



Optimal Transport Between GMM for Texture Synthesis

Arthur Leclaire, Julie Delon, Agnès Desolneux

► To cite this version:

Arthur Leclaire, Julie Delon, Agnès Desolneux. Optimal Transport Between GMM for Texture Synthesis. 2022. hal-03613622v1

HAL Id: hal-03613622

<https://hal.science/hal-03613622v1>

Preprint submitted on 18 Mar 2022 (v1), last revised 1 Jun 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMAL TRANSPORT BETWEEN GMM FOR TEXTURE SYNTHESIS

*Julie Delon , Agnès Desolneux, Arthur Leclaire **

MAP5, CNRS, Université de Paris, France
Centre Borelli, CNRS, ENS Paris-Saclay, France
IMB, CNRS, Université de Bordeaux, France

ABSTRACT

Using optimal transport in image processing tasks has become very popular. However, it still faces difficult computational issues when dealing with high dimensional distributions. We propose here to use the recently introduced GMM-OT formulation, which consists in restricting the optimal transport problem to the set of Gaussian mixture models. As a proof of concept, we use it to revisit the texture model Texto based on optimal transport between distributions of image patches. Using GMM-OT in this texture model allows to deal with larger patches, hence providing results with better geometric details. This new model allows for synthesis, mixing, and style transfer.

Index Terms— Optimal transport, Gaussian mixture models, texture synthesis

1. INTRODUCTION

Numerical optimal transport (OT) has undergone spectacular progress in the last ten years, and is now used in a large variety of applications [1, 2, 3, 4]. In particular, important advances have been made in the numerical approximations of optimal transport, with the emergence of efficient tools like regularized optimal transport [5] or the sliced optimal transport [6]. Nevertheless, it remains complex to compute optimal transport distances between empirical distributions when the dimension (the number of samples n or the space dimension d) of the problem increases too much.

In this context, several questions were raised concerning the ability to compute numerical solutions of high-dimensional problems or the sample complexity of the different transport approximations [7, 8, 9]. However, for several applications, targeting exact solutions of optimal transport might not be desirable, whereas proxy formulations sharing similar properties might deliver more relevant solutions in practice. Among these alternative formulations, an OT-like distance between Gaussian Mixture Models (GMM) has been

introduced in [10]. It consists in restricting the set of possible couplings to GMM in the product space. Solutions of this formulation are easy to compute and merely require to calculate Bures distances between Gaussian measures and solve a small-scale discrete OT problem. When applied to discrete data, the dependency on the dimension d and the number of samples n lies only in the GMM fitting step on the data, and in the computation of Bures distances. This makes the approach very versatile and robust to dimension in practice. In this paper we explore the use of GMM-OT for texture modeling, as a proof of concept. To this aim, we revisit the texture model Texto [11], which is based on semi-discrete OT on image patches. We end up with a lighter and simpler formulation of the same problem. More precisely, in this context, the computing time of GMM-OT is at least one order of magnitude faster than the ones of semi-discrete OT or regularized OT, for similar (or even better) quality of synthesized images. This permits to use larger patch dimensions, and much more patches than the original Texto model.

2. REMINDERS ON OPTIMAL TRANSPORT BETWEEN GAUSSIAN MIXTURE MODELS

This section recalls the main results on optimal transport between GMM [10]. The quadratic Wasserstein distance between two probability measures μ_0 and μ_1 on \mathbb{R}^d with finite second moments is defined as

$$W_2^2(\mu_0, \mu_1) := \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|y_0 - y_1\|^2 d\gamma(y_0, y_1), \quad (1)$$

where $\Pi(\mu_0, \mu_1)$ is the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ_0 and μ_1 . A solution γ^* for the problem (1) is called an OT plan between μ_0 and μ_1 . This distance has been extensively used for various applications in data science, and especially to define Wasserstein barycenters of probability measures, which are defined similarly to Euclidean barycenters, replacing the Euclidean distance by W_2 .

2.1. Definition of MW_2

Let us denote by GMM_d the set of probability distributions which can be written as finite GMM on \mathbb{R}^d . Optimal transport

*The authors acknowledge support from the French Research Agency through the PostProdLEAP project (ANR-19-CE23-0027-01) and the MISTIC project (ANR-19-CE40-005). A. Leclaire acknowledges support of GdR Isis with Project Rémoga. ©Copyright 2022 IEEE – All rights reserved.

plans and Wasserstein barycenters between GMM are usually not GMM themselves, which can be annoying if we rely on this modeling to analyse or generate data. For this reason, the authors of [10] propose to modify the formulation of the classical Wasserstein distance by restricting the set of possible coupling measures to GMM on $\mathbb{R}^d \times \mathbb{R}^d$.

More precisely, let $\mu_0, \mu_1 \in GMM_d$, we can define

$$MW_2^2(\mu_0, \mu_1) := \inf_{\gamma \in \Pi^{\text{GMM}}(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|y_0 - y_1\|^2 d\gamma(y_0, y_1), \quad (2)$$

where $\Pi^{\text{GMM}}(\mu_0, \mu_1)$ is the set of probability measures in GMM_{2d} with marginals μ_0 and μ_1 . It is shown in [10] that MW_2 defines a distance between elements of GMM_d . Moreover, if $\mu_0 = \sum_{k=1}^{K_0} \pi_0^k \mu_0^k$ and $\mu_1 = \sum_{l=1}^{K_1} \pi_1^l \mu_1^l$, where the π_0^k, π_1^l are non-negative scalars and where the μ_0^k, μ_1^l are Gaussian measures, it can be shown [10] that

$$MW_2^2(\mu_0, \mu_1) = \min_{w \in \Pi(\pi_0, \pi_1)} \sum_{k,l} w_{kl} W_2^2(\mu_0^k, \mu_1^l), \quad (3)$$

where $\Pi(\pi_0, \pi_1)$ is the set of $K_0 \times K_1$ matrices with non negative entries and discrete marginals π_0 and π_1 . This discrete expression makes MW_2 easy to compute in practice, even in large dimension. Indeed, the distance W_2 between two Gaussian measures $\mu = \mathcal{N}(m, \Sigma)$ and $\tilde{\mu} = \mathcal{N}(\tilde{m}, \tilde{\Sigma})$ has the following closed-form expression

$$W_2^2(\mu, \tilde{\mu}) = \|m - \tilde{m}\|^2 + \text{tr} \left(\Sigma + \tilde{\Sigma} - 2 \left(\Sigma^{\frac{1}{2}} \tilde{\Sigma} \Sigma^{\frac{1}{2}} \right)^{\frac{1}{2}} \right), \quad (4)$$

where we denote by $M^{\frac{1}{2}}$ the unique semi-definite positive square-root of a symmetric semi-definite positive matrix M . If the different parameters of the GMM μ_0 and μ_1 are known, computing (3) boils down to compute $K_0 \times K_1$ Wasserstein distances between Gaussians and then to solve a $K_0 \times K_1$ discrete OT problem.

Similarly to the Wasserstein distance, it is possible to define barycenters for MW_2 , and this also gives rise to a simple discrete formulation if the GMM are known.

2.2. Using MW_2 in practice

Optimal plans for MW_2 are not supported on the graph of a function and hence do not directly yield a transport map between the mixtures μ_0 and μ_1 . In order to define a transport map from the optimal plan γ^* we can for instance use

$$T(x) = \mathbb{E}_{(X,Y) \sim \gamma^*}(Y|X=x). \quad (5)$$

As shown in [10], the closed-form formula for T is given by

$$T(x) = \frac{\sum_{k,l} w_{k,l}^* g_{m_0^k, \Sigma_0^k}(x) T_{k,l}(x)}{\sum_k \pi_0^k g_{m_0^k, \Sigma_0^k}(x)}, \quad (6)$$

where w^* is the optimal solution of the discrete problem (3), $T_{k,l}$ are the optimal affine maps between Gaussians μ_0^k and μ_1^l , and $g_{m, \Sigma}$ is the density of $\mathcal{N}(m, \Sigma)$.

3. TEXTO-GMM, A MULTISCALE TEXTURE SYNTHESIS APPROACH WITH OPTIMAL TRANSPORT BETWEEN PATCHES

First, the Texto-GMM model is defined at one single scale, as a transformation of a stationary Gaussian random field. Next, it is extended to a multi-scale framework by using successive exemplar-based upsampling. The construction is similar to the Texto model [11], except that the semi-discrete OT maps are replaced with transport maps obtained with GMM-OT.

Let $u : \Omega \rightarrow \mathbb{R}^d$ be the exemplar texture defined on a rectangular domain $\Omega \subset \mathbb{Z}^2$. Let also $\omega = \{0, \dots, w-1\}^2$ be the patch domain, and $u_{|a+\omega} \in \mathbb{R}^\omega$ the patch at position a .

3.1. Monoscale Model

The monoscale model can be decomposed as a coarse Gaussian synthesis followed by a patch-based enhancement. The Gaussian model [12] used for coarse synthesis is defined by

$$\forall a \in \mathbb{Z}^2, U(a) = \bar{u} + \frac{1}{\sqrt{|\Omega|}} \sum_{b \in \Omega} (u(b) - \bar{u}) W(a - b) \quad (7)$$

where $\bar{u} = \frac{1}{|\Omega|} \sum_{a \in \Omega} u(a)$ and where W is a Gaussian $\mathcal{N}(0, 1)$ white noise on \mathbb{Z}^2 . The field U is a stationary Gaussian random field with mean \bar{u} and whose covariance equals the empirical covariance of u . Then, a patch transformation $T : \mathbb{R}^\omega \rightarrow \mathbb{R}^\omega$ is applied to get a new field V given by

$$\forall a \in \mathbb{Z}^2, V(a) = \frac{1}{|\omega|} \sum_{h \in \omega} T(U_{|a-h+\omega})(h). \quad (8)$$

The patch transformation T is adjusted so that it sends the distribution μ of the Gaussian patch $U_{|a+\omega}$ to the patch distribution of u . This allows to reimpose local features observed in u in a statistically coherent way. We here exploit GMM-OT to get a relevant patch transformation. Indeed, we can first use the EM algorithm to fit a GMM distribution ν to the patches of u , and then solve the GMM-OT between the Gaussian measure μ and the GMM ν . Once the GMM-OT problem solved, we get the patch transformation T given by (5).

3.2. Multiscale Model

We will now see how the texture model can be extended in a multiscale fashion. For $0 \leq s \leq S-1$, we consider a sub-sampled version u_s of u defined on a subdomain $\Omega_s \subset 2^s \mathbb{Z}^2$. We will also denote by ν_s a GMM distribution that is fitted by EM to the empirical distribution of the $w \times w$ patches of u_s .

The multiscale Texto-GMM model is defined in a coarse-to-fine manner. At the coarsest scale $s = S-1$, U_{S-1} is initialized as the Gaussian field (7) estimated from u_{S-1} . Suppose now that U_s at scale $s \in \{1, \dots, S-1\}$ is given. Then, a GMM model μ_s is fitted to the patches of U_s , and an optimal transport plan γ_s^* for $MW_2(\mu_s, \nu_s)$ is computed as explained

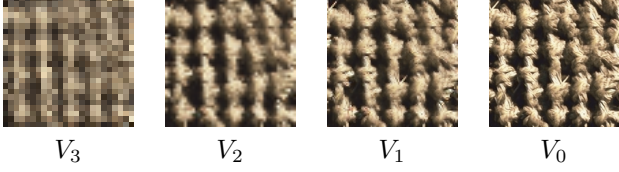


Fig. 1. Illustration of the coarse-to-fine synthesis process.

in Section 2.2. Then, U_s is transformed using the map T_s related to γ_s^* (Eq.(5)). However, for the need of the upcoming upsampling step, we need to compose with a L^2 nearest-neighbor projection on the exemplar patches at scale s . Thus, denoting by $C_s(a)$ the coordinate of the nearest neighbor of $T_s(U_{s|a+2^s\omega})$ in u_s , we set

$$V_s(a) = \frac{1}{|\omega|} \sum_{b \in 2^s \omega} u_s(C_s(a-b) + b), \quad a \in 2^s \mathbb{Z}^2. \quad (9)$$

This allows to upsample by taking twice larger patches at the same positions, thus initializing U_{s-1} for the next finer scale: for all $a \in 2^s \mathbb{Z}^2$ and for all $k \in \{0, \dots, 2^{s-1}\}^2$, we set

$$U_{s-1}(a+k) = \frac{1}{|\omega|} \sum_{b \in 2^s \omega} u_{s-1}(C_s(a-b) + b + k). \quad (10)$$

At the end, we obtain the sample texture image V_0 . This coarse-to-fine synthesis process is illustrated on Fig. 1. Let us finally mention that, once the model estimation has been done (i.e. when all GMM models are estimated and all transport plans γ_s^* have been computed using one synthesis), then the model can be synthesized offline to produce (possibly-larger) images on-demand.

4. EXPERIMENTS

Several synthesis examples are displayed on Fig. 2. In the captions, the output of the multiscale process explained in Section 3.2 is referred to as Textto-GMM. In all the results of Textto-GMM shown here, the estimated GMM models have 4 components. We empirically observed that increasing further this parameter only brings minor improvements to the visual results (while considerably slowing the estimation of GMM).

In practice, when working with patches larger than 5×5 , we do not use all patches in the EM algorithm and the nearest neighbor projection steps of Textto-GMM, but only $N_p = 10^4$ randomly chosen patches. Such subsampling of the patch distribution was observed to be harmless for the texture synthesis application and speed-up these two steps. It also represents a major improvement compared to the original Textto method [11], which was proven successful only for discrete distributions of 10^3 patches of size 3×3 (since it relies on a costly stochastic algorithm for semi-discrete OT).

Regarding the computation time, Textto-GMM also allows for a much faster estimation than the original Textto

model [11], even though Textto-GMM handles 10 times more patches. Indeed, the GMM-OT algorithm runs faster than the stochastic algorithm used in [11] for semi-discrete OT or even faster than the Sinkhorn algorithm [5]. For instance, on a transport problem with 10^4 points in source and target distributions in dimension 147 (for 7×7 color patches), with a modern laptop, solving GMM-OT takes $\approx 1'$, to be compared with $2.5h$ for 10^5 iterations of the stochastic OT algorithm, or $40'$ to perform 10^3 iterations of Sinkhorn algorithm. With 5×5 patches, for an image of size 256×256 , the whole synthesis algorithm for Textto-GMM (with $N_p = 10^4$) takes $\approx 1'$ while the synthesis algorithm (including the model estimation) for Textto (with $N_p = 10^3$) takes more than $1h$. Moreover, the most costly step of Textto-GMM is actually the nearest neighbor projection that is needed for upsampling, and which could be greatly accelerated with dedicated algorithms [13].

GMM-OT allows to partially cope with the curse of dimensionality, and thus permits to use much larger patches in Textto-GMM than Textto (also because GMM-OT can handle distributions with much more points). This leads to more faithful synthesis of structured textures, with a better preservation of the sharp details, as illustrated in Fig. 2. The second column of Fig. 2 (NN) confirms the importance of using OT for patch transformation and not only simple nearest-neighbor matching. Also, the third and fourth columns confirm that the previous Textto model works only with small patches. In contrast, Textto-GMM produces remarkable results on these textures with the proper choice of parameters w and S . However, we do not claim here to reach state-of-the-art results for texture synthesis. Our aim is only to show how the original Textto model can be greatly improved by relying on GMM-OT instead of semi-discrete OT.

Finally, let us mention that the Textto-GMM model can be easily adapted for style transfer and texture mixing. The adaptation to style transfer is a straightforward extension of the technique explained in [14] where the texture information can be treated with GMM-OT maps and then blended with the geometric features. For mixing, we exploit the explicit formula of [10] for mixing with GMM-OT. For a mixing parameter $\alpha \in [0, 1]$, starting from a mixed Gaussian random field U_{S-1} as in [15], at each scale s we apply a patch transformation that targets the mixed patch distribution ν_s^α obtained by mixing GMM patch distributions ν_s^0, ν_s^1 of two source images u^0, u^1 . For exemplar-based upsampling, we rely on a nearest-neighbor projection on patches $(1-\alpha)p + \alpha T(p)$ where p is a patch of ν_s^0 and T the map (6) for the GMM transport between ν_s^0, ν_s^1 . On Fig. 3 and Fig. 4, one can see that Textto-GMM leads to relevant results for both applications. However it is still unclear how the mixed texture model could be formulated globally (thus avoiding the ad-hoc projection on mixed patches).

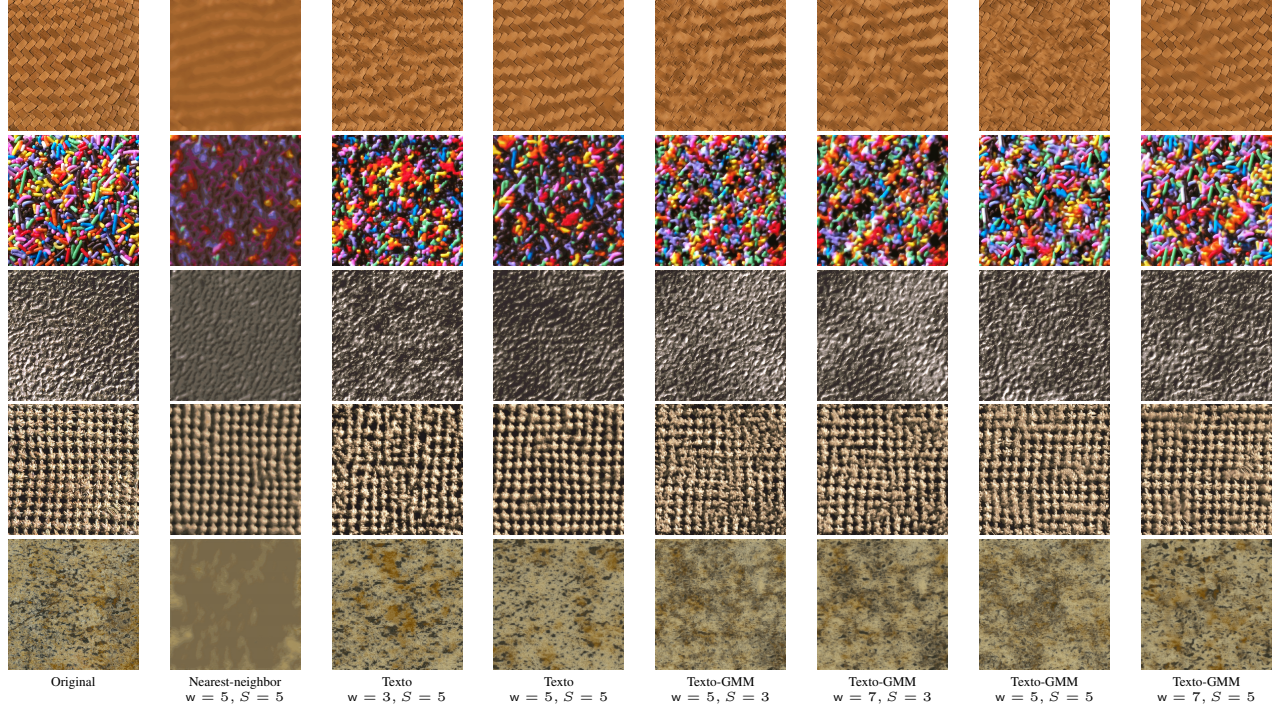


Fig. 2. In this figure, we display several 512×512 original texture images (column 1), and synthesized images obtained with different models explained in the paper, with different parameters w (patch size) and S (number of scales). In column 2, the synthesis is obtained by using only patch nearest-neighbor projections at each scale. And then, we display the results obtained with the previous Texto model [11] (columns 3,4) and the Texto-GMM model proposed here (columns 5–8).

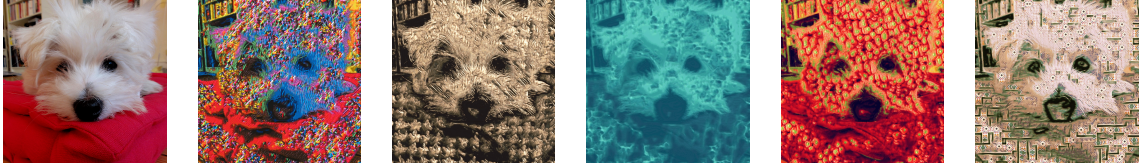


Fig. 3. In this figure, we show that adapting the style transfer technique from [14] to the Texto-GMM model can produce convincing style transfer results.

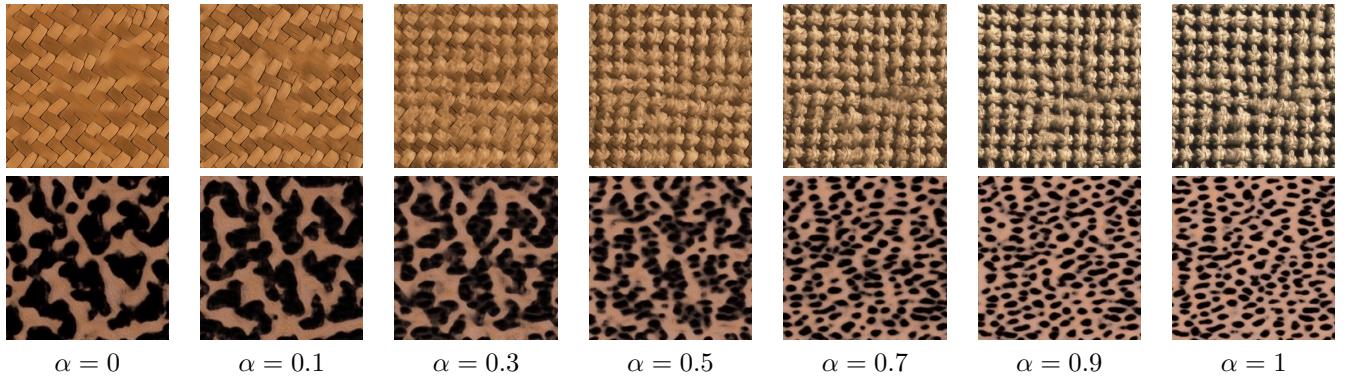


Fig. 4. We display here examples of texture mixing computed from two source images u_0, u_1 (not shown in this figure). From left to right, we display a sample of the mixed texture model with mixing parameter $\alpha \in [0, 1]$. For $\alpha = 0$ (resp. 1), we get a sample of the Texto-GMM model associated with u_0, u_1 . It is interesting to see how the mixed Texto-GMM model is able to combine the geometric structures of u_0, u_1 . Parameters are set to $(w, S) = (7, 5)$ for the first row and $(5, 4)$ for the second row.

5. REFERENCES

- [1] Nicolas Bonneel, Michiel Van De Panne, Sylvain Paris, and Wolfgang Heidrich, “Displacement interpolation using lagrangian mass transport,” in *Proceedings of the 2011 SIGGRAPH Asia conference*, 2011, pp. 1–12.
- [2] Nicolas Courty, Rémi Flamary, Amaury Habrard, and Alain Rakotomamonjy, “Joint distribution optimal transportation for domain adaptation,” *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [3] Tero Karras, Samuli Laine, and Timo Aila, “A style-based generator architecture for generative adversarial networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4401–4410.
- [4] Jean Feydy, Pierre Roussillon, Alain Trounev, and Pietro Gori, “Fast and scalable optimal transport for brain tractograms,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 636–644.
- [5] Marco Cuturi, “Sinkhorn distances: Lightspeed computation of optimal transport,” in *Advances in neural information processing systems*, 2013, pp. 2292–2300.
- [6] Nicolas Bonnotte, *Unidimensional and evolution methods for optimal transportation*, Ph.D. thesis, Paris 11, 2013.
- [7] Jonathan Weed and Francis Bach, “Sharp asymptotic and finite-sample rates of convergence of empirical measures in wasserstein distance,” *Bernoulli*, vol. 25, no. 4A, pp. 2620–2648, 2019.
- [8] Aude Genevay, Lénaïc Chizat, Francis Bach, Marco Cuturi, and Gabriel Peyré, “Sample complexity of Sinkhorn divergences,” in *The 22nd international conference on artificial intelligence and statistics*. PMLR, 2019, pp. 1574–1583.
- [9] Lénaïc Chizat, Pierre Roussillon, Flavien Léger, François-Xavier Vialard, and Gabriel Peyré, “Faster wasserstein distance estimation with the sinkhorn divergence,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 2257–2269, 2020.
- [10] Julie Delon and Agnès Desolneux, “A Wasserstein-type distance in the space of gaussian mixture models,” *SIAM Journal on Imaging Sciences*, vol. 13, no. 2, pp. 936–970, 2020.
- [11] Bruno Galerne, Arthur Leclaire, and Julien Rabin, “A texture synthesis model based on semi-discrete optimal transport in patch space,” *SIAM Journal on Imaging Sciences*, vol. 11, no. 4, pp. 2456–2493, 2018.
- [12] B. Galerne, Y. Gousseau, and J.-M. Morel, “Random phase textures: Theory and synthesis,” *IEEE Trans. Image Proc.*, vol. 20, no. 1, pp. 257 – 267, 2011.
- [13] Lin Liang, Ce Liu, Ying-Qing Xu, Baining Guo, and Heung-Yeung Shum, “Real-time texture synthesis by patch-based sampling,” *ACM Transactions on Graphics (ToG)*, vol. 20, no. 3, pp. 127–150, 2001.
- [14] Arthur Leclaire and Julien Rabin, “A stochastic multi-layer algorithm for semi-discrete optimal transport with applications to texture synthesis and style transfer,” *Journal of Mathematical Imaging and Vision*, vol. 63, no. 2, pp. 282–308, 2021.
- [15] G. Xia, S. Ferradans, G. Peyré, and J. Aujol, “Synthesizing and Mixing Stationary Gaussian Texture Models,” *SIAM Journal on Imaging Sciences*, vol. 7, no. 1, pp. 476–508, 2014.