Statistical Consistency of Ranking Methods in A Rank-Differentiable Probability Space

Yanyan Lan[1], Jiafeng Guo[1], Xueqi Cheng[1], Tie-Yan Liu[2]

[1]Institute of Computing Technology, Chinese Academy of Sciences & [2]Microsoft Research Asia

{lanyanyan, guojiafeng, cxq}@ict.ac.cn, Tie-Yan.Liu@microsoft.com

1. MOTIVATION

- Statistical Consistency of Ranking Methods

  Surrogate Loss → Pairwise Ranking Methods —Consistency→ Weighted Pairwise Disagreement Loss ← True Loss
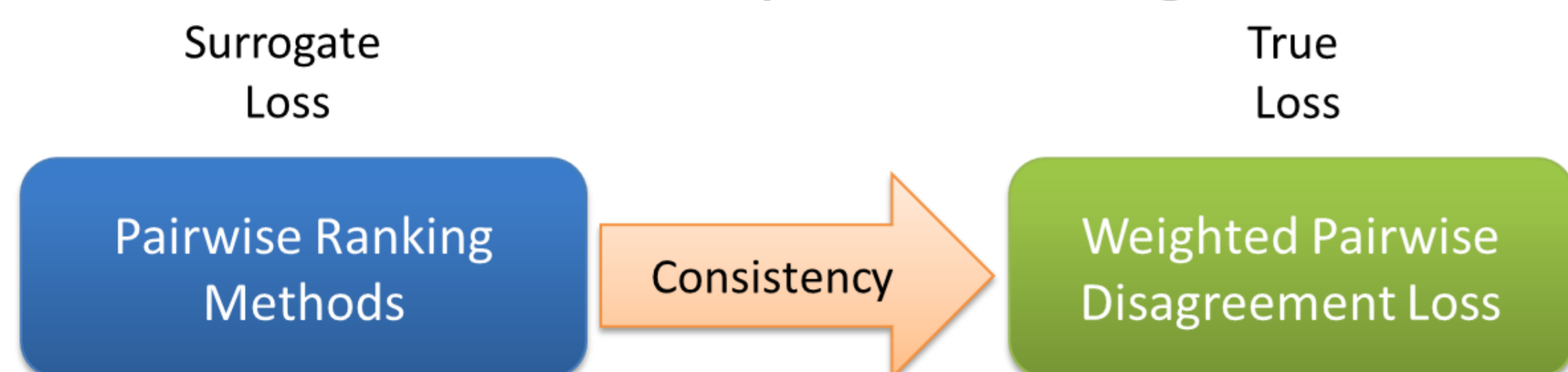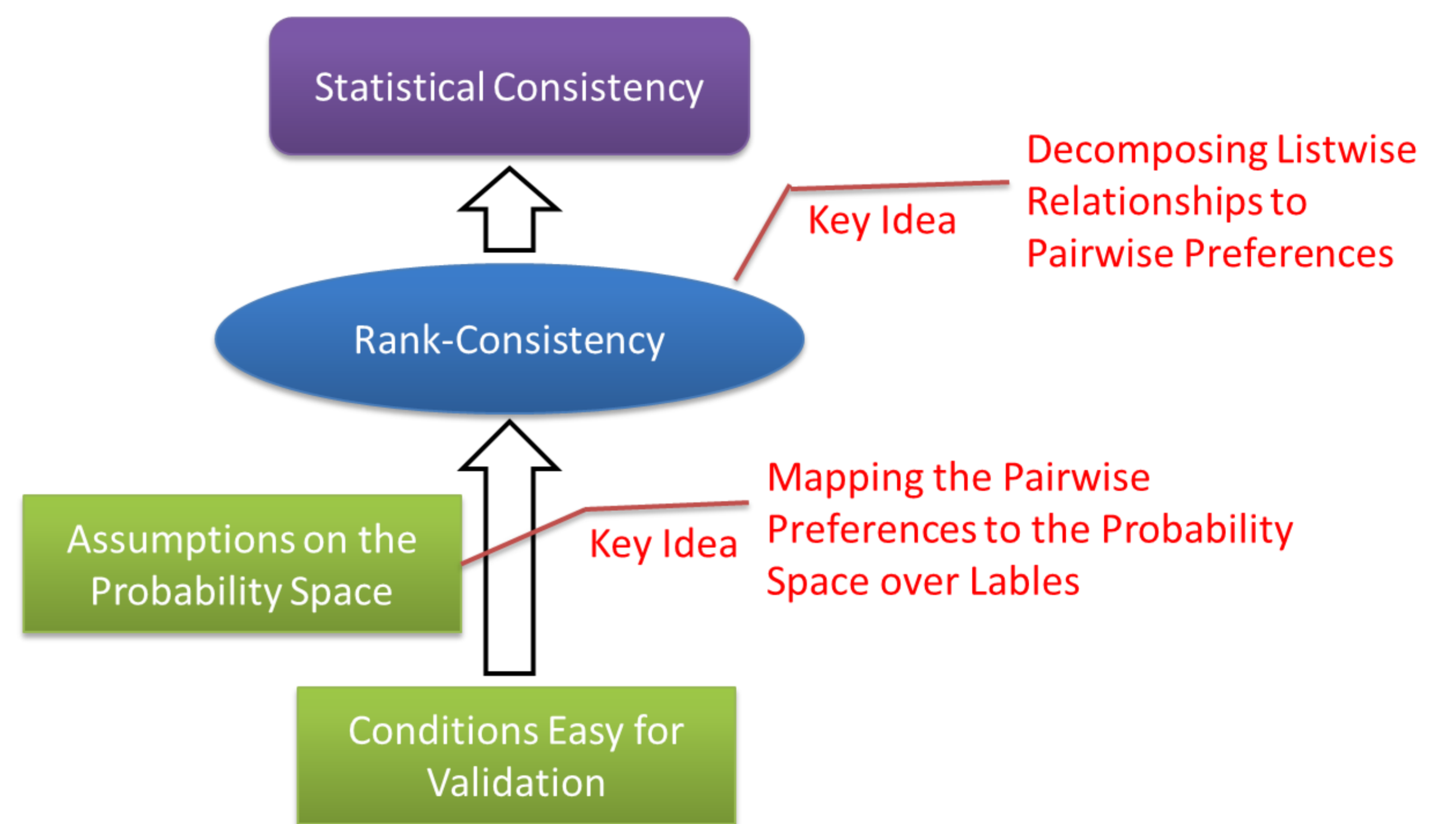
  — Theoretically, not consistent, even in low-noise setting [Duchi et. al. ICML 2010].

    Contradiction!

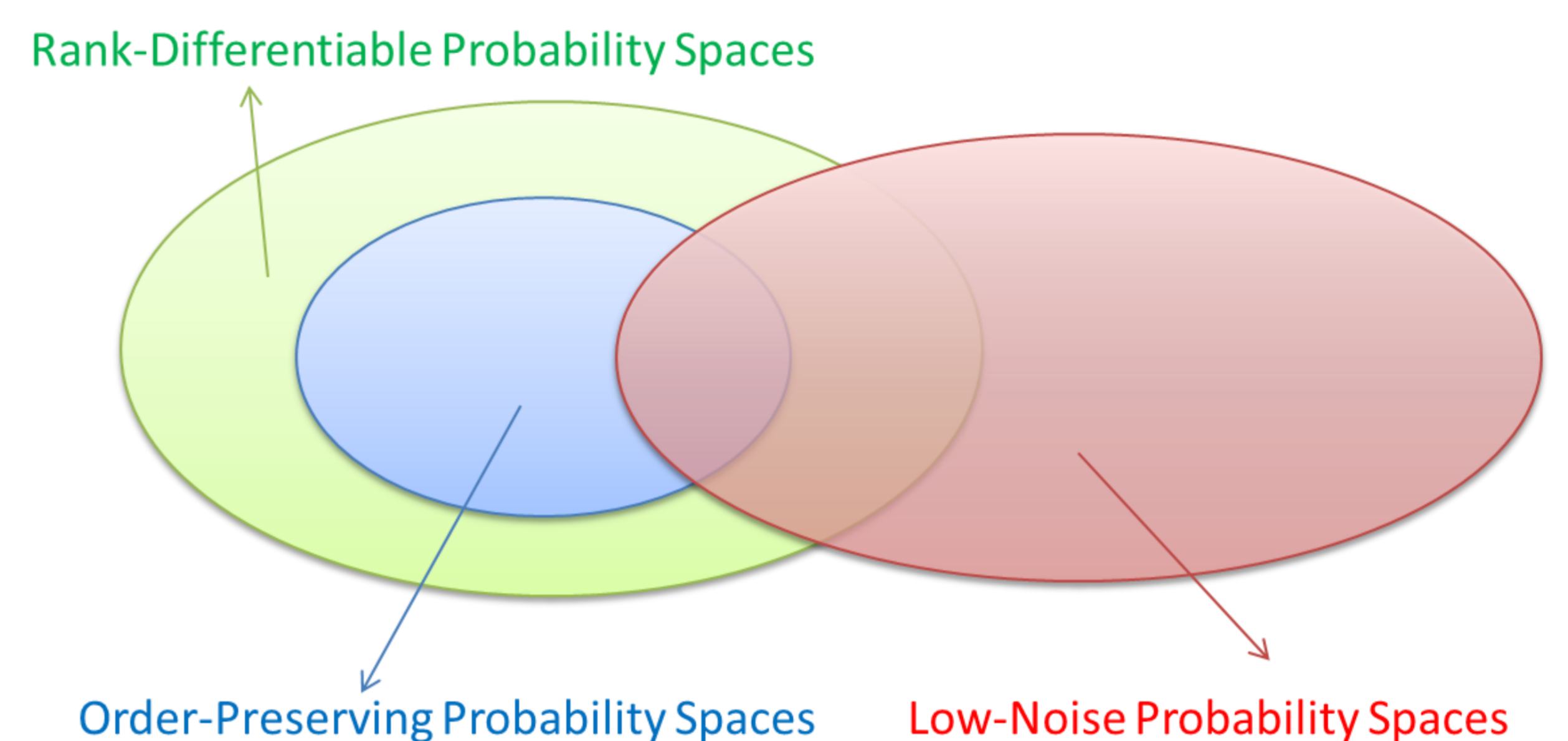 — Practically, very effective, still the-state-of-the-art.

2. KEY IDEA

Statistical Consistency

↑ — Key Idea — Decomposing Listwise Relationships to Pairwise Preferences

Rank-Consistency

↑ — Key Idea — Mapping the Pairwise Preferences to the Probability Space over Lables

Assumptions on the Probability Space

Conditions Easy for Validation

3. RANK-DIFFERENTIABLE PROBABILITY SPACE

Rank-Differentiable Probability Space

Dual Ratings

$$\begin{pmatrix} r_1 \\ \vdots \\ r_i \\ \vdots \\ r_j \\ \vdots \\ r_n \end{pmatrix} \quad \begin{pmatrix} r_1 \\ \vdots \\ r_j \\ \vdots \\ r_i \\ \vdots \\ r_n \end{pmatrix}$$

With the Same Preference in the Probability Space

Separability on Pairs

the optimal function will induce the same preference relationship

4. RELATIONSHIPS WITH PREVIOUS WORK

Rank-Differentiable Probability Spaces

Order-Preserving Probability Spaces   Low-Noise Probability Spaces
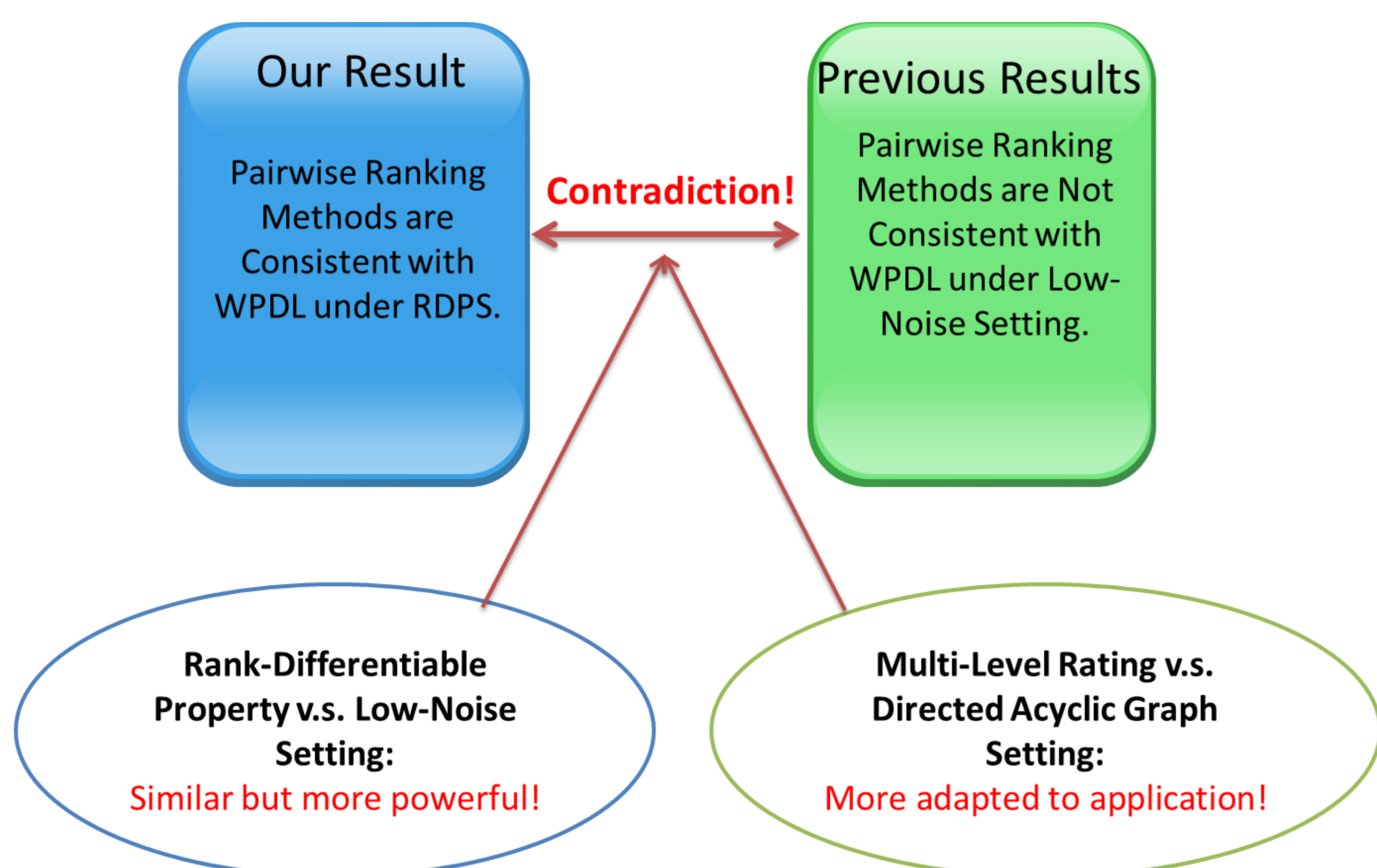
➤ Rank-Differentiable property is much weaker than order-preserving property.
➤ Rank-Differentiable property is very similar to low-noise setting.

5. WHY CONTRADICTION

Our Result

Pairwise Ranking Methods are Consistent with WPDL under RDPS.

Contradiction!

Previous Results

Pairwise Ranking Methods are Not Consistent with WPDL under Low-Noise Setting.

Rank-Differentiable Property v.s. Low-Noise Setting: Similar but more powerful!

Multi-Level Rating v.s. Directed Acyclic Graph Setting: More adapted to application!

6. CONCLUSIONS

Conclusions
- The previous results on the inconsistency of commonly-used pairwise ranking methods are not conclusive, depending on the assumptions about the probability space.
- We prove that the pairwise ranking methods are consistent with WPDL under the rank-differentiable assumption.
- Pairwise ranking methods are still good choices since their effective performance in application.

Future Work
- Statistical consistency of ranking methods with respect to the IR evaluation measures.
- Statistical consistency of listwise ranking methods.