The Ideal Continual Learner: An Agent That Never Forgets

Liangzu Peng[1] Paris V. Giampouras[1] René Vidal[2,3]

[1]Johns Hopkins University [2]University of Pennsylvania [3]NORCE Norwegian Research Centre, Norway

Continual Learning: Problem Setup

$$\mathcal{G}_t \coloneqq \operatorname*{argmin}_{w \in \mathcal{W}} L_t(w; D_t)$$

- L_t: Loss function of task t
- D_t: Data for task t
- $\mathcal{W}$: Search space
- $\mathcal{G}_t$: The set of global minimizers
- Tasks $1, \ldots, T$ are presented sequentially
- **Goal**: learn a model $\hat{w}_T$ that solves all tasks
- **Challenge** ("Catastrophic Forgetting"): model $\hat{w}_T$ may perform poorly on previous tasks

Methods to Prevent Forgetting

- *Regularization-based*, e.g.,

$$\min_{w \in \mathcal{W}} L_t(w; D_t) + \delta \cdot \left|\left| w - \hat{w}_{t-1} \right|\right|_2$$

- *Memory-based*, e.g., rehearsal: train with current data and part of previous data
- *Expansion-based*: tasks ↑ ⇒ parameters ↑, only train on new parameters

These methods greatly improved empirical performance in the deep learning context!

Theory that can explain their empirical success has been few and far between!!!

The Ideal Continual Learner (ICL)

With $\mathcal{K}_t \coloneqq \mathcal{W}$, **ICL** is a method that solves

$$\mathcal{K}_t \coloneqq \operatorname*{argmin}_{w \in \mathcal{K}_{t-1}} L_t(w; D_t)$$

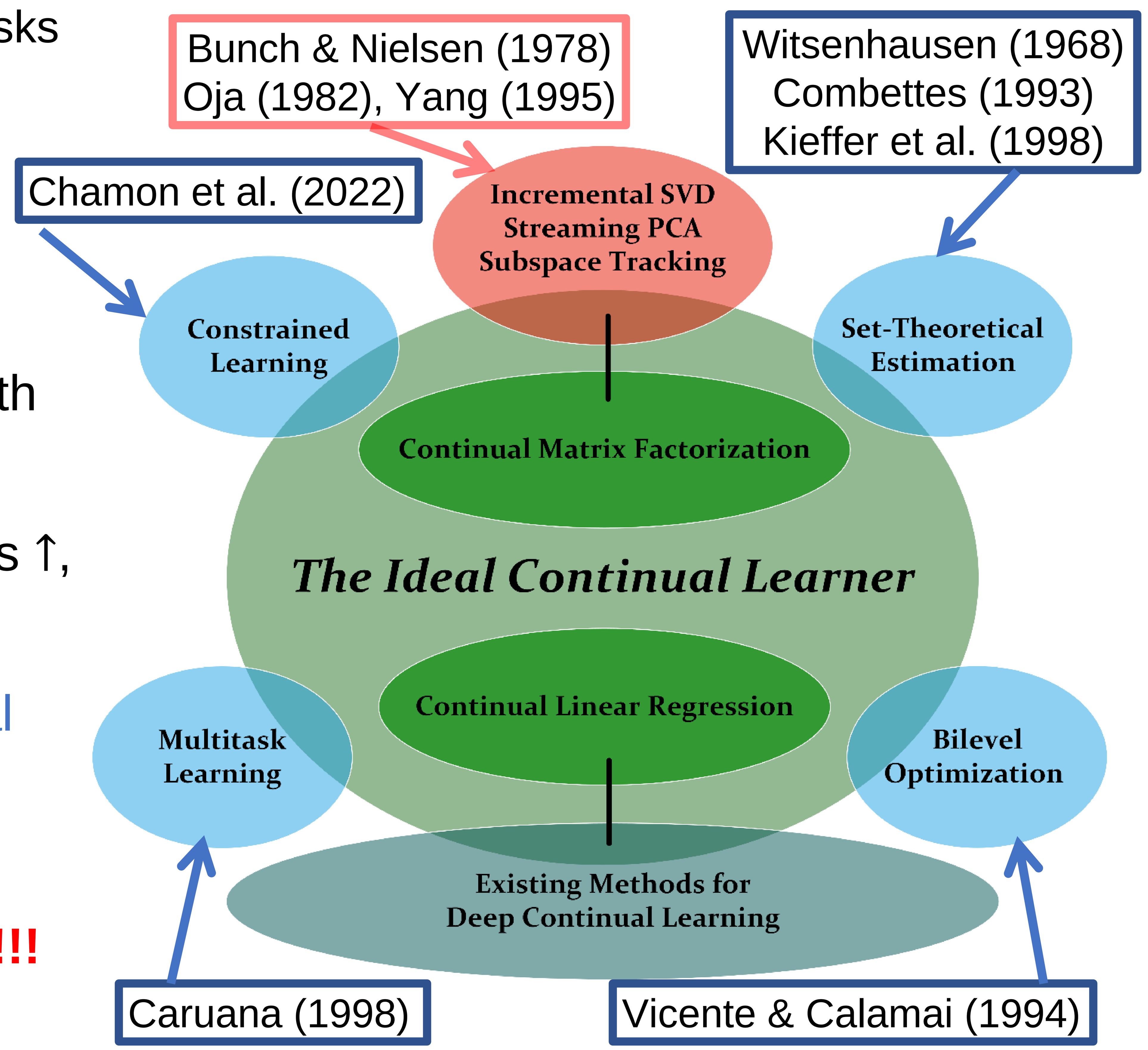
sequentially, for $t = 1, 2, \ldots, T$.

ICL Never Forgets by Design

Suppose $\cap_{t=1}^{T} \mathcal{G}_t \neq \emptyset$. Then:

✓ **ICL** is **sufficient** to prevent forgetting

✓ The storage consumption of **ICL** is **minimal** to prevent catastrophic forgetting



Take-Away Messages from ICL

History & Future (Figure in the middle)

Connections of **ICL** to other topics shed light on historical remarks & research avenues

All Roads Lead to Rome

ICL can be viewed as *a regularization-based*, *memory-based* (*projection-based*), or *expansion-based* method. Put differently, these methods, designed via engineering insights, are all **Ideal Continual Learners.**

Rehearsal is an Ideal Continual Learner

We prove the **first generalization guarantee for rehearsal**, a highly performant memory-based method

Wider Networks ⇒ Less Forgetting

By analyzing **ICL**, we rigorously show, **for the first time**, that wider neural networks forget less catastrophically.

Conclusion & Limitation

- **ICL** is the first general theoretical framework that never forgets by design
- Implementing ICL is challenging, but approximating it is possible

Acknowledgements: Work supported by the project ULEARN "Unsupervised Lifelong Learning" and co-funded under the grant number 316080 of the Research Council of Norway.