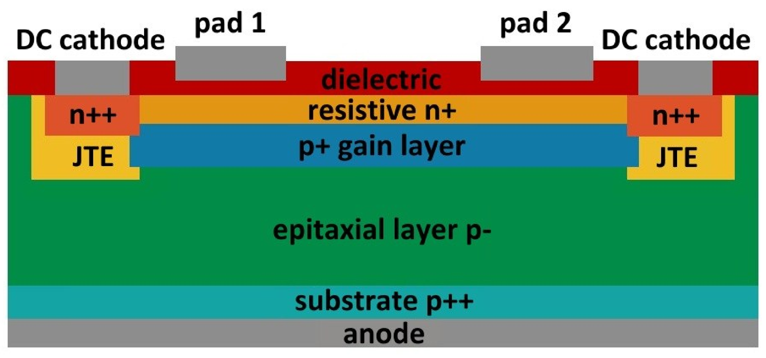
KNN-Based Position Reconstruction Algorithm for AC-Coupled Low Gain Avalanche Diodes

Xiaoxu ZHANG[1,2], Mei ZHAO[1], Lei ZHANG[2], Mengzhao LI[1,4], Weiyi SUN[1,3], Zhijun LIANG[1], João Guimaraes da Costa[1]

[1]Insitute of High Energy Physics, CAS, Beijing, China [2]Nanjing University, Nanjing, China [3]University of Chinese Academy of Sciences [4]Spallation Neutron Source Science Center, Dongguan, China

Introduction

To meet the particle detection requirements of next-generation high-energy physics experiments, the AC-Coupled Low Gain Avalanche Diodes (AC-LGADs) have emerged as a breakthrough technology. The figure presents a cross-sectional schematic of an AC-LGAD structure. The intentional resistivity, achieved through deliberate low doping concentration of the n+ layer, causes the path length and direction of electron diffusion to vary with the particle hit position. As a result, the signal amplitudes induced on different metal pads exhibit a spatial dependence, enabling reconstruction of the hit position based on their relative signal strengths.



The complex nonlinear relationship between the signal ratios of metal pads and the particle hit position limits the effectiveness of analysis models, which motivated the adoption of K-Nearest Neighbors (KNN) algorithm for position reconstruction.

Modeling and Simulating Charge Diffusion

Developing reconstruction algorithms requires simulating the charge diffusion and collection process within the resistive readout structure to establish a mapping between particle hit positions and pad signal ratios.

The proposed 2D simplified model preserves the key physics of charge diffusion in the resistive readout structure while significantly reducing computational complexity. Its construction is relied on three fundamental principles (Ohm's law, current conservation and Dirichlet boundary conditions) :

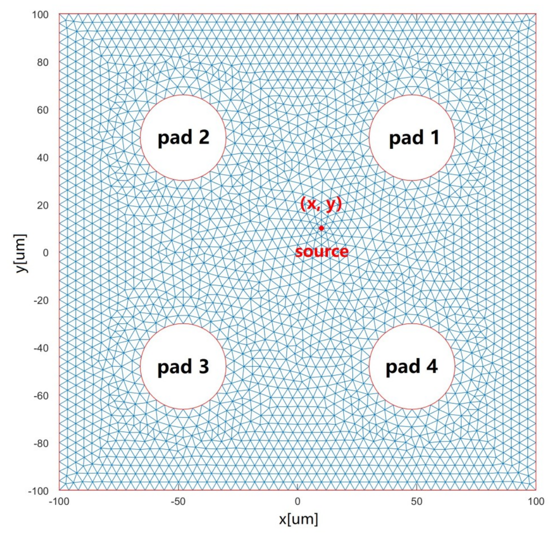
$$\vec{j}(x,y,t) = -\sigma \nabla V(x,y,t) \tag{1}$$

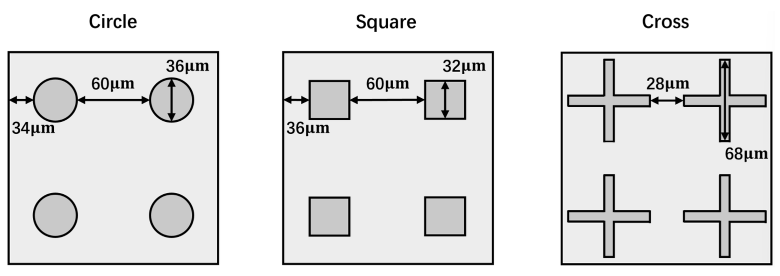
$$\frac{\partial V(x.y.t)}{\partial t} + \frac{1}{c}\nabla \vec{j}(x,y,t) = \frac{1}{c} I(x,y,t) \tag{2}$$

$$V(x,y,t) = 0 \tag{3}$$

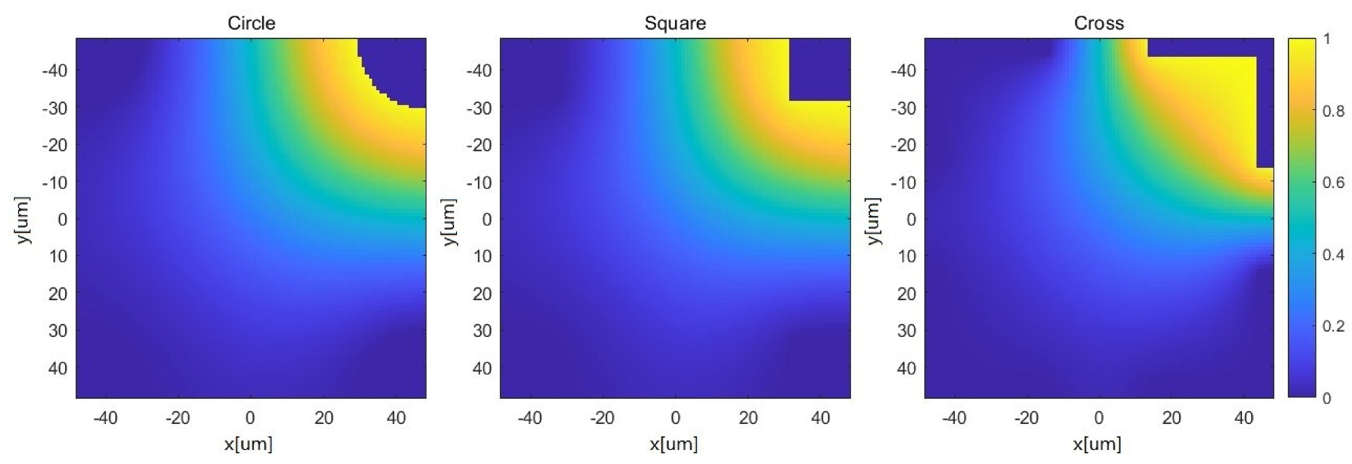
Finite element analysis (FEA) and iterative solvers are employed as the core numerical methods to solve the 2D model. The computational workflow was implemented using MATLAB's PDE Toolbox.

The simulation includes an annular DC cathode enclosing an inner area of 200 μm × 200 μm and four metal pads. Three different metal pad patterns (circle, square and cross), were simulated with equal pad areas and center-to-center spacing. Taking the circle-shaped metal pads as an example, the geometric configuration is illustrated. The n+ layer is defined as the entire square region, subtracted by the four metal pads and a transient source point modeled as a 1 μm-radius circle. The position of the source point is dynamically parameterized by (x,y) to simulate diverse particle hit positions.





By varying the position of the transient source, the particle hit points are scanned across the whole area at 1 μm intervals. The results are shown below, where the x and y axes represent the particle hit position and the color axis indicates the normalized signal proportion of the upper-right metal pad (Pad 1), relative to the total signals from all four pads.



When the particle hit around Pad 1, the signal proportion of this pad exhibits strong positional sensitivity, manifested as a steep gradient of signal proportion across the area. Near the pad edges, where one pad collects nearly all charges, position reconstruction becomes more challenging. This effect is particularly pronounced for cross-shaped pads, whose extended perimeter has larger ambiguous regions than circular or square pads.

KNN Algorithm and Feature Optimization

Let Q_1, Q_2, Q_3, Q_4 denote the normalized charge proportions of the four metal pads. For a test data point in the four-dimensional feature space (Q_1, Q_2, Q_3, Q_4), the KNN algorithm identifies the k closest training samples based on the Euclidean distance defined as:

$$d = \sqrt{\sum_{i=1}^{4} (Q_i^{test} - Q_i^{train})^2} \tag{4}$$

The coordinates of the unknown hit position are then reconstructed as the weighted average of the k neighbors' spatial coordinates, with the weights inversely proportional to their distances:

$$x^{rec} = \frac{\sum_{i=1}^{k} w_i x_i^{train}}{\sum_{i=1}^{k} w_i},\ y^{rec} = \frac{\sum_{i=1}^{k} w_i y_i^{train}}{\sum_{i=1}^{k} w_i} \tag{5}$$

The training data are generated using the MATLAB PDE 2D simulation, where particle hit positions were scanned at 1 μm intervals. The test points are randomly generated. Their corresponding charge proportions Q_1, Q_2, Q_3 and Q_4 are also from simulation. These charge proportions serve as input for position reconstruction using the trained KNN algorithm. The positional residual of each test point is defined as the distance between the reconstructed position and the true position as:

$$Residual = \sqrt{(x_{rec} - x_{true})^2 + (y_{rec} - y_{true})^2} \tag{6}$$

The reconstruction accuracy was evaluated using the Root Mean Square Error (RMSE) of N test points as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} Residual_i^2} \tag{7}$$

The optimal k-value is determined by the k-curve. Figure (a) (b) (c) show the RMSE as a function of k-value for different pad patterns. Small k-values may overfit noise from individual samples, while large k-values may oversmooth the local geometry and degrade spatial resolution.

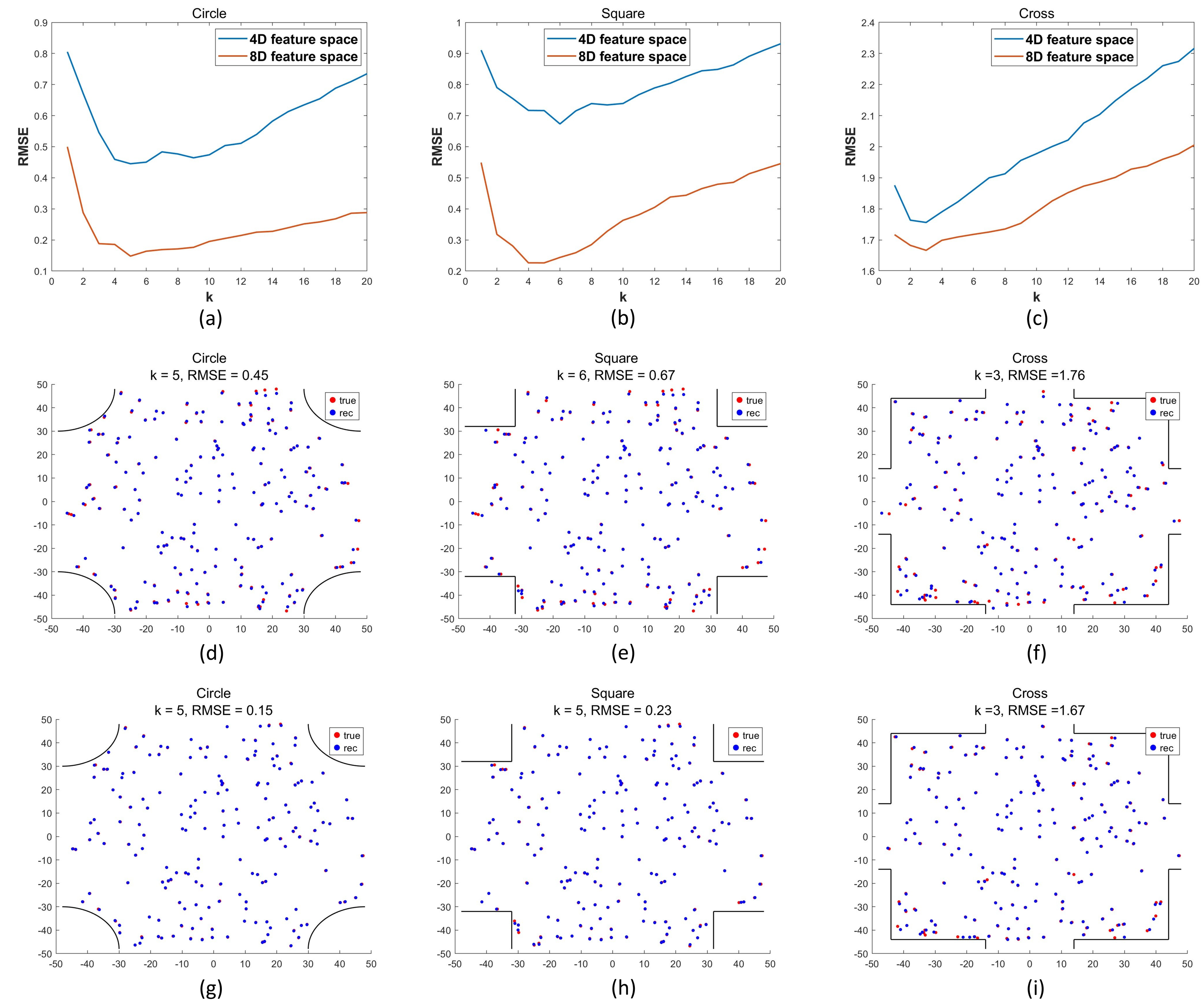
The reconstruction results for different pad patterns are shown in Figure (d) (e) (f). As expected, the reconstruction accuracy degrades significantly near the edges of metal pads, consistent with the locally predominant signal distribution observed in simulation. To correct the non-linearity, four additional features were incorporated:

$$R_1 = \frac{Q_1 + Q_4}{Q_2 + Q_3},\ R_2 = \frac{Q_2 + Q_3}{Q_1 + Q_4},$$
$$R_3 = \frac{Q_1 + Q_2}{Q_3 + Q_4},\ R_4 = \frac{Q_3 + Q_4}{Q_1 + Q_2} \tag{8}$$

The original four-dimensional feature space was extended to an eight-dimensional feature space ($Q_1, Q_2, Q_3, Q_4, R_1, R_2, R_3, R_4$) and the Euclidean distance is modified to:

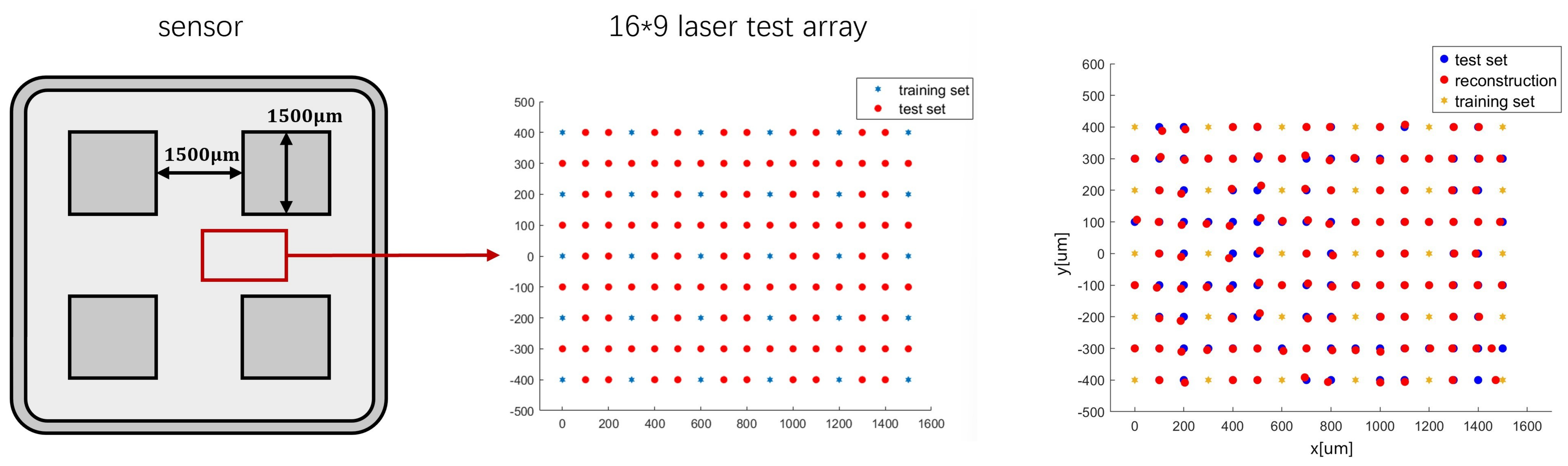
$$d = \sqrt{\sum_{i=1}^{4} (Q_i^{test} - Q_i^{train})^2 + \sum_{i=1}^{4} \left(1 - \frac{R_i^{train}}{R_i^{test}}\right)^2} \tag{9}$$

This optimized feature space improves performance by combining direct charge proportions and their combinatorial relationships. The improved reconstruction results are shown in Figure (g) (h) (i).



Experimental Validation

The AC-LGAD sensor is fabricated on 8-inch wafers by the Institute of Microelectronics (IME). A 1064 nm laser with a focused spot size of ~10 μm (3σ) was used in conjunction with a three-dimensional translation platform with ~1 μm precision. A 16×9 laser hit array was scanned with 100 μm step between points. At each position, 1000 events were recorded by the oscilloscope and the corresponding signals of four metal pads were averaged for position reconstruction.



To validate the reconstruction algorithm, the 144 scanning points were divided into 30 for training and 114 for testing. Employing the KNN algorithm with the 8-dimensional feature space under the optimal k-value (k=3), the reconstruction results are shown above. As seen, the reconstruction positions match well with actual test positions. With 3000 μm pitch size and 300(x)/200(y) μm training grid intervals, the achieved RMSE is 11.2 μm.

Conclusion

In summary, this work has addressed several key aspects:

① A computationally efficient 2D charge diffusion model for resistive readout structure, grounded in fundamental physical principles, was developed and numerically implemented using MATLAB's PDE Toolbox.

② Comprehensive simulations scanning particle hit position were performed, generating essential datasets that revealed the complex, position-dependent signal distribution characteristics of the pixel-type AC-LGADs for various pad geometries.

③ To address the nonlinear position-signal mapping, a KNN-based algorithm was proposed and optimized. By expanding the feature space and selecting the optimal neighborhood size (k-value), the optimized algorithm significantly improved the reconstruction accuracy, particularly reducing residuals near pad edges for circular and square configurations.

④ The optimized algorithm was validated using experimental data. A polynomial surface fitting and interpolation effectively use sparse training data, enabling the algorithm to achieve a satisfactory positional reconstruction accuracy.

By integrating modeling, simulation, algorithm development and experimental validation, this study establishes a systematic framework for validating AC-LGADs design and optimizing the position reconstruction algorithm, exploiting the potential of AC-LGADs for next-generation high-energy physics experiments.

Email: xiaoxuzhang@smail.nju.edu.cn