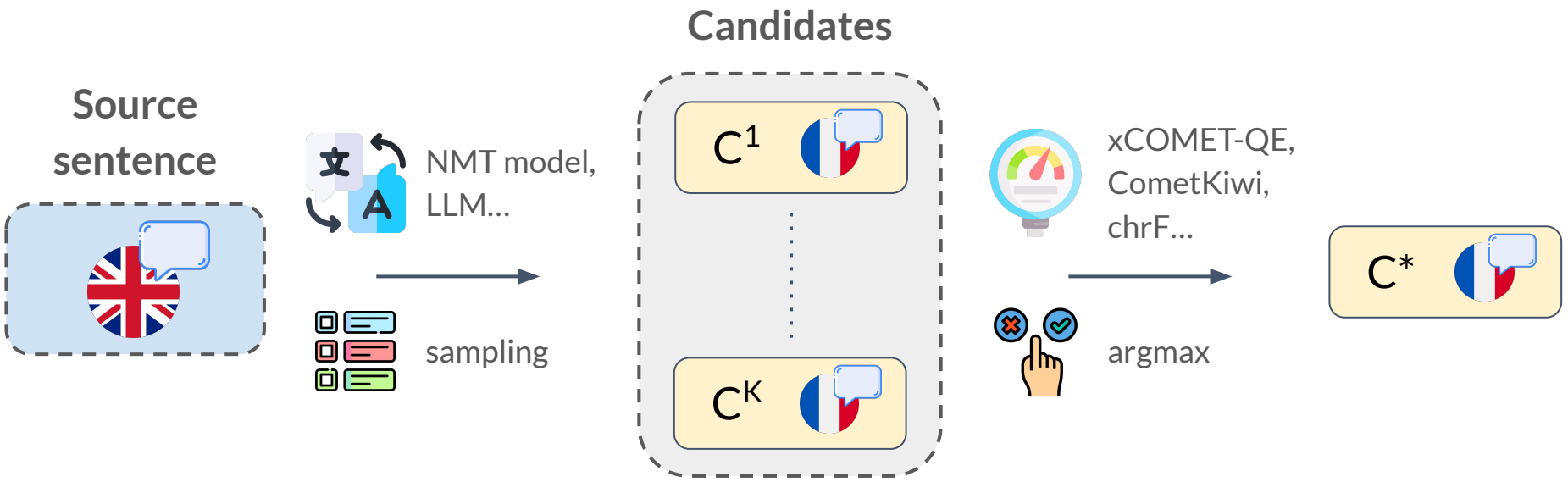
Is Preference Alignment Always the Best Option to Enhance LLM-Based Translation? An Empirical Analysis

NINTH CONFERENCE IN MACHINE TRANSLATION
NOVEMBER 2024

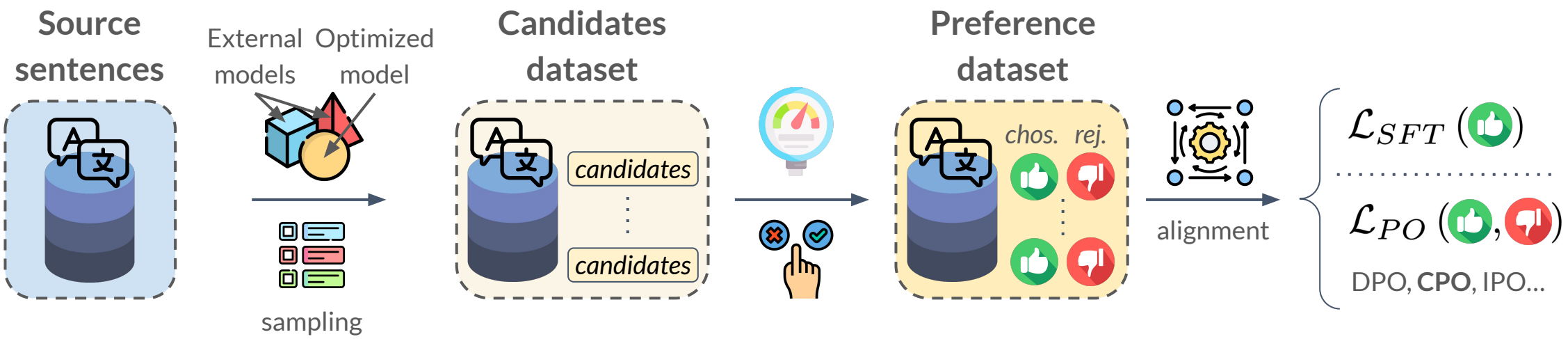
Motivations

- **Likelihood-based decoding** often falls short of matching human preferences, especially in MT
- Intuitive approach to maximize translation quality at inference time: **quality-aware decoding**



Motivations (cont'd)

- **Alignment techniques for MT** have arised for 2 main reasons:
  - Success of such approaches for general purpose models
  - Great metrics/evaluators that can be used to get preferences/rewards, but are expensive to use during decoding

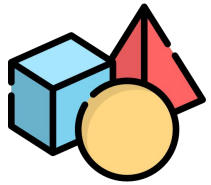


What is the best way to align a model for translation?

Experimental setup

Alignment data

- Subset of FLORES-200
- Languages: en, cs, de, is, ru, zh
- Directions: en-xx and xx-en



Translation systems

- Optimized model: *ALMA-13B-LoRA*
- Others systems: GPT-4, gold reference

Metrics

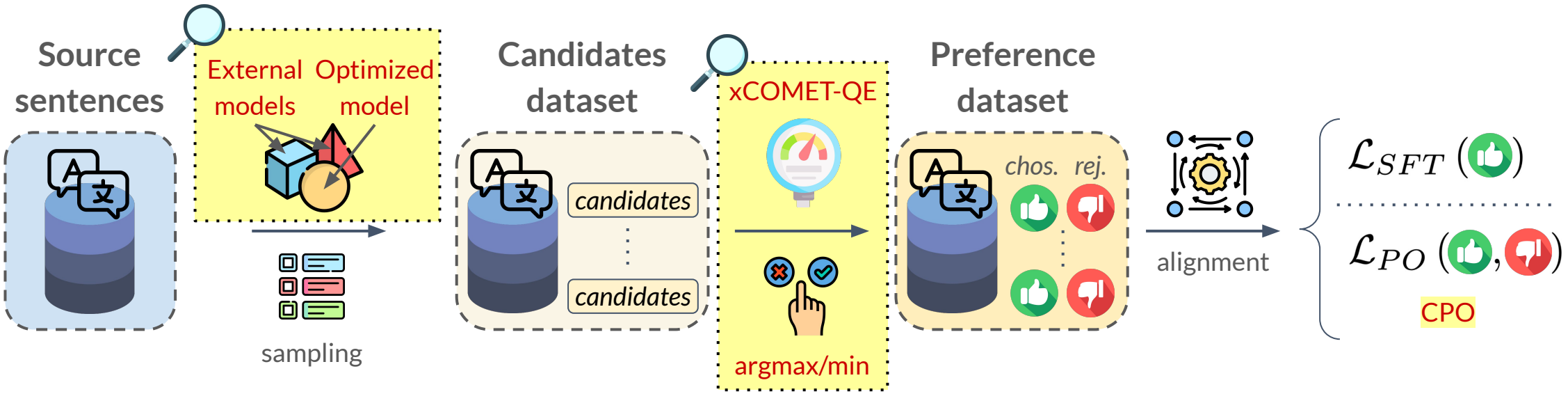
- Neural: xCOMET-QE, CometKiwi
- Lexical: chrF

Evaluation data

- WMT'22 (en, cs, de, is, ru, zh)
- WMT'23 (en, cs, de, ru, zh)

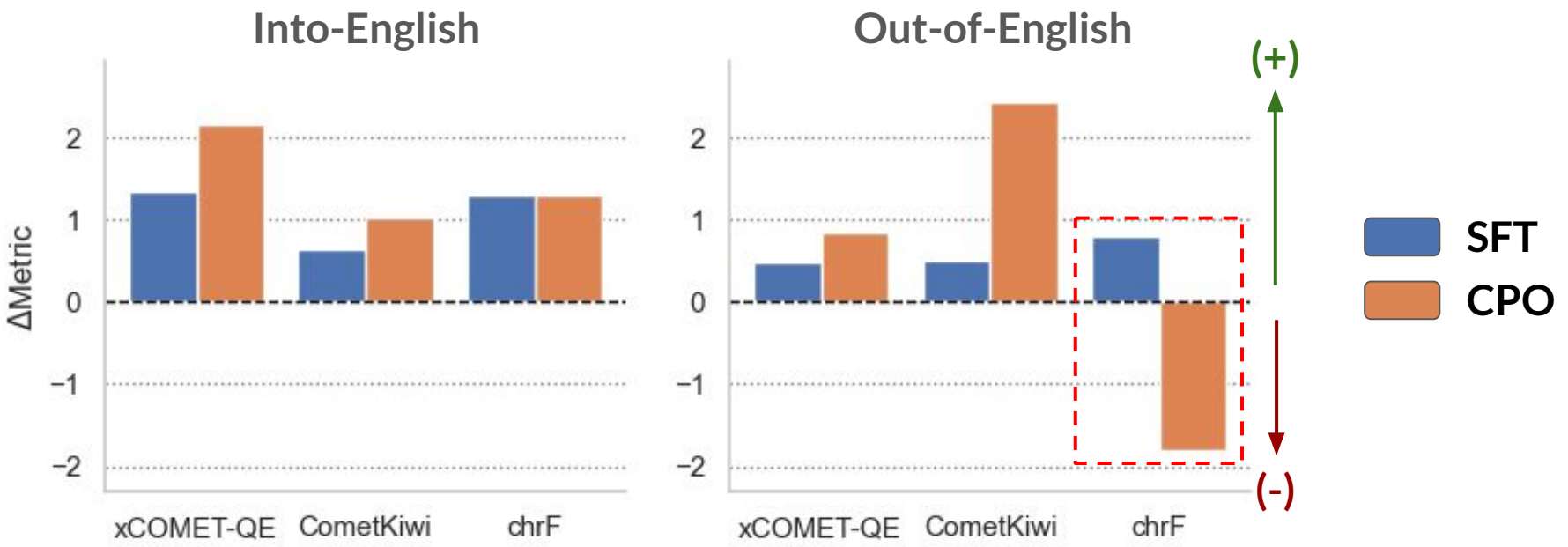
General results in the multi-system setting

Aligning on neural metrics



General results in the multi-system setting

Aligning on neural metrics

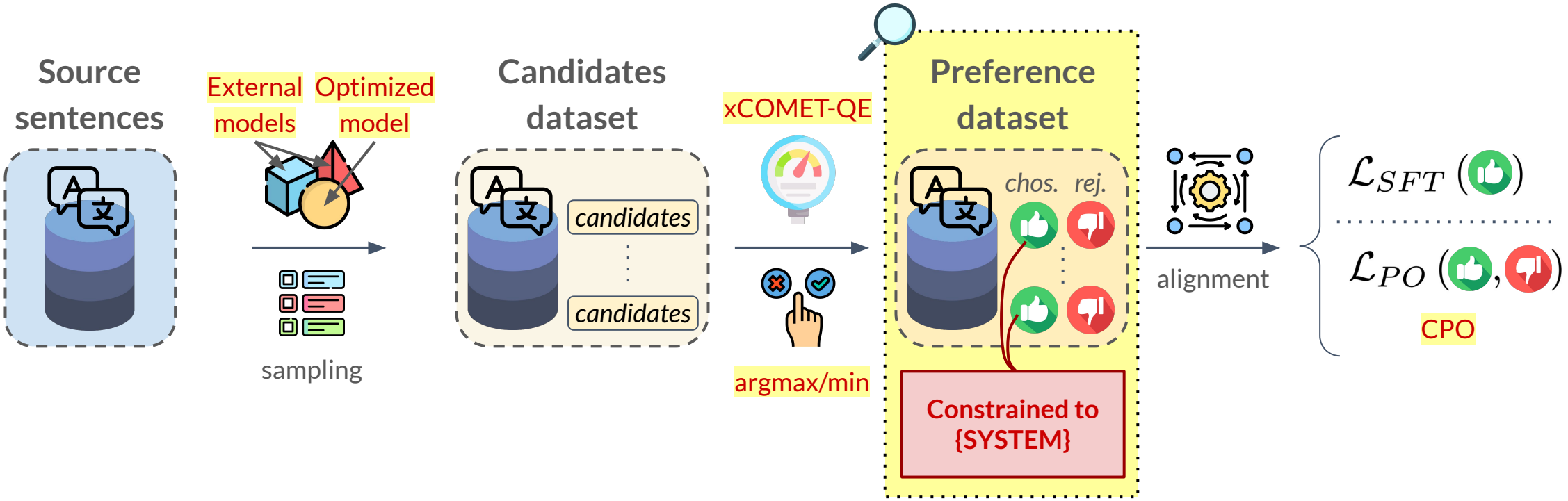


- **CPO outperforms SFT** on neural metrics when aligning on neural metrics

- CPO induces **adverse lexical effects** for out-of-English translations

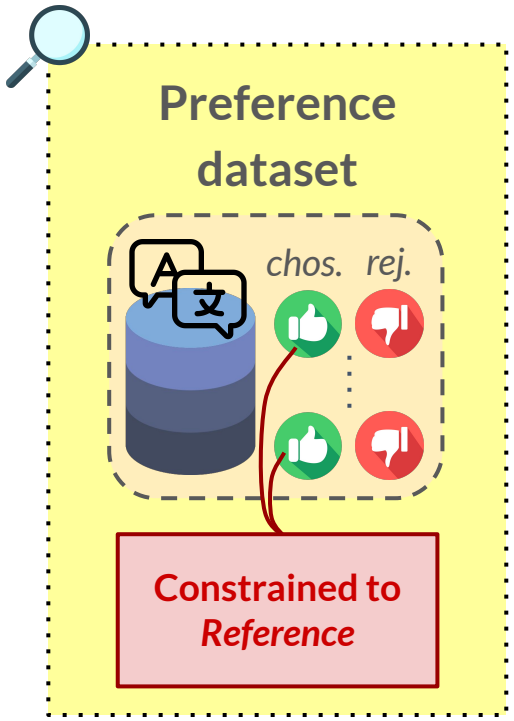
General results in the multi-system setting

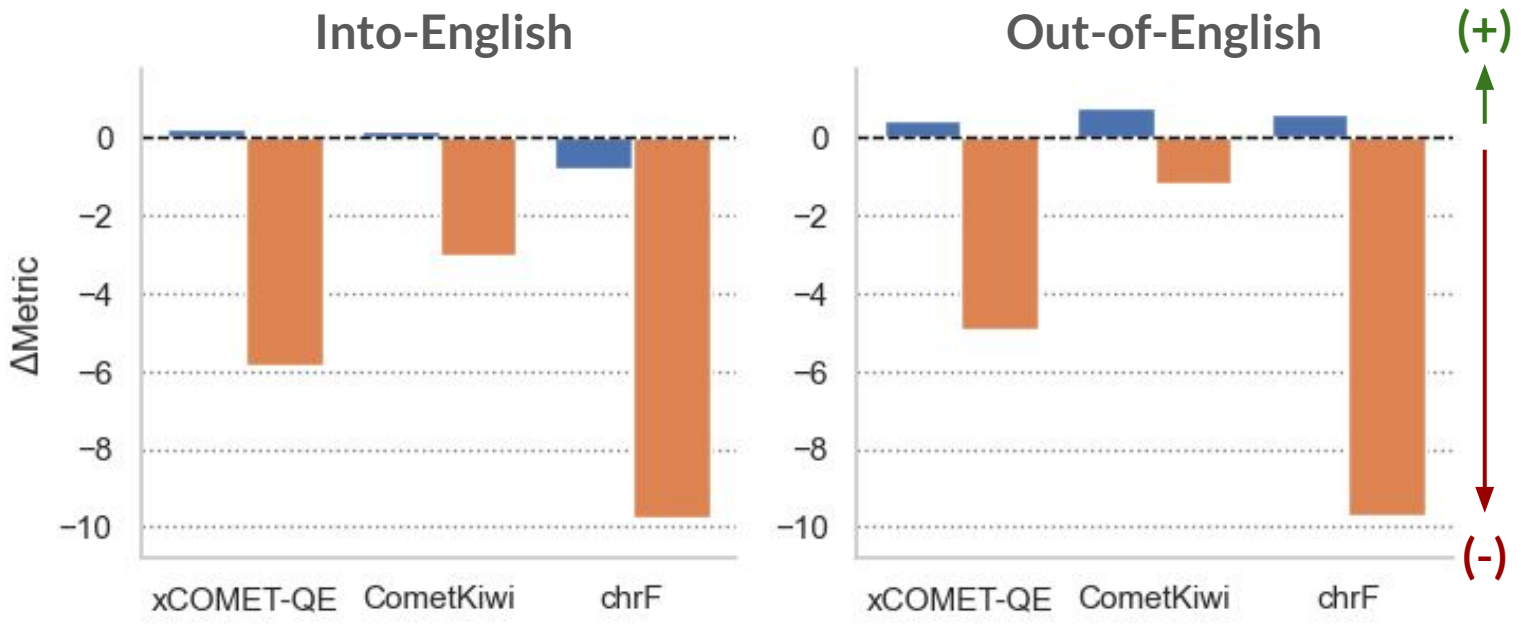
Constraining the preferred system



General results in the multi-system setting

Constraining the preferred system



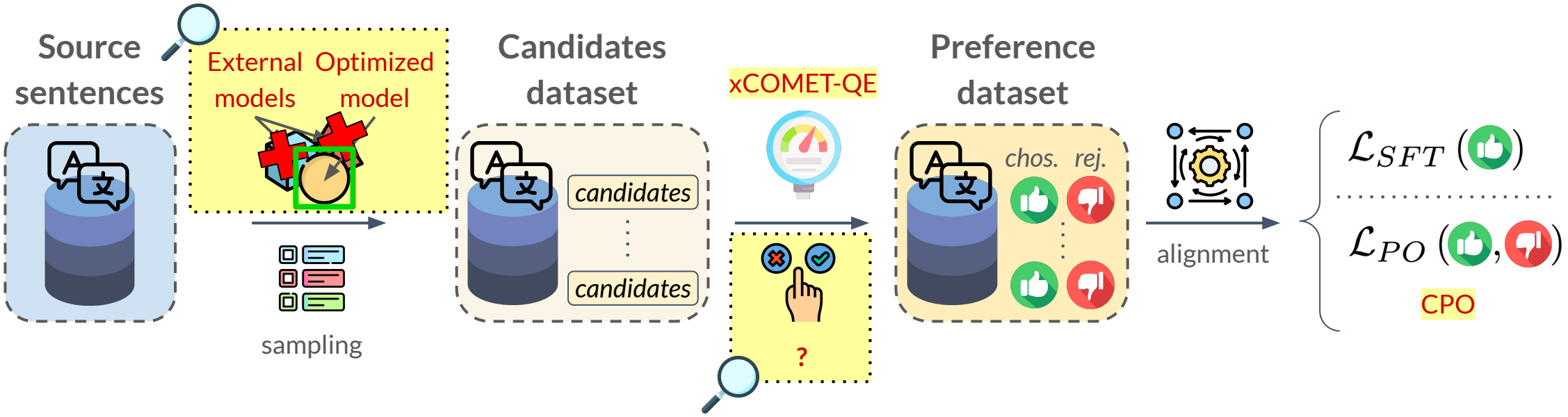


- CPO is highly **non robust** to the preference setting

- SFT demonstrates significantly g**reater stability** in this unfavorable setting

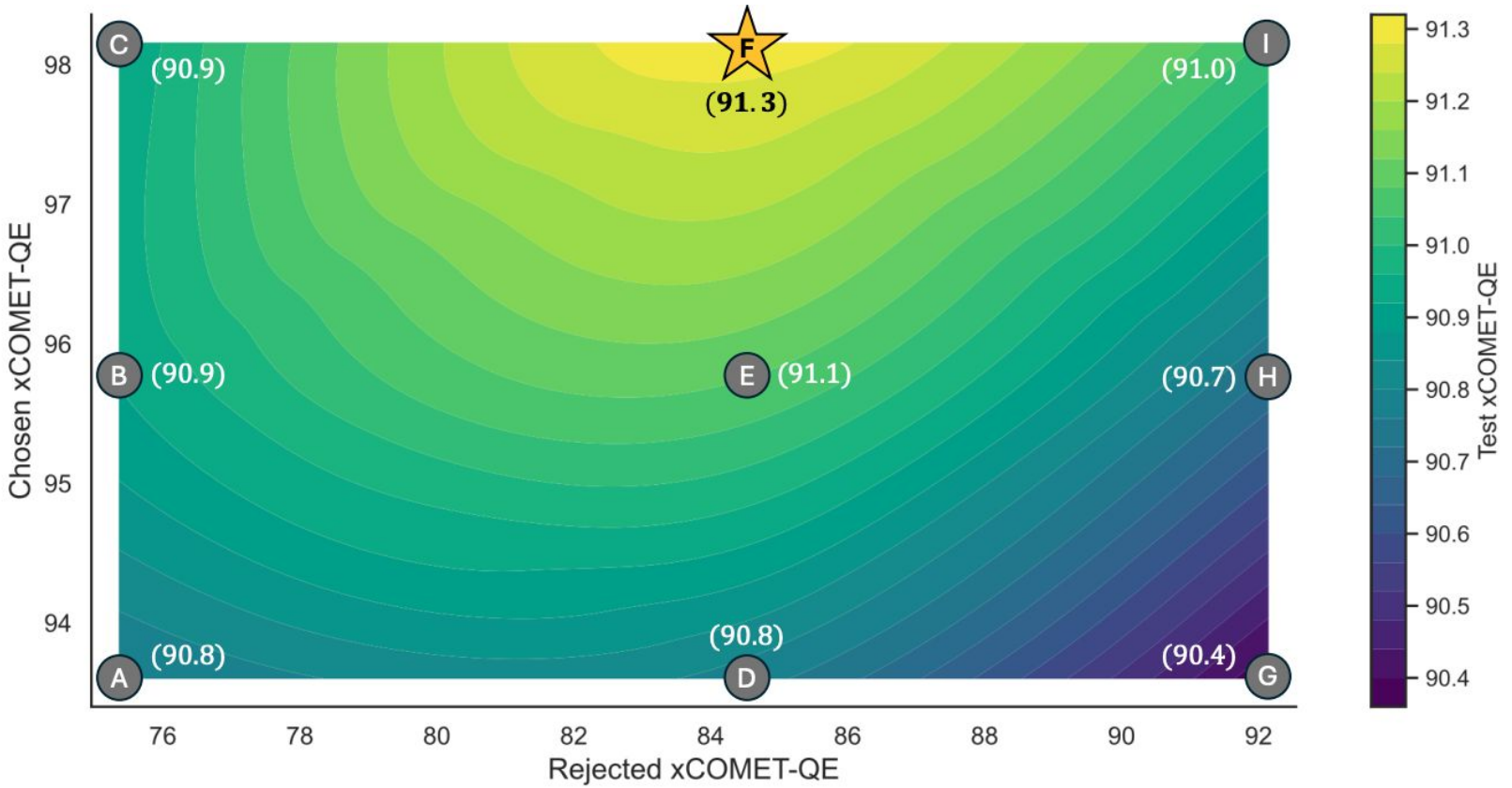
Investigating the mono-system approach

Impact of the qualities of rejected and chosen hypotheses



Investigating the mono-system approach

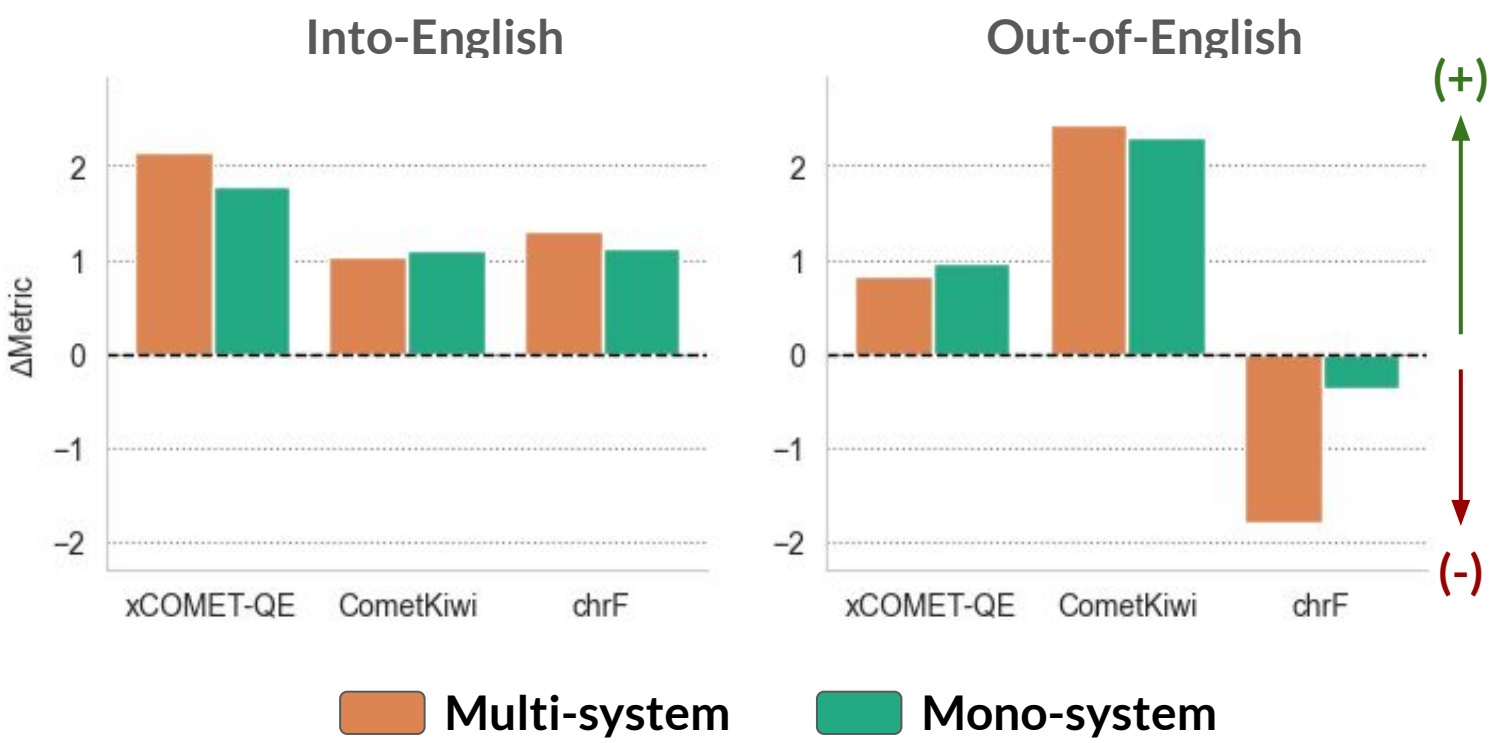
Impact of the qualities of rejected and chosen hypotheses



- Rejected/chosen qualities have a **significant impact** on downstream performance
- **Chosen:** highest quality as possible
- **Rejected:** neither too high nor too low

Investigating the mono-system approach

Comparing with the multi-system approach

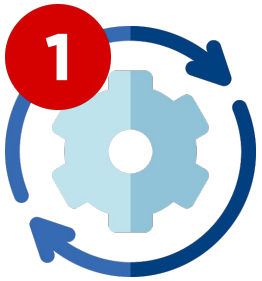


- Optimizing preference data **delivers competitive performance** with the multi-system setting…
- … while **reducing adverse lexical effects**

Takeaways

Beware of preference optimization

- PO seems to be **effective** when aligning on neural metrics
- But it also seems to **lack robustness** when modifying the preference setting



No need to have access to high-quality external models to achieve effective alignment

- Generating hypotheses using the model intended for alignment produces results **on par** with the multi-system approach
- However, this requires **carefully balancing** the quality of chosen and rejected examples when constructing the preference dataset

ARTEFACT RESEARCH CENTER

Unbabel

CentraleSupélec
université PARIS-SACLAY

Thank you for your attention!

Hippolyte Gisserot-Boukhlef

Ricardo Rei

Emmanuel Malherbe

Céline Hudelot

Pierre Colombo

Nuno M. Guerreiro

NINTH CONFERENCE IN MACHINE TRANSLATION
NOVEMBER 2024