Reviewer #1

[Issue] Limited novelty – the proposed method is very similar to CF-VAE.

- CF-VAE sets the noise variance to a fixed value to regularize the maximum amount of contraction of the base density. However, SoftFlow uses different noise variances for every training data and sets the noise variance to a very small value or zero when sampling. Our novelty comes from a two step noise sampling procedure: (1) sample a noise variance c_i, (2) sample a noise ν_i from $\mathcal{N}(0, c_i I)$. Both the purpose and method of SoftFlow are quite different from those of CF-VAE.

[Issue] It is unclear why the performance wrt to PointFlow decreases in case of the Car class.

- We believe that the actual performance is not degraded even in the case of the Car class. First of all, we want to point out that there's still an ongoing discussion of evaluation metrics for point cloud generation and the 1-NNA is not a perfect metric. To evaluate the perceptual quality, we conducted the preference test and SoftPointFlow outperformed in the ALL classes. We provide various samples of SoftPointFlow in the appendix to support the performance improvement.

Reviewer #2

[Issue] Could this be useful for more general density estimation (e.g., MNIST/CIFAR-10 images)?

- A conventional normalizing flow (e.g., Masked Autoregressive Flow, Flow++) dequantize image by adding noise. This allows image data to have continuous values and volume components in data space. Instead, we focused on point clouds. Unlike pixels and voxels, point clouds have innately continuous values and cannot apply the same method as in image (the original point clouds cannot be retrieved directly from the noise added one). Leveraging the proposed method for more general density estimation can be studied in future work.

[Issue] Why use the uniform for sampling the standard deviation of the noise? Why not exponential, gamma or chi-squared?

- We agree and plan to study the effect of noise variance choice. We mentioned that point in the conclusion section, 'Our framework can be further improved by theoretically identifying which noise distribution is more useful for training.'.

[Issue] Why sample c from [0, 0.1] and then multiply? Why not just sample from [0, 2]?

- This is just rescaling to use a noise variance as a good input to neural networks. If we sample from [0, 2], the effect of perturbation will be very high and the perturbed data will become too blurry.

Reviewer #3

[Issue] (Sec. 4.1) Why is the number of training iterations different in Glow and CNF?

- We trained them until they converged. Since CNF is more expressive and has much lower number of parameters than Glow, CNF converges with relatively lower iterations.

[Issue] (Sec. 4.1) It is not explained what kind of distribution $p_Z(z)$ is. Is this isotropic Gaussian?

- We apply the proposed method to FFJORD. FFJORD employs isotropic Gaussian for $p_Z(z)$.

[Issue] (Sec. 4.1) Why is c_i multiplied by 20 when conditioning on the CNF network?

- Since the value of c_i is too small, we rescaled it to be [0, 2] for neural network input.

[Issue] (Sec. 5.1) Why is the range of samples from the uniform distribution very narrower than that of Sec 4.1?

- The range is determined taking into account the data distribution. Since data points of the ShapeNet is more compactly distributed than artificial data (Sec 4.1), we narrowed the uniform distribution interval.

[Issue] (Sec. 5.1) In this experiment, the authors did not evaluate the performance by estimating the likelihood. Why?

- While PointFlow estimates log-likelihood by solving ODEs, SoftPointFlow computes the explicit conditional log-likelihood in a deterministic way (i.e., a closed form solution). We think the direct likelihood comparison is not appropriate.

[Issue] (Sec. 5.1) I am not sure that point clouds in 3D space meet the challenge addressed in this paper.

- Point clouds are usually produced by 3D scanners and are scattered over the surface of an object. If not perturbed, point clouds often contain line or plane components. We believe that learning the distribution of point clouds is a good application of the proposed method.