GNN for Water Cherenkov Detector

Junjie Xia, 5.28.2021

Fixing Correlation Implementation

Last time showed:

2. Say that

This can be interesting?

$$X_1 = Z_1 + Z_c$$
$$X_2 = Z_2 + Z_c$$

where $Z_i \sim \mathcal{N}(0, e^{O_i})$ are "private information" and $Z_c \sim \mathcal{N}(0, e^{O_3})$ is a "shared" gaussian information. Then we will have that

$$v_1 = e^{O_1} + e^{O_3}$$
$$v_2 = e^{O_2} + e^{O_3}$$
$$\rho = \frac{e^{O_3}}{\sqrt{v_1 \cdot v_2}}$$

But in this case, the correlation can only be positive!
Instead:

$$f_{\mathbf{X}}(x_1, \ldots, x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}}$$

$$\Sigma^{-1} = A^{\mathrm{T}}A = \begin{pmatrix} \alpha_{11} & 0 & 0 & 0 \\ \alpha_{12} & \alpha_{22} & 0 & 0 \\ \alpha_{13} & \alpha_{23} & \alpha_{33} & 0 \\ \alpha_{14} & \alpha_{24} & \alpha_{34} & \alpha_{44} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ 0 & 0 & \alpha_{33} & \alpha_{34} \\ 0 & 0 & 0 & \alpha_{44} \end{pmatrix}$$

The matrix $A^T A$ has to be positive definite, which imposes restriction on the network parameter output. Specifically:

1. The max value has to be on diagonal.
2. The average of any 2 diagonal elements must be larger than the real part of corresponding off-diagonal term $a_{ii}+a_{jj} > |2\mathrm{Re}(a_{ij})|$.

But since we are only using 2-dimensional Hermitian matrices, this can be relatively simple to apply.

Fixing Correlation Implementation

$$f_{\mathbf{X}}(x_1,\ldots,x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}}$$

$$\Sigma^{-1} = A^{\mathrm{T}}A = \begin{pmatrix} \alpha_{11} & 0 & 0 & 0 \\ \alpha_{12} & \alpha_{22} & 0 & 0 \\ \alpha_{13} & \alpha_{23} & \alpha_{33} & 0 \\ \alpha_{14} & \alpha_{24} & \alpha_{34} & \alpha_{44} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ 0 & 0 & \alpha_{33} & \alpha_{34} \\ 0 & 0 & 0 & \alpha_{44} \end{pmatrix}$$

1. The max value has to be on diagonal.
2. The average of any 2 diagonal elements must be larger than the real part of corresponding off-diagonal term $a_{ii}+a_{jj} > |2\mathrm{Re}(a_{ij})|$.

In NN, apply the constraint on off-diagonal terms by

- A_{12} = (Sigmoid(A_{12}) - 0.5) * (A_{11} + A_{22})

So that the resulted diagonal term is within (-0.5, 0.5)*(A_{11}+A_{22}), and hence satisfying the requirements.

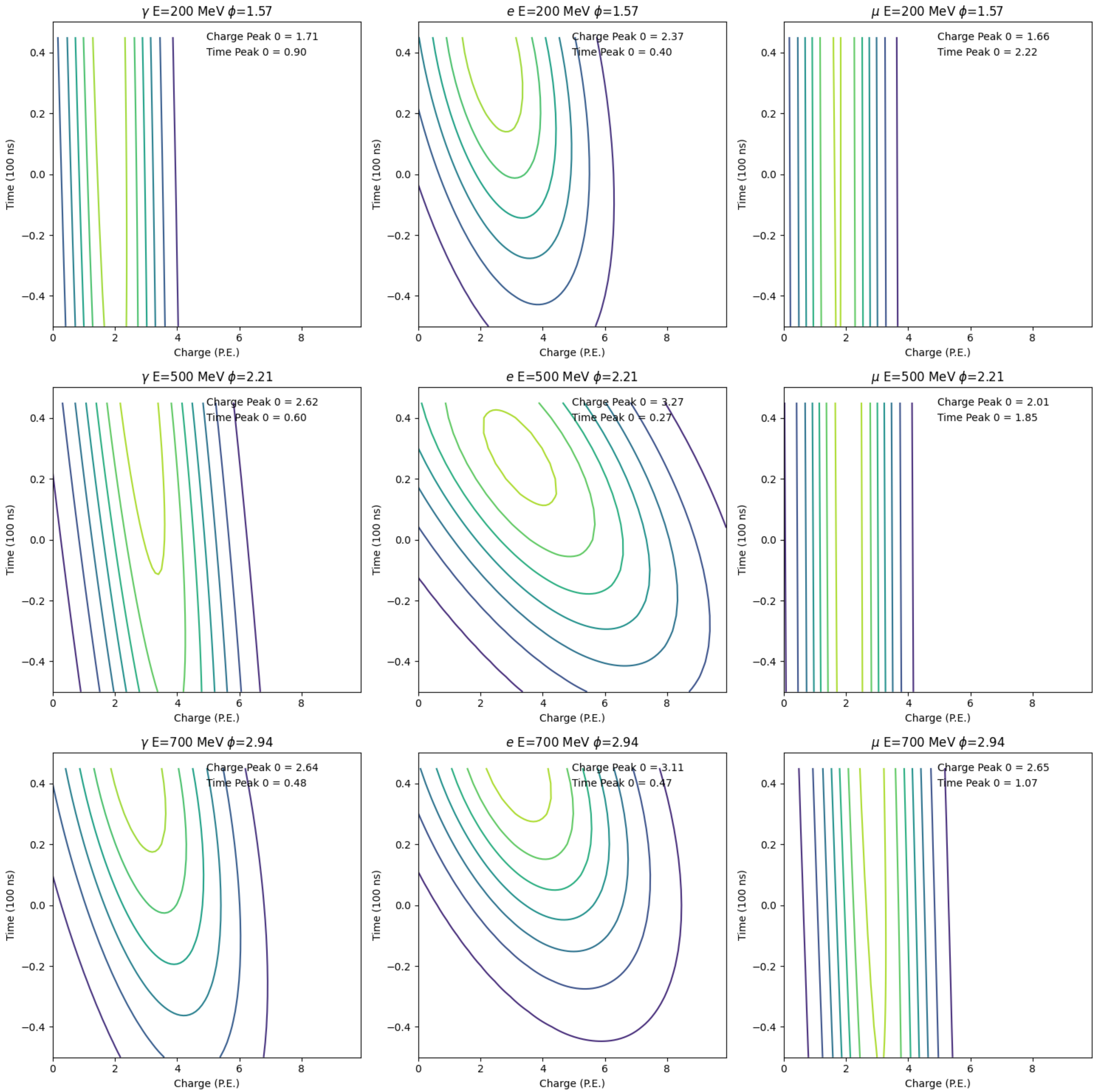
However, the more correct way of this should be done on the squared terms, i.e. A_{11}^2, etc.

But with the square, there is an extra complication to determine the sign of A_{12}. The current implementation can result in a wider range of A_{12} than what's mathematically allowed (since $a^2 + b^2 < (a+b)^2$), but hopefully it will be a small effect…

Example of The New Implementation

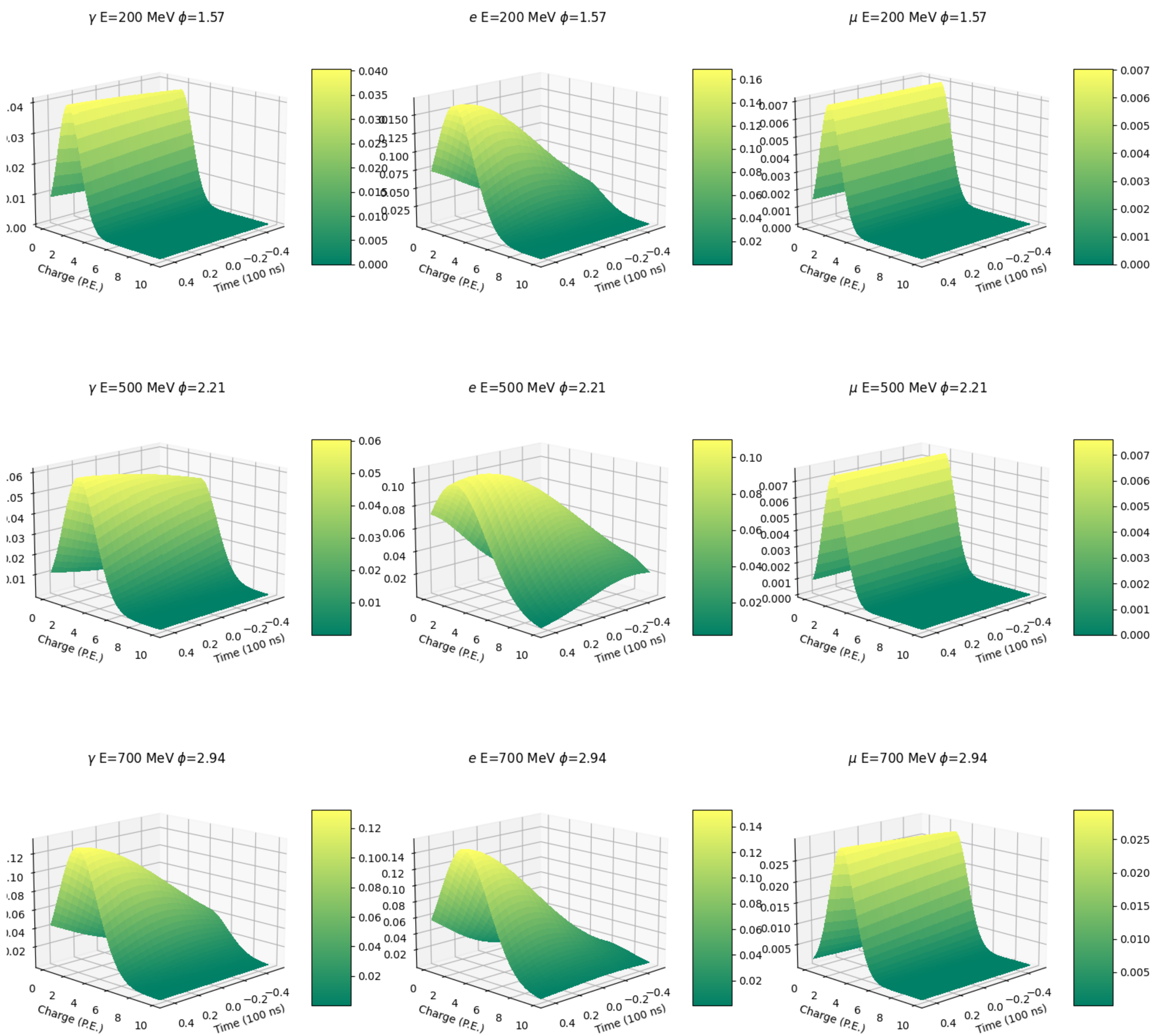
1 Gauss

Sorry time unit should be 10 ns, not 100 ns



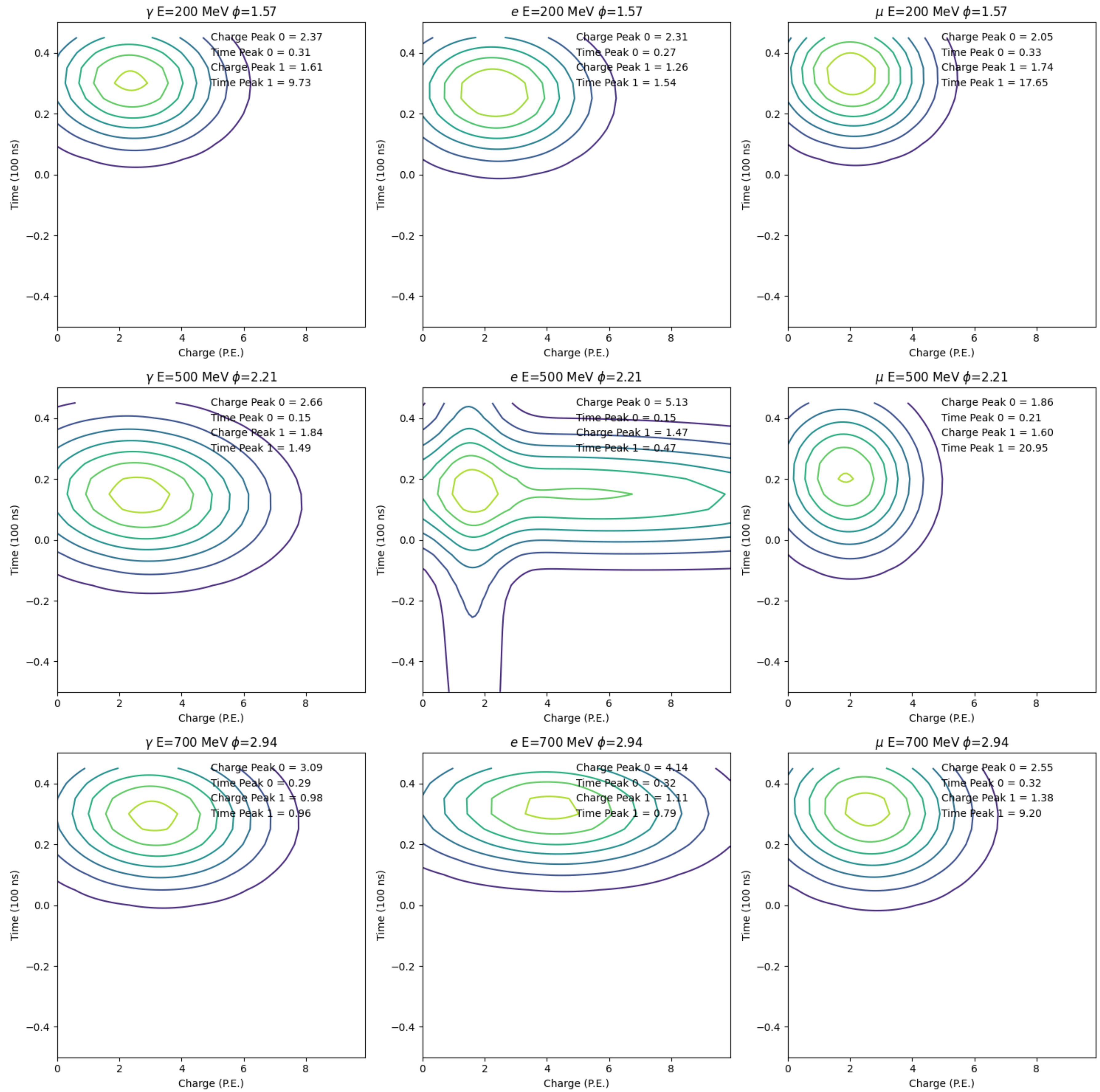
Example of The New Implementation

1 Gauss



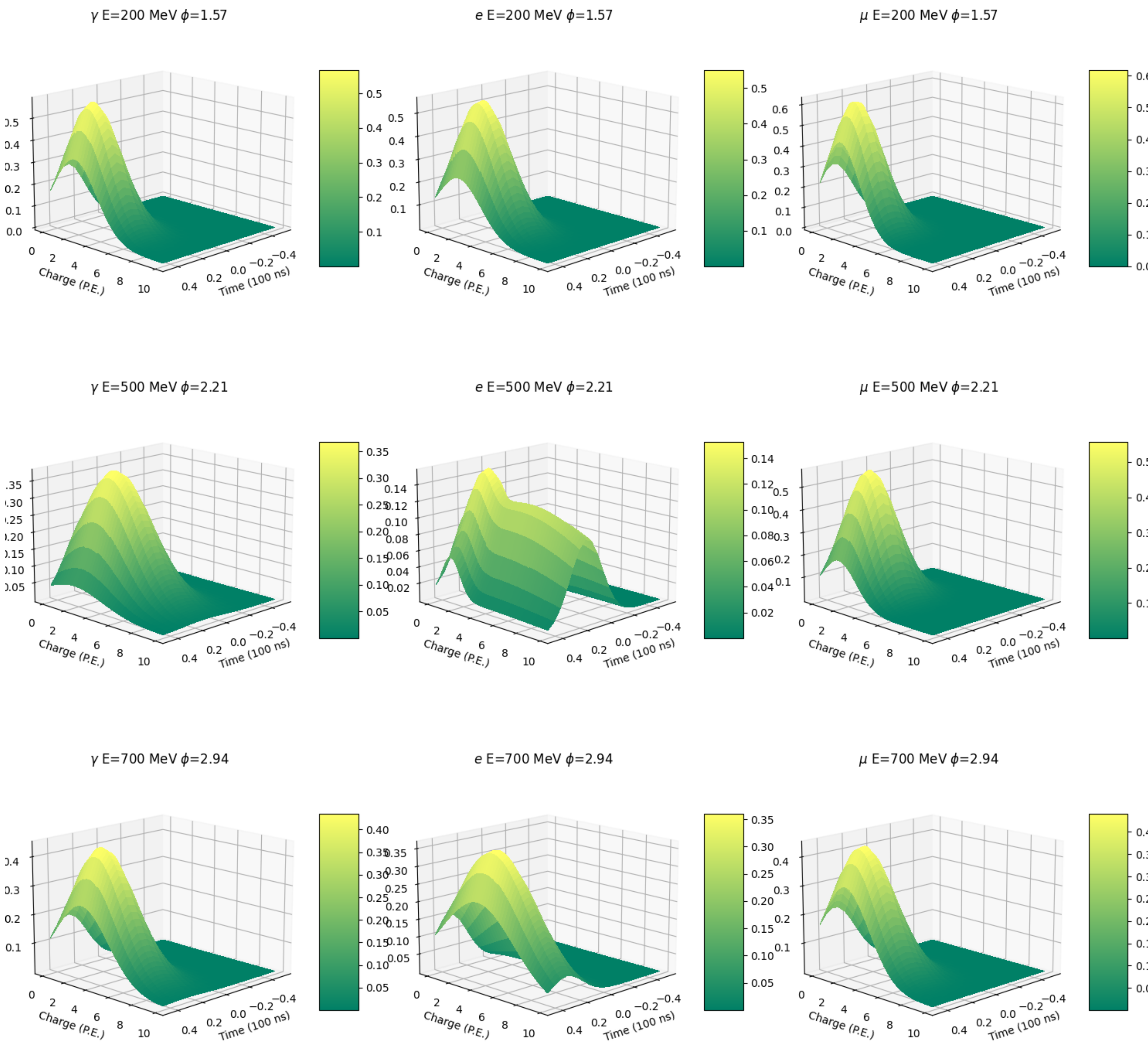
Example of The New Implementation

2 Gauss



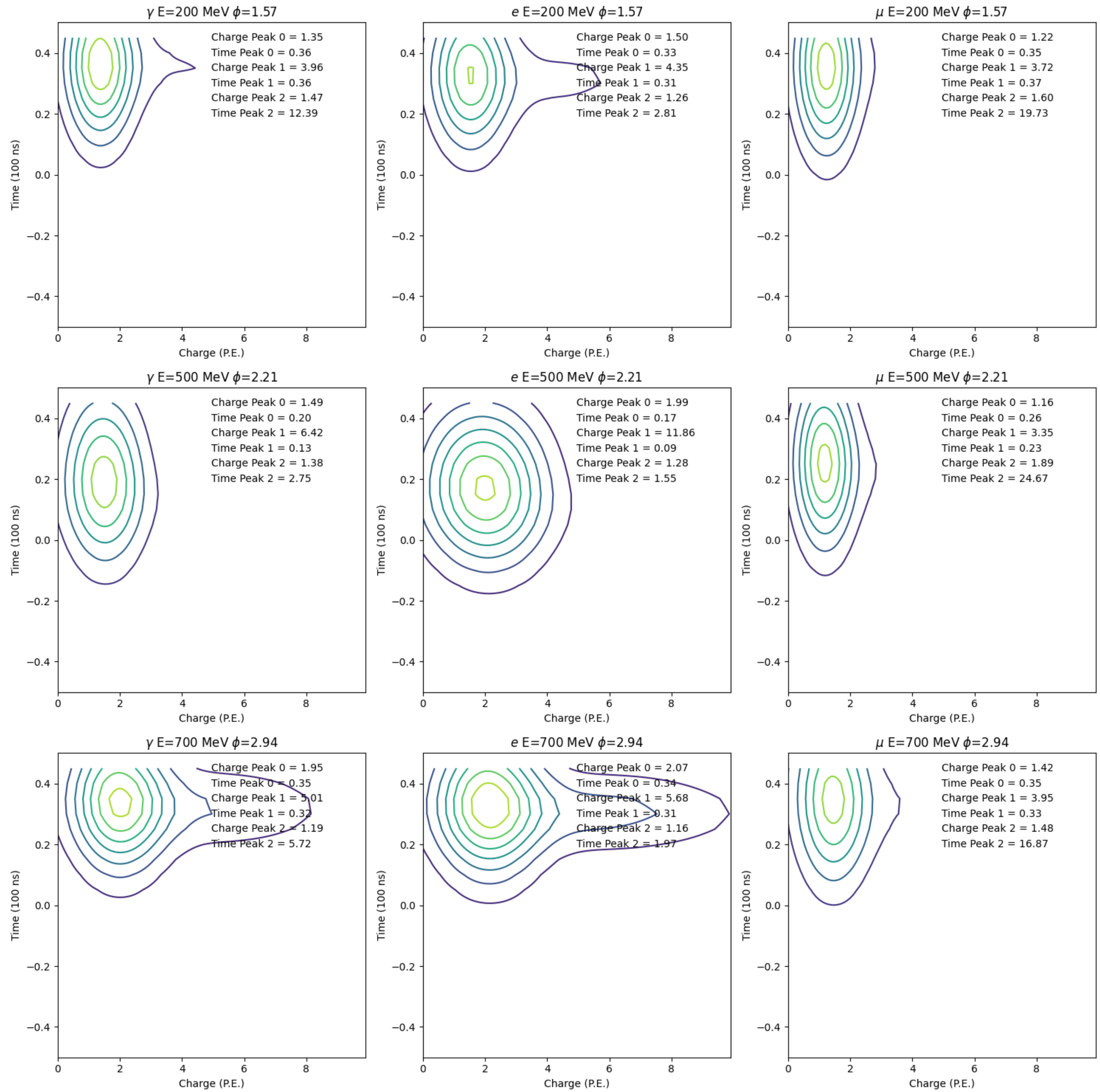
Example of The New Implementation

2 Gauss



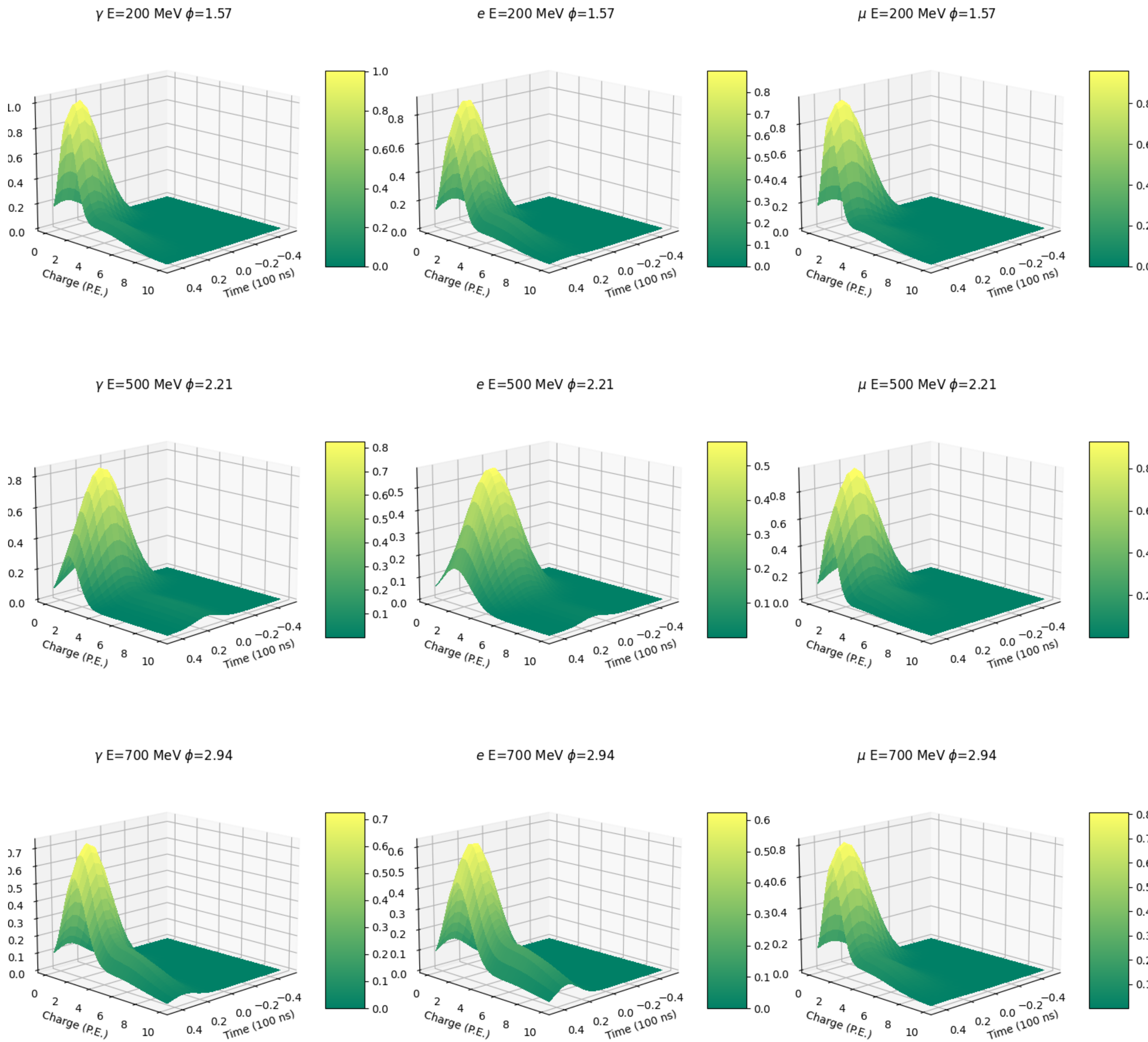
Example of The New Implementation

3 Gauss



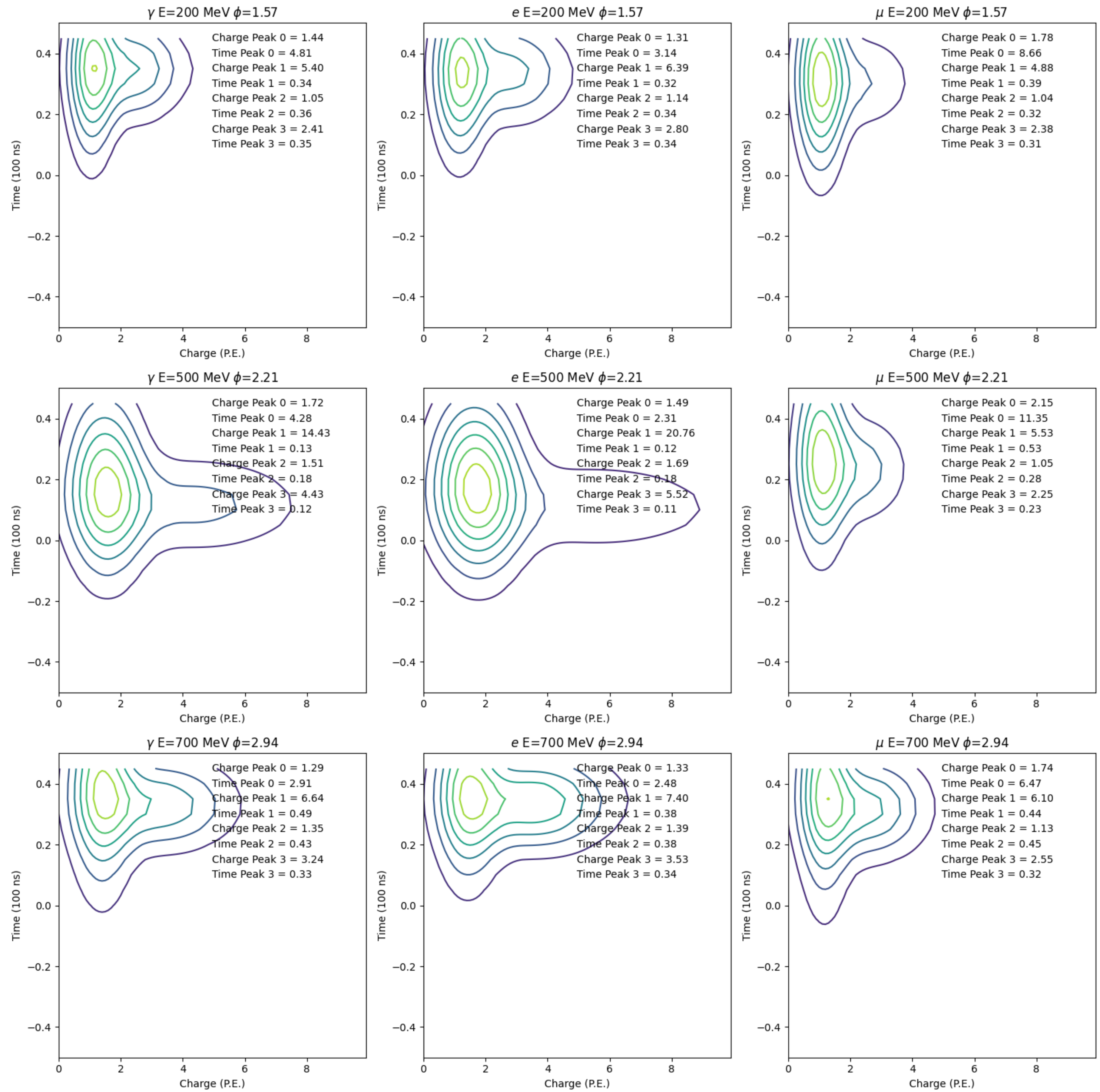
Example of The New Implementation

3 Gauss



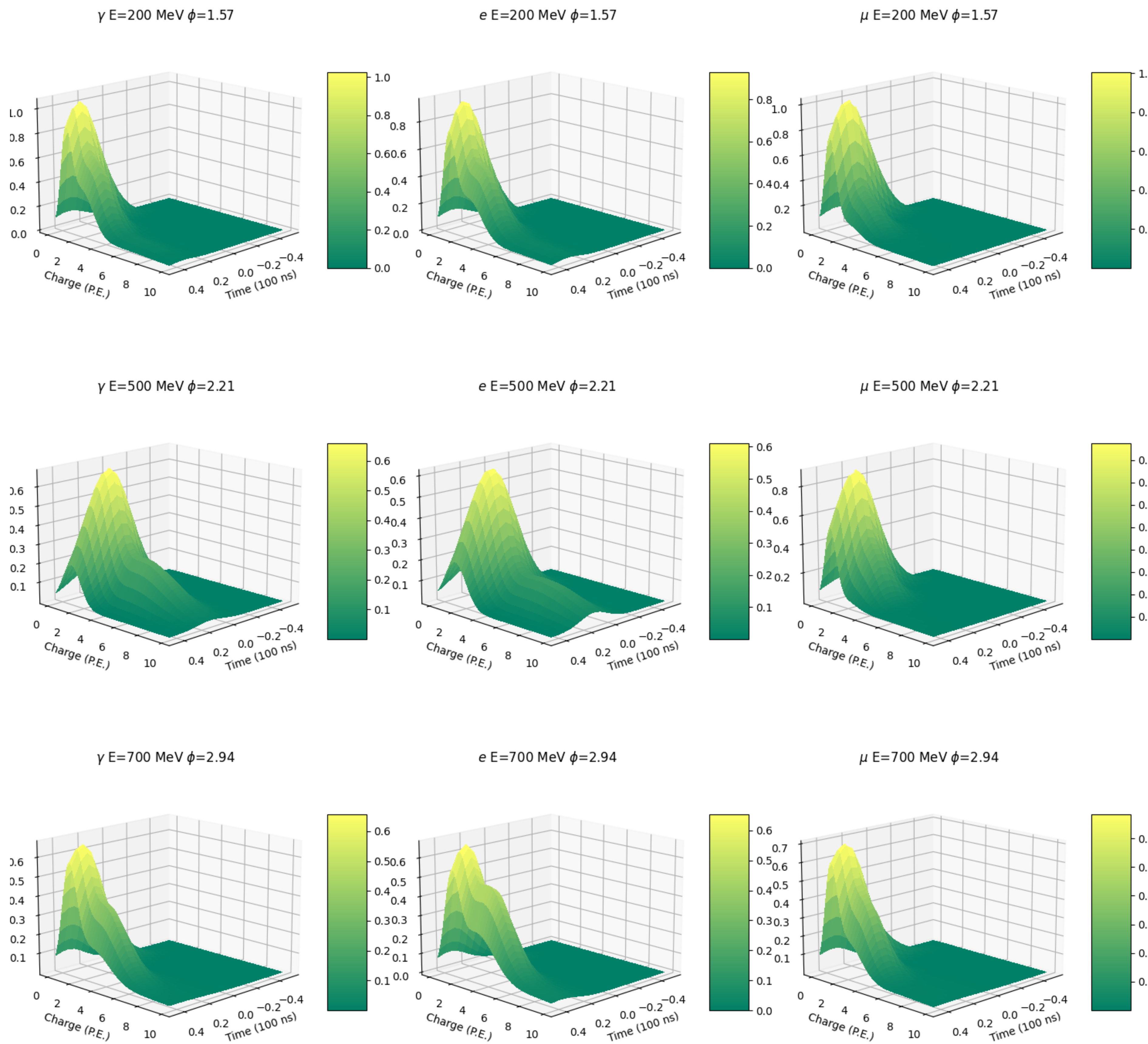
Example of The New Implementation

4 Gauss



Example of The New Implementation

4 Gauss



- 5 Gauss also runs fine. But the job lasted longer than 12 hrs and got terminated, 87% done.

- Still have intermediate outputs.

Summary

$$f_{\mathbf{X}}(x_1,\ldots,x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}}$$

$$\Sigma^{-1} = A^{\mathrm{T}}A = \begin{pmatrix} \alpha_{11} & 0 & 0 & 0 \\ \alpha_{12} & \alpha_{22} & 0 & 0 \\ \alpha_{13} & \alpha_{23} & \alpha_{33} & 0 \\ \alpha_{14} & \alpha_{24} & \alpha_{34} & \alpha_{44} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ 0 & 0 & \alpha_{33} & \alpha_{34} \\ 0 & 0 & 0 & \alpha_{44} \end{pmatrix}$$

Implemented this “matrix method” in the NN, with changed definition of network outputs. The network now outputs a_{11}, a_{12}, a_{22}, which in terms of sigma and correlation is

$$a_{11} = \frac{1}{\sigma_t\sqrt{1-\rho^2}}$$

$$a_{22} = \frac{1}{\sigma_q}$$

$$a_{12} = \frac{-\rho}{\sigma_q\sqrt{1-\rho^2}}$$

Working on the maths for the LogNorm case.

Back Ups