Supplementary Material for Gaussian-Smoothed Optimal Transport: Metric Structure and Statistical Efficiency

A Non-Uniform Results

Figure 2 shows results for a non-uniform μ, specifically for μ an isotropic $d = 100$ Gaussian. Note that the behavior is qualitatively the same as the results for uniform μ in the main text.

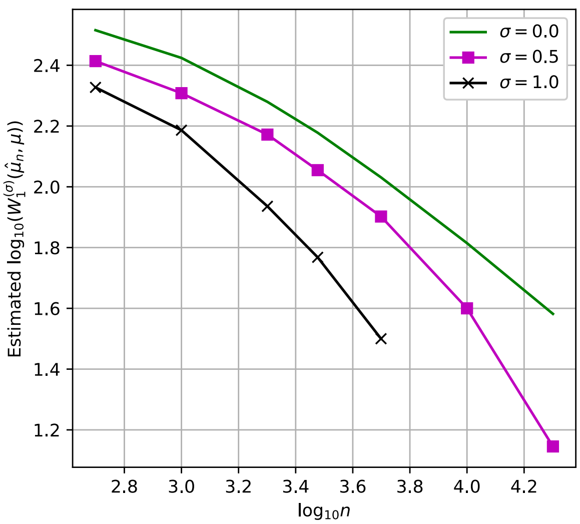

Figure 2: Non-uniform experiment. Convergence of $\mathsf{W}_1^{(\sigma)}(\hat{\mu}_n, \mu)$ as a function of n for various values of σ, shown in log-log space. The measure μ is the d-dimensional standard normal distribution, where $d = 100$. The $\sigma = 0$ case corresponds to the vanilla 1-Wasserstein distance, which converges slower than GOT (note the difference in slopes).

B Proof of Lemma 2

Recall that $g_\sigma(t) = \prod_{j=1}^d \tilde{g}_\sigma(t_j)$, where $\tilde{g}_\sigma$ is σ-subgaussian, zero mean, bounded, and monotonically decreasing as t_j moves away from zero. We first analyze the one-dimensional densities $\tilde{g}_\sigma$, and show that there exists a constant $c > 0$, such that

$$\tilde{g}_\sigma(t) \leq c e^{2\delta|t| - \delta^2 - \log\delta} \tilde{\varphi}_\sigma(t), \quad \forall t \in \mathbb{R}, \tag{28}$$

Note that the σ-subgaussianity of $\tilde{g}_\sigma$ (Def. 3) implies that

$$\mathbb{E}_{\tilde{g}_\sigma}\left[e^{\alpha X}\right] \leq e^{\frac{1}{2}\sigma^2\alpha^2}, \quad \forall \alpha \in \mathbb{R}, \tag{29}$$

which by [31] yields

$$\mathbb{P}_{\tilde{g}_\sigma}\big((-\infty, t) \cup (t, \infty)\big) \leq \exp(1 - t^2/(2\sigma^2)) = c' \tilde{\varphi}_\sigma(t), \tag{30}$$

where $\tilde{\varphi}_\sigma$ is a scalar Gaussian density (zero mean and σ^2 variance). We prove (28) for $t > 0$; the $t < 0$ case is identical.

where $c' = \sqrt{2\pi\sigma^2 e^2}$. Consequently, for any $t^\star$,

$$\begin{aligned} \mathbb{P}_{\tilde{g}_\sigma}\big((t^\star - \delta, t^\star]\big) &\leq \mathbb{P}_{\tilde{g}_\sigma}\big((t^\star - \delta, \infty)\big) \\ &\leq c' \tilde{\varphi}_\sigma(t^\star - \delta) \\ &= c' e^{(t^\star)^2 - (t^\star - \delta)^2} \tilde{\varphi}_\sigma(t^\star) \\ &= c' e^{2\delta t^\star - \delta^2} \tilde{\varphi}_\sigma(t^\star). \end{aligned} \tag{31}$$

Now, since $\tilde{g}_\sigma(t)$ monotonically decreases as t moves away from zero, for any $t^\star \geq \delta$ we have $\mathbb{P}_{\tilde{g}_\sigma}\big((t^\star - \delta, t^\star]\big) \geq \delta \tilde{g}_\sigma(t^\star)$. Substituting this into (31), we have for all $t^\star \geq \delta$ that

$$\begin{aligned} \delta \tilde{g}_\sigma(t^\star) &\leq c' e^{2\delta t^\star - \delta^2} \tilde{\varphi}_\sigma(t^\star), \\ \tilde{g}_\sigma(t^\star) &\leq c' e^{2\delta t^\star - \delta^2 - \log\delta} \tilde{\varphi}_\sigma(t^\star). \end{aligned}$$

Repeating the argument for $t < 0$ then yields

$$\tilde{g}_\sigma(t) \leq c' e^{2\delta|t| - \delta^2 - \log\delta} \tilde{\varphi}_\sigma(t)$$

for all $|t| \geq \delta$. Since $\tilde{g}_\sigma$ is bounded, $\sup_{|t| \leq \delta} \tilde{g}_\sigma(t) \left(e^{2\delta t - \delta^2 - \log\delta} \tilde{\varphi}_\sigma(t)\right)^{-1}$ exists, and hence (28) holds (for all $t \in \mathbb{R}$) with

$$c = \max\left[c', \sup_{|t| \leq \delta} \tilde{g}_\sigma(t) \left(e^{2\delta t - \delta^2 - \log\delta} \tilde{\varphi}_\sigma(t)\right)^{-1}\right].$$

Extending to the full d-dimensional distribution, note that since $t^2 + 1 > |t|$ for all t, we have that $\tilde{g}_\sigma(t) \leq c e^{2\delta t^2 + 2\delta - \delta^2 - \log\delta} \tilde{\varphi}_\sigma(t)$ for all t. We can then write

$$g_\sigma(t) \leq (c')^d e^{2\delta\|t\|^2 + 2d\delta - d\delta^2 - d\log\delta} \varphi_\sigma(t), \tag{32}$$

which establishes the lemma after collecting terms.