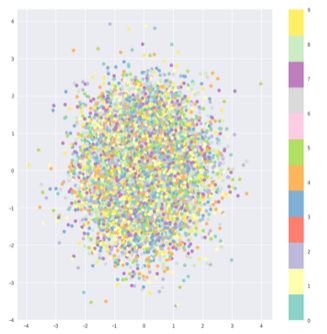
# Latent Constraints – Generating Conditionally from Unconditional Models

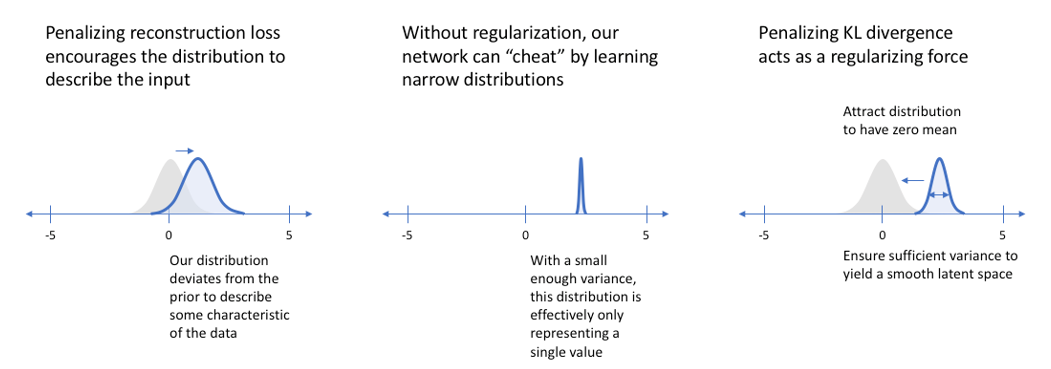
### **GAN and VAE disadvantages**

* GANs suffer from mode-collapse: The network is required to learn a certain number of different samples to generate, but only manages to generate one subset of the samples – the mode within the distribution collapses to one region



* VAEs suffer from sample-reconstruction tradeoff due to the Evidence Lower Bound: They tend to produce blurry reconstructions and samples, or sharp reconstructions but bizarre samples



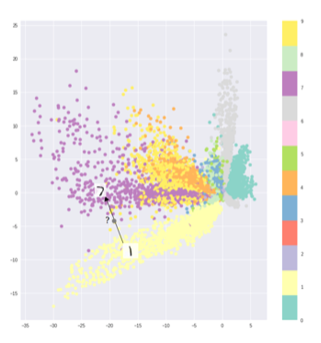


Left – blurry reconstruction, very blurry samples; Middle – mode collapse; Right – Sharper reconstruction, blurry samples

* CGAN and CVAEs – labels must be provided for entire training set; New training data with new labels requires retraining of entire model. They also suffer from the same problems as GANs and VAEs

### **Realism Constraint – VAE samples**

* A critic, **D**realism, trained to differentiate samples from p(z) and q(z) – instead of generating samples from the entire distribution p(z), samples are generated within a constraint-defined region that lie within the probability distributions of the inputs, q(z)



* Downside – if KL divergence between p(z) and q(z) is high, chances are low for producing a sample from constrained p(z) that has high probability in q(z) – blurry samples
  + Solution: Use Gradient Descent on **D**realism to move p(z) towards q(z) [**G**opt]
  + Critic loss:



* + Regularization: distance penalty introduced to encourage nearby solutions
  + Regularized loss:



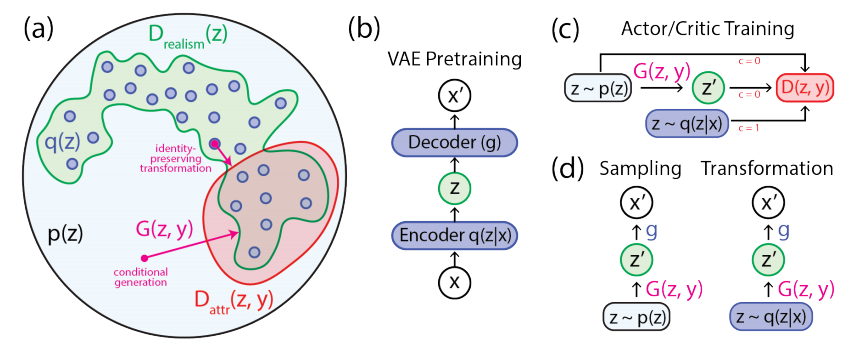
* + Alternative to speed up training: use neural network as function approximator (e.g. encoder of a VAE)

### **Attribute Constraint**

* Use a CGAN in latent space to give attribute labels y for a dataset, in order to control the attributes that generated samples produce – This thereby controls the distance penalty in the regularization, to shift p(z) to the desired q(z)
  + **D**(z)**, G**(z) **🡺 D**(z, y), **G**(z, y)
  + Without the distance penalty, generated samples are more realistic with more attributes, but are farther from the desired attributes
* VAE is used as the base generative model to train the attribute constraint critic, **D**attr(z, y)
* Using a VAE that can produce good reconstructions, **D**attr(z, y) is trained with existing samples and labels to predict attribute labels p(y|z) in z ~ q(z|x)

### **Zero-Shot Conditional Generation**

* With the model pretrained using labels, samples can be generated without labels using a reward function with the actor **G** and critic **D. D** approximates the true value of the original inputs x, and G shifts samples from the prior inputs to high-value states (desired attributes)



References:

* Main paper:
  + <https://arxiv.org/pdf/1711.05772.pdf>
  + <https://colab.research.google.com/notebooks/latent_constraints/latentconstraints.ipynb#scrollTo=WZ75KBiNiflS>
  + <https://medium.com/@jrodthoughts/deepmind-is-bringing-love-to-tensorflow-with-sonnet-531b7a32c2cb>
* Basics:
  + <https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf>
  + <https://www.jeremyjordan.me/variational-autoencoders/>
  + <https://jaan.io/what-is-variational-autoencoder-vae-tutorial/>
  + <http://kvfrans.com/variational-autoencoders-explained/>
* Implementations:
  + <https://github.com/kvfrans/variational-autoencoder/blob/master/main.py>
  + <https://github.com/shaohua0116/VAE-Tensorflow/blob/master/demo.ipynb>
* Other stuff:
  + Balancing learning and inference in VAEs: <https://arxiv.org/pdf/1706.02262.pdf>
  + <https://medium.com/@jonathan_hui/gan-why-it-is-so-hard-to-train-generative-advisory-networks-819a86b3750b>
  + <https://medium.com/@jonathan_hui/gan-cgan-infogan-using-labels-to-improve-gan-8ba4de5f9c3d>
* Further research:
  + <https://www.google.ca/search?q=echo-state+conditional+variational+autoencoder+for+anomaly+detection&rlz=1C1GCEA_enCA819CA820&oq=echo+state+conditi&aqs=chrome.1.69i57j0.5008j0j7&sourceid=chrome&ie=UTF-8>
  + <https://pdfs.semanticscholar.org/3f25/e17eb717e5894e0404ea634451332f85d287.pdf>