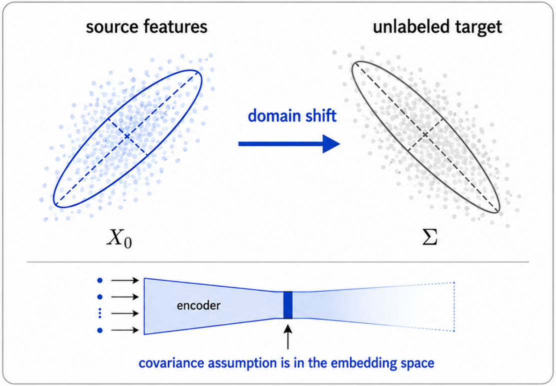
ITSPACE: Monotone Gaussian Optimal Transport Updates

Woojoo Na[1], Jennifer Dy[1]

[1]Northeastern University

ICML 2026

Motivation: Why Few-step Covariance Updates?



short update loop

$$X_0 \to X_1 \to \cdots \to X_K$$

K small, $r \ll d$

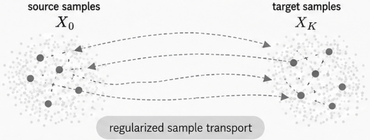
downstream module uses X_k

1. $X_k \geq 0$
2. $\mathrm{rank}(X_k) \leqslant r$
3. useful progress at each step

Short-budget covariance adaptation needs intermediate iterates that are already usable.

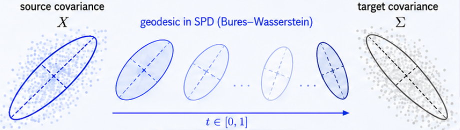
Problem: Why Exact Gaussian OT?

Sample-level OT / Sinkhorn



- **entropic regularization** (bias toward diffuse couplings)
- **sample-level coupling** (depends on sample size)
- not the exact covariance objective

Gaussian OT on covariances



- closed form for centered Gaussians
- covariance-level geometry (geodesic interpolation)
- OT-consistent (true Bures–Wasserstein / exact W_2)

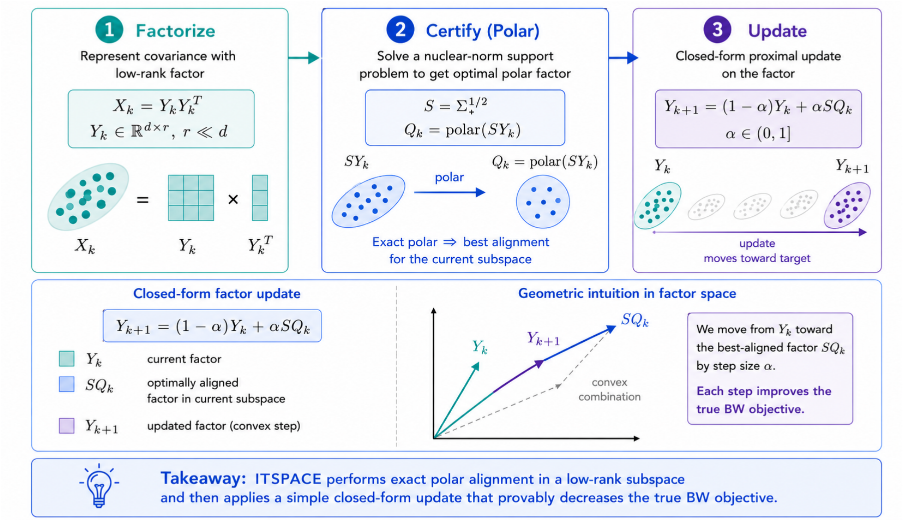
For centered Gaussian covariances, exact W_2 is available in closed form.

$$W_2^2(X, \Sigma) = \operatorname{tr}(X) + \operatorname{tr}(\Sigma) - 2 \operatorname{tr}\left(\left(\Sigma^{1/2} X \Sigma^{1/2}\right)^{1/2}\right)$$

1. PSD-valid
2. rank $\leq = r$
3. K small

Goal: make useful progress on the true BW objective in a few rank-constrained steps.

ITSPACE: Factorize, Certify, Update



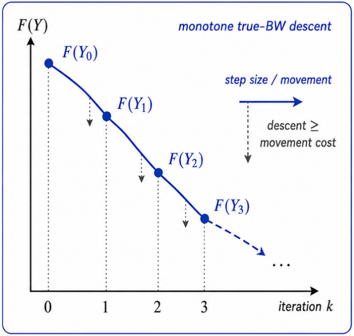
Guarantee: monotone descent of the true BW objective

Theorem 4.2: sufficient BW descent

$$F(Y_k) - F(Y_{k+1}) \geq \frac{1}{\alpha} \|Y_{k+1} - Y_k\|_F^2$$

$$\Rightarrow \quad W_2^2(X_{k+1}, \Sigma_\star) \leq W_2^2(X_k, \Sigma_\star)$$

Under exact polar computation,
each step decreases the true BW objective.

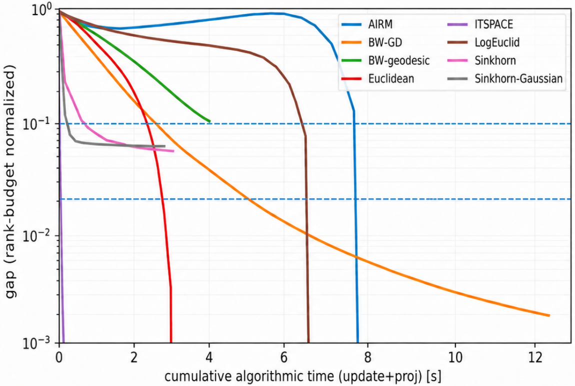


PSD-valid iterates
$\mathrm{rank}(X_k) \leq r$

monotone true BW descent
under exact polar

finite-budget control
of cumulative movement

Few-step Low-rank Optimization Results



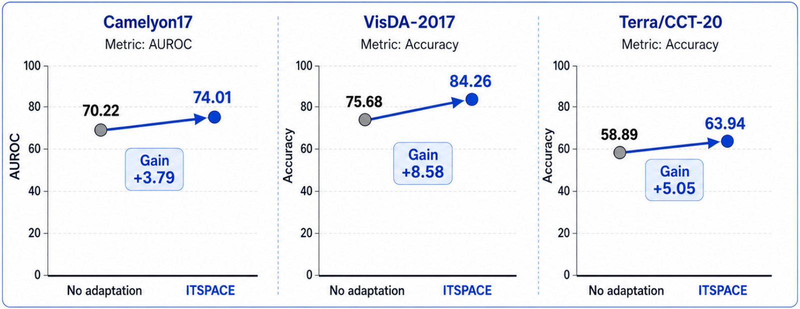
Time to reach gap ≤ 0.02

Dataset	ITSPACE (time)	BW-GD (time)	Speedup (×)
Camelyon17	0.0575s	13.8s	239.7x
VisDA-2017	0.0550s	13.2s	239.8x
Terra/CCT-20	0.0369s	11.5s	310.8x

Under the same rank and step budget, ITSPACE reaches the near-floor BW gap in tens of milliseconds.

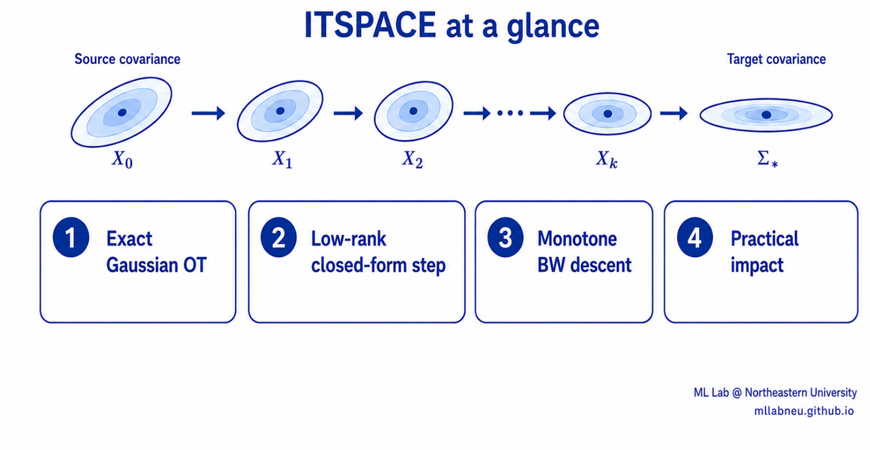
Downstream Task Results

CovDrift-MR: freeze the source head and adapt using unlabeled target moments.



ITSPACE improves downstream performance using only unlabeled target moments.

Conclusion


ITSPACE at a glance
Source covariance
Target covariance
X_0
X_1
X_2
X_k
Σ_*
1 Exact Gaussian OT
2 Low-rank closed-form step
3 Monotone BW descent
4 Practical impact
ML Lab @ Northeastern University
mllabneu.github.io