1. Which component method do you find most compelling from all of these papers? Justify what's interesting to you.

"Analyzing and Improving the Image Quality of StyleGAN" is the most compelling method. In the paper, the authors redesign the architecture of StyleGAN, along with making other improvements. The authors note that StyleGAN systematically produces water droplet-like artifacts. Although these artifacts are not always obvious, they are present in almost every image (99.9% of images). When they are not present, the images are severely corrupted. The authors pinpoint the cause to the AdaIN operation "that normalizes the mean and variance of each feature map separately, thereby destroying any information found in the magnitudes of the features relative to each other." They hypothesize that the artifact "is a result of the generator intentionally sneaking signal strength information past instance normalization: by creating a strong, localized spike that dominates the statistics, the generator can effectively scale the signal as it likes elsewhere." The authors redesign the "style blocks" inside of the network, such as applying bias and noise outside of the style block where they operate on normalized data, rather than inside of the style block, where bias and noise will be inversely proportional to the style magnitude. Additionally, with this change, the normalization and modulation components of the style block no longer requires the mean, only requiring the standard deviation. The authors also remove the application of bias, noise, and normalization to the input.

The authors test a new architecture utilizing skip connections. They found that the new generator "behaves similar to progressive growing… without changing the network topology." Notably, the network does not "fully utilize" the highest resolution near the end of training. This was due to a capacity problem with the number of feature maps in the network. Doubling the number of feature maps caused the network to perform as expected.

The redesign reminded me of Pre-Activation for ResNets; It is a seemingly minor change to the architecture (rearranging the order of some components), but has a significant effect, nonetheless. It fixes fundamental problems with the original architecture. In this case, it removes the water droplet-like artifacts.

2. Are there methods that you think are unreasonable?

MSG-GANs can achieve suboptimal results on high-resolution images, while also taking as long, if not longer, than an alternative model. On the CelebA-HQ and FFHQ datasets, the MSG-ProGAN got a slightly higher FID score than the ProGAN. MSG-StyleGAN achieved a significantly higher FID score than StyleGAN on the CelebA-HQ and FFHQ dataset. MSG-GANs also cannot take advantage of mixing regularization, which is done to "allow the mixing of different styles at different levels at test time" and "also improves overall quality." MSG-GANs achieve stable convergence for higher resolutions, but in my opinion, converging (slightly) more stably to a less optimal result is not very impressive.

"Smoothness and Stability in GANs" also seems unreasonable. The authors discuss a pair of generator constraints, but do not discuss a realistic implementation that will satisfy those constraints. Instead, they propose a very simple particle-based generator. It serves as a concrete example to explain their propositions, but the significant work of finding a realistic generator that satisfies their constraints remains. It is an interesting starting point and can serve as a guide for designing new architectures but does not seem practical at this point.