Introduction

We present an approximate inference method, based on a synergistic combination of Rényi α-divergence variational inference (RDVI) and rejection sampling (RS). RDVI is based on minimization of Rényi α-divergence $D_\alpha(p||q)$ between the true distribution $p(x)$ and a variational approximation $q(x)$; RS draws samples from a distribution $p(x) = \frac{\tilde{p}(x)}{Z_p}$ using a proposal $q(x)$, s.t. $Mq(x) \geq \tilde{p}(x), \forall x$. Our inference method is based on a crucial observation that $D_{\infty}(p||q)$ equals $\log M(\theta)$ where $M(\theta)$ is the optimal value of the RS constant for a given proposal $q_\theta(x)$. This enables us to develop a two-stage hybrid inference algorithm.

There is an increasing interest in developing more expressive variational posteriors for (shallow/deep) latent variable models and Bayesian neural networks [8, 9, 4]. In particular, the combination of MCMC and variational methods have been used in recent work to learn expressive variational posteriors [9] having the best of both worlds. Rejection Sampling [3], which we use as a subroutine (with learned $M$) in our algorithm $\alpha$-DRS, is a popular sampling technique that generates independent samples from a complex distribution indirectly through a simple distribution. In addition to being a useful sampling algorithm in its own right, recently approximations of Rejection Sampling have also been used for designing variational inference algorithms. In particular, Variational Rejection Sampling (VRS) [6], which uses rejection sampling to learn a better variational approximation. Recently Rejection sampling has also been used to improve the generated samples from GAN (Generative Adversarial Nets) [1] and improve priors for variational inference [2].

Connecting Rejection Sampling with Rényi α-Divergence

We now show how Rényi α-divergence is related to rejection sampling, and how this connection can be leveraged to finetune the $q_\theta$ estimated by RDVI using $q_\theta$ as a proposal distribution of a rejection sampler, and generating a sample-based approximation of the exact distribution. The connection between Rényi α-divergence and rejection sampling is made explicit by the following result

**Theorem 1.** When $\alpha \to \infty$, the Rényi α divergence becomes equal to the worst-case regret [10, Theorem 6].

$$\lim_{\alpha \to \infty} D_\alpha(p||q_\theta) = \log \max_{x \in X} \frac{p(x)}{q_\theta(x)}$$

It is interesting to note that $\lim_{\alpha \to \infty} D_\alpha(p||q_\theta)$ in Eq. (1) is equal to the log of the optimal $M(\theta)$ value used in Rejection Sampling. It is easy to show that $q_\theta(x) \left( \max_{x \in X} \frac{p(x)}{q_\theta(x)} \right) \geq p(x), \forall x \in \text{supp}(p(x))$. 
In Rényi $\alpha$-divergence variational inference [7], we learn the variational parameters $\theta$ such that the value of $\alpha$ divergence is minimized. Therefore, minimizing Rényi $\alpha$ divergence of $\infty$ order can serve the following purposes:

- We can learn the optimal variational distribution $q_\theta(x)$.
- We can learn the optimal value $M(\hat{\theta})$ (expected number of iterations needed to generate one sample) such that rejection sampling could be performed with fewer rejections.
- The above rejection sampler can be used to “refine” $q_\theta$ using a sample-based approximation.

Although the above idea seems like an appealing prospect, optimizing Rényi $\alpha$ divergence of $\infty$ order is problematic. Instead of using Rejection Sampling for $\infty$ order $\alpha$-divergence, we will develop an approximate version of Rejection sampling for finite order $\alpha$-divergence.

### 2.1 $\alpha$-Divergence Rejection Sampling

In this section, we summarize our algorithm $\alpha$-Divergence Rejection Sampling ($\alpha$-DRS) which augments the $\alpha$ divergence [7] method. The algorithm requires an input $\alpha$, the target distribution $p(x) = \tilde{p}(x)/Z_p$, and the variational distribution $q_\theta(x)$. Our algorithm $\alpha$-DRS consists of two stages.

- In stage-1, given an input $\alpha$, we minimize the Monte-Carlo estimate of the exponentiated version of finite order $\alpha$-divergence [5] with respect to the variational parameters $\theta$, i.e.,

\[
\hat{\theta} = \arg \min_\theta \frac{1}{S} \sum_{s=1}^S \left( \frac{\tilde{p}(x_s)}{q_\theta(x_s)} \right)^\alpha,
\]

where $x_s$ are iid samples drawn from $q_\theta(x)$.

- From stage-1, we learned the optimal $\theta$. For the second stage we will learn $T$ from equation [5] and perform approximate Rejection Sampling [9] to learn a refined distribution $r_\hat{\theta}(x)$.

The acceptance probability for approximate RS is as follows:

\[
a_\theta(x|T) = 1/ \left[ 1 + \left( \frac{q_\theta(x)e^{-T}}{\tilde{p}(x)} \right) \right],
\]

where $T$ is a hyperparameter controlling the acceptance rate.

**Theorem 2.** For a fixed $\theta$, the approximate Rejection sampling always improves the Rényi $\alpha$ divergence between the estimated and actual posterior. The acceptance probability is approximated by equation [9]. The proof of the theorem can be found in the supplementary material.

\[
D_\alpha(p||r) \leq D_\alpha(p||q)
\]

### 2.2 Choosing the hyperparameter $T$

Although $D_\alpha(p||q)$ is a lower bound on $\log M(\hat{\theta})$ (property of $\alpha$-divergence), for high dimensions even this may be too large. The hyperparameter $T$ should be defined such that we can control the acceptance rate. Let’s define $L_\theta(x) = - \log \tilde{p}(x) + \log q_\theta(x)$ where $x \sim q_\theta(x)$, and redefine $T$ as

\[
T = \begin{cases} 
-D_\alpha(p||q) & \text{For low dimensions} \\
Q L_\theta(x)(\gamma) & \text{For high dimensions} 
\end{cases}
\]

where $Q$ is quantile function defined over the random variable $L_\theta(x)$ with hyperparameter $\gamma \in [0, 1]$. The quantile function $Q$ approach [6] allows us to select samples that have high-density ratios (similar to Rejection sampling) along with a well-defined acceptance rate (around $\gamma$ for most samples). Note that a similar methodology has been recently employed in Variational Rejection Sampling (VRS) [6] as well.

### 3 Experiments

In this section, we evaluate our proposed $\alpha$-DRS algorithm on synthetic as well as real-world datasets. In particular, we are interested in assessing the performance of $\alpha$-DRS as a method that can improve the variational approximation learned by RDVI.
3.1 Gaussian Mixture Model Toy Example

In this experiment, we have chosen \( p(x) \) to be a mixture of four Gaussian distributions:

\[
p(x) = \frac{1}{4} \mathcal{N}(-12, 0.64) + \frac{1}{4} \mathcal{N}(-6, 0.64) + \frac{1}{4} \mathcal{N}(0, 0.64) + \frac{1}{4} \mathcal{N}(6, 0.64)
\]

The variational distribution \( q_\theta(x) \) is assumed to be a \( t \)-distribution with 10 degrees of freedom and parameters \( \mu \) and \( \log \sigma^2 \). We have generated 3000 samples from \( t \)-distribution to approximate \( D_\alpha(p||q) \). The hyperparameter \( \alpha \) was learned using Eq. (5). Table 1 compares the \( \alpha \)-divergence with RS step (\( D_\alpha(p||q) \)) and without RS step (\( D_\alpha(p||q) \)).

In this case, as evident from Fig. 1, with the RS step, we are able to get a very good approximation of the target density \( p(x) \) despite it having multiple modes. Table 1 compares the \( \alpha \)-divergence with RS step (\( D_\alpha(p||q) \)) and without RS step (\( D_\alpha(p||q) \)).

3.2 Bayesian Neural Network

In this section, we will perform approximate inference for Bayesian Neural Network regression. The datasets are collected from the UCI data repository. We have used a single layer NN with 50 hidden units and ReLU activation to model the regression task. Let’s denote the neural network weights by \( \delta \) having a Gaussian prior \( \delta \sim \mathcal{N}(\delta; 0, I) \). The true posterior distribution of NN weights (\( \delta \)) is approximated by a fully factorized Gaussian distribution \( q(\delta) \).

All the datasets are randomly partitioned 20 times into 90% training and 10% test data. The stochastic gradients are approximated by 100 samples from \( q(\delta) \) and a minibatch of size 32 from the training set. We summarize the average RMSE and test log-likelihood in Table 1. For \( \alpha \)-DRS method we have chosen acceptance rate to be around 10% (\( \gamma = 0.1 \) in equation (5)). We have compared the results of \( \alpha \)-DRS method with RDVI and adaptive f-divergence \( \alpha \)-DRS (\( \beta = -1 \)).

![Figure 1: Black Plot: Empirical p.d.f. of the generated samples from \( \alpha \)-DRS algorithm, Red plot: \( p(x) \), Blue plot: learned \( t \)-distribution by RDVI](image)

| dataset     | \( \alpha = 1.0 \) | \( \alpha = 2.0 \) | \( \alpha = 1.0 \) | \( \alpha = 2.0 \) |
|-------------|------------------|------------------|------------------|------------------|
| Boston      | 2.881±0.177      | 2.991±0.198      | 3.099±0.196      | 3.099±0.196      |
| Concrete    | 5.343±0.116      | 5.425±0.121      | 5.424±0.105      | 5.424±0.105      |
| Kin8nm      | 0.085±0.001      | 0.084±0.001      | 0.083±0.001      | 0.083±0.001      |
| Yacht       | 0.810±0.064      | 1.193±0.082      | 1.192±0.089      | 1.192±0.089      |

Table 1: Test RMSE and Test LL

4 Conclusion

We have presented a two-stage approximate inference method to generate samples from a target distribution. Our approach, essentially a hybrid of Rényi divergence variational inference [2] and rejection sampling, leverages a new connection between Rényi \( \alpha \)-divergences and the parameter \( M \) controlling the acceptance probabilities of the rejection sampler. Therefore our method can be seen as a rejection sampling-based algorithm that can finetune the variational approximation produced by RDVI into a more expressive sample-based estimate. Our experimental results demonstrate the clear benefits of these improvements in the context of improving variational approximations via rejection sampling.
References
[1] Azadi, S., C. Olsson, T. Darrell, I. Goodfellow, and A. Odena (2018). Discriminator rejection sampling. *arXiv preprint arXiv:1810.06758*.

[2] Bauer, M. and A. Mnih (2018). Resampled priors for variational autoencoders. *arXiv preprint arXiv:1810.11428*.

[3] Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

[4] Chen, X., D. P. Kingma, T. Salimans, Y. Duan, P. Dhariwal, J. Schulman, I. Sutskever, and P. Abbeel (2016). Variational lossy autoencoder. *arXiv preprint arXiv:1611.02731*.

[5] Dieng, A. B., D. Tran, R. Ranganath, J. Paisley, and D. Blei (2017). Variational inference via χ upper bound minimization. In *Advances in Neural Information Processing Systems*, pp. 2732–2741.

[6] Grover, A., R. Gummadi, M. Lazaro-Gredilla, D. Schuurmans, and S. Ermon (2018). Variational rejection sampling. *arXiv preprint arXiv:1804.01712*.

[7] Li, Y. and R. E. Turner (2016). Rényi divergence variational inference. In *Advances in Neural Information Processing Systems*, pp. 1073–1081.

[8] Rezende, D. J. and S. Mohamed (2015). Variational inference with normalizing flows. *arXiv preprint arXiv:1505.05770*.

[9] Salimans, T., D. Kingma, and M. Welling (2015). Markov chain monte carlo and variational inference: Bridging the gap. In *International Conference on Machine Learning*, pp. 1218–1226.

[10] Van Erven, T. and P. Harremos (2014). Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory* 60(7), 3797–3820.

[11] Wang, D., H. Liu, and Q. Liu (2018). Variational inference with tail-adaptive f-divergence. In *Advances in Neural Information Processing Systems*, pp. 5737–5747.
5 Supplementary Material

In this section, we will show that the approximate Rejection sampling step can further reduce the \(\alpha\)-divergence between an exact distribution and approximate posterior distribution.

Notations:

- True distribution \(p(x) = \frac{\tilde{p}(x)}{Z_p}\), where \(Z_p\) is the normalization constant.
- Let’s denote the learned distribution from \(\alpha\)-DRS by \(r_\theta(x)\). We can write this learned distribution as follows:

\[
    r(x) = \frac{q_\theta(x)a_\theta(x)T}{Z_R(x, T)},
\]

where \(Z_R(x, T)\) is a normalization constant. For the sake of clarity we will denote \(r(x) = \frac{\tilde{r}(x)}{Z_R}\), where \(Z_R\) is a normalization constant.

We are making the following assumptions:

- The acceptance probability for every sample can be denoted by \(a_\theta(x|T)\), where \(T = -\log M\), \(M\) is the constant used for approximate rejection sampling. \(T\) can be learned through equation \(5\).

\[
    a_\theta(x|T) = \min \left[ 1, \frac{\tilde{p}(x)}{e^{-T}q_\theta(x)} \right] \equiv \frac{1}{\left[ 1 + \left( \frac{e^{-T}q_\theta(x)}{\tilde{p}(x)} \right)^t \right]^{1/t}}
\]

- Take \(t=1\) for getting a differentiable approximation of the acceptance probability.

**Theorem 2:** For a fixed \(\theta\), the approximate Rejection sampling always improves the Rényi \(\alpha\) divergence between the estimated and actual posterior for \(\alpha \in (0, \infty)\). The following equation approximates the acceptance probability.

\[
    a_\tilde{\theta}(x|T) = 1 / \left[ 1 + \left( \frac{q_\theta(x)e^{-T}}{\tilde{p}(x)} \right) \right],
\]

\[
    D_\alpha(p||r) \leq D_\alpha(p||q)
\]

- \(T \to \infty\) implies \(r_\theta(x) \to q_\theta(x)\)
- \(T \to -\infty\) implies \(r_\theta(x) \to p(x)\)

**Proof:** We are using the above notations.

\[
    D_\alpha(p||R) = \frac{1}{\alpha - 1} \log \left[ \int (\tilde{p}(x)/r(x))^\alpha r(x)dx \right] - \frac{\alpha}{\alpha - 1} \log Z_p
\]

\[
    = \frac{1}{\alpha - 1} \left( \alpha \log Z_R + \log \left[ \int (\tilde{p}(x)/\tilde{r}(x))^\alpha r(x)dx \right] \right) - \frac{\alpha}{\alpha - 1} \log Z_p
\]

\[
    = \frac{\alpha}{\alpha - 1} \log Z_R + \frac{1}{\alpha - 1} \log \left[ \int (\tilde{p}(x)/\tilde{r}(x))^\alpha r(x)dx \right] - \frac{\alpha}{\alpha - 1} \log Z_p
\]

Now we will take the derivative of \(D_\alpha(p||R)\) with respect to \(T\) such that variable \(T = -\log M\).

\[
    \nabla_T D_\alpha(p||R) = \frac{\alpha}{\alpha - 1} \nabla_T \log Z_R + \frac{1}{\alpha - 1} \nabla_T \log \left[ \int (\tilde{p}(x)/\tilde{r}(x))^\alpha r(x)dx \right]
\]

\[
    = \frac{\alpha}{\alpha - 1} \nabla_T \log Z_R + \frac{1}{\alpha - 1} \nabla_T \left[ \int (\tilde{p}(x)/\tilde{r}(x))^\alpha r(x)dx \right]
\]
We will take the derivative of numerator separately now for more clarity. Let’s denote the numerator by $D_1$. Note that the $Z_R$ term would be canceled out.

$$D_1 = \nabla_T \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha r(x) dx$$  \hspace{1cm} (16)

$$= -\alpha \int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha \nabla_T \log \tilde{r}(x) r(x) dx + \int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha \nabla_T \log r(x) r(x) dx$$  \hspace{1cm} (17)

$$= -\alpha \nabla_T \log Z_R \int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha r(x) dx + (1 - \alpha) \int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha \nabla_T \log r(x) r(x) dx$$  \hspace{1cm} (18)

By substituting the above result, we will finally get the following equation.

$$\nabla_T D_\alpha (P || R) = -\frac{\int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha \nabla_T \log \tilde{r}(x) r(x) dx}{\int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha r(x) dx}$$  \hspace{1cm} (19)

Since we know that $E_R[\nabla_T \log r(x)] = 0$ we can directly change the numerator above into a covariance function. Also we know that covariance function is unaffected by adding a constant, hence we will add $\nabla \log Z_R$ to $\nabla_T \log r(x)$ in order to convert it into $\nabla_T \log \tilde{r}(x)$. The final derivative would come out to be:

$$\nabla_T D_\alpha (P || R) = -\frac{\text{COV}_R \left[ \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha, \nabla_T \log \tilde{r}(x) \right]}{\int \left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha r(x) dx}$$  \hspace{1cm} (20)

$$\geq 0$$  \hspace{1cm} (22)

Note that in above equation we are taking covariance of a random variable $\left( \frac{\tilde{p}(x)}{\tilde{r}(x)} \right)^\alpha$ with its monotonic transformation $-\left( e^{-T \frac{\tilde{r}(x)}{\tilde{p}(x)}} \right)$, $\alpha > 0$ which is always positive. Hence, we can conclude that for any general $T$, $D_\alpha (P || R) \leq D_\alpha (P || Q)$. 

6