Finding Meaning in the Void or (Finding Lidar Characteristics in Noisy Observations)

or (How I spent my summer ♡)

NCAR

Sandy Urazayev[1, 2], Matthew Hayman[1]

[1]Earth Observing Laboratory (EOL) at the National Center for Atmospheric Research (NCAR)

[2]Department of Electrical Engineering and Computer Science, University of Kansas

Background & Goal

The MicroPulse DIAL or MPD is a low transmit power lidar system capable of retrieving profiles of water vapor (absolute humidity), aerosol and cloud scattering properties (backscatter coefficient) and temperature. MPD collects backscattered photons across a wavelength range. Due to the system characteristics, the received signal can be very noisy. **Our goal is to analyze the MPD receiver to minimize the effect of noise on system characterization.**

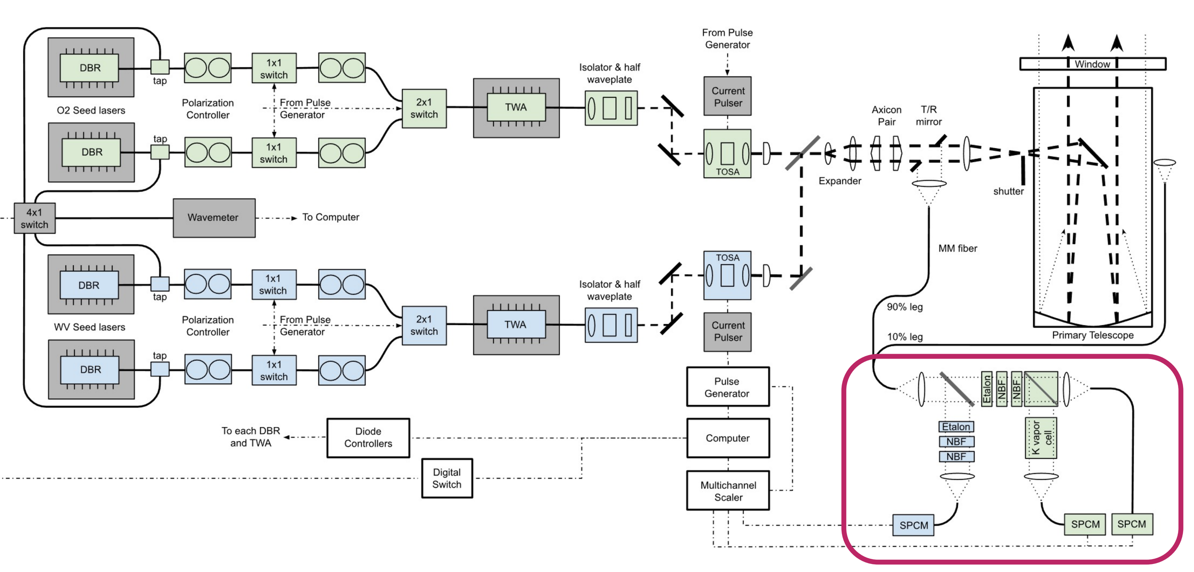

Figure 1: MPD Block Diagram (Pink box indicates the receiver)

Motivation

The accuracy of the characterization of our instrument directly depends on the way we fit the signal photon counts. The previous method used a traditional smoothing spline. This requires a user to make a qualitative assessment of fit quality. Attempting to minimize noise by reducing the resolution of the spline results in degradation of captured features.

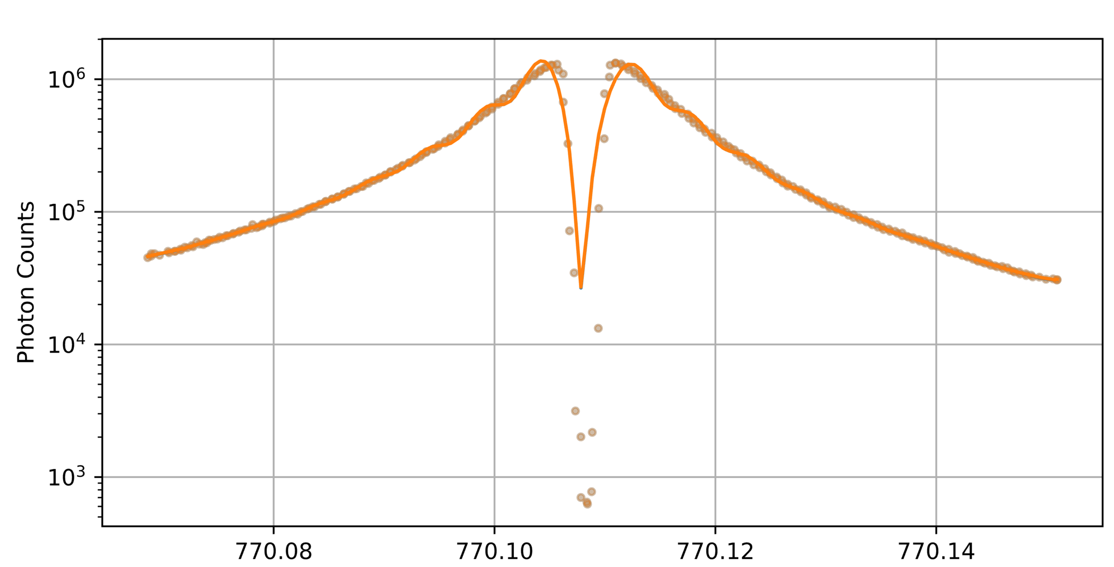

Figure 2: Spline fit of the molecular channel from MPD from Figure 1 (Notice it doesn't perfectly capture the sharp edge in the bottom)

This problem motivates us to research new ways of understanding the signal within noisy observations. Similarly, we could define better constraints, such as fitting the data onto our receiver function. It would grant us a better understanding of how observed data can confirm our theoretical optical models.

Methodology

The main optical filter that exists in every MPD scan is a Fabry-Perot etalon. A Fabry-Perot etalon is made out of two highly reflective parallel surfaces. We have to account for light traveling at multiple angles θ through the filter. Let $x = \cos\theta$, then the transmission function for the etalon is

$$f(\lambda, x) = \frac{1}{1 + F\sin^2\left({}^{\delta_0 x}/_{2\lambda}\right)}$$

where F is the coefficient of finesse, δ_0 is the optical path length experienced by light at normal incidence, and λ is the wavelength. The actual observed transmission is then

$$f(\lambda) = \int_1^0 w(x) f(\lambda, x)\, dx$$

where we must also fit the weights for each incidence angle $w(x)$.

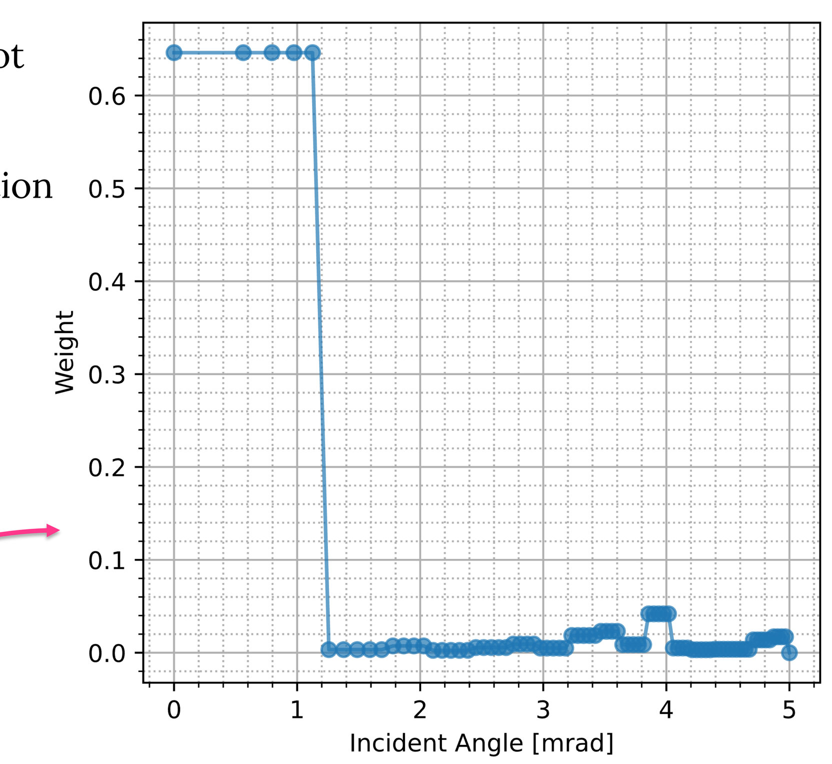

Figure 3: Fitted Angle Weights

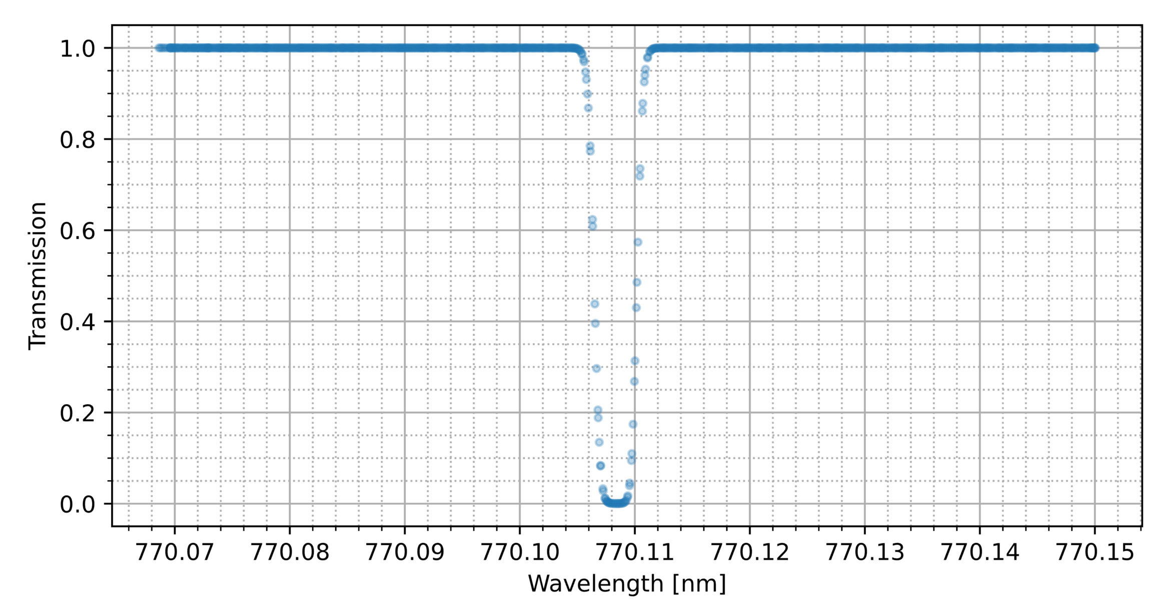

Figure 4: Potassium Cell Transmission

Molecular types of scans use an additional alkali metal vapor cell as a narrow band notch filter. This uses a potassium (K) gas as a filter. In general, the transmission of these cells obey the Beer-Lambert law, so the transmission is

$$h(\lambda) = (1 - \epsilon)\exp(-\kappa(\lambda, T, P)L) + \epsilon$$

where T is the temperature of the cell, P is the pressure, L is the length, κ is the cell extinction factor, and ϵ is the spectral purity of the laser. During the training we are fitting the temperature and spectral purity.

Incorporating both the Fabry-Perot etalon and potassium transmission models and our fit terms, we find our full model to be

$$\tilde{\alpha}_n(\lambda_i) = g_n f(\lambda_i) h(\lambda_i) + b$$

where g_n is a scalar efficiency coefficient for the n^{th} channel and b is the background noise. We assume photon arrivals are described by Poisson statistics. We improve the accuracy of our full model by minimizing the negative log-likelihood of the noise model

$$\mathcal{L}(\tilde{\alpha}, k) = \sum_i (\tilde{\alpha}_i - k_i \ln \tilde{\alpha}_i)$$

where $\tilde{\alpha}$ is the full model of MPD channels that we fit to noisy observations k.

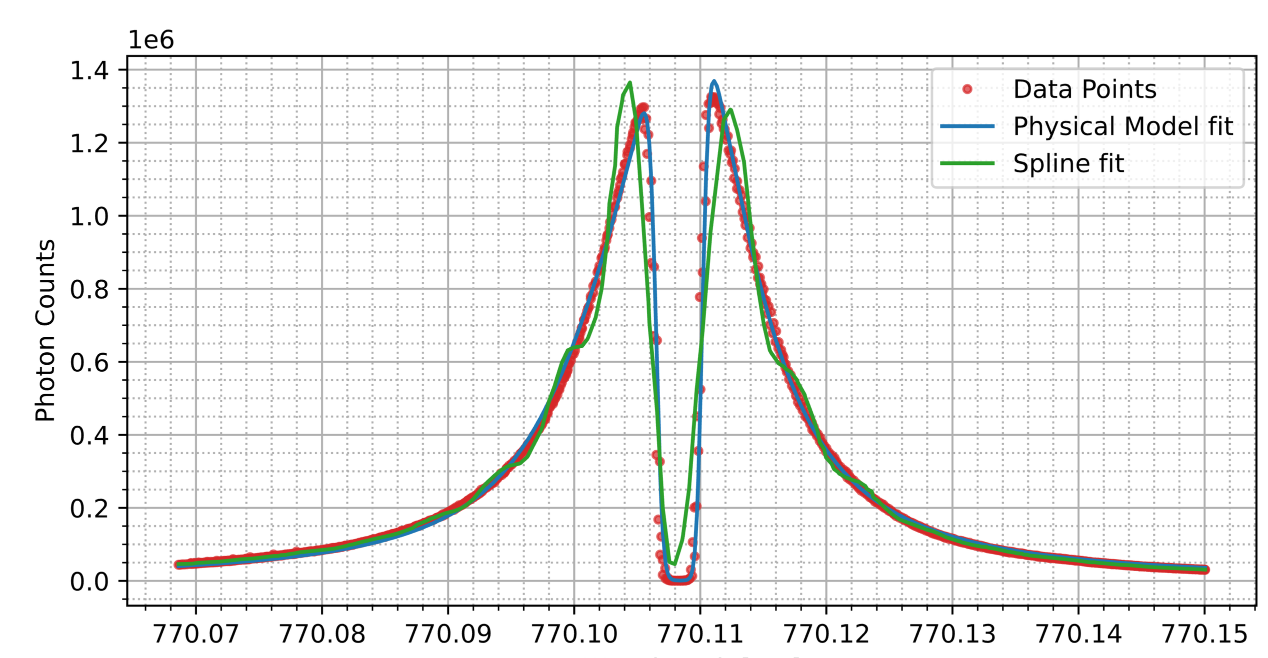

Figure 7: Spline fit and physical model fit for an etalon and potassium cell

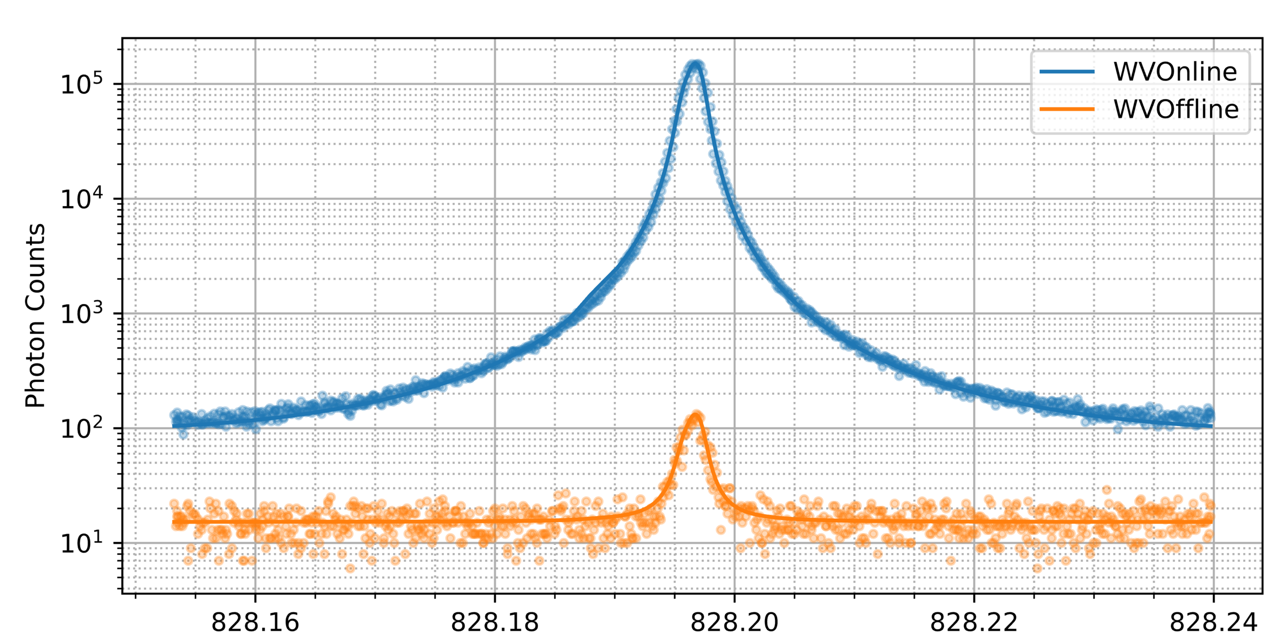

Figure 5: Fabry-Perot etalon fit

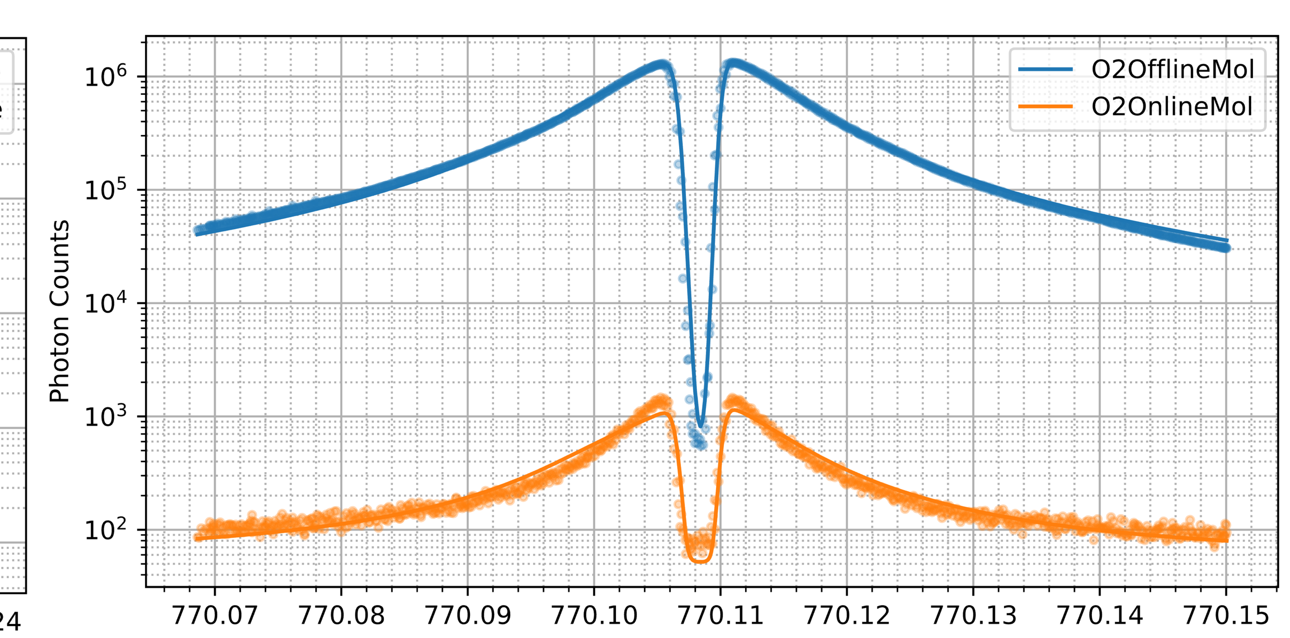

Figure 6: Fabry-Perot etalon and potassium cell fit

Results

The resulting characterization using constraints yields better accuracy and allows us to confirm the theoretical foundations of the photon receiver function. We compared popular numerical optimization algorithms, **where Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm was observed to be the most accurate one, yet the slowest one.**

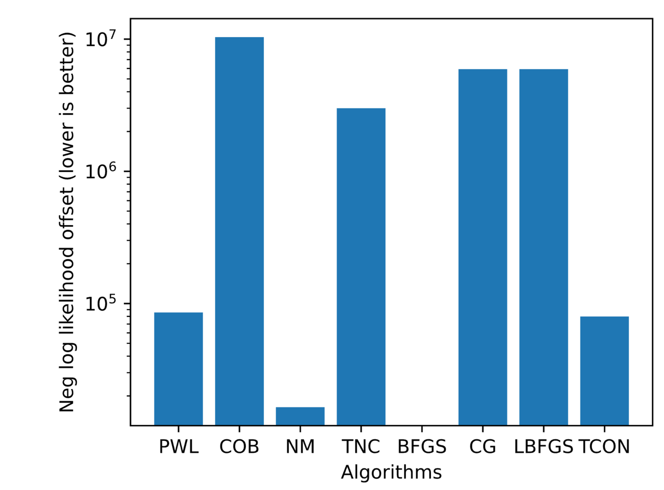

Figure 8: Relative inverse accuracy

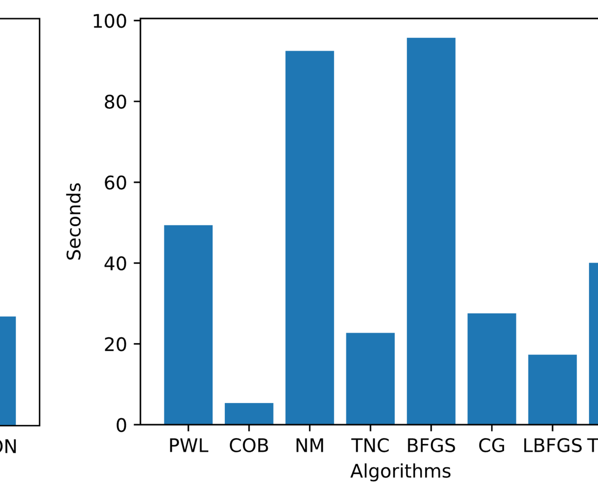

Figure 9: Training time [2]

In order to speedup the fitting routine, **we used Dask [1] to manually parallelize the optimization jobs**, which helped us to reduce the time taken to fit the most complicated type of scan, Molecular Temperature scans, **from 1 hour to 3 minutes**.

Conclusion

While applying a physical model of the receiver, we have successfully minimized the effect of noise on system characterization. We found that by applying constraints to the noisy data when fitting it onto an optical model not only returns more accurate results, but also helps us to confirm our theoretical understanding of MPD's optics and etalons. We later applied advanced multiprocessing techniques, such as job parallelization and memoization to drastically reduce the time taken for the model training.

Acknowledgments

This work has been gratefully funded by the Summer Undergraduate Program for Engineering Research (SUPER) at the National Center for Atmospheric Research's (NCAR) Earth Observing Laboratory (EOL).

Notes

[1] Dask is a flexible library for parallel computing in Python. See dask.org for further details.

[2] Training times are computed from our testing suites.