Towards Bridging the Performance Gaps of Joint Energy-based Models
Supplementary Material

Xiulong Yang, Qing Su, and Shihao Ji
Georgia State University
{xyang22,qsu3,sji}@gsu.edu

A. Experimental Details

To have a fair comparison, we largely follow the settings of JEM [2] and JEM++ [5], and train our models based on the Wide-ResNet 28x10 architecture [6] for 200 epochs. We use SGD for CIFAR10 and CIFAR100 with an initial learning rate of 0.1 and 0.01, respectively, and decay the learning rate by 0.2 at epoch [60, 120, 180] for most cases. Apart from this, we find that the cosine learning rate scheduler can be adopted for SADA-JEM, which achieves much better accuracy and FID on CIFAR10.

The hyper-parameters used in our experiments are listed in Table 1.

Table 1. Hyper-parameters of SADA-JEM for CIFAR10 and CIFAR100.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of SGLD steps $K$</td>
<td>5, 10, 20</td>
</tr>
<tr>
<td>Buffer size $</td>
<td>B</td>
</tr>
<tr>
<td>Reinitialization freq. $\gamma$</td>
<td>5%</td>
</tr>
<tr>
<td>SGLD step-size $\alpha$</td>
<td>1</td>
</tr>
<tr>
<td>SGLD noise $\sigma$</td>
<td>0</td>
</tr>
<tr>
<td>SAM noise radius $\rho$</td>
<td>0.2</td>
</tr>
</tbody>
</table>

B. Visualizing Generated Images

Table 1 in the main text reports the quantitative performance comparison of different stand-alone generative models and hybrid models. Here in Figure 1 we provide a qualitative comparison of generated images from (a) SADA-JEM, (b) VERA [3], and (c) DiffuRecov [1]. As we can see, the perceived image qualities of them are comparable even though DiffuRecov has a much better FID score than that of VERA (9.58 vs. 30.5), indicating that visualizing generated images is less effective to evaluate image quality.

C. Energy Landscapes

Figure 2 illustrates the energy landscapes of different models trained on CIFAR10. The energy landscape is generated by visualizing $E(\theta) = \sum_{x \in X} E_{\theta}(x)$ with the technique introduced in [4], where $X$ is a 10% random samples from CIFAR10 training data. As we can see, SADA-JEM’s energy landscapes are much smoother than those of the competing methods (see different scales of the y-axes).

D. Out-of-Distribution Detection

Table 2 reports the OOD detection performances of different models and SADA-JEM with different $K$s, where the input density $\log p_{\theta}(x)$ is used as $s_{\theta}(x)$ for OOD detection on CIFAR10.

E. Additional Generated Samples

Additional SADA-JEM generated class-conditional (best and worst) samples of CIFAR10 are provided in Figures 3-12.

References

Figure 1. Generated images from SADA-JEM, VERA, and DiffuRecov.

Figure 2. Energy landscapes of different models trained on CIFAR10. Please note the different scales of the y-axes.
<table>
<thead>
<tr>
<th>Dataset</th>
<th>JEM</th>
<th>JEM++ (M=10)</th>
<th>SADA-JEM (K=5)</th>
<th>SADA-JEM (K=10)</th>
<th>SADA-JEM (K=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cifar10</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
</tr>
<tr>
<td>svhn</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
</tr>
<tr>
<td>cifar100</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
</tr>
<tr>
<td>celeba</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
<td>![Histogram]</td>
</tr>
</tbody>
</table>

Table 2. Histograms of $\log p_\theta(x)$ for OOD detection. Green corresponds to in-distribution dataset, while red corresponds to OOD dataset.
Figure 3. SADA-JEM generated class-conditional samples of **Plane**.

Figure 4. SADA-JEM generated class-conditional samples of **Car**.

Figure 5. SADA-JEM generated class-conditional samples of **Bird**.

Figure 6. SADA-JEM generated class-conditional samples of **Cat**.
Figure 7. SADA-JEM generated class-conditional samples of Deer.

Figure 8. SADA-JEM generated class-conditional samples of Dog.

Figure 9. SADA-JEM generated class-conditional samples of Frog.

Figure 10. SADA-JEM generated class-conditional samples of Horse.
Figure 11. SADA-JEM generated class-conditional samples of Ship.

Figure 12. SADA-JEM generated class-conditional samples of Truck.