A PROOFS

Lemma 1. *Let X be a non-negative random variable and Φ be a continuous function on $[0, \infty)$. If Φ' is integrable on all closed intervals in $[0, \infty)$,*

$$\mathbb{E}\left[\Phi(X)\right] = \Phi(0) + \int_0^{\infty} \Phi'(t) \Pr(X \geq t) dt$$

Proof.

$$\begin{aligned}
\Phi(0) + \int_0^{\infty} \Phi'(t) \Pr(X \geq t) dt &= \Phi(0) + \int_0^{\infty} \int_t^{\infty} \Phi'(t) p(x) dx dt \\
&= \Phi(0) + \int_{\{x \geq t, t \geq 0\}} \Phi'(t) p(x) d \begin{pmatrix} x \\ t \end{pmatrix} \\
&= \Phi(0) + \int_{\{t \leq x, t \geq 0\}} \Phi'(t) p(x) d \begin{pmatrix} x \\ t \end{pmatrix} \\
&= \Phi(0) + \int_0^{\infty} \int_0^{x} \Phi'(t) p(x) dt dx \\
&= \Phi(0) + \int_0^{\infty} \left(\int_0^{x} \Phi'(t) dt \right) p(x) dx \\
&= \Phi(0) + \int_0^{\infty} \left(\Phi(x) - \Phi(0) \right) p(x) dx \quad \text{(2nd FTC)} \\
&= \Phi(0) + \int_0^{\infty} \Phi(x) p(x) dx - \int_0^{\infty} \Phi(0) p(x) dx \\
&= \Phi(0) + \int_0^{\infty} \Phi(x) p(x) dx - \Phi(0) \int_0^{\infty} p(x) dx \\
&= \Phi(0) + \mathbb{E}\left[\Phi(X)\right] - \Phi(0) \\
&= \mathbb{E}\left[\Phi(X)\right]
\end{aligned}$$

□

Lemma 2. *Under the choice of $\Phi_\tau(\cdot)$ above and its associated $\Phi'_\tau(\cdot)$,*

$$\mathbb{E}_{z_1,\ldots,z_m \sim \mathcal{N}(0,I)} \left[\Phi_{\tau_i}(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)))\right] = \tau_i - \int_{\delta\tau_i}^{\tau_i} \Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) < t) dt.$$

Proof. By definition, $\Phi_{\tau_i}(0) = \delta\tau_i$.

$$\begin{aligned}
\mathbb{E}_{z_1,\ldots,z_m \sim \mathcal{N}(0,I)} \left[\Phi_{\tau_i}(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)))\right] &= \Phi_{\tau_i}(0) + \int_0^{\infty} \Phi'_{\tau_i}(t) \Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) \geq t) dt \quad \text{(Lemma 1)} \\
&= \delta\tau_i + \int_{\delta\tau_i}^{\tau_i} \Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) \geq t) dt \\
&= \delta\tau_i + \int_{\delta\tau_i}^{\tau_i} \left(1 - \Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) < t)\right) dt \\
&= \delta\tau_i + (\tau_i - \delta\tau_i) - \int_{\delta\tau_i}^{\tau_i} \Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) < t) dt \\
&= \tau_i - \int_{\delta\tau_i}^{\tau_i} \Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) < t) dt
\end{aligned}$$

□

Lemma 3. *Under the choice of $\Phi_\tau(\cdot)$ above and its associated $\Phi'_\tau(\cdot)$,*

$$\mathcal{L}_{\{\tau_i\}_i}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{mw_i}\sum_{j=1}^{m}\int_{\delta\tau_i}^{\tau_i}\Pr(d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)) < t)dt.$$

Proof.

$$\begin{aligned}
\mathcal{L}_{\{\tau_i\}_i}(\theta) &= \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \frac{1}{m}\sum_{j=1}^{m}\Phi_{\tau_i}(d(\tilde{\mathbf{x}}_j, \mathbf{x}_i))\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \frac{1}{m}\sum_{j=1}^{m}\mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\Phi_{\tau_i}(d(\tilde{\mathbf{x}}_j, \mathbf{x}_i))\right]\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \frac{1}{m}\sum_{j=1}^{m}\left(\tau_i - \int_{\delta\tau_i}^{\tau_i}\Pr(d(\tilde{\mathbf{x}}_j, \mathbf{x}_i) < t)dt\right)\right) \qquad \text{(Lemma 2)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{mw_i}\sum_{j=1}^{m}\int_{\delta\tau_i}^{\tau_i}\Pr(d(\tilde{\mathbf{x}}_j, \mathbf{x}_i) < t)dt
\end{aligned}$$

□

Equation 3.

Proof.

$$\begin{aligned}
\mathcal{L}_{\{\tau_i\}_i}(\theta) &= \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \frac{m-1}{m}\tau_i - \frac{1}{m}\max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right)\right] \\
&= \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{w_i}\left(\frac{1}{m}\tau_i - \frac{1}{m}\max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right)\right] \\
&= \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\frac{1}{nm}\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right)\right]
\end{aligned}$$

□

Equation 4.

Proof.

$$\begin{aligned}
\arg\max_\theta \mathcal{L}_{\{\tau_i\}_i}(\theta) &= \arg\max_\theta \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\frac{1}{nm}\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right)\right] \\
&= \arg\max_\theta \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\sum_{i=1}^{n}\frac{1}{w_i}\left(\tau_i - \max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right)\right] \\
&= \arg\max_\theta \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\sum_{i=1}^{n}\frac{\tau_i}{w_i} - \sum_{i=1}^{n}\frac{1}{w_i}\max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right] \\
&= \arg\max_\theta \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[-\sum_{i=1}^{n}\frac{1}{w_i}\max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right] \\
&= \arg\min_\theta \mathbb{E}_{z_1,\ldots,z_m\sim\mathcal{N}(0,I)}\left[\sum_{i=1}^{n}\frac{1}{w_i}\max(\min_{j\in[m]} d(\mathbf{x}_i, T_\theta(\mathbf{z}_j)), \delta\tau_i)\right]
\end{aligned}$$

□

Lemma 4. *Under the choice of* $w_i = \int_{\delta\tau_i}^{\tau_i} \mathrm{vol}(B_t(\mathbf{x}_i))dt := \int_{\delta\tau_i}^{\tau_i} \int_{B_t(\mathbf{x}_i)} d\mathbf{x}dt$*, where* $B_r(\mathbf{x}) = \{\mathbf{y}|d(\mathbf{y},\mathbf{x}) < r\}$ *is an open ball of radius* r *centred at* $\mathbf{x}$*,*

$$\lim_{\{\tau_i \to 0^+\}_i} \mathcal{L}_{\{\tau_i\}_i}(\theta) = \frac{1}{n}\sum_{i=1}^{n} p_\theta(\mathbf{x}_i)$$

Proof.

$$\begin{aligned}
\mathcal{L}_{\{\tau_i\}_i}(\theta) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{mw_i}\sum_{j=1}^{m}\int_{\delta\tau_i}^{\tau_i} \Pr(d(\tilde{\mathbf{x}}_j, \mathbf{x}_i) < t)dt \quad \text{(Lemma 3)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{mw_i}\sum_{j=1}^{m}\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x}dt \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{w_i}\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x}dt \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x}dt}{\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} d\mathbf{x}dt}
\end{aligned}$$

$$\begin{aligned}
\lim_{\{\tau_i \to 0^+\}_i} \mathcal{L}_{\{\tau_i\}_i}(\theta) &= \frac{1}{nm}\sum_{i=1}^{n}\left(\lim_{\tau_i\to 0^+}\left(\sum_{j=1}^{m}\frac{\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x}dt}{\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} d\mathbf{x}dt}\right)\right) \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\lim_{\tau_i\to 0^+}\frac{\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x}dt}{\int_{\delta\tau_i}^{\tau_i}\int_{B_t(\mathbf{x}_i)} d\mathbf{x}dt}\right) \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\lim_{\tau_i\to 0^+}\frac{\int_{B_{\tau_i}(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x} - \delta\int_{B_{\delta\tau_i}(\mathbf{x}_i)} p_\theta(\mathbf{x})d\mathbf{x}}{\int_{B_{\tau_i}(\mathbf{x}_i)} d\mathbf{x} - \delta\int_{B_{\delta\tau_i}(\mathbf{x}_i)} d\mathbf{x}}\right) \quad \text{(L'Hôpital and 2nd FTC)} \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\lim_{\tau_i\to 0^+}\frac{\int_{B_{\tau_i}(\mathbf{x}_i)} p_\theta(\mathbf{x})(1-\delta\mathbf{1}_{B_{\delta\tau_i}(\mathbf{x}_i)}(\mathbf{x}))d\mathbf{x}}{\int_{B_{\tau_i}(\mathbf{x}_i)} 1-\delta\mathbf{1}_{B_{\delta\tau_i}(\mathbf{x}_i)}(\mathbf{x})d\mathbf{x}}\right) \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\lim_{\tau_i\to 0^+}\frac{\int_0^{\tau_i}(1-\delta\mathbf{1}_{\{r<\delta\tau_i\}}(r))\int_{\{\mathbf{x}|d(\mathbf{x},\mathbf{x}_i)=r\}} p_\theta(\mathbf{x})d\mathbf{x}dr}{\int_0^{\tau_i}(1-\delta\mathbf{1}_{\{r<\delta\tau_i\}}(r))\int_{\{\mathbf{x}|d(\mathbf{x},\mathbf{x}_i)=r\}} d\mathbf{x}dr}\right) \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\lim_{\tau_i\to 0^+}\frac{\int_{\{\mathbf{x}|d(\mathbf{x},\mathbf{x}_i)=\tau_i\}} p_\theta(\mathbf{x})d\mathbf{x}}{\int_{\{\mathbf{x}|d(\mathbf{x},\mathbf{x}_i)=\tau_i\}} d\mathbf{x}}\right) \quad \text{(L'Hôpital and 2nd FTC)} \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} p_\theta(\mathbf{x}_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} p_\theta(\mathbf{x}_i)
\end{aligned}$$

□

Note that under common metrics like ℓ_p distances, w_i can be found in closed form, i.e., $\mathrm{vol}(B_t(\mathbf{x}_i)) = (2t)^d\frac{\Gamma(1+1/p)^d}{\Gamma(1+d/p)}$, and so $w_i = \int_{\delta\tau_i}^{\tau_i}\mathrm{vol}(B_t(\mathbf{x}_i))dt = \int_{\delta\tau_i}^{\tau_i}(2t)^d\frac{\Gamma(1+1/p)^d}{\Gamma(1+d/p)}dt = \frac{(2(1-\delta)\tau_i)^{d+1}}{2(d+1)} \cdot \frac{\Gamma(1+1/p)^d}{\Gamma(1+d/p)}$, where $\Gamma(\cdot)$ denotes the gamma function.

B Pseudo Code for IMLE

Algorithm 2 Implicit maximum likelihood estimation (IMLE) procedure

Require: The set of inputs $\{\mathbf{x}_i\}_{i=1}^{n}$

Initialize the parameters θ of the generator T_θ
for $k = 1$ **to** K **do**
  Pick a random batch $S \subseteq [n]$
  Draw latent codes $Z \leftarrow \mathbf{z}_1, ..., \mathbf{z}_m$ from $\mathcal{N}(0, \mathbf{I})$
  $\sigma(i) \leftarrow \arg\min_{j \in [m} d(\mathbf{x}_i, T_\theta(\mathbf{z_j})) \; \forall i \in S$
  for $l = 1$ **to** L **do**
    Pick a random mini batch $\tilde{S} \subseteq S$
    $\theta \leftarrow \theta - \eta \nabla_\theta \left(\sum_{i \in \tilde{S}} d\left(\mathbf{x}_i, T_\theta\left(\mathbf{z}_{\sigma(\mathbf{i})}\right)\right)\right) / |\widetilde{S}|$
  end for
end for
return θ

C Additional Results

We show more interpolation results for Adaptive IMLE on FFHQ subset and Obama in Fig. 5. We also show randomly generated samples for FFHQ subset and Obama in Fig. 7 and 6.

Figure 5: Interpolation results for Adaptive IMLE (Ours). Each row shows a different interpolation.

Figure 6: Adaptive IMLE (Ours): *FFHQ subset* randomly generated samples.

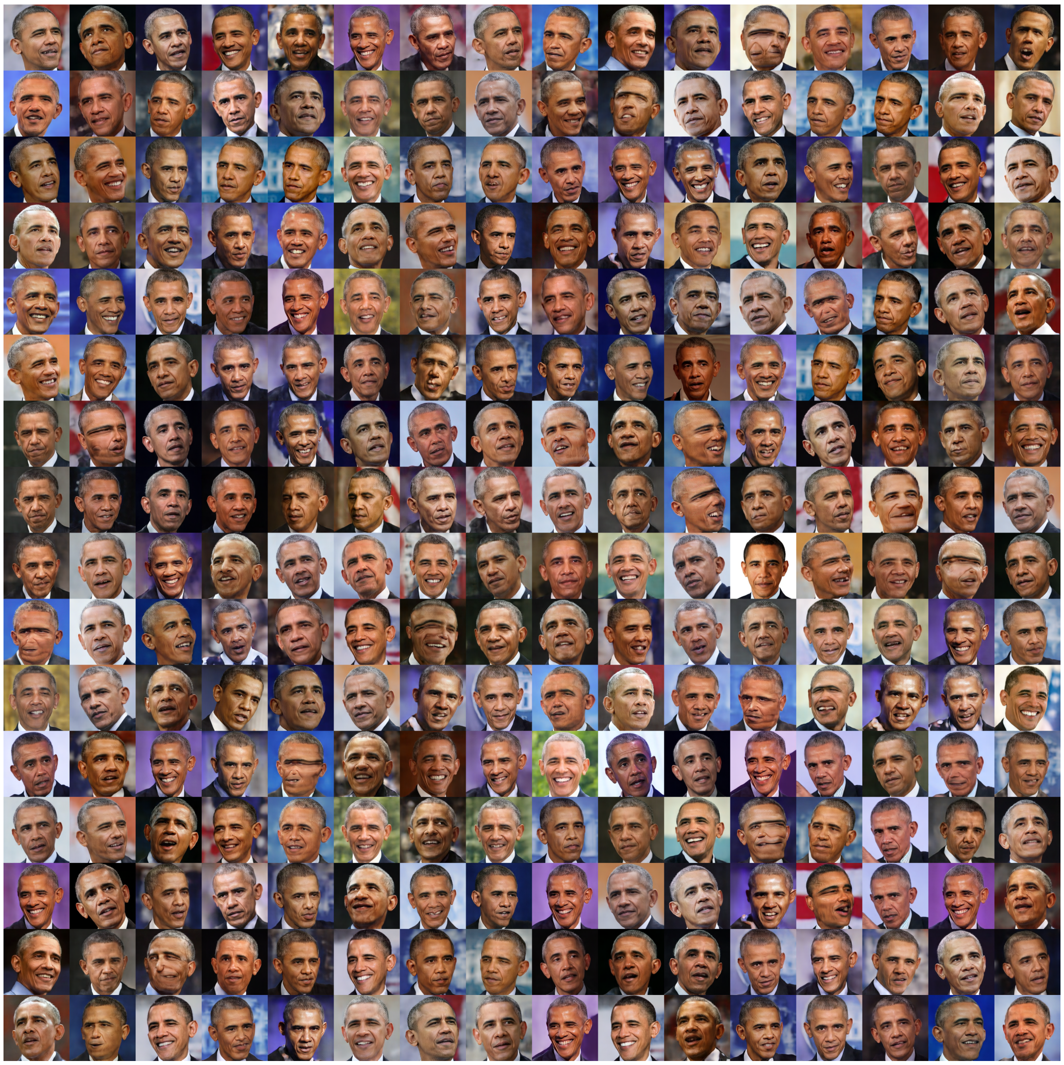

Figure 7: Adaptive IMLE (Ours): *Obama* randomly generated samples.