Summary of the Workshop

Chuanhui Hao

Tsinghua University

haoch23@mails.tsinghua.edu.cn

April 18, 2025

Probe

For vertex $\{E, \boldsymbol{r}, t_0\}$, the PE counts received on the J_{th} PMT follows an inhomogeneous Poisson process from a time-dependent function R_j.

- Probe function: $R_j(t; E, \boldsymbol{r}, t_0) \to R(x, y, z, t)$
- Expected PE counts: $\lambda(x, y, z) = \int_0^T R(x, y, z, t)\mathrm{d}t$. (marginal distribution)

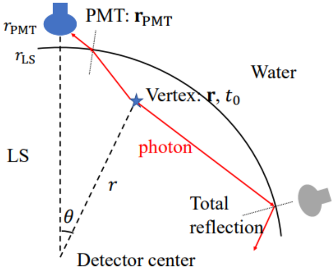

Fig 1: Geometric model for the probe

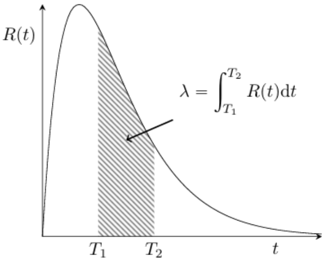

Fig 2: Schematic curve of $R(t)$

Generalized Additive Model(GAM) fitting

regression formulation of the expected PE counts $\lambda(x, y)$ and Probe function $R(x, y, t)$

$$\mathrm{nPE} \sim \mathrm{Poisson}(\lambda), \mathrm{nPE} \sim \mathrm{Poisson}(\lambda_{\mathrm{tbin}})$$

$$\log \lambda = \mathrm{te}(x, y, z) + \log(\mathrm{nEV})$$

$$\log R = \mathrm{te}(x, y, z, t) + \log(\mathrm{nEV}) + \log(\mathrm{tbinwidth})$$

- We use big MC dataset via **generalized additive model (GAM)** to fit the Probe function.
- We use the **bam** function of the **mgcv** package in R to perform the fitting.
- **bam** is a function that supports GAM for very large datesets.

limitation

As shown in the regression formulation, λ and R require separate fitting procedures, which may lead to insufficient accuracy in the fitting results.

Problem & Goal

Problem

- According to the mathematical relationship between λ and R, if we fit the integral of R, we only need one fitting to get them.
- The integral $\int_0^t R(\vec{r}, \tau)\mathrm{d}\tau$ must increase strictly monotonically. But the mgcv package does not support monotonic spline fitting for tensor product functions.

Goal

- Realize the monotonic spline for tensor product in **bam** function by modifying the source code.

Progress during the workshop

About **bam**

- Read two articles, Understood how it supports the GAM for very big datasets. Briefly, they used **iteration algorithm** for **blocked** matrix to deal with the big matrix.
- Reviewed key code blocks about the iteration algorithm.

About **scam**

- Found another package similar to mgcv, named scam. It provides the shape constrained GAM model.
- Got the constraint principles of scam via reading article and code.
- Found that scam supports tensor product smoothing constructor for a bivariate function monotone increasing in a certain covariate.
- Had successful experiments on 1-D and 2-D datasets.

2-D experiment results

$$\text{nPE} = \text{te}(z, t).$$

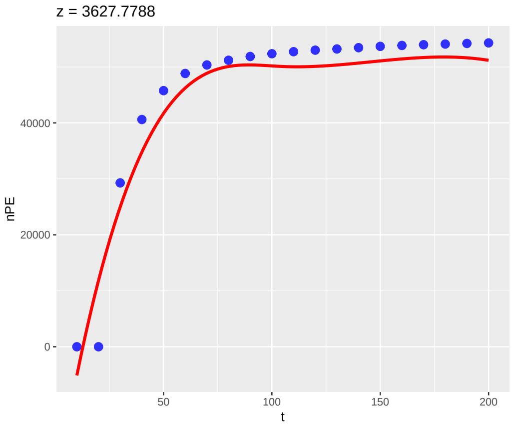

Fig 3: B-spline

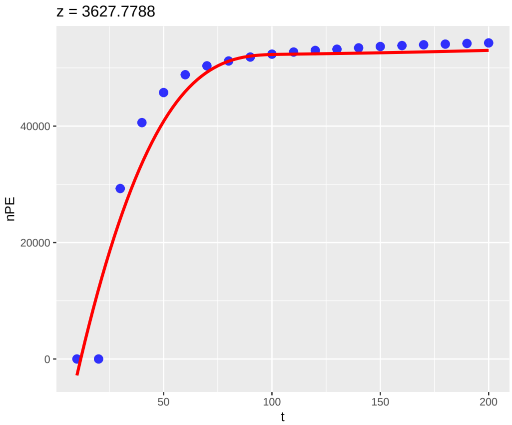

Fig 4: Monotonic spline in t-direction with **scam**

TODO

1. Focus on modifying the code of **scam** to realize the multivariate form.
2. Integrate **scam** modifications with **mgcv**'s **bam** function to support very large datasets.

Thank you for listening!
Q & A