We report the proofs of the different propositions (App. A), an instantiation of online Sinkhorn for discrete measures (App. B), and a supplementary experiment (App. C).

We advise the reader that we provide and prove a slightly modified version of Prop. 4. As is, the online Sinkhorn algorithm only converges approximately, as became apparent after carefully checking our derivations.

- Prop. 4 will therefore be replaced by Prop. 4 bis in future revisions. Prop. 4 bis is an approximate (but global) convergence result.
- We describe a slight modification of the online Sinkhorn algorithm that does converge (Prop. 5). This result will be integrated to the main text in future revisions.

A. Proofs

We prove the propositions in their order of appearance. We make the following classic assumptions on the cost regularity and distribution compactness.

Assumption 1. *The cost $C : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is L-lipschitz, and $\mathcal{X}$ is assumed to be compact.*

A.1. Proof of Prop. 2

Proof. We use Theorem 1 from Diaconis & Freedman (1999). For this, we simply note that the space $\mathcal{C}(\mathcal{X}) \times \mathcal{C}(\mathcal{X})$ in which the chain $x_t \triangleq (f_t, g_t)_t$, endowed with the metric $\rho((f_1, g_1), (f_2, g_2)) = \|f_1 - f_2\|_{\text{var}} + \|g_1 - g_2\|_{\text{var}}$, is complete and separable (the countable set of polynomial functions are dense in this space, for example). We consider the operator $A_\theta \triangleq T_{\hat\beta}(T_{\hat\alpha}(\cdot))$. $\theta \triangleq (\hat\alpha, \hat\beta)$ denotes the random variable that is sampled at each iteration. We have the following recursion:

$$x_{t+2} = A_{\theta_t}(x_t).$$

For all $\hat\alpha \in \mathcal{P}(\mathcal{X})$, $\hat\beta \in \mathcal{P}(\mathcal{X})$, A_θ with $\theta = (\hat\alpha, \hat\beta)$ is contracting, with module $\kappa_\theta < \kappa < 1$ (see for instance Proposition 19 from Vialard (2019)), as we assume $\mathcal{X}$ to be compact (Ass. 1). Therefore

$$\int_\theta \kappa_\theta \mathrm{d}\mu(\theta) < 1, \qquad \int_\theta \log \kappa_\theta \mathrm{d}\mu(\theta) < 0.$$

Finally, we note, for all $f \in \mathcal{C}(\mathcal{X})$

$$\|T_\beta(T_{\hat\alpha}(f))\|_\infty \leqslant \|f\|_\infty + 2 \max_{x,y\in\mathcal{X}} C(x,y),$$

therefore $\rho(A_\theta(x_0), x_0) \leqslant 2\|x_0\|_\infty + 2\max_{x,y\in\mathcal{X}} C(x,y)$ for all θ $(\hat\alpha, \hat\beta)$. The regularity condition of the theorem are therefore met. Each of the induced Markov chains $(f_{2t}, g_{2t})_t$ and $(f_{2t+1}, g_{2t+1})_t$ has a unique stationary distribution. These stationary distributions are the same: the stationary distribution is independent of the initialisation and both sequences differs only by their initialisation. Therefore $(f_t, g_t)_t$ have a unique stationary distribution (F_∞, G_∞). □

A.2. Proof of Prop. 3

The "slowed-down" Sinkhorn iterations converge toward an optimal potential couple, up to a constant factor: this stems from the fact that we apply contractions in the space $(\mathcal{C}(\mathcal{X}), \|\cdot\|_{\text{var}})$ with a contraction factor that decreases sufficiently slowly.

Proof. We write $(f_t, g_t)_t$ the sequence of iterates. Given a pair of optimal potentials $(f^\star, g^\star)$, we write $u_t \triangleq f_t - f^\star$, $v_t \triangleq g_t - g^\star$, $u_t^C \triangleq T_\alpha(f_t) - g^\star$ and $v_t^C \triangleq T_\alpha(g_t) - f^\star$ For all $t > 0$, we observe that

$$\begin{aligned}
\max u_{t+1} &= -\log \min \exp(-u_{t+1}) \\
&= -\log \big(\min \big((1-\eta_t)\exp(-u_t) + \eta_t \exp(-v_t^C)\big)\big) \\
&\leqslant -\log \big((1-\eta_t)\min\exp(-u_t) + \eta_t \min\exp(-v_t^C)\big) \\
&\leqslant -(1-\eta_t)\log\min\exp(-u_t) - \eta_t \log\min\exp(-v_t^C) \\
&= (1-\eta_t)\max u_t + \eta_t \max v_t^C,
\end{aligned}$$

where we have used the algorithm recursion on the second line, $\min f + g \geqslant \min f + \min g$ on the third line and Jensen inequality on the fourth line. Similarly

$$\min u_{t+1} \geqslant (1-\eta_t) \min u_t + \eta_t \min v_t^C,$$

and mirror inequalities hold for v_t. Summing the four inequalities, we obtain

$$\begin{aligned} e_{t+1} \triangleq \|u_{t+1}\|_{\text{var}} + \|v_{t+1}\|_{\text{var}} &= \max u_{t+1} - \min u_{t+1} + \max v_{t+1} - \min v_{t+1} \\ &\leqslant (1-\eta_t)(\|u_t\|_{\text{var}} + \|v_t\|_{\text{var}}) + \eta_t(\|u_t^C\|_{\text{var}} + \|v_t^C\|_{\text{var}}), \\ &\leqslant (1-\eta_t)(\|u_t\|_{\text{var}} + \|v_t\|_{\text{var}}) + \eta_t \kappa (\|u_t\|_{\text{var}} + \|v_t\|_{\text{var}}), \end{aligned}$$

where we use the contractance of the soft-C-transform, that guarantees that there exists $\kappa < 1$ such that $\|v_t^C\|_{\text{var}} \leqslant \kappa \|v_t\|_{\text{var}}$ and $\|u_t^C\|_{\text{var}} \leqslant \kappa \|u_t\|_{\text{var}}$ (Peyré & Cuturi, 2019).

Unrolling the recursion above, we obtain

$$\log e_t = \sum_{s=1}^{t} \log(1 - \eta_t(1-\kappa)) + \log(e_0) \to -\infty,$$

provided that $\sum \eta_t = \infty$. The proposition follows. □

A.3. Corrected version and proof of Prop. 4

We make the following Robbins-Monroe assumptions on weights.

Assumption 2. *The weight sequence $(\eta_t)_t$ is square summable but not summable: $\sum \eta_t = \infty$ and $\sum \eta_t^2 < \infty$.*

Finally, we recall that

$$\|f\|_{\text{var}} \triangleq \sup_{x \in \mathcal{X}} f(x) - \inf_{y \in \mathcal{X}} f(y).$$

We prove the following modified proposition, which establishes approximate convergence of the online Sinkhorn algorithm.

Proposition 4 bis. *We assume Ass. 1 and 2. The online Sinkhorn algorithm (Alg. 1) yields a sequence (f_t, g_t) that reaches a ball centered around $f^\star, g^\star$ for the variational norm $\|\cdot\|_{\text{var}}$. Namely, there exists $T > 0$, $C > 0$ such that for all $t > T$,*

$$\|f_t - f^\star\|_{\text{var}} + \|g_t - g^\star\|_{\text{var}} \leqslant \frac{C}{\sqrt{n}}$$

Proof. For discrete realizations $\hat\alpha$ and $\hat\beta$, we define the perturbation terms

$$\varepsilon_{\hat\beta}(\cdot) \triangleq f^\star - T_{\hat\beta}(g^\star), \qquad \iota_{\hat\alpha}(\cdot) \triangleq g^\star - T_{\hat\alpha}(f^\star),$$

so that the updates can be rewritten as

$$\begin{aligned} \exp(-f_{t+1} + f^\star) &= (1-\eta_t)\exp(-f_t + f^\star) + \eta_t \exp(-T_{\hat\beta_t}(g_t) + T_{\hat\beta_t}(g^\star) + \varepsilon_{\hat\beta_t}) \\ \exp(-g_{t+1} + g^\star) &= (1-\eta_t)\exp(-g_t + g^\star) + \eta_t \exp(-T_{\hat\beta_t}(f_t) + T_{\hat\beta_t}(f^\star) + \iota_{\hat\beta_t}). \end{aligned}$$

We denote $u_t \triangleq -f_t + f^\star$, $v_t \triangleq -g_t + g^\star$, $u_t^C \triangleq T_{\hat\beta_t}(f_t) - T_{\hat\beta_t}(f^\star)$, $v_t^C \triangleq T_{\hat\beta_t}(g_t) - T_{\hat\beta_t}(g^\star)$. Reusing the same derivations as in the proof of Prop. 3, we obtain

$$\begin{aligned} \|u_{t+1}\|_{\text{var}} &\leqslant (1-\eta_t)\|u_t\|_{\text{var}} + \eta_t \log\big(\max_{x,y \in \mathcal{X}} \exp(\varepsilon_{\hat\beta_t}(x) - \varepsilon_{\hat\beta_t}(y)) \exp(v_t^C(x) - v_t^C(y))\big) \\ &\leqslant (1-\eta_t)\|u_t\|_{\text{var}} + \eta_t \|v_t^C\|_{\text{var}} + \eta_t \|\varepsilon_{\hat\beta_t}\|_{\text{var}}, \end{aligned}$$

where we have used $\max_x f(x)g(x) \leqslant \max_x f(x) \max_x f(x)$ on the second line. Therefore, using the contractance of the soft C-transform,

$$e_{t+1} \leqslant (1-\tilde\eta_t) e_t + \tilde\eta_t (\|\varepsilon_{\hat\beta_t}\|_{\text{var}} + \|\iota_{\hat\alpha_t}\|_{\text{var}}), \tag{10}$$

where we set $e_t \triangleq \|u_t\|_{\text{var}} + \|v_t\|_{\text{var}}$, $\tilde\eta_t = \frac{\eta_t}{1-\kappa}$ and κ is set to be the minimum of the contraction factor over all empirical realizations $\hat\alpha_t$, $\hat\beta_t$ of the distributions α and β. It is upper bounded by $1 - e^{-L\text{diam}(\mathcal{X})}$, thanks to Ass. 1 and according to Proposition 19 in Vialard (2019).

The realizations $\hat\beta_t$ and $\hat\alpha_t$ are sampled according to the same distribution $\hat\alpha$ and $\hat\beta$. We define the sequence r_t to be the running average of the variational norm of the functional error terms

$$r_{t+1} = (1-\tilde\eta_t) r_t + \tilde\eta_t (\|\varepsilon_{\hat\beta_t}\|_{\text{var}} + \|\iota_{\hat\alpha_t}\|_{\text{var}}),$$

so that we have, for all $t > 0$, $e_t \leqslant r_t$. Using Lemma B.7 of Mairal (2013) on running averages, this sequence converges towards the scalar expected value

$$R = \mathbb{E}_{\hat\alpha,\hat\beta}[\|\varepsilon_{\hat\beta}\|_{\text{var}} + \|\iota_{\hat\alpha}\|_{\text{var}}] > 0.$$

We now relate R to the number of samples n using a uniform law of large number result on parametric functions. We indeed have, by definition

$$\begin{aligned}
\mathbb{E}_{\hat\beta}\|\varepsilon_{\hat\beta}\|_{\text{var}} &\leqslant \mathbb{E}_{\hat\beta}\|T_\beta(g^\star) - T_{\hat\beta}(g^\star)\|_\infty \\
&= \mathbb{E}_{Y_1,\dots Y_n\sim\beta} \sup_{x\in\mathcal{X}} |\frac{1}{n}\sum_{i=1}^n \exp(g^\star(Y_i)) - C(x,Y_i)) - \mathbb{E}_{Y\sim\beta}[\exp(g^\star(Y)) - C(x,Y)]| \\
&= \mathbb{E}\sup_{x\in\mathcal{X}} |\frac{1}{n}\sum_{i=1}^n f_i(x) - f(x)|,
\end{aligned} \tag{11}$$

where we have defined $f_i : x \to \exp(g^\star(Y_i - C(x,Y_i))$ and set f to be the expected value of each f_i. The compactness of $\mathcal{X}$ ensures that the functions are square integrable and uniformly bounded. From Example 19.7 and Lemma 19.36 of Van der Vaart (2000) (see also Lemma B.6 from Mairal (2013)), there exists C (that depends only on α and β) such that

$$\mathbb{E}_{\hat\beta}\|\varepsilon_{\hat\beta}\|_{\text{var}} \leqslant \frac{C}{\sqrt{n}}.$$

A similar results holds for $\|\iota_\alpha\|_{\text{var}}$. The result follows by a simple comparison of sequence. □

Dependency on the mini-batch size. The online Sinkhorn algorithm thus exhibits a rate of error in $\mathcal{O}(\frac{1}{\sqrt{n}})$ for estimating potentials. This is similar to the regular Sinkhorn algorithm (Genevay et al., 2019), whose sample complexity is also in $\mathcal{O}(\frac{1}{\sqrt{n}})$. The proof is simpler in our case, but we assume repeated sampling. We show in the experiment section (§5) that online Sinkhorn strongly outperforms the regular Sinkhorn algorithm thanks to repeated sampling. This suggests that the constants appearing in the bounds of online Sinkhorn are much better than the ones appearing in the sample complexity of the regular Sinkhorn algorithm.

A.4. Convergence of online Sinhkorn

In the proof of Prop. 4 bis and in particular Eq. (10), the term that prevents the convergence of e_t is the term

$$\eta_t(\|\varepsilon_{\hat\beta_t}\|_{\text{var}} + \|\iota_{\hat\alpha_t}\|_{\text{var}}),$$

which is not summable in general. We can control this term by increasing the size of $\hat\alpha_t$ and $\hat\beta_t$ with time, at a sufficient rate. This rate will be controlled by a second sequence $(w_t)_t$. We set a reference batch-size $n \in \mathbb{N}^*$, and denote $n(t)$ the batch size at time t. We assume that it is increasing sufficiently fast.

Assumption 3. *For all $t > 0$, $n(t) = \frac{n}{w_t^2}$ is an integer and the realizations $\hat\alpha_t$ and $\hat\beta_t$ are defined as*

$$\hat\alpha_t \triangleq \frac{1}{n(t)}\sum_{i=1} n(t)\delta_{x_i^t}, \quad \hat\beta_t \triangleq \frac{1}{n(t)}\sum_{i=1} n(t)\delta_{y_i^t}$$

where $(x_i^t)_i$ and $(y_i^t)_i$ are i.i.d samples of α and β and

$$\sum w_t\eta_t < \infty, \quad \textit{as well as} \quad \sum \eta_t = \infty.$$

We then have the following result of global convergence.

Proposition 5. *We assume Ass. 1, and 3. Almost surely, the iterates of online Sinkhorn (Alg. 1) converge, and we have*

$$\|f_t - f^\star\|_{\text{var}} + \|g_t - g^\star\|_{\text{var}} \to 0.$$

Proof. From (10), for all $t > 0$, we have

$$e_{t+1} \leqslant (1 - \tilde{\eta}_t) e_t + \tilde{\eta}_t (\|\varepsilon_{\hat{\beta}_t}\|_{\text{var}} + \|\iota_{\hat{\alpha}_t}\|_{\text{var}}).$$

Taking the expectation and using the uniform law of large number (11), there exists $C, C' > 0$ such that

$$\begin{aligned}\mathbb{E} e_{t+1} &\leqslant (1 - \tilde{\eta}_t)\mathbb{E} e_t + \tilde{\eta}_t \frac{C}{\sqrt{n(t)}} \\ &\leqslant (1 - \tilde{\eta}_t)\mathbb{E} e_t + C' \tilde{\eta}_t w_t.\end{aligned}$$

Writing $E_t = \mathbb{E} e_t$, for all $t > 0$, $E_{t+1} - E_t \leqslant C' \tilde{\eta}_t w_t$. From Ass. 3, $(E_{t+1} - E_t)_t$ is summable and $E_t \to_{t\to\infty} \ell \geqslant 0$. Finally, we write

$$E_T \leqslant E_1 - \sum_{s=1} \tilde{\eta}_s E_s + C \sum_{s=1} \tilde{\eta}_s w_s.$$

Assuming $\ell > 0$ leads to $E_t \to -\infty$ which is absurd. Therefore $\mathbb{E} e_t \to_{t\to\infty} 0$. Since $e_t \geqslant 0$ for all $t > 0$, this implies that $e_t \to_{t\to\infty} 0$ almost surely. $\square$

Choosing appropriate weights. Online Sinkhorn thus works for $\eta_t = \frac{1}{t^a}$ with $a \in [0, 1]$, provided we use batch-sizes of size $n\, t^{2b}$, with $b > 1 - a$. Slowing down Sinkhorn iterations permits to work with batches whose size increases more slowly. The limit case when $a = 1$ requires only $b > 0$, i.e. batch-sizes growing arbitrarily slowly. With these simple weights, the complexity of the iteration t of the algorithm is then $\mathcal{O}(\frac{n^2}{t^{1+2b}})$.

Reusing samples from the past. The distributions α_t and β_t do not have to contain only new samples: they may contained already seen ones. This allows to reuse memorized costs $C(x_i, y_j)$. We consider as in the main text a stream of batches of size n of constant step sizes. We set $n_t = n\, t$ and index samples of batch t as $(x_i)_{(n_t, n_{t+1}]}$. At each iteration, we can set the empirical distribution of size $n(t)$ to be

$$\hat{\alpha}_t = \frac{1}{n(t)} \Big(\sum_{i=n_t}^{n_{t+1}} \delta_{x_i}) + \sum_{i=n_t - n(\frac{1}{w_t^2} - 1)}^{n_t} \delta_{x_i} \Big),$$

and similarly for $\hat{\beta}_t$. The past starting point $i = n_t - n(\frac{1}{w_t^2} - 1)$ remains positive provided that $w_t = \mathcal{O}(\frac{1}{t^{1/2}})$, which is compatible with Ass. 3 whenever $\eta_t = \frac{1}{n^a}$, $a \in (\frac{1}{2}, 1]$. Using samples from the past is thus possible within the range of weights meeting Ass. 2, i.e. in the setting of Prop. 4 bis.

Concretely, reusing samples from the past amounts to correct a fraction of the weights $(p_i)_i$ and $(q_i)_i$ at each iteration.

B. Online Sinkhorn for discrete distributions

The online Sinkhorn algorithm takes a simpler form when dealing with discrete distributions. We derive it in Alg. 2. We take two distributions of the same size for the sake of simplicity. In this case, we evaluate the potentials as

$$g_t(y) = -\log \sum_{j=1}^{n} \exp(p_j - C(x_j, y))$$

$$f_t(x) = -\log \sum_{j=1}^{n} \exp(q_j - C(x, y_j)).$$

Note that the computations derived in Alg. 2 are proposed in log-space so as to prevent numerical overflows. The sets $|I|$ and $|J|$ can have varying sizes along the algorithm, which allows for example to speed-up the initial Sinkhorn iteration (§5.2). In such case, the cost matrix $(C(x_i, y_j))_{i,j}$ is progressively computed along the iterations.

Algorithm 2 Online Sinkhorn potentials in the discrete setting

Input: Distribution $\alpha \in \triangle^n$ and $\beta \in \triangle^n$, $x \in \mathbb{R}^{n\times d}$, $y \in \mathbb{R}^{n\times d}$, learning weights $(\eta_t)_t$
Set $p = q = -\infty \in \mathbb{R}^n$
for $t = 0, \dots, T-1$ **do**
  $q \leftarrow q + \log(1-\eta_t)$, $p \leftarrow p + \log(1-\eta_t)$,
  Sample $I \subset [1, n]$, $J \subset [1, n]$
  for $i \in I$ **do**
    $q_i \leftarrow \log\big(\exp(q_i) + \exp(\log(\eta_t) - \log \frac{1}{n}\sum_{j=1}^n \exp(p_j - C(x_j, y_i))\big)$
  for $i \in J$ **do**
    $p_i \leftarrow \log\big(\exp(q_i) + \exp(\log(\eta_t) - \log \frac{1}{n}\sum_{j=1}^n \exp(q_j - C(x_i, y_j))\big)$
  Optional: refit all $q_i = g_t(y_i) - \log(n)$
    $p_i = f_t(x_i) - \log(n)$
Returns $f_T : (q, y)$ and $g_T : (p, x)$

C. Experiments

We report the performance of online+full Sinkhorn for $\varepsilon \sim 10^{-4} \max C$ in Fig. 5. Although the gains are less important than with higher ε, they remain significant in this low regularization regime.

Grids in §5.1. We run the online Sinkhorn algorithm with step-sizes $\eta_t = \frac{1}{t^a}$, $a \in \{\frac{1}{2}, 1\}$ and $w_t = \frac{1}{t^b}$, $b \in \{\frac{1}{2}, 1\}$. In all experiments, $a = b = \frac{1}{2}$ turned out to perform best.

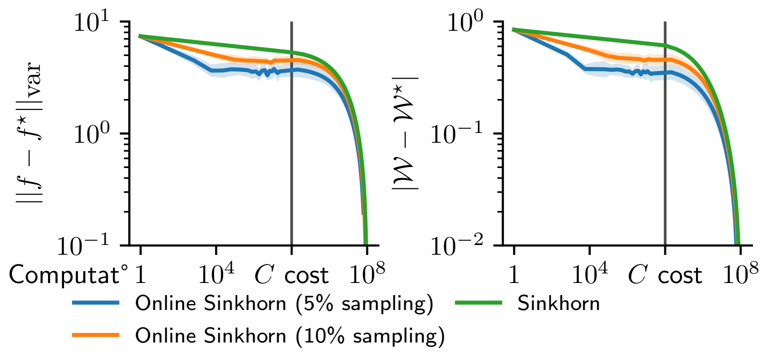

Figure 5. Online Sinkhorn accelerates the first Sinkhorn iterations even for low regularization. $\varepsilon = 10^{-4} \max C$.

References

Diaconis, P. and Freedman, D. Iterated random functions. *SIAM Review*, 41(1):45–76, 1999.

Genevay, A., Chizat, L., Bach, F., Cuturi, M., and Peyré, G. Sample complexity of sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, 2019.

Mairal, J. Stochastic majorization-minimization algorithms for large-scale optimization. In *Advances in Neural Information Processing Systems*, 2013.

Peyré, G. and Cuturi, M. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6): 355–607, 2019.

Van der Vaart, A. W. *Asymptotic statistics*. Cambridge University Press, 2000.

Vialard, F.-X. An elementary introduction to entropic regularization and proximal methods for numerical optimal transport. 2019.