



WWW.DAUPHINE.PSL.EU

13/11/2024

GAN improvements by rejection sampling methods

Authors :

Mohamed Amine Belkasmi, Emmarius DELAR, Rithy Sochet

Rejection Sampling Methods

- **Intuition:**
 - Keep only good draw from the GAN's generator
- **Problem:**
 - How to determine what's a good draw ?
 - How to assure the diversity of the sample ?
 - What about the reject sampling time during generation ?
- **Approach:**
 - *Discriminator Rejection Sampling*
 - *Metropolis-Hastings GAN*
 - *Optimal Budgeted Rejection Sampling*

Rejection Sampling:

- We firstly train a VanillaGAN
- We assume our discriminator is optimal, i.e.:

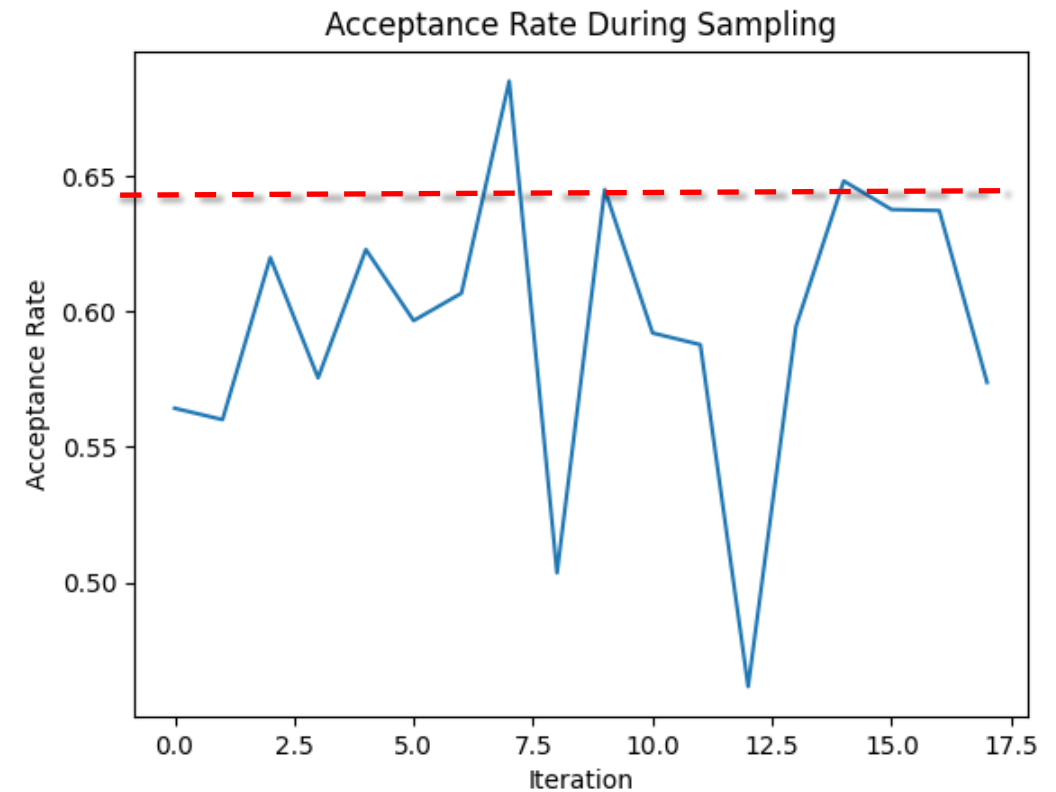
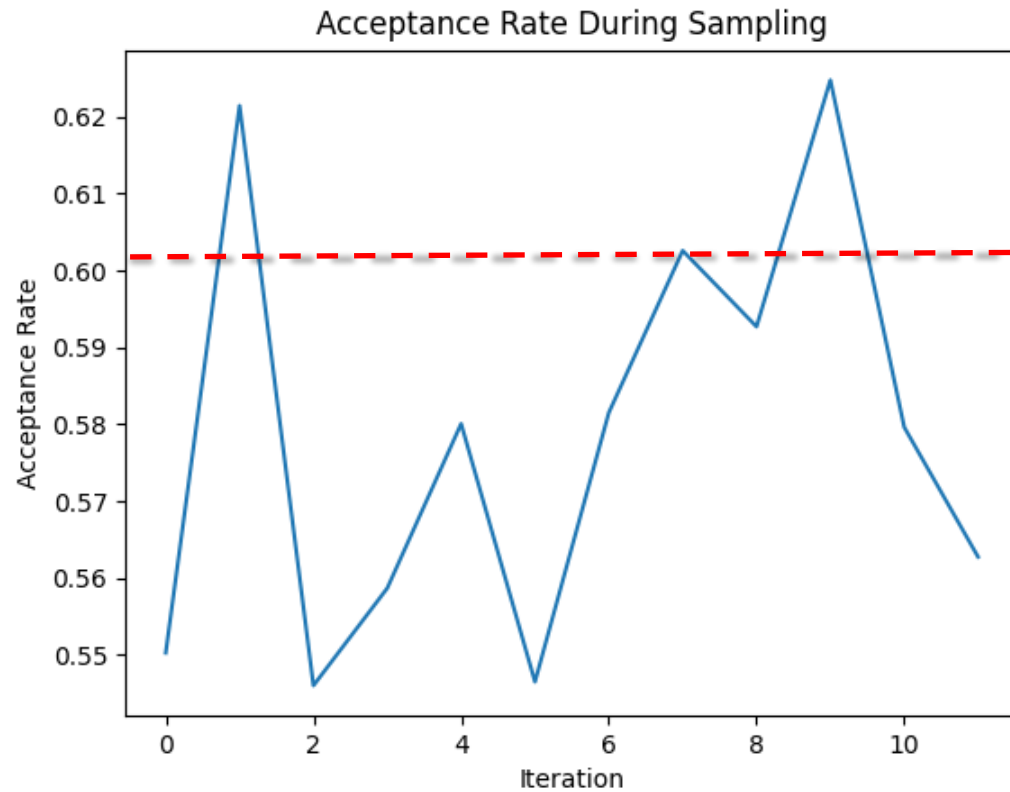
$$D(x) = \frac{p_d(x)}{p_d(x) + p_g(x)}$$

- We then obtain the acceptation rate:

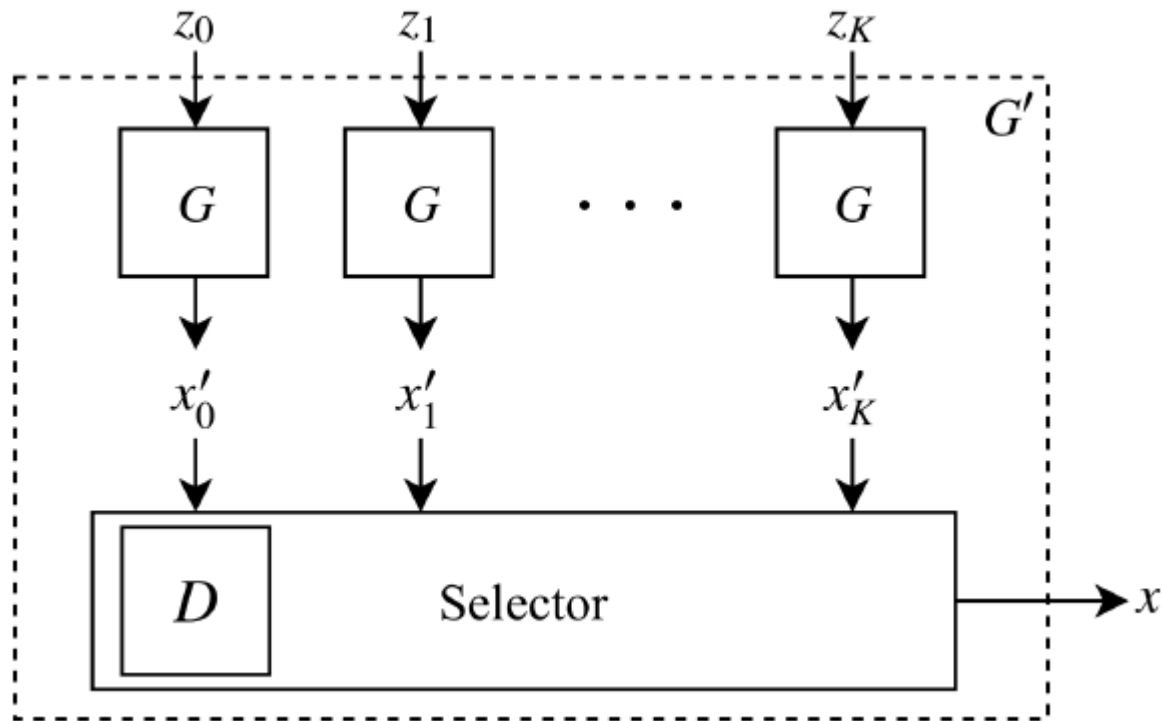
$$\frac{p_d(x)}{p_g(x)} = \frac{D(x)}{1 - D(x)}$$

Source : [1]

A figure of acceptance rate:



Metropolis-Hastings GAN



Algorithm 1 MH-GAN

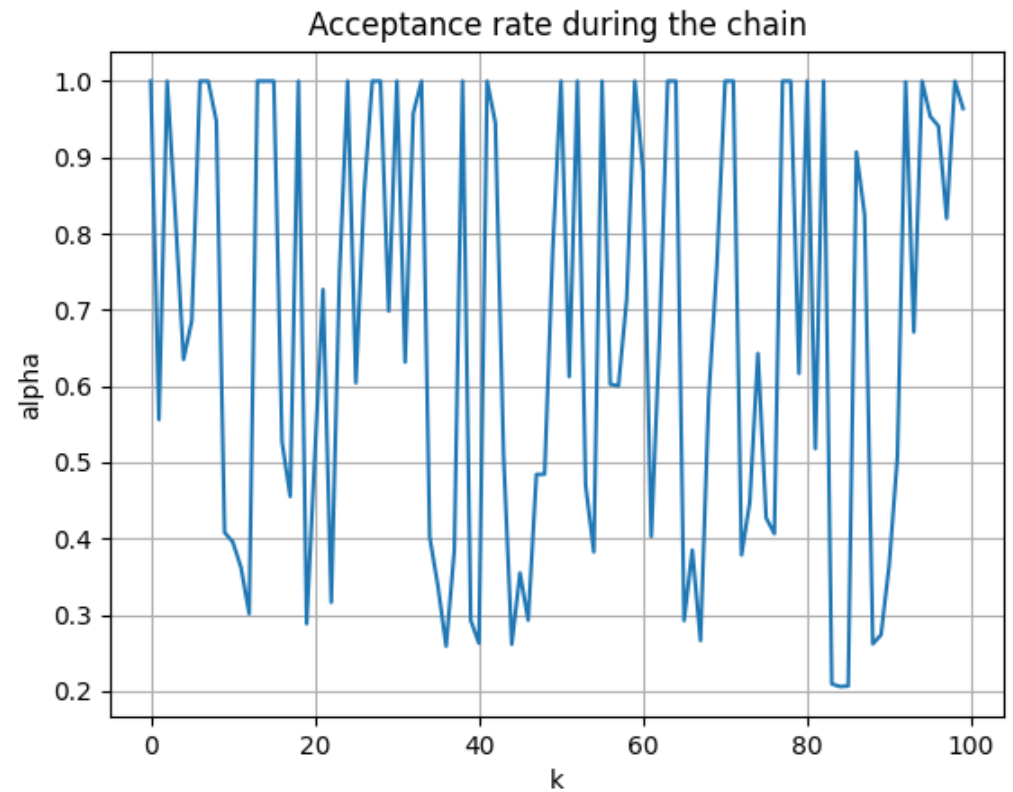
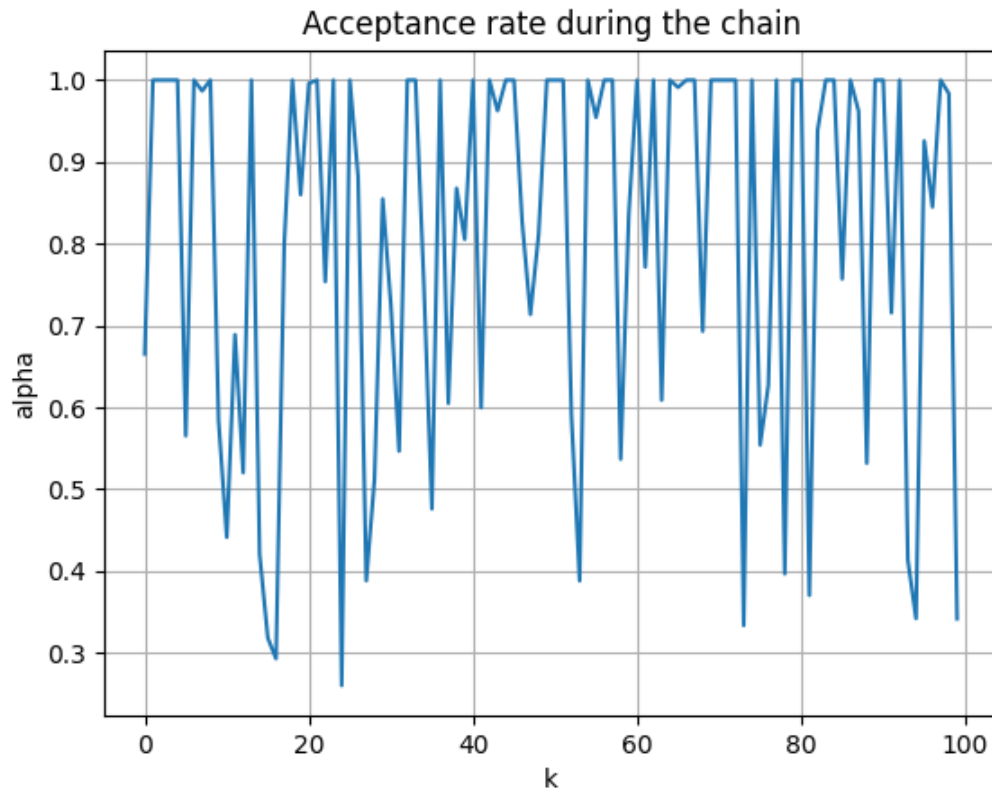
Input: generator G , calibrated disc. D , real samples
Assign random real sample x_0 to x
for $k = 1$ **to** K **do**
 Draw x' from G
 Draw U from Uniform(0, 1)
 if $U \leq (D(x)^{-1} - 1) / (D(x')^{-1} - 1)$ **then**
 $x \leftarrow x'$
 end if
end for
If x is still real sample x_0 restart with draw from G as x_0
Output: sample x from G'

Source : [2]

Metropolis-Hastings GAN

- Calibrator C, classifier where: $D(x_i) = C(\tilde{D}(x_i))$
- Acceptance rate : $\alpha(x', x_k) = \min(1, \frac{D(x_k)^{-1} - 1}{D(x')^{-1} - 1})$
- High computation time

A figure of acceptance rate:



Optimal Budgeted Rejection Sampling

- Optimal Acceptance function:

$$a_O(x) = \min\left(\frac{p(x)}{\hat{p}(x)} \frac{c_K}{M}, 1\right)$$

Algorithm B.1 Dichotomy to compute c_K .

Input: N generated samples $\mathbf{x}_1^{\text{fake}}, \dots, \mathbf{x}_N^{\text{fake}} \sim \hat{P}$

Parameter: Budget K , Threshold ϵ

Output: Constant c_K

```
1: Let  $c_{\min} = 1e^{-10}$  and  $c_{\max} = 1e^{10}$ .
2:  $c_K = (c_{\max} + c_{\min})/2$ 
3: Define the loss  $\mathcal{L}(c_K) = \sum_{i=1}^N a(\mathbf{x}_i^{\text{fake}}, c_K) - \frac{1}{K}$ 
4: while  $|\mathcal{L}(c_K)| \geq \epsilon$  do
5:   if  $\mathcal{L}(c_K) > \epsilon$  then
6:      $c_{\max} = c_K$ 
7:   else if  $\mathcal{L}(c_K) < -\epsilon$  then
8:      $c_{\min} = c_K$ 
9:   end if
10:  Update:  $c_K = (c_{\max} + c_{\min})/2$ 
11:  Update:  $\mathcal{L}(c_K)$ 
12: end while
```

Source : [3]

Optimal Budgeted Rejection Sampling

Algorithm 2 GAN Tw/OBRS

repeat

Update T by ascending the gradient of

$$\mathbb{E}_{\mathbf{x} \sim P} [T(\mathbf{x})] - \mathbb{E}_{\mathbf{x} \sim \hat{P}_G} [f^*(T(\mathbf{x}))].$$

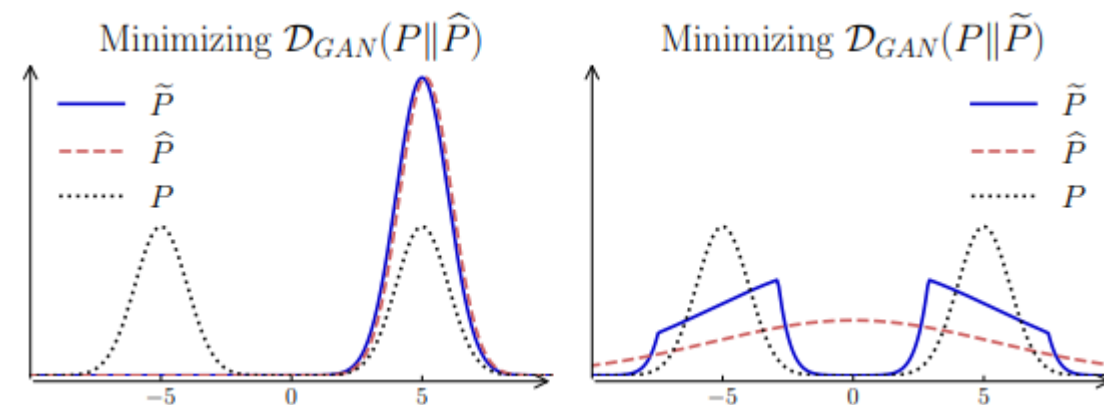
Update c_K such that $\mathbb{E}_{\hat{P}_G} [a_O(\mathbf{x})] \leq 1/K$.

(See Alg B.1 in App B.2) for details.)

Update G by descending the gradient of

$$\mathbb{E}_{\mathbf{x} \sim \hat{P}_G} \left[K a_O(\mathbf{x}) f \left(\frac{r(\mathbf{x})}{K a_O(\mathbf{x})} \right) \right].$$

until convergence.



Source : [3]

Optimal Budgeted Rejection Sampling

Algorithm 2 GAN Tw/OBRS

repeat

Update T by ascending the gradient of

$$\mathbb{E}_{\mathbf{x} \sim P} [T(\mathbf{x})] - \mathbb{E}_{\mathbf{x} \sim \tilde{P}_G} [f^*(T(\mathbf{x}))].$$

Update c_K such that $\mathbb{E}_{\tilde{P}_G} [a_O(\mathbf{x})] \leq 1/K$.

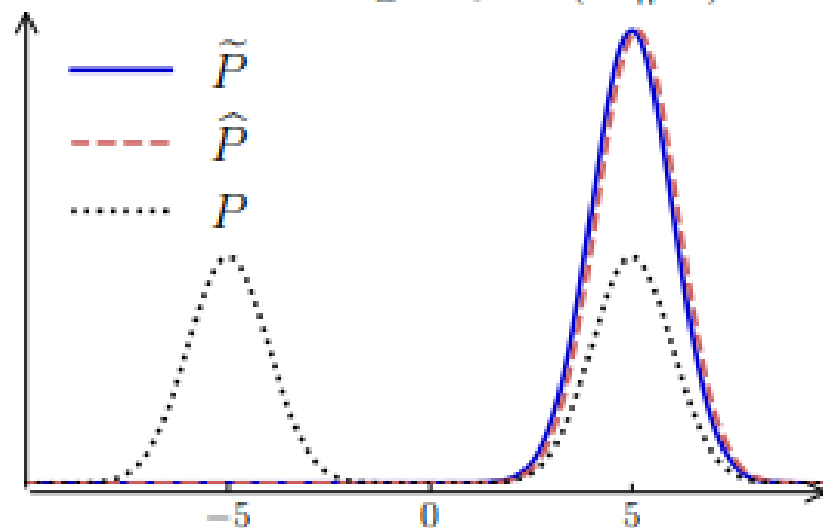
(See Alg B.1 in App B.2) for details.)

Update G by descending the gradient of

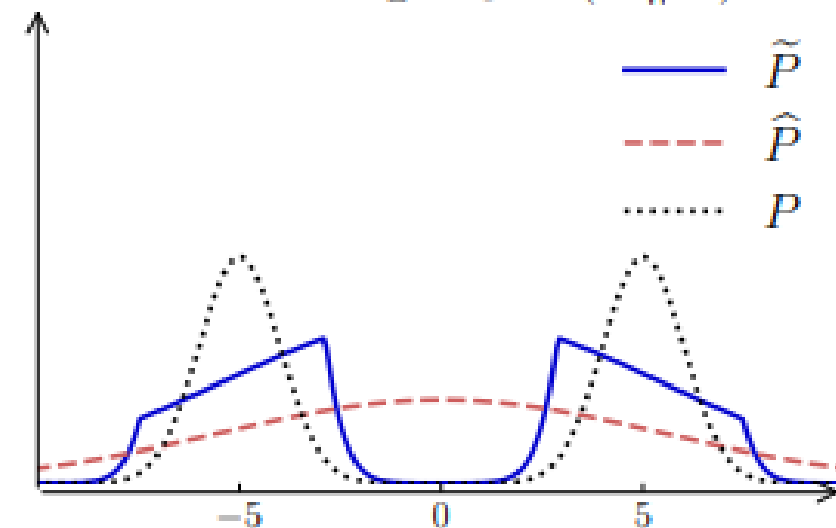
$$\mathbb{E}_{\mathbf{x} \sim \tilde{P}_G} \left[K a_O(\mathbf{x}) f \left(\frac{r(\mathbf{x})}{K a_O(\mathbf{x})} \right) \right].$$

until convergence.

Minimizing $\mathcal{D}_{GAN}(P \parallel \hat{P})$



Minimizing $\mathcal{D}_{GAN}(P \parallel \tilde{P})$



Source : [3]

EXPERIMENTAL RESULTS

Model	FID	Precision	Recall	Time (in s)
Vanilla GAN (100 epochs)	26.86	0.53	0.24	90
DRS	30.04	0.52	0.26	106.27
Vanilla GAN + MH-GAN (K=10)	25.36	0.53	0.22	1044
Vanilla GAN Tw/ OBRS	27.69	0.51	0.2	101

Table 1: Results on testing platform

Model	FID	Precision	Recall	Time (in s)
Vanilla GAN (100 epochs)	27.62	0.328	0.176	33
Vanilla GAN + MH-GAN (K=10)	25.42	0.324	0.177	942
Vanilla GAN + MH-GAN (K=100)	25.43	0.322	0.196	6201
Vanilla GAN + MH-GAN (K=100)	24.77	0.332	0.190	8243
Vanilla GAN Tw/ OBRS	27.4	0.352	0.21	33.52

Table 2: Results with our own testing

DISCUSSION

- Computing M , the maximum of the quotient
- The assumption on the optimal discriminator
- The two distributions don't necessarily have the same support
- Computation time may be long (especially for MH-GAN)

THANKS FOR LISTENING

Time for your questions



UNIVERSITÉ PARIS DAUPHINE - PSL

Place du Maréchal de Lattre de Tassigny – 75775 Paris cedex 16

References

- [1] Samaneh Azadi, Catherine Olsson, Trevor Darrell, Ian Goodfellow, and Augustus Odena. ***Discriminator rejection sampling***. arXiv preprint arXiv:1810.06758, (2018)
- [2] Ryan Turner, Jane Hung, Eric Frank, Yunus Saatchi, and Jason Yosinski. ***Metropolis-hastings generative adversarial networks***. (2019)
- [3] Alexandre Verine, Muni Sreenivas Pydi, Benjamin Negrevergne, and Yann Chevaleyre. ***Optimal budgeted rejection sampling for generative models***. (2024)