

ANN: Explanation of Problem

NOTE: As you are reading this document, it will be helpful to look at a data file. You can find these in the ANN GitHub Repo. Just look at any of the CSV files in any of the Pseudodata folders in Liliet's folder.

The goal of this study is to find a method to accurately predict the values of functions called Compton Form Factors (CFFs) using data from Deep-Virtual Compton Scattering (DVCS) experiments.

Each experiment, or set, consists of measuring the cross-section (F) as an angle (ϕ) is changed. Since ϕ is an angle measurement, its range is from 0° to 360° . At each angle that is used, the cross-section is measured multiple times. The mean of these values is recorded in the 'F' column of the data file, and the standard deviation is recorded in the 'errF' column.

Different experiments represent different environments in which these measurements are performed. The environment of an experiment is determined by the kinematic variables k , QQ , x_b , and t . Each set of values of these variables represents an environment (also called a kinematic setting) for an experiment to be run in.

The cross-section is a known function of the Compton Form Factors, the kinematic variables, ϕ , and a couple other values. The other values are F_1 , F_2 , and $dvcs$, which are all known functions of the kinematic variables, and their values are provided in each data file. The CFFs are all unknown functions of the kinematic variables: we know that they are functions of the kinematic variables, but we do not know their exact formulations. In each experiment, we are fixing the kinematic variables, varying ϕ , and measuring F . The values of F_1 , F_2 , and $dvcs$ are known, so the only unknowns left are the Compton Form Factors. Our goal in this experiment is to determine the values of the CFFs using the other values. We are focusing on three Compton Form Factors: ReH , ReE , and $Re\tilde{H}$.

There are two kinds of fits that are used to determine the Compton Form Factors: local fits and global fits. A local fit's end goal is to find the numerical value of one of the CFFs in a certain kinematic setting. A global fit's end goal is to find the equation that relates the values of the kinematic variables to the value of one of the Compton Form Factors. A local fit is a guess of the value of a CFF in a certain kinematic setting, while a global fit is a guess of the equation for a CFF that can be used to calculate the value of the CFF in any kinematic setting.

One important aspect of this exploration is correctly propagating the error from the cross-section to the Compton Form Factors. We are not given exact values for F , so we cannot possibly get exact values for the CFFs. Every output from our models is given as a guess and an error range. Our goal is to have the guesses be as close as possible to the true values of the CFFs, but it is also important that the error ranges consistently enclose the true values.

The most popular way of propagating error is to use the replica method. In this method, the F value from the data file is not inputted directly into the model. Instead, we use several values that are sampled from a Gaussian distribution with mean equal to F and standard deviation equal to $errF$. By using input values that represent the distribution of the cross-section, we hope to be able to get outputs that represent the distributions of the Compton Form Factors (meaning that the true value of the CFF should lie in the error bars produced by the model).

Right now, we are using pseudo-data where we know the true values of the CFFs instead of real experimental data where we don't. This way, we can compare the CFF guesses to their true values and have a good idea of a model's performance. Once we develop a method that can consistently and accurately predict the CFFs, we will move on to experimental data.