BOSTON UNIVERSITY

Resolving Voltage Peaks in Particle Beam Microscopy

Alexander Mehta[1,2], Ruangrawee Kitichotkul[2], Vaibhav Choudhary[2], Akshay Agarwal[2], Vivek Goyal[2]
Fremont High School, Sunnyvale, CA[1], Boston University, Electrical and Computer Engineering Department, Boston, MA[2]

Introduction

- The goal of Particle Beam Metrology (PBM) is to estimate the SE Yield (η) of a pixel through its voltage.
- Commonly used techniques use estimators that assume SE Count (X_i) is known to find η, but X_i is not known, and follows the distribution $X_i \overset{iid}{\sim} \mathcal{P}(\eta)$ where $\mathcal{P}$ indicates Poisson distribution.
- Recent works [1] address this by introducing estimators that can go straight from voltage to η by relying on knowledge of all $\widetilde{M}$ peaks, and their heights $\{\widetilde{U}_1, \dots, \widetilde{U}_{\widetilde{M}}\}$.
- If 2 peaks in voltage data overlap, it can lead to a misrepresentation of peaks and their heights for the estimators. A heuristic correction factor $\gamma_\tau(\Lambda, \eta)$ to account for this.
- We aim to get a more accurate measurement by accurately estimating the number of peaks and their heights instead of using the correction factor.

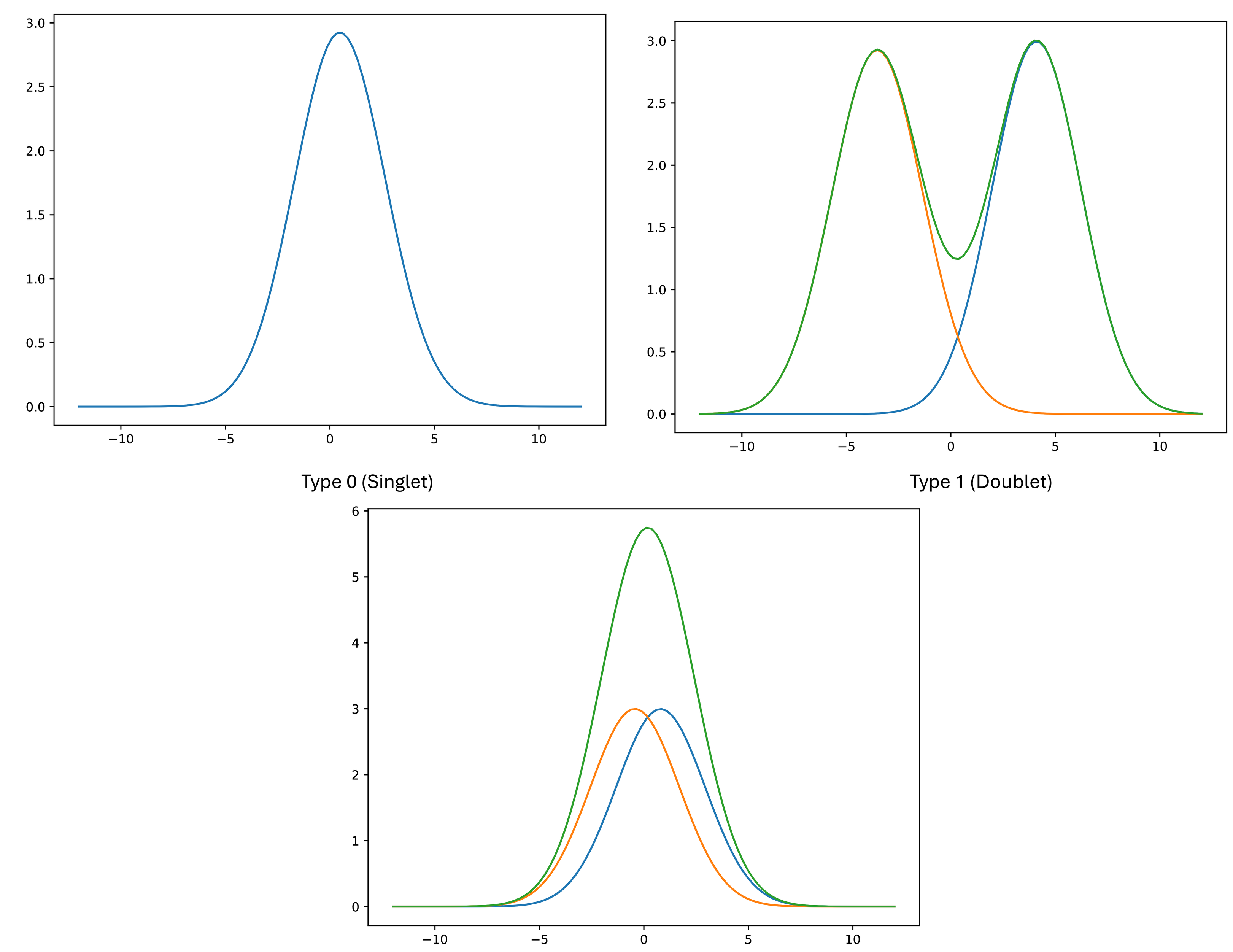
Estimators [1], used by finding the unique root, and where $H = \sum_{i=0}^{\widetilde{M}} \widetilde{U}_i$:

$$\hat{\eta}_{QM} = \frac{H/c_1}{\widetilde{M}\gamma_\tau(\Lambda, \eta_{QM})}$$

$$\hat{\eta}_{MLI} = \frac{H/c_1}{\widetilde{M}\gamma_\tau(\Lambda, \eta_{MLI}) + \lambda \exp(-\eta_{MLI})}$$

$$\gamma_\tau(\Lambda, \eta) = \exp(-\lambda(1 - e^{-\eta})\tau)$$

We notate a fixed time window of peaks into the following types:



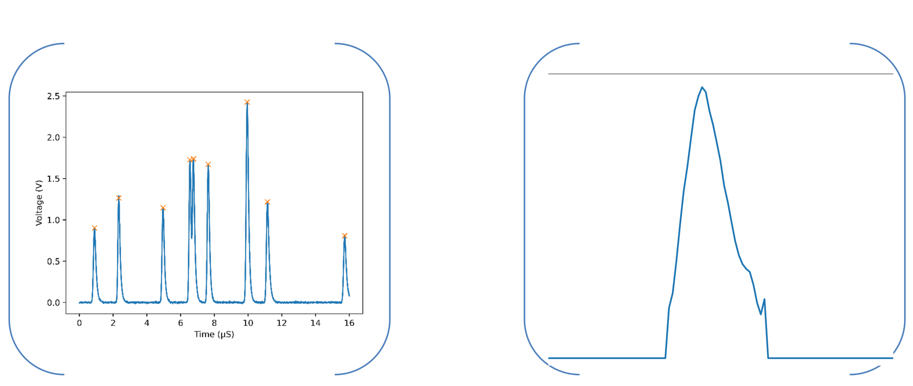
Type 0 (Singlet)
Type 1 (Doublet)
Type 2 (Doublet)

Peak Resolution

For peak resolution, we first use synthetic data with known ground truths in order to train classifiers. This synthetic data is from a known probabilistic model [1]. With real data, we employ a multi-step approach to classify peaks by type:

1. We first use a handcrafted windowing method to detect peaks, and if another peak is detected within the same window, we then classify it as Type 1. If there is no second peak, we mark it as either a Type 0 or 2 peak that requires additional steps. In the case of synthetic data for training, we can extract ground truths by comparing detected versus true generated peaks.
2. For unclassified peaks, we feed a tensor with f(x), f'(x), and Savitzky–Golay filtered F(x) into a convolutional neural network trained on synthetic data.
3. Finally, we determine final height resolutions of peaks as outlined in the figure below.

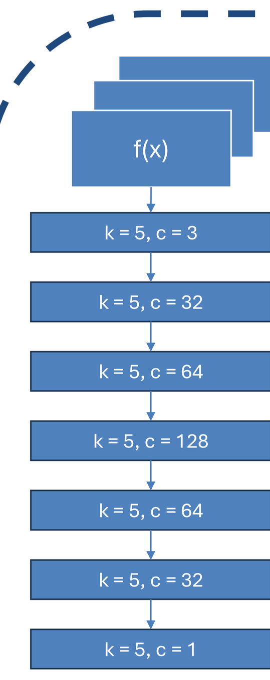
Through this process, we have determined $\{U_0, U_1, \dots, U_M\}$, which can then be used to estimate the SE Yield.



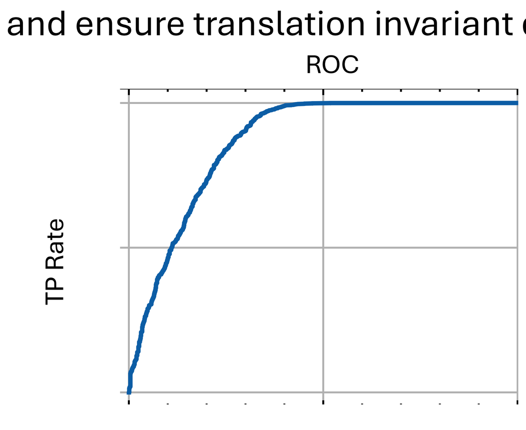
Original Signal with Detected Peaks
Denoised + Windowed Single Peak (Type 0)

To detect peaks, we first smoothed the signal using the Savitzky–Golay filter. We check if adjacent peaks are within the same time window of 300 microseconds. This allows us to know already type 1 peaks, as we can use the extracted peak heights and window to determine the number of peaks.

Peak Detection

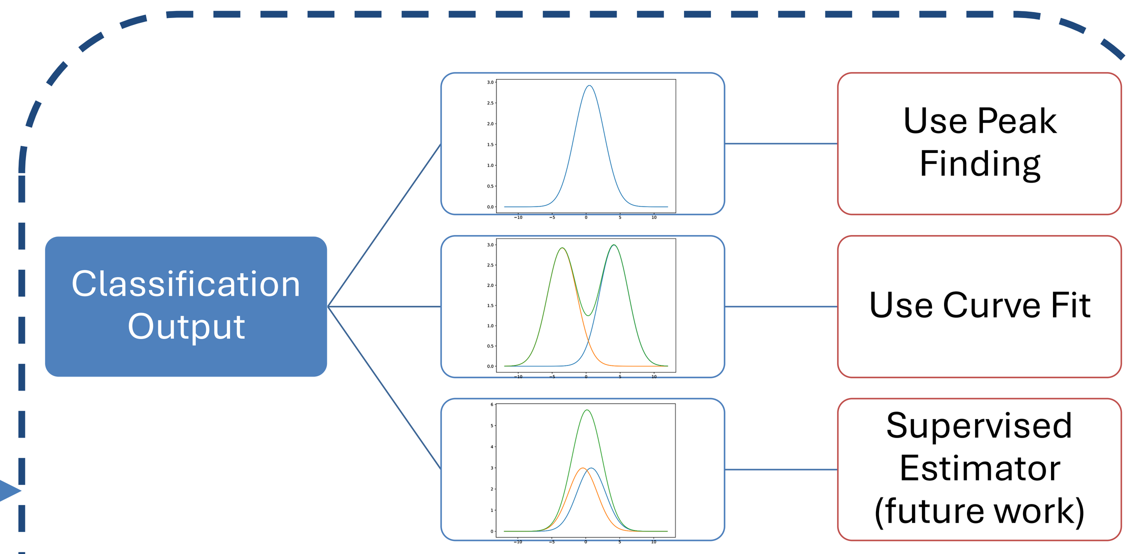


A ResNet inspired architecture was used. We opt to use convolution downscaling (opposed to global pooling or flattening), as we can then optimize selection of most important features and ensure translation invariant design.



Kernel Size	Accuracy (Binary)	AUC-ROC	Class Accuracy
11	94.87%	0.8663	99.57%/55.44%
5	92.32%	0.8221	99.67%/33.87%

CNN Classification



Resolving heights in singlet peaks (type 0) and simpler doublets (type 1) can be done through simple feature extraction techniques. Harder doublets (type 2) pose a unique issue, as the sum of EMGs can look like another EMG (in this case, it can look like just a larger voltage spike). This makes our future work into a supervised height extraction for type 2 very important.

Height Resolution

SE Yield

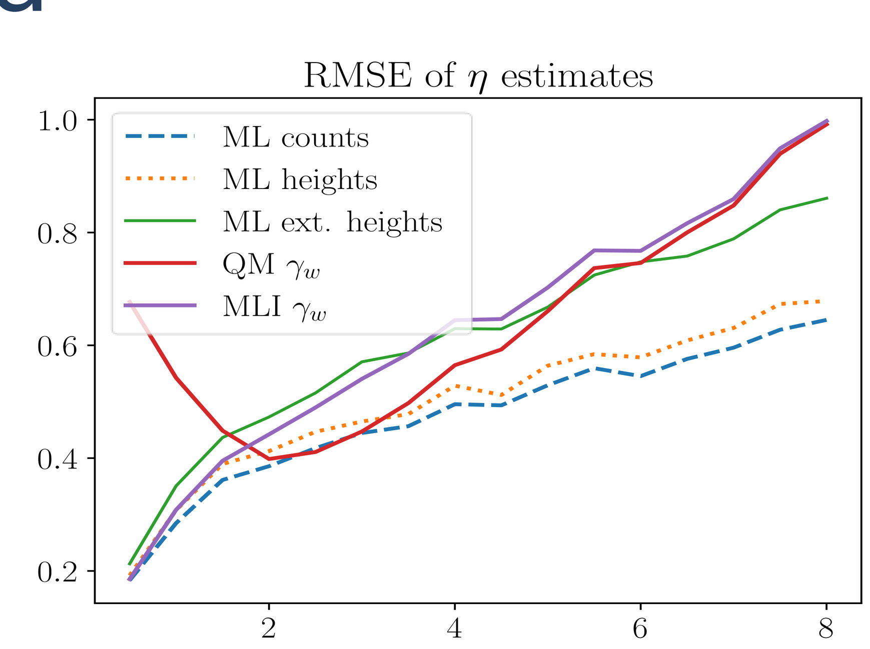
- Prior work has proposed Quotient Mode (QM) and Maximum Likelihood inspired (MLI) estimators that combines naïve peak extraction with a heuristic correction factor $\gamma_\tau(\Lambda, \eta)$.
- The ML counts estimator relies on the true SE Counts, serving as the lower bound on our RMSE.
- The ML heights estimator relies on the true heights and maximizes the following likelihood:

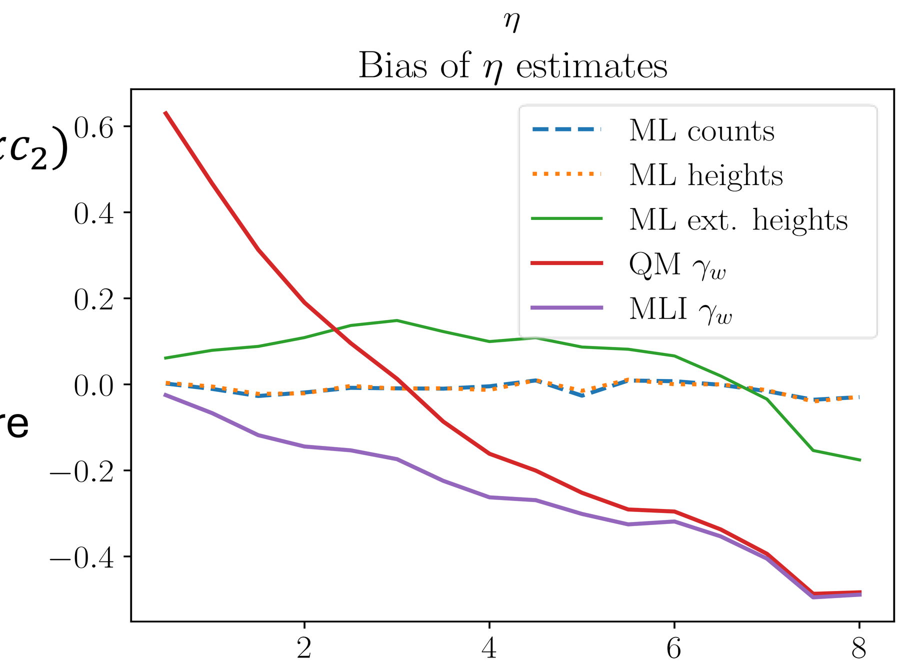
$$\hat{\eta}_{ML} = \arg\max_{\eta} \log p(\widetilde{M}, \widetilde{U}_1, \dots, \widetilde{U}_{\widetilde{M}}; \eta)$$

$$\hat{\eta}_{ML} = \arg\max_{\eta} \log \mathcal{P}(\widetilde{M}; (-e^{-\eta})\lambda) + \sum_{i=1}^{\widetilde{M}} \log \sum_{x=1}^{\infty} \mathcal{ZP}(x; \eta)\mathcal{N}(\widetilde{U}; xc_1, xc_2)$$

Where $\mathcal{ZP}(x; \eta)$ denotes the zero-truncated Poisson Distribution.

In our benchmarking, we reported RMSE and Bias for QM, MLI, and ML estimators. For the ML with height estimators, we reported results where the estimators knew either true heights or the resolved peak heights.





Conclusion

Contributions:
We have proposed a method of getting from a voltage peak to SE Yield through resolving peaks. The method uses deep learning on synthetic data combined with peak resolution to get accurate classification of the peak type. Additionally, an ML estimator to get from heights to SE Yield is shown to improve over prior works in RMSE and more significantly, bias.

Future Work:
One of the main improvements that can be made to make the resolved heights estimator have similar performance to the true heights is proper resolution of type 2 peaks. Future work into a supervised method for type 2 peak height extraction could yield better results.

Overall, the goal of future work should be to have the ML estimator that uses extracted/resolved heights tend towards the true heights estimator.

References:
[1] Agarwal, A.; Peng, M.; Goyal, V. K. Continuous-Time Modeling and Analysis of Particle Beam Metrology. 2023. https://doi.org/10.48550/arxiv.2303.04100.

I would like to thank everyone in the STIR Group who have supported me this summer.