PROCEEDINGS OF SCIENCE

Bayesian method for waveform analysis with GPU acceleration

Yuyi Wang[a,*] **and Benda Xu**[a]

[a] *Department of Engineering Physics & Center for High Energy Physics, Tsinghua University, Beijing, China*

E-mail: wangyy21@mails.tsinghua.edu.cn

Typically, in the reconstruction of events recorded by liquid scintillators, we need to extract the photo-electron (PE) hit times and PE charges from waveforms. We developed a new method called Fast Scholastic Matching Pursuit (FSMP). It is based on Bayesian principles, and the possible solutions are sampled with Markov Chain Monte Carlo (MCMC). To accelerate the method, we ported it to GPU, and could analyze the waveforms with 0.01s per waveform. This method will benefit event reconstruction. The position and energy resolution will be improved, as the method extracts all the information in the waveforms.

41st International Conference on High Energy physics - ICHEP2022
6-13 July, 2022
Bologna, Italy

*Speaker

PoS(ICHEP2022)923

1. Bayesian waveform analysis

We use liquid scintillator detector in neutrino experiments. Our final goal is the neutrino mass ordering (NMO). We need high resolution to achieve that goal. Therefore, we need very accurate waveform analysis and reconstruction method, to make use of total information in waveforms, and gain high energy resolution.

In Bayesian method, for waveform $\boldsymbol{w}$, expectation of number of PEs μ, and event time t_0, we have

$$p(\mu, t_0|\boldsymbol{w}) \propto p(\boldsymbol{w}|\mu, t_0)p(\mu, t_0) \tag{1}$$

where $p(\mu, t_0)$ is a Bayesian prior. With maximum likelihood estimation (MLE), to estimate the event energy and position $(E, \boldsymbol{r})$:

$$(\hat{E}, \hat{\boldsymbol{r}}) = \arg\max_{E,\boldsymbol{r}} p(E, \boldsymbol{r}|\mu, t_0, \boldsymbol{w}) = \arg\max_{E,\boldsymbol{r}} \frac{p(\mu, t_0|E, \boldsymbol{r})p(E, \boldsymbol{r})}{p(\mu, t_0|\boldsymbol{w})} \tag{2}$$

Therefore, it is important to estimate μ and t_0 with high resolution, to guarantee the resolution of latter steps. It is a Poisson process from μ, t_0 to PE sequence z, the expectation of this process is $\mu\phi(t - t_0)$. ϕ is a normalized shape function. $z = (t_1, t_2, \ldots)$ represents the times of PEs.

$$p(\boldsymbol{w}|\mu, t_0) = \widetilde{\sum_{z}} p(\boldsymbol{w}|z)p(z|\mu, t_0) \tag{3}$$

The coefficient space of z explodes because there are any number of possible z. We need Markov chain Monte-Carlo (MCMC) to sample the most possible ones.

2. The MCMC steps in FSMP

We sample t_0, z with MCMC [1, 2]. μ is estimated with MLE from posterior distribution. This algorithm is called fast stochastic matching pursuit (FSMP) [3]. Fast Bayesian methods are used in calculating [4]. Figure 1 gives a sketch of the MCMC chain steps.

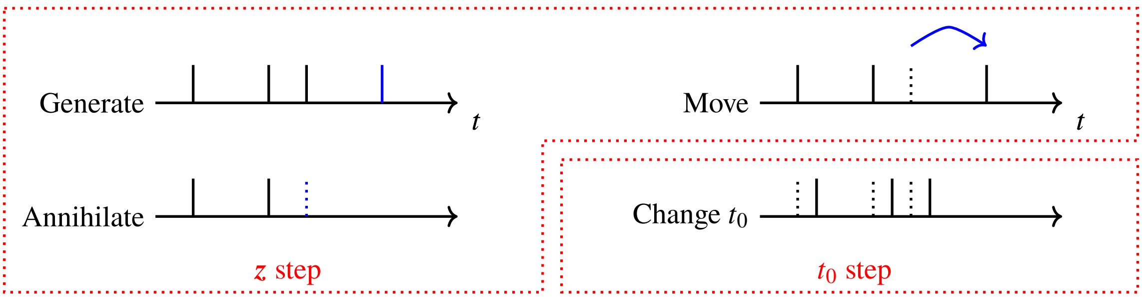

Figure 1: Sketch of the MCMC chain steps.

FSMP opens the opportunity to boost energy resolution (×1.07) and particle identification in PMT-based neutrino experiments. Figure 2 and 3 shows bias and resolution for μ and t_0 [3].

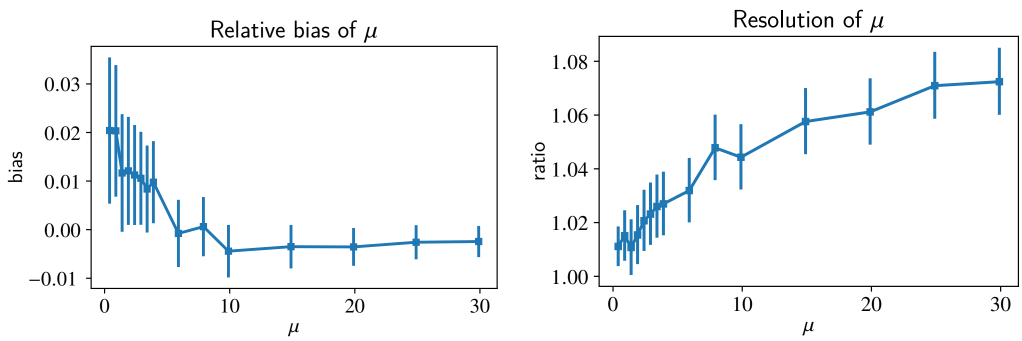

Figure 2: The resolution of μ is defined as $\frac{\sqrt{\mathrm{Var}[\hat{\mu}]}/\mathrm{E}[\hat{\mu}]}{\sqrt{\mathrm{Var}[N_{\mathrm{PE}}]}/\mathrm{E}[N_{\mathrm{PE}}]}$, N_{PE} is number of PEs.

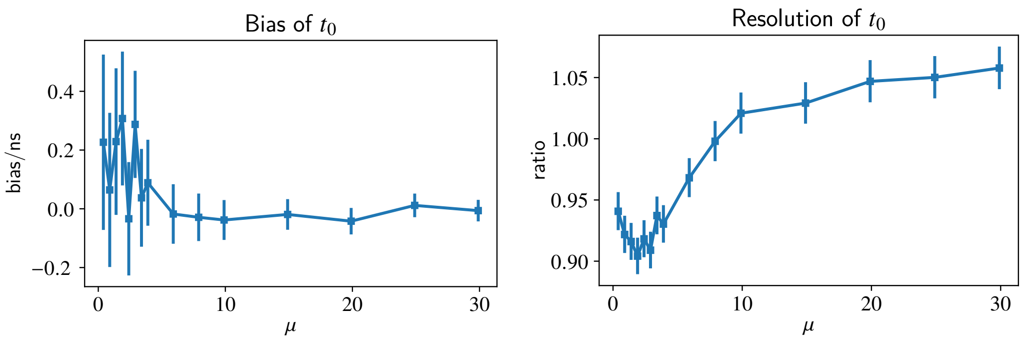

Figure 3: The resolution of t_0 is defined as $\sqrt{\frac{\mathrm{Var}[\hat{t}-t_0]}{\mathrm{Var}[\hat{t}_{\mathrm{ALL}}-t_0]}}$, $\hat{t}_{\mathrm{ALL}}$ is ideal estimator for t_0 by truth PE times.

3. GPU acceleration

The FSMP algorithm is accelerated with batched method on GPU. Figure 4 shows that in the batched algorithm, a lot of waveforms are operated together, instead of analyzing them one by one.

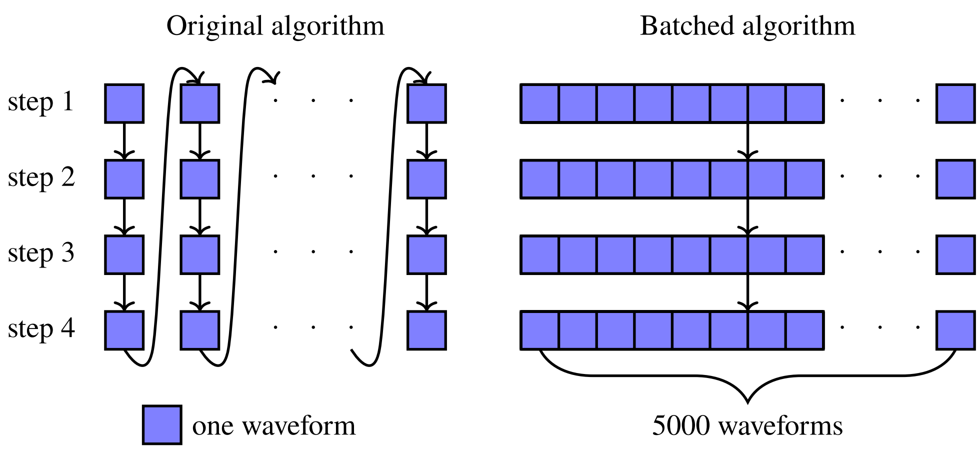

Figure 4: A sketch of comparison between original algorithm and batched algorithm.

Figure 5 shows that the batched method performs 0.01s per waveform with batched size 5000 on NVIDIA® A100, and it is faster than original algorithm on CPU by 2 orders of magnitude.

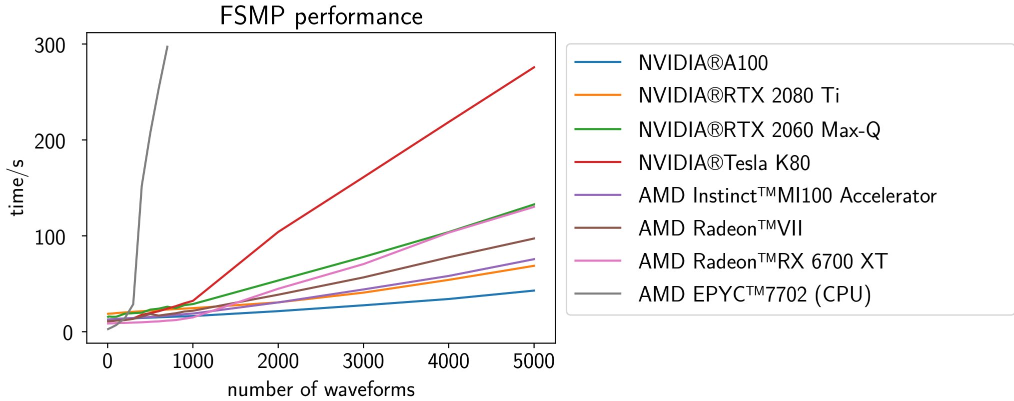

Figure 5: Comparison of performance of FSMP on different devices.

4. Summary

FSMP is a method which makes use of total information in waveforms. It performs fast on consumer GPUs, with high precision results. FSMP method proves the practicability of the Bayesian method, and we are willing to extend it to event reconstruction.

References

[1] W.K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, *Biometrika* **57** (1970) 97.

[2] L. Tierney, *Markov chains for exploring posterior distributions*, *The Annals of Statistics* **22** (1994) 1701.

[3] D. Xu, B. Xu, E. Bao, Y. Wu, A. Zhang, Y. Wang et al., *Towards the ultimate PMT waveform analysis for neutrino and dark matter experiments*, *Journal of Instrumentation* **17** (2022) P06040.

[4] P. Schniter, L.C. Potter and J. Ziniel, *Fast Bayesian matching pursuit*, in *2008 Information Theory and Applications Workshop*, pp. 326–333, Jan., 2008, DOI.