

# User Guide for AIM

In our algorithm referred to as Adaptive Intersection Maximization (AIM), three parameters can potentially affect the final drift tracking performance: intersection distance (D), tracking time interval (m), and the radius of the local region (R). However, the impact of these parameters varies.

The intersection distance (D) is primarily determined by the localization precision of the dataset. Typically, its value can be set to around three times the localization precision for optimized performance. However, unlike model-based algorithms, our AIM is quite robust to unoptimized parameters. As shown in Supplementary Figure S2B, AIM can achieve sub-nanometer tracking precision with a fixed intersection distance ( $D=20\text{nm}$ ) for datasets with resolutions ranging from 10 nm to 40 nm. In practice, if users have little information about the localization precision of the dataset, they can simply set D to 20 nm for near-optimal performance.

The radius of the local region (R) is primarily determined by the magnitude of the drift. Typically, its value can be set to around three times the magnitude of the drift to handle extreme jumps in a short period. Since users may not have access to drift information before drift estimation, we have set a fixed value of 60 nm in practice. This value can handle sudden drifts of up to 120 nm in 0.2 seconds. With  $R = 60\text{ nm}$ , AIM maintains its state-of-the-art performance for all numerical and biological experiments in this study.

The tracking interval (m) is primarily determined by the number of localizations per frame. Typically, it can be set to 20 to 50 frames for large datasets with a large field of view (e.g.,  $2048 \times 2048$  pixels), and 50 to 100 frames for datasets with a small field of view (e.g.,  $128 \times 128$  pixels). For users with limited knowledge of localization density, opting for  $m = 100$  frames carries minimal risk for most datasets.

For less experienced users, choosing  $D = 20\text{nm}$ ,  $R = 60\text{nm}$ , and  $m = 100$  frames should yield satisfactory results across various image sizes and imaging conditions without the need for parameter fine-tuning.