Thomas Cannon
MATH540

Kernel Smoothing on BATSE TTE Data

Kernel density estimation (KDE) is a method to estimate the probability density function of a random variable without relying on parameters. Commonly used on time-series data in fields such as signal processing and econometrics, KDE is used as a way to smooth data with kernel smoothers. A kernel smoother is the weighted average of neighboring observed data. The weight of the average is defined by the kernel. Kernels can take all manner of shapes, but will typically give weight to the closest points. For the scope of this paper, we will be using kernels in the shape of a Gaussian curve. Although, underlying physical properties may prompt us to use a different curve in the future.

We will be attempting to smooth time-series Gamma Ray Burst (GRB) data with a Guassian kernel. GRBs are transient astrophysical events that occur about once per day from anywhere in the sky. If humans had a pair of goggles that allowed us to visually see GRBs, we would see an incredibly bright event, localized to a point in the sky. GRB have been detected since the 1970s, but the progenitor physics has yet to be solved. The data takes the form of microsecond-binned, time-stamped arrival of photons to the detector if the Burst and Transient Source Experiment (BATSE) on board the Compton Gamma Ray Observatory (CGRO). We commonly refer to the timestamped data as BATSE TTE (Time-Tagged Event) data. The fundamental unit of a GRB is called a pulse. Pulses typically take on a specific shape, and can be easily visually identified at different time binnings. They are characterized by distinct rises above the background over a GRB light curve. Pulses from one burst to another can often vary wildly; some lasting minutes, while others last milliseconds. However at any duration, a pulse normally has a predictable and measurable shape.

It has been proposed that GRB pulses themselves contain underlying structure (Hakkila 2014). For example, some of these features present themselves as periodic oscillations, overlaid on top of the intrinsic pulse shape found in the binned data. These oscillations are often difficult to discern and obscured by a lot of noise. As mentioned above, the pulses will have a predictable shape regardless of

the time scale, so we are justified in performing operations on the data such as rebinning or smoothing. The motivation for smoothing the data with kernel smoothing is therefore to attempt to reduce noise of the GRB data to a point where an analysis of the underlying structure in the pulse is discernible.

We start the kernel smoothing with the python programming language, using the KernelDensity functions from the Scikit-Learn library. To get a sense of how this process applies to the BATSE TTE data, we will test the KernelDensity class on the first GRB detected with TTE resolution: BATSE trigger number 00138. The data has been cleaned before hand for faster processing. The entire GRB takes place over about 3 million microseconds. Because of the long range in **micro**seconds, we will test a variety of different kernel bandwidths at 5, 15, 25, and 50 **milli**seconds.

The results of this test can be seen in Figure 1. We can immediately see that there may be two distinct pulses in this burst. While this is not uncommon to see at the typical binning of 64ms that we see in other datasets in BATSE, the evolution from two distinct pulses to overlapping pulses as the kernel bandwidth increases calls into question the definition of a pulse with respect to its binning. For years, structure has been proposed with not much regard to the binning size in these experiments. This is mostly because a fundamental pulse shape that is independent of a time scale is still a relatively new concept. Because of this, Kernel smoothing might be an interesting process to play around with to see what sort of GRB pulse properties are maintained at different bandwidths.

There is also the idea of using a different kernel. Some members of the community believe that the pulse shape – which is defined by an empirical function that looks similar to a log-normal curve – can actually be explained by phenomenon called interstellar reddening and pulse broadening. This is commonly reference in pulsar physics, but has not successfully been applied to GRBs due to the difference in energies: pulsars are radio and GRBs are gamma. However, if these ideas were to hold to GRBs as well, then it might be possible to take the pulse shape function and use that as the kernel instead of a Gaussian kernel. Doing so might be able to "undo" the pulse broadening properties and show us a cleaner picture of what GRBs are really supposed to looks like.

Figure 1: Kernel Smoothing of BATSE TTE PULSE 00130
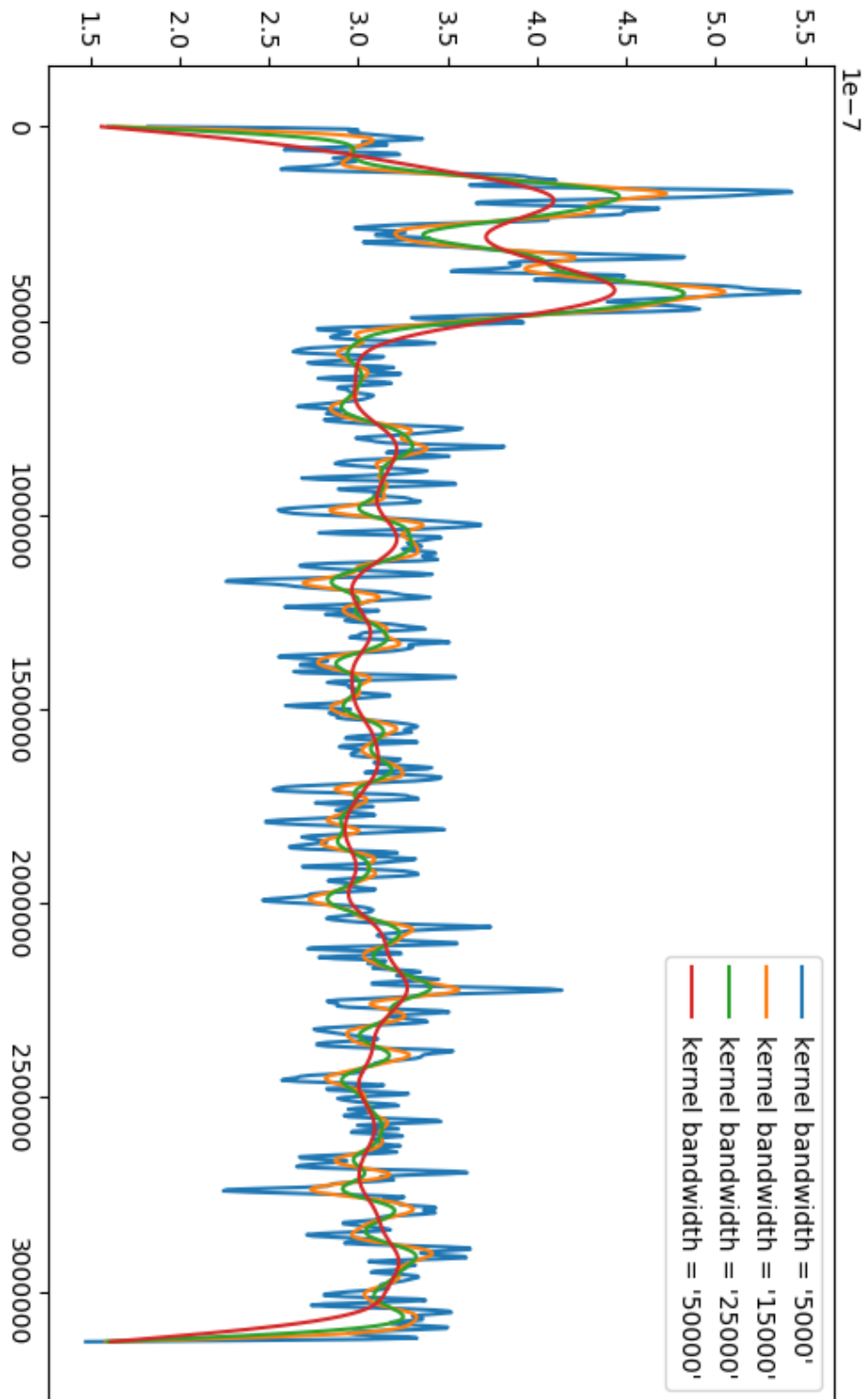
Python Code

```python
import numpy as np
import matplotlib.pyplot as plt
from sklearn.neighbors import KernelDensity
# from grbpy.burst import Burst
import csv

file_path = 'tte_00138_clean.txt'

tte_list = []
with open(file_path, newline='') as csvfile:
    reader = csv.reader(csvfile, delimiter=' ', quotechar='|')
    for row in reader:
        tte_list.append(row)
tte_array = np.asarray(tte_list)

bandwidths = [5000,15000,25000,50000]
fig, ax = plt.subplots()

line = np.linspace(int(tte_array[0][0]),
    int(tte_array[-1][0]), 1000)[:, np.newaxis]
for bandwidth in bandwidths:
    kde = KernelDensity(kernel='gaussian',
bandwidth=bandwidth).fit(tte_array)
    log_dens = kde.score_samples(line)

    ax.plot(line[:, 0], np.exp(log_dens),linestyle='-',
        label="kernel bandwidth = '{0}'".format(str(bandwidth)))
ax.legend(loc='upper right')
plt.show()
```

References

Hakkila, J., Preece, R. 2014. "Gamma-Ray Burst Pulse Shapes: Evidence for Embedded Shock Signatures?". arXiv:1401.4047.