Figure 3: STCHbook GUI Interface

STCHbook was built using the versatile Jupyter Notebook platform. The notebook is a python run-time environment where code fragments, or "cells", can be independently executed. Leveraging Jupyter Notebook Widgets, an interactive GUI was built to configure the various parameters needed for analysis. All back-end code was wrapped in the STCHpy package, so users can simply execute a single line in a cell to call upon the interactive functionalities. One interactive feature is the ability to select different regions of a chromatogram using the mouse. This is adapted from the PySpecKit package[3].

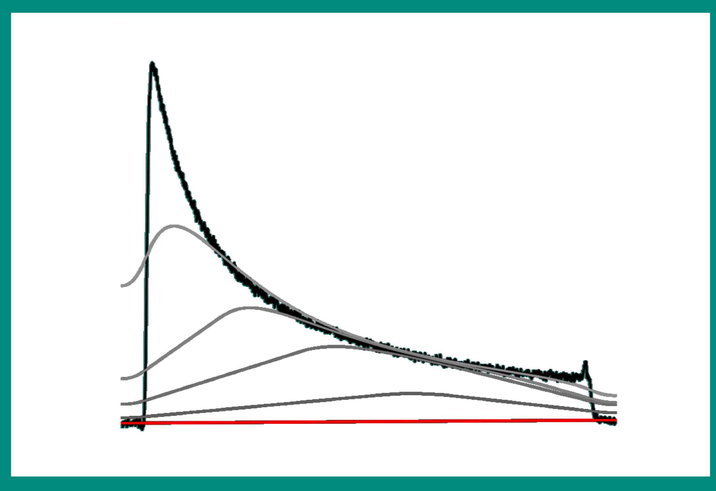
Figure 4: Interactive Region Selection

Baselining Algorithm

The data stream from Sandia's node records temporal signatures of gases that evolve during the water splitting reaction, like hydrogen and oxygen. These raw signals are directly proportional to volumetric flow rates, so accurately detecting the peaks and areas within the temporal profile can provide useful information about reaction kinetics. An important first step to analysis is the removal of background signals. STCHpy automates this process by extending the algorithm proposed by Zhang et. al.[4], which itself extends the idea proposed by Eilers et. al.[2] The basic premise is as follows:

- The quality of a smoothed result can be quantitatively assessed by its fidelity to the original data and its overall smoothness. A parameter λ, which the user adjusts using a slider, controls how important smoothness is compared to fidelity.
- Regions with extremely low fidelity to the original data are assumed to be peaks. Here, fidelity is completely disregarded and smoothness is prioritized.
- By minimizing the penalty (i.e. maximizing the quality), a smoothed result is obtained.

By repeatedly smoothing a smoothed result, the peaks are essentially "shaved out" resulting in the desired baseline.



A baseline is essentially a smoothed version of a spectrum, sans peaks

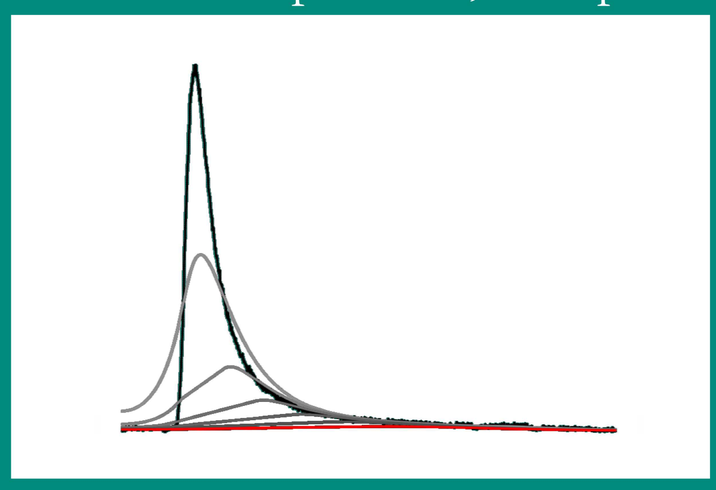


Figure 5: Baseline Correction Process

Future Work

STCHpy and STCHbook will be made available to the public as an open-source GitHub repository, allowing for future extensions to be created and the ability for current functionalities to be further refined. Other nodes can also use this project as a model to develop similar analysis tools.

References

[1] Dinh, H. N., McDaniel, A.H., Weber A. Z., Boardman, R., Ogitsu, T. (2019, April). HydroGEN Overview: A Consortium on Advanced Water Splitting Materials. 2019 DOE Annual Merit Review. Crystal City, VA.
[2] Eilers, Paul H. C. & Boelens, Hans. (2005). Baseline Correction with Asymmetric Least Squares Smoothing. Unpubl. Manuscr.
[3] Ginsburg, A., & Mirocha, J. (2011). PySpecKit: Python Spectroscopic Toolkit. Astrophysics Source Code Library.
[4] Zhang, Z. M., Chen, S., & Liang, Y. Z. (2010). Baseline correction using adaptive iteratively reweighted penalized least squares. Analyst, 135(5), 1138-1146.

This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program

SULI Science Undergraduate Laboratory Internships

U.S. DEPARTMENT OF ENERGY | Office of Science

PURDUE UNIVERSITY

Sandia National Laboratories