Trajectory-Matching Meta Pseudo-Labeling for Semi-Supervised Learning

Minh-Duc Le[1,2], Minh-Duong Nguyen[1], Dung D. Le[1]

[1]VinUniversity, [2]VNU University of Engineering and Technology

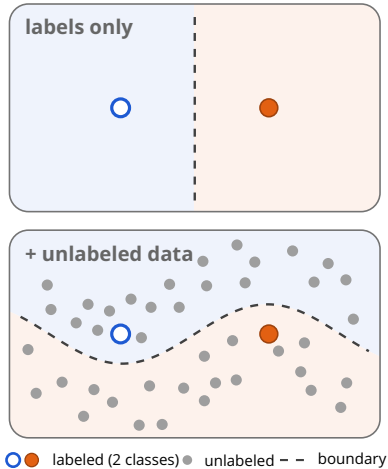
NeurIPS 2026 · Sydney, Australia · December 2026

1. Introduction: Semi-Supervised Learning

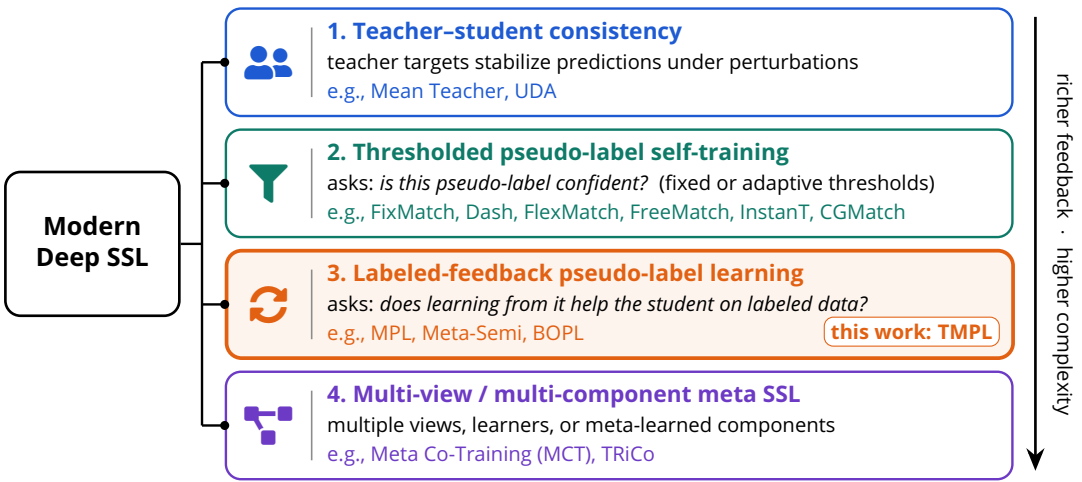
- **Motivation:**
  - Labels are costly, slow, or domain-dependent
  - Raw (unlabeled) data are abundant and cheap
  - e.g., driving around with a camera

- Can unlabeled points tell us where a **better decision boundary** lies?

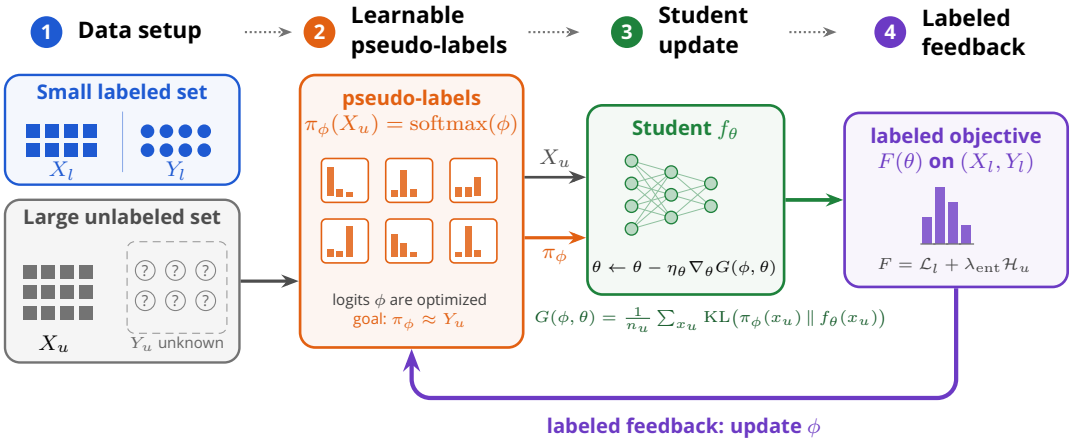
- Modern SSL must decide **which unlabeled supervision is useful** for generalization.



2. Related Work: Modern Deep SSL

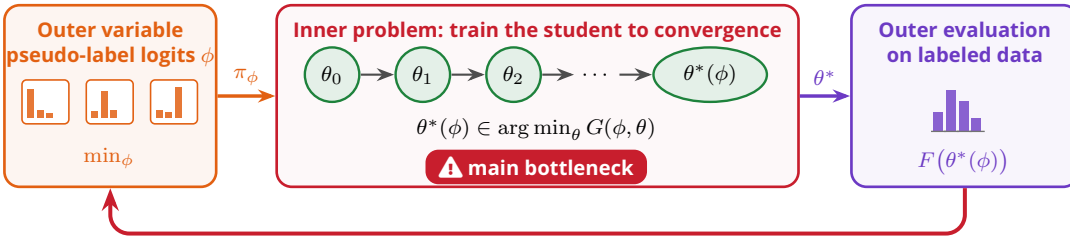


3. Problem Statement: Meta Pseudo-Labeling



$$\min_{\phi} \; F\big(\theta^*(\phi)\big) \quad \text{s.t.} \quad \theta^*(\phi) \in \arg\min_{\theta} G(\phi, \theta)$$

3. Problem Statement: The Main Difficulty



Exact bilevel

differentiate through the whole trajectory / argmin map ⇒ **intractable** for deep nets

One-step MPL (practical)

$\theta'(\phi) = \theta - \eta_\theta \nabla_\theta G(\phi, \theta)$; the teacher gradient needs the mixed $\nabla^2_{\phi\theta} G$ ⇒ **extra memory**, **noisy minibatch meta-gradients**

4. Methodology: TMPL in a Nutshell

We propose **Trajectory-Matching Meta Pseudo-Labeling (TMPL)**: a **first-order** alternative that extracts labeled-data feedback from **recent student optimization trajectories**.

- **Insight:** one-step meta pseudo-labeling is locally a **gradient-alignment** problem (Prop. 1).
- **Method:** match **finite differences** of the outer and pseudo-label objectives along **recent student displacements**; no differentiation through the student update.
- **Implementation:** a FIFO cache of $O(k)$ **detached scalars**; no student graphs, no past networks.
- **Theory:** small matching error controls the gradient mismatch within the span of recent displacements (Prop. 2), plus $O(T^{-1/2})$ stationarity.
- **Results:** improves MPL in **8/8** settings, best in **5/8**; -39.7% **peak memory** vs. MPL on CIFAR100.

4. Overview: From MPL to Trajectory Matching

1. One-step MPL

$\min_\phi F(\theta'(\phi))$

$\theta' = \theta - \eta_\theta \nabla_\theta G$

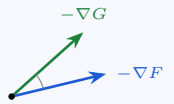
- needs $\nabla^2_{\phi\theta} G$
- graph kept
- extra memory
- noisy grads

2nd-order

2. Prop. 1: alignment

$\mathcal{F}_{\eta_\theta} \approx F - \eta_\theta \mathcal{A}$

$\mathcal{A} = \langle \nabla_\theta F, \nabla_\theta G \rangle$

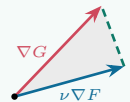


good labels: step descends F

$\nabla_\phi \mathcal{A}$: **2nd-order**

3. Gradient matching

$\min_{\nu>0} \|\nabla_\theta G - \nu \nabla_\theta F\|$



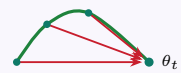
step $\approx$ rescaled $-\nabla_\theta F$

still 2nd-order

4. Finite differences

$v_i = \theta_t - \theta_{t-i}$

$\Delta_i^F \approx \langle \nabla_\theta F, v_i \rangle$



free directions from the optimizer

zeroth-order

5. TM loss + scalar cache

$\sum_i \left(\nu \Delta_i^F - \Delta_i^G\right)^2$

+ scale anchor



no graph, no past networks

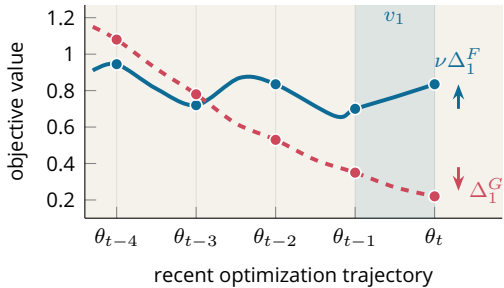
first-order, $O(k)$

Takeaway: **TMPL** replaces meta-gradient backpropagation through the student update with a cheap trajectory-based zeroth-order surrogate.

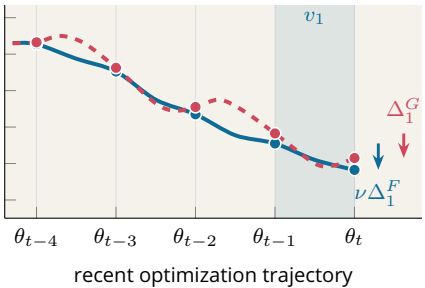
Intuition: Matching Descent Along the Trajectory

TM penalizes $\left(\nu\Delta_i^F - \Delta_i^G\right)^2$: pseudo-label descent ($G$) must follow outer-objective descent (F) along recent displacements $v_i = \theta_t - \theta_{t-i}$.

Without TM: G drops, F stalls



With TM: F and G descend together



$\nu F(\theta)$: outer (labeled) objective

$G(\phi, \theta)$: pseudo-label objective

4.1 Prop. 1: One-Step MPL ≈ Gradient Alignment

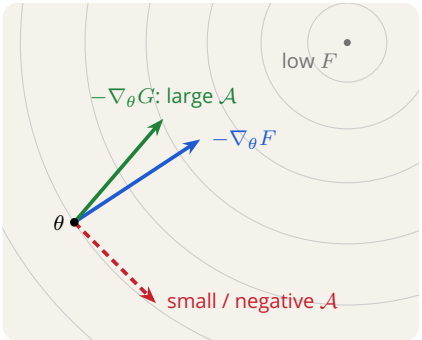
Alignment score:
$\mathcal{A}(\phi;\theta) := \langle \nabla_\theta F(\theta),\, \nabla_\theta G(\phi,\theta) \rangle$

Proposition 1

If F is L_F-smooth and $\|J_\theta \log f_\theta(x)\|_{\mathrm{op}} \le B_f$, then for every ϕ

$$\mathcal{F}_{\eta_\theta}(\phi;\theta) = F(\theta) - \eta_\theta\, \mathcal{A}(\phi;\theta) + O(\eta_\theta^2),$$

uniformly in ϕ: minimizing the one-step meta objective $\equiv$ **maximizing alignment**.



⚠ Ascending $\mathcal{A}$ directly is still second-order:
$\nabla_\phi \mathcal{A} = \left[\nabla^2_{\phi\theta} G(\phi,\theta)\right]^\top \nabla_\theta F(\theta)$.

4.2 From Gradient Matching to Finite Differences

Surrogate matching: make the student gradient a positive rescaling of the outer gradient

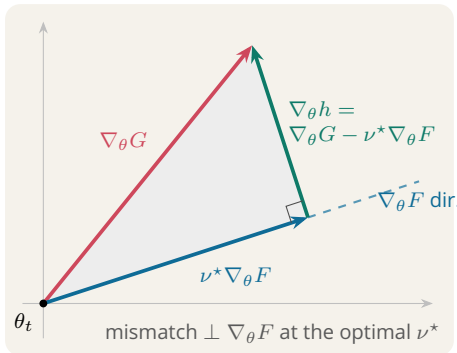
$$\min_{\nu>0} \left\| \nabla_\theta G(\phi,\theta) - \nu \nabla_\theta F(\theta) \right\|$$

⚠ matching gradients w.r.t. ϕ still needs $\nabla_\phi \nabla_\theta G$.

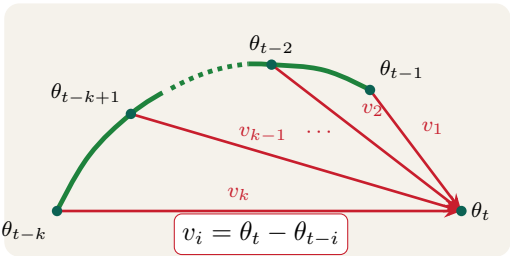
Zeroth-order proxy: for a small displacement v,

$$F(\theta) - F(\theta - v) = \langle \nabla_\theta F(\theta), v \rangle + O(\|v\|^2)$$
$$G(\phi,\theta) - G(\phi,\theta - v) = \langle \nabla_\theta G(\phi,\theta), v \rangle + O(\|v\|^2)$$

$\Rightarrow$ matching **function-value changes** along several directions matches the **directional derivatives**.



4.2 Reusing Recent Optimization Displacements



Instead of random probes, reuse the **last** k **student displacements**:

$$v_i := \theta_t - \theta_{t-i}, \qquad i = 1, \dots, k$$
$$\Delta_i^F := F(\theta_t) - F(\theta_{t-i})$$
$$\Delta_i^G := G(\phi, \theta_t) - G(\phi, \theta_{t-i})$$

- $\nu\Delta_i^F \approx \Delta_i^G$ for all i: $G(\phi, \cdot)$ behaves locally like a positive rescaling of F.
- Directions come **for free** from the optimizer.
- SGD trajectories tend to stay in **low-dimensional subspaces** (Gur-Ari et al. 2019; Li et al. 2023; Song et al. 2025).

4.2 The Trajectory-Matching (TM) Loss

Train the pseudo-labeler ϕ and a scale $\nu > 0$ by minimizing

$$\mathcal{L}_{\mathrm{TM}}(\phi, \nu) = \underbrace{\lambda_{\mathrm{mag}} \big(\nu F(\theta_t) - G(\phi, \theta_t)\big)^2}_{\text{magnitude anchor}} + \underbrace{\lambda_{\mathrm{dir}} \sum_{i=1}^{k} \left(\frac{\nu \Delta_i^F - \Delta_i^G}{\sqrt{\|v_i\|^2 + \varepsilon^2}} \right)^2}_{\text{directional matching}}$$

- **Magnitude anchor:** fixes the scale between F and G at the current iterate.
- **Directional term:** $G(\phi, \cdot)$ must change like νF along recent student motions: the alignment signal of Prop. 1.

Alternating first-order updates

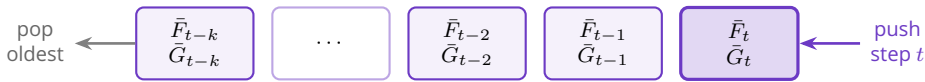
$\theta \leftarrow \theta - \eta_\theta \nabla_\theta G(\phi, \theta)$

$(\phi, \nu) \leftarrow (\phi, \nu) - \eta_\phi \nabla_{(\phi,\nu)} \mathcal{L}_{\mathrm{TM}}(\phi, \nu)$

$\theta_t, \ldots, \theta_{t-k}$ are **constants** in the teacher step: **no backprop through the student update**.

4.3 Scalar-Cache Implementation

FIFO queue $\mathcal{Q}_t$ of detached scalars ($\bar{F}_s = \mathrm{sg}[F(\theta_s)]$, $\bar{G}_s = \mathrm{sg}[G^{\mathrm{teacher}}(\phi_s, \theta_s)]$)



$$\widetilde{\mathcal{L}}_{\mathrm{TM}}(\phi, \nu) = \lambda_{\mathrm{mag}} \big(\nu \bar{F}_t - G_t(\phi)\big)^2 + \lambda_{\mathrm{dir}} \sum_{i=1}^{k_t} \left(\frac{\nu(\bar{F}_t - \bar{F}_{t-i}) - \big(G_t(\phi) - \bar{G}_{t-i}\big)}{i} \right)^2$$

$G_t(\phi) := G^{\mathrm{teacher}}(\phi, \theta_t)$ is the only live term; the denominator i is a temporal normalization replacing $\|\theta_t - \theta_{t-i}\|$ (practical online approximation).

MPL / RawMPL
student-update graph
or frozen students

Idealized TM
past student parameters: $O(kd)$

TMPL scalar cache
$O(k)$ scalars, first-order

4.4 Prop. 2: TM Controls the Gradient Mismatch

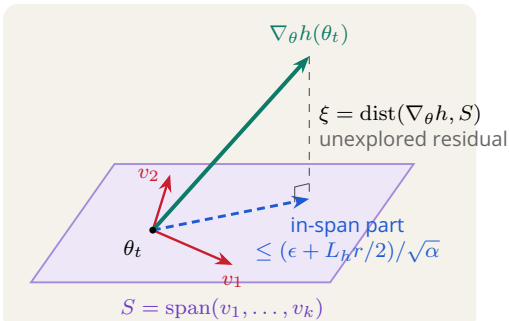
$h(\theta) := G(\phi, \theta) - \nu F(\theta)$, so
$\nabla_\theta h(\theta_t) = \nabla_\theta G(\phi, \theta_t) - \nu \nabla_\theta F(\theta_t)$ is the mismatch.

$u_i = v_i / \|v_i\|$, $r = \max_i \|v_i\|$, $S = \operatorname{span}(v_1, \ldots, v_k)$,
$\alpha = \lambda_{\min}\big(\frac{1}{k} \sum_i u_i u_i^\top |_S\big)$,
$\xi = \operatorname{dist}\big(\nabla_\theta h(\theta_t), S\big)$, TM error
$\epsilon = \max_i |h(\theta_t) - h(\theta_t - v_i)| / \|v_i\|$.

Proposition 2

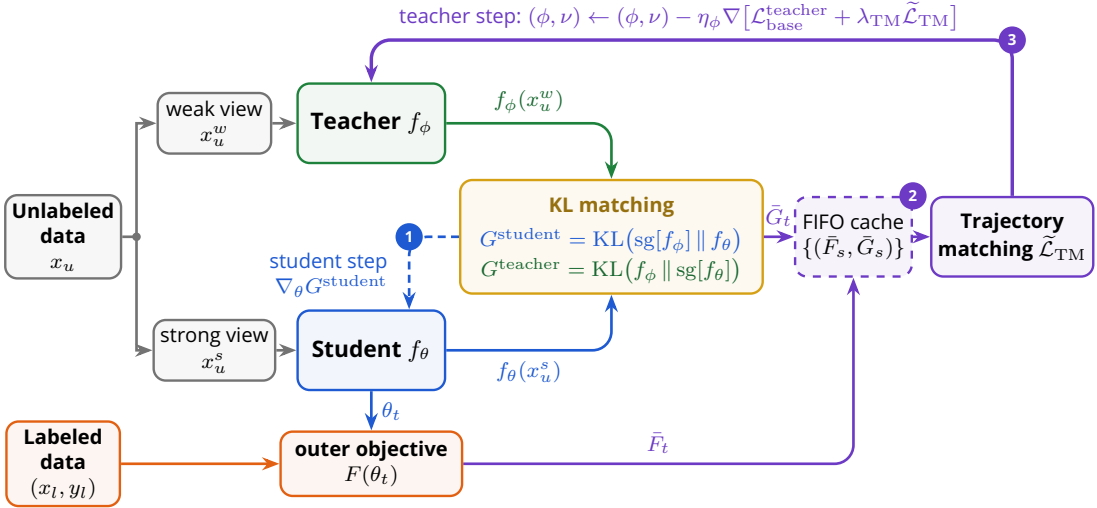
If $\nabla_\theta h$ is L_h-Lipschitz near the segments $\{\theta_t - s v_i\}_{s \in [0,1]}$, then

$$\big\|\nabla_\theta h(\theta_t)\big\| \le \sqrt{\frac{(\epsilon + L_h r/2)^2}{\alpha} + \xi^2}.$$



In-span error: TM error ϵ, radius r, conditioning α.
Residual ξ**:** directions the trajectory has not explored.

4.5 Practical Teacher–Student Instantiation



$\mathcal{L}_{\text{base}}^{\text{teacher}}$: e.g., supervised + FixMatch-style consistency. No gradient ever flows through the student update.

4.6 Convergence Guarantee (Idealized TMPL)

Theorem 1 (coupled stationarity)

Full-batch, label-space TMPL; $\eta_{\theta,t} \asymp t^{-1/4}$, $\eta_{\phi,t} \asymp t^{-1/2}$, square-summable $\lambda_{\mathrm{mag},t}$:

$$\frac{1}{T}\sum_{t=0}^{T-1}\Big(\|\nabla_\theta G(\phi_t,\theta_t)\|^2 + \|\nabla_z \Phi_t(z_t)\|^2\Big) = O\big(T^{-1/2}\big), \qquad z=(\phi,\nu)$$

Theorem 2 (outer-stationarity certificate)

If the retained directions are well-conditioned ($\alpha_t \geq \underline{\alpha} > 0$):

$$\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla_\theta F(\theta_t)\|^2 \leq O\big(T^{-1/2}\big) + O\Big(\underbrace{\tfrac{1}{T}\textstyle\sum_t \Phi^{\mathrm{dir}}_{t,\mathrm{raw}}(z_t)}_{\text{TM error}}\Big) + O\Big(\underbrace{\tfrac{1}{T}\textstyle\sum_t \xi_t^2}_{\text{unexplored residual}}\Big)$$

Interpretation: small TM error + good coverage $\Rightarrow$ approximate stationarity of F (a local trajectory-subspace certificate).

5.1 Experiments: Setup

Benchmarks (number of labels)

- CIFAR10 (250 / 4000), SVHN (250 / 1000)
- CIFAR100 (2500 / 10000), STL10 (250 / 1000)

Protocol (unified USB codebase)

- WRN-28-2 (CIFAR, SVHN), WRN-VAR-37-2 (STL10)
- SGD, lr 0.03, momentum 0.9, cosine decay, $T = 2^{20}$
- 64 labeled + 7×64 unlabeled per batch
- top-1 error, mean $\pm$ std over 3 seeds

TMPL (same hyperparameters everywhere)

- $k = 100$, $\lambda_{\mathrm{TM}} = 1$, $\lambda_{\mathrm{mag}} = 1/(t+1)$, $\lambda_{\mathrm{dir}} = 1$
- scale ν learned, projected to $[0.1, 100]$

Baselines

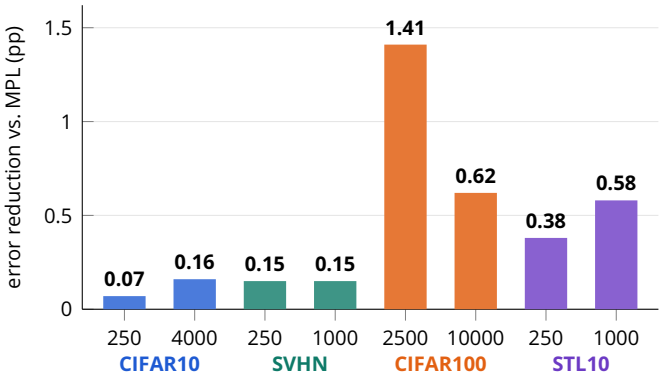
- PseudoLabel, MeanTeacher, MixMatch, FixMatch, FlexMatch, SoftMatch, FreeMatch, FullFlex, CGMatch
- **MPL**: one-step meta pseudo-labels
- **RawMPL**: exact gradient of the soft-label one-step meta objective (autograd/VJP)

5.2 Main Results: Top-1 Error (%) ↓

Method	CIFAR10		SVHN		CIFAR100		STL10	
	250	4000	250	1000	2500	10000	250	1000
PseudoLabel	48.28 ±2.10	14.90 ±0.20	16.60 ±1.13	9.33 ±0.58	59.09 ±0.61	38.86 ±0.09	49.21 ±1.54	34.30 ±0.78
MeanTeacher	56.06 ±2.03	15.47 ±0.43	25.10 ±3.17	12.29 ±0.45	61.13 ±0.57	39.05 ±0.12	44.51 ±2.02	33.39 ±1.24
MixMatch	37.28 ±0.61	7.38 ±0.06	3.71 ±0.20	3.12 ±0.09	49.58 ±0.62	32.10 ±0.13	42.34 ±0.99	25.12 ±0.71
FixMatch	4.97 ±0.05	4.29 ±0.10	**1.99 ±0.05**	2.00 ±0.04	34.29 ±0.43	28.28 ±0.18	9.81 ±1.10	6.61 ±0.36
FlexMatch	5.00 ±0.05	4.28 ±0.05	5.02 ±1.20	5.43 ±0.46	33.35 ±0.33	27.12 ±0.04	8.71 ±0.17	6.17 ±0.17
SoftMatch	4.84 ±0.10	4.27 ±0.12	2.15 ±0.05	2.08 ±0.04	33.05 ±0.05	27.26 ±0.03	**8.01 ±0.32**	**5.89 ±0.20**
FreeMatch	4.85 ±0.10	4.14 ±0.19	4.09 ±0.66	4.31 ±0.00	32.79 ±0.21	27.17 ±0.11	8.78 ±0.88	6.07 ±0.28
FullFlex	4.98 ±0.04	4.21 ±0.10	4.63 ±1.26	4.18 ±0.08	32.85 ±0.10	26.58 ±0.03	8.83 ±0.36	6.26 ±0.12
CGMatch	4.80 ±0.07	–	2.14 ±0.05	–	32.51 ±0.07	–	9.48 ±0.20	–
MPL	4.84 ±0.12	4.24 ±0.04	2.18 ±0.15	2.12 ±0.04	33.79 ±0.51	27.08 ±0.11	9.11 ±0.29	6.76 ±0.02
TMPL (ours)	**4.77 ±0.08**	**4.08 ±0.05**	2.03 ±0.07	**1.97 ±0.03**	**32.38 ±0.21**	**26.46 ±0.08**	8.73 ±0.19	6.18 ±0.07

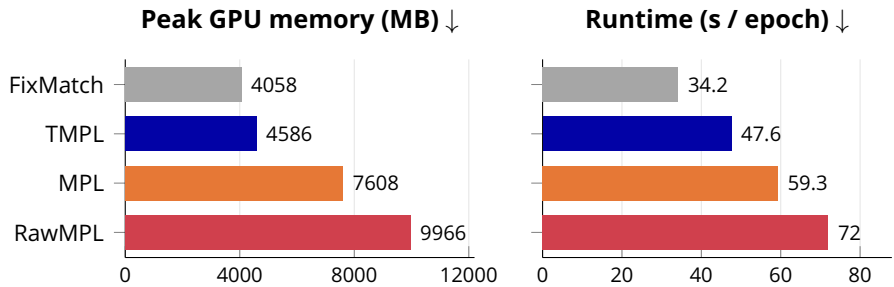
Mean ± std over 3 seeds; **bold** = best, underline = second best. TMPL improves MPL in **8/8** settings and is best in **5/8**.

5.2 TMPL vs. MPL: Gains in All Eight Settings



- **8/8** settings improved over MPL; best in **5/8**.
- Largest on **CIFAR100**: $33.79 \rightarrow 32.38$ (2500), $27.08 \rightarrow 26.46$ (10000).
- Small but consistent on CIFAR10 and SVHN.
- STL10: SoftMatch stays strongest; TMPL is not uniformly dominant.

5.3 Memory and Runtime (CIFAR100, 2500 labels)



- TMPL needs only **1.13**× FixMatch memory.
- -39.7% vs. MPL, -54.0% vs. RawMPL.

- MPL / RawMPL are **1.25**× **/ 1.51**× **slower**.
- No student-update graph or frozen students: only $O(k)$ scalars.

5.4 Ablation Study (CIFAR100, 2500 labels)

Trajectory-matching components

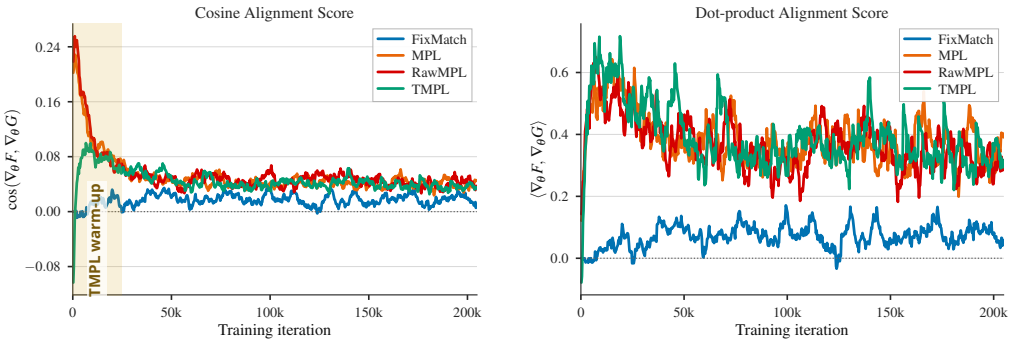
Configuration	Top-1 error (%) ↓
$\lambda_{\text{TM}} = 0$ (no TM)	34.29 ± 0.43
$\lambda_{\text{mag}} = 0,\ \lambda_{\text{dir}} = 1$	32.63 ± 0.75
$\lambda_{\text{mag}} = 1,\ \lambda_{\text{dir}} = 1$	$\mathbf{32.38 \pm 0.21}$

Averaging RawMPL meta-gradients

Method	Top-1 error (%) ↓
MPL	33.79 ± 0.51
RawMPL, $M = 1$	33.45 ± 0.54
RawMPL, $M = 2$	33.28 ± 0.29
RawMPL, $M = 4$	33.21 ± 0.25
TMPL	$\mathbf{32.38 \pm 0.21}$

- The **directional** term gives most of the gain ($34.29 \rightarrow 32.63$); the magnitude anchor adds more and cuts variance ($\pm 0.75 \rightarrow \pm 0.21$).
- Averaging M RawMPL meta-gradients helps only modestly ($33.45 \rightarrow 33.21$): TMPL is not just a cheaper RawMPL but a **more stable first-order surrogate**.

5.5 TMPL Recovers MPL-like Gradient Alignment



CIFAR100 (2500 labels), first 200k iterations: cosine (left) and dot-product (right) alignment between $\nabla_\theta F$ and $\nabla_\theta G$.

Takeaway: MPL / RawMPL align immediately (they differentiate through the student step); TMPL needs a **short warm-up**, then reaches **comparable alignment** without any meta-gradient. FixMatch stays near zero.

6. Conclusion

- **Takeaways:**
  - One-step meta pseudo-labeling $\approx$ gradient alignment of $\nabla_\theta F$ and $\nabla_\theta G$.
  - TMPL matches finite differences along recent displacements: first-order, $O(k)$ cache.
  - Improves MPL in 8/8 settings (best in 5/8), 39.7% less peak memory than MPL.
- **Limitations:**
  - Signal only along recent directions; with poor coverage the residual ξ stays large.
  - Theory covers an idealized full-batch, label-space dynamic.
  - Not uniformly best (e.g., SoftMatch on STL10).
- **Future work:**
  - Better trajectory buffers and direction selection.
  - Larger benchmarks, stronger backbones, dense prediction, other bilevel SSL.