# A   Experiments

In this section, we provide further experimental visualization. Our implementation codes can be found in SM.

In Fig. 5, we plot the temperature as well as the accuracy curve for CIFAR10-DVS experiments. The spiking ResNte-18 was trained for 300 epochs with cosine learning rate decay. We initialize b to 1.0 and use Δb=0.2 to optimize both temperature for gradient calculation and the weights parameters for network learning. For the initial 25 epochs, the temperature remains under 1. We think this is primarily due to the complex loss landscape at the starting stage of the training. During the 20 to 80 epoch, the temperature continues to be increasing. In this stage of learning the FDG desires to learn a sharp, narrow gradient with low bias but high variance. We also find the accuracy curve becomes more unstable in this stage of the training, as a consequence of the sharp gradient. After 80 epochs, the temperature becomes decreasing. Although the decreasing process is not aggressive as the former increasing stage, the trend is still obvious. From 80-250, the temperature decreases from 4.0 to 1.6 with some fluctuations. Then, at the end of the training stage, the temperature again becomes increasing.

The intuition behind this tuning mechanism may remain unexplainable right now, however, we think it is the intrinsic nature of the gradient estimation in SNN. Through FDG, we are able to optimize the gradient calculation along with the weights learning.
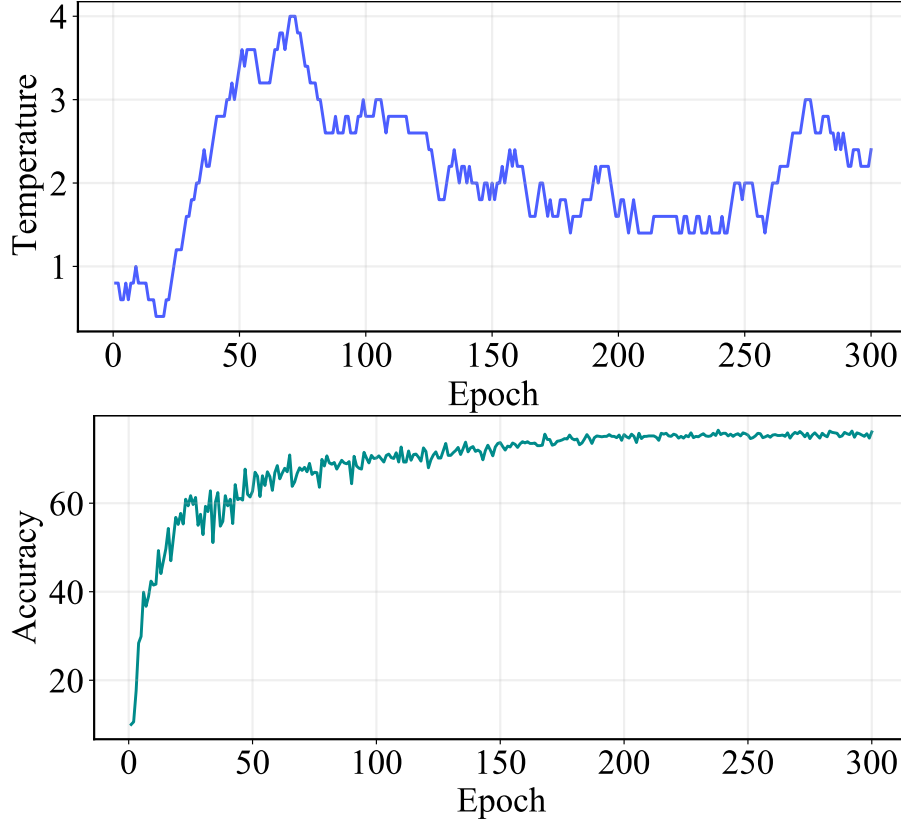
Figure 5: Uper side: the evolutution of temperature b during optimization. Lower side: the test accuracy curve.