SAND2019-14018C

Fine-Grained Analysis of Communication Similarity between Real and Proxy Applications

PRESENTED BY

Omar Aaziz (Sandia National Labs)

Courtenay Vaughan (Sandia National Labs)

Jonathan Cook (New Mexico State University)

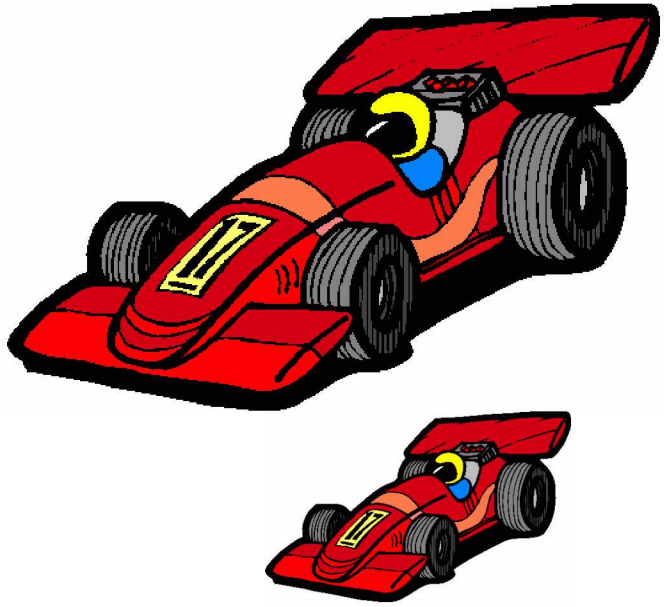
Jeanine Cook (Sandia National Labs)

Jeffery Kuehn (Los Alamos National Lab)

David Richards (Lawrence Livermore National Laboratory)

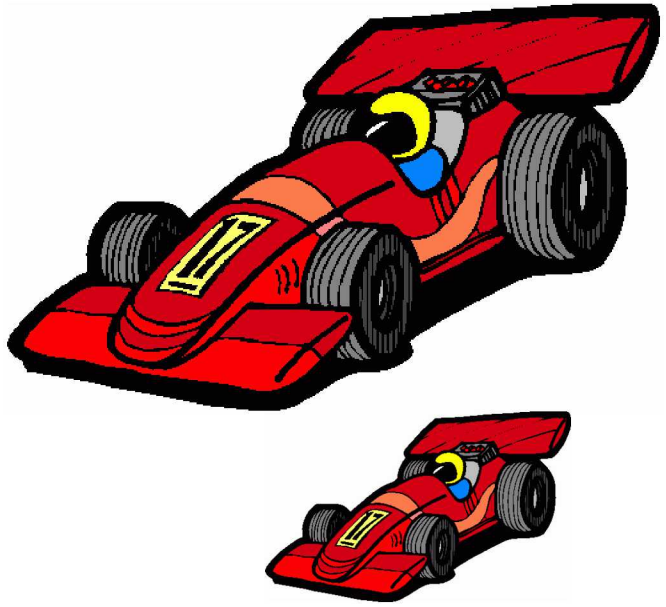
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Proxy Apps are Wonderful!



Easy to Build!

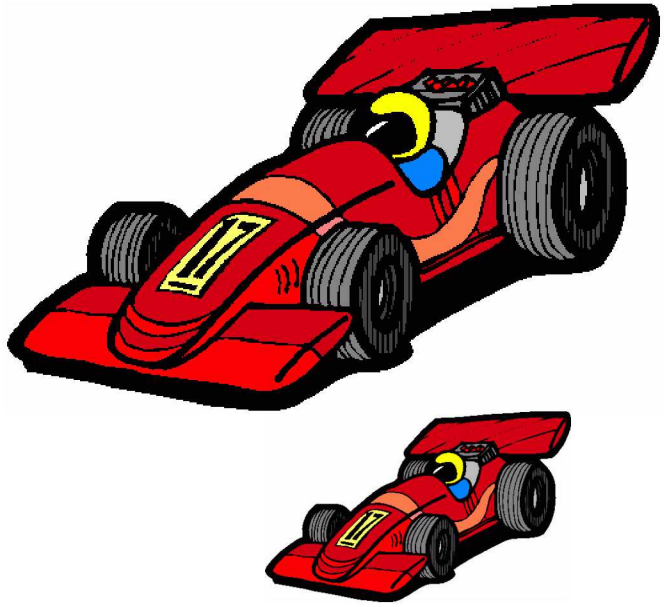
Proxy Apps are Wonderful!



Easy to Modify!

Easy to Build!

Proxy Apps are Wonderful!



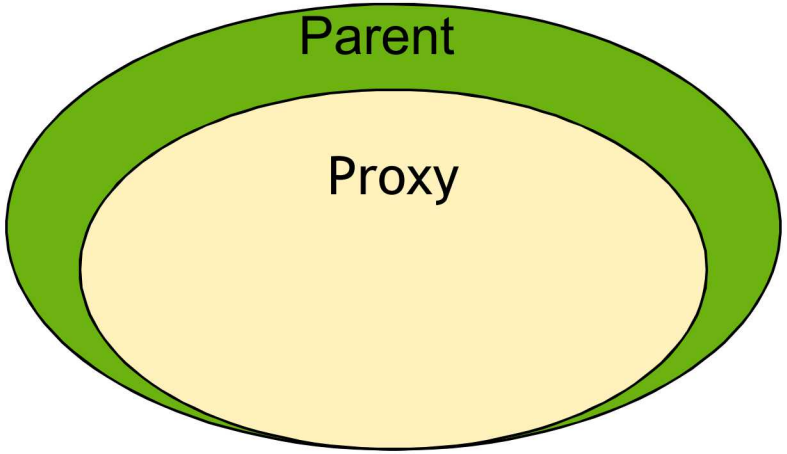
Easy to Modify!

Easy to Build!

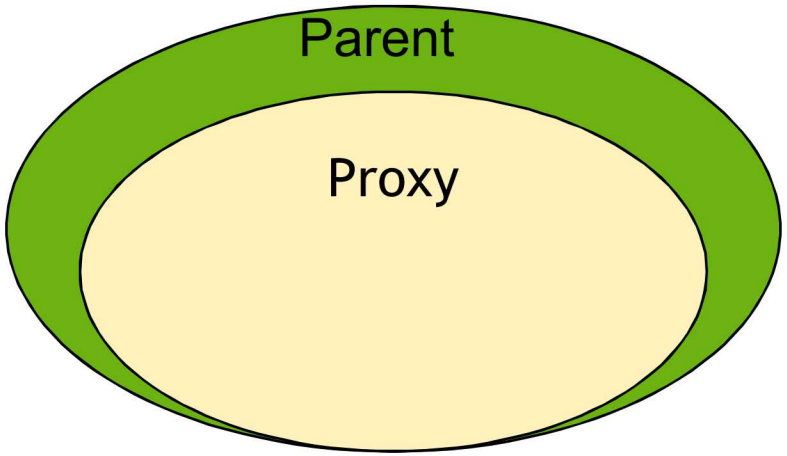
Easy to Run!

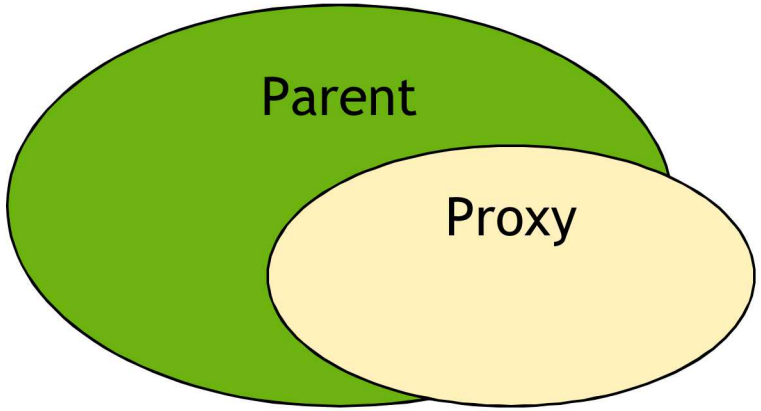
Explore Ideas!

Do Proxies Match the Real Thing?

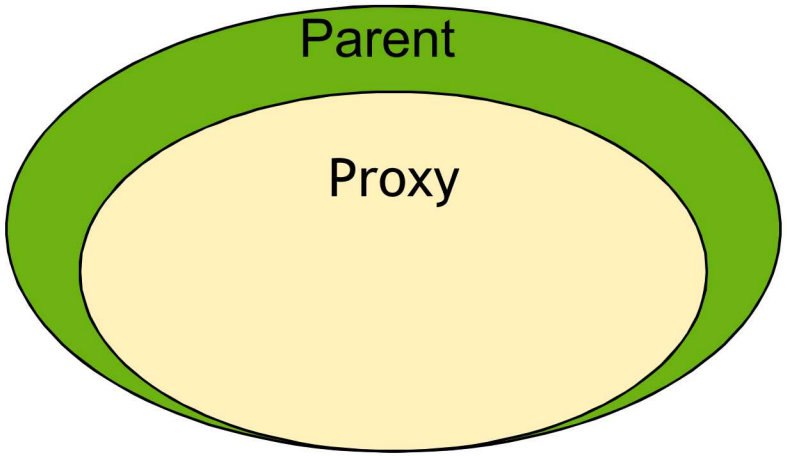


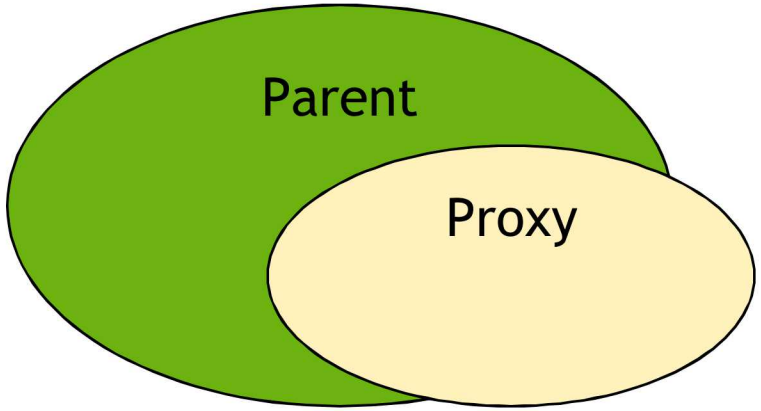
Do Proxies Match the Real Thing?

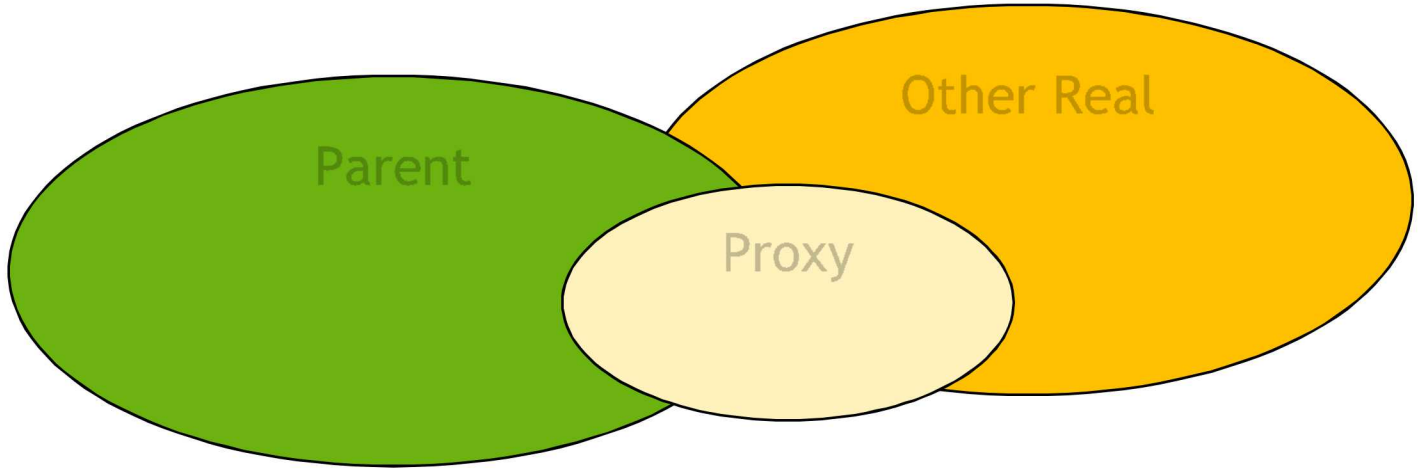




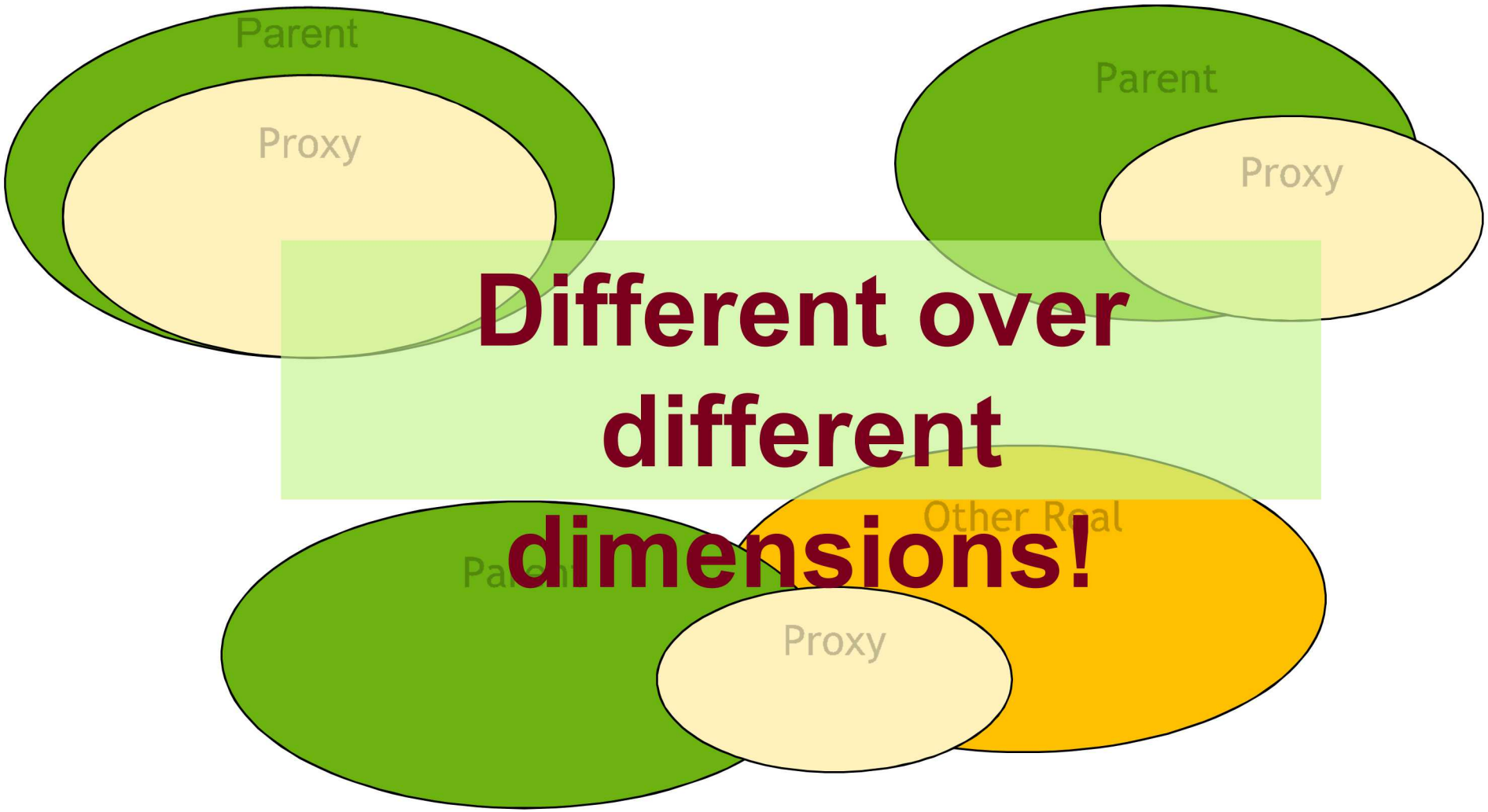
Do Proxies Match the Real Thing?





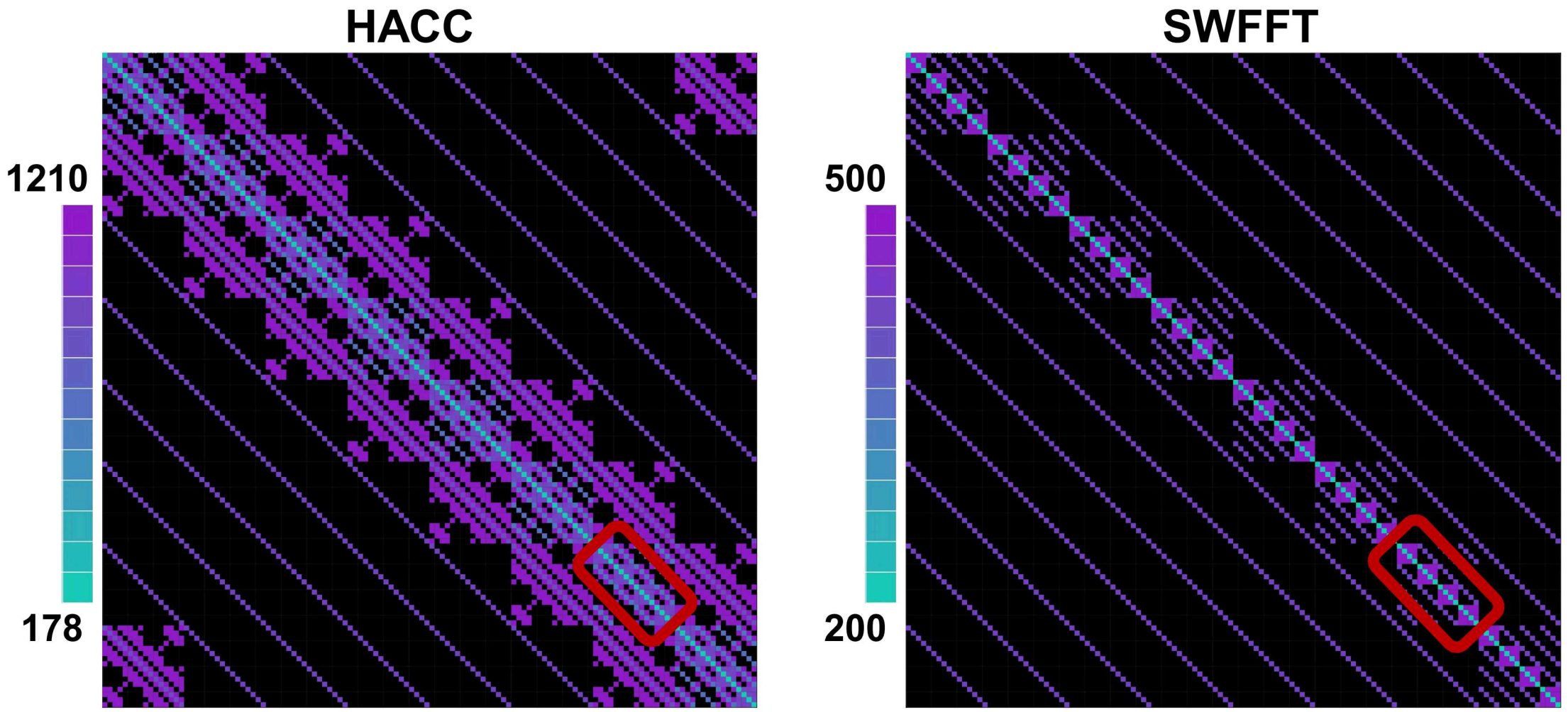


Do Proxies Match the Real Thing?

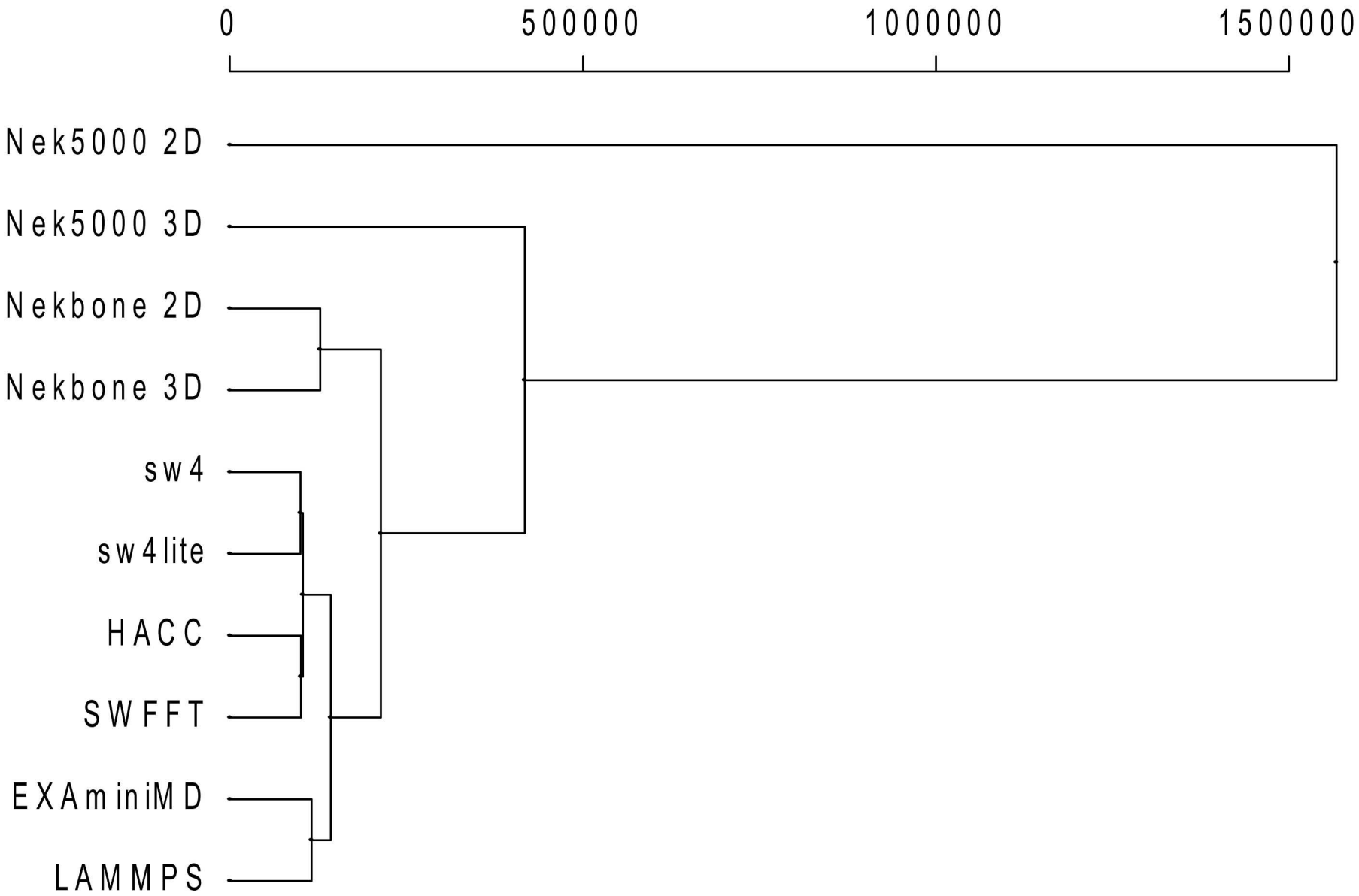


Different over different dimensions!

Previous work, Communication



Previous work, Communication



What about this work?

Motivation

Explored the dynamic communication behavior

Evaluated time-varying communication behavior

Improvement to miniAMR that makes it more closely match its parent CTH

Methodology

Define a reusable approach to evaluating parent/proxy correspondences

Preference towards simplicity

- But still effective

Cumulative vs time-varying communication behavior

Experiment with different statistical analysis methods

- Cosine similarity
- Person
- Spearman

Target: Time-varying Dynamic Communication Behavior

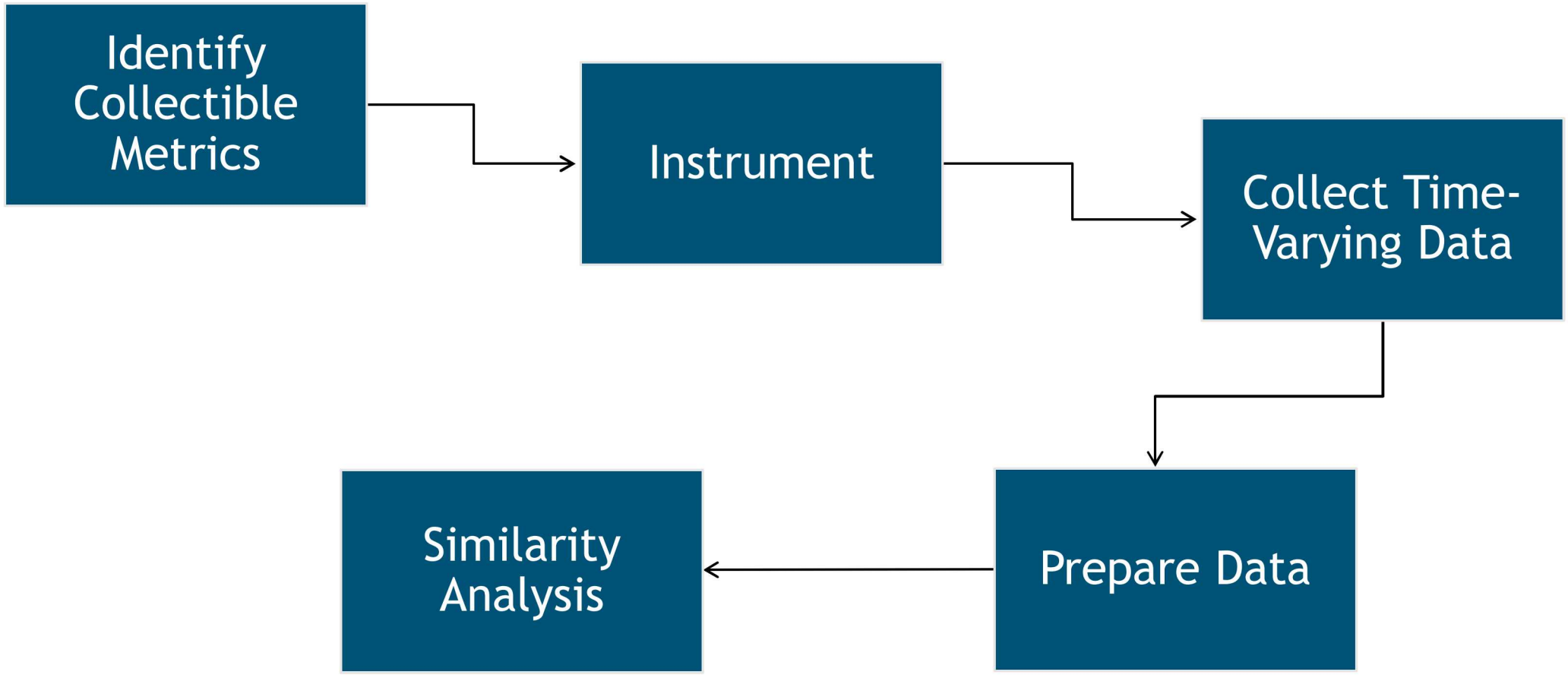
Goal is to evaluate if proxy exercises the communication similar to the parent

Measurements of target time-varying dynamic behavior

Currently: assume that parent and proxy configurations are similar

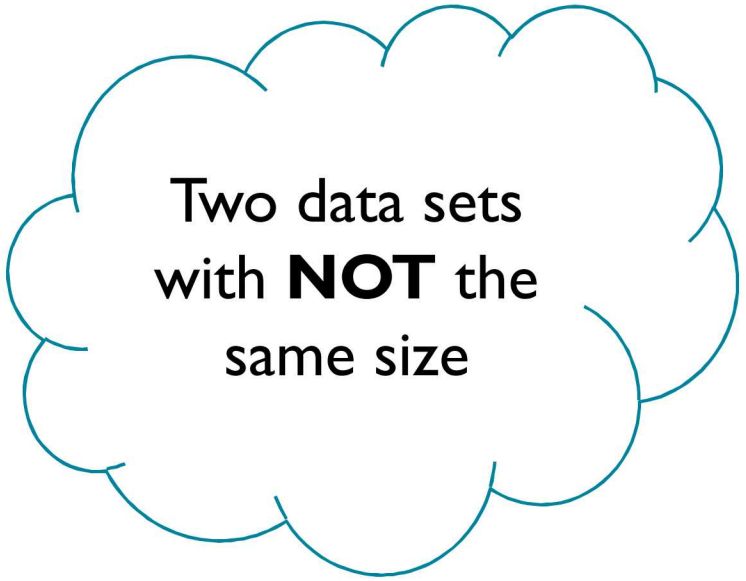
- Build configuration
- Run configuration

Methodology Flow



Pairwise Communication Data

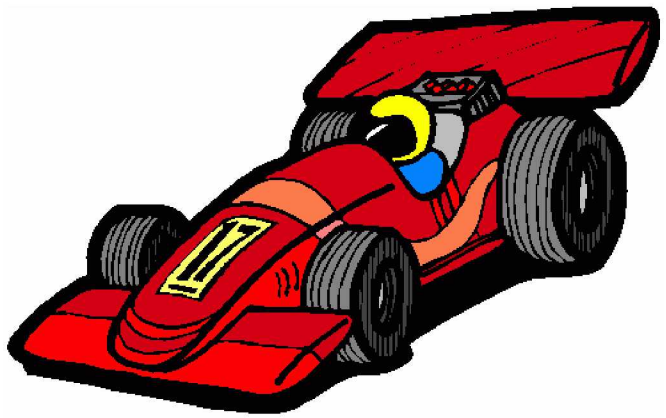
How to compare the Parent/Proxy patterns?



Parent		
Src	Dst	#Msg
0	1	35046
68	64	86
68	67	62
68	69	5887
68	70	24
68	71	29090
68	75	34984
68	91	5916

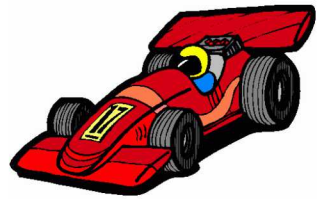
Proxy		
Src	Dst	#Msg
0	1	152120
0	10	153422
0	100	1302
68	64	13020
68	65	13020
68	67	153422
68	69	1302

Full View

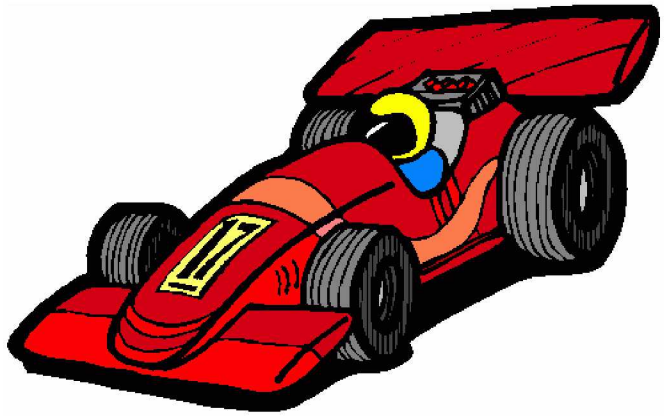


Parent		
Src	Dst	#Msg
0	1	35046
0	10	0
0	100	0
68	60	0
68	64	86
68	65	0
68	67	62
68	70	24
68	71	29090
68	75	34984
68	91	5916

Proxy		
Src	Dst	#Msg
0	1	152120
0	10	153422
0	100	1302
68	60	13020
68	64	0
68	65	13020
68	67	153422
68	70	0
68	71	0
68	75	0
68	91	0

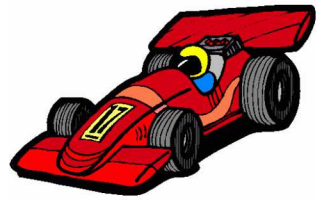


Parent View

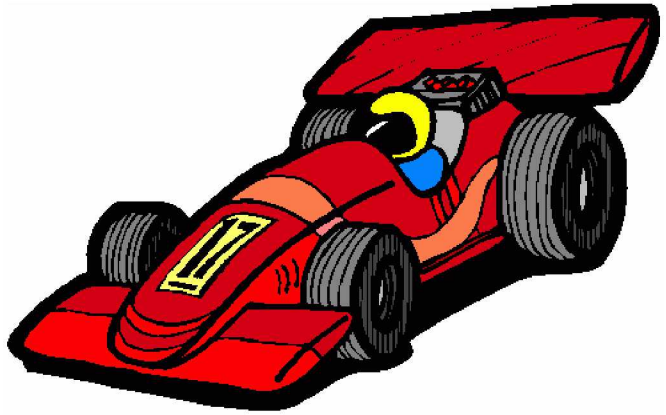


Parent		
Src	Dst	#Msg
0	1	35046
68	64	86
68	67	62
68	69	5887
68	70	24
68	71	29090
68	75	34984
68	91	5916

Proxy		
Src	Dst	#Msg
0	1	152120
68	64	0
68	67	153422
68	69	1302
68	70	0
68	71	0
68	75	0
68	91	0

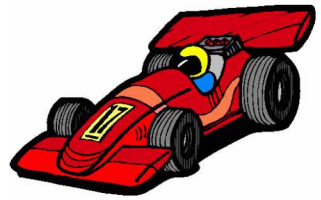


Proxy View

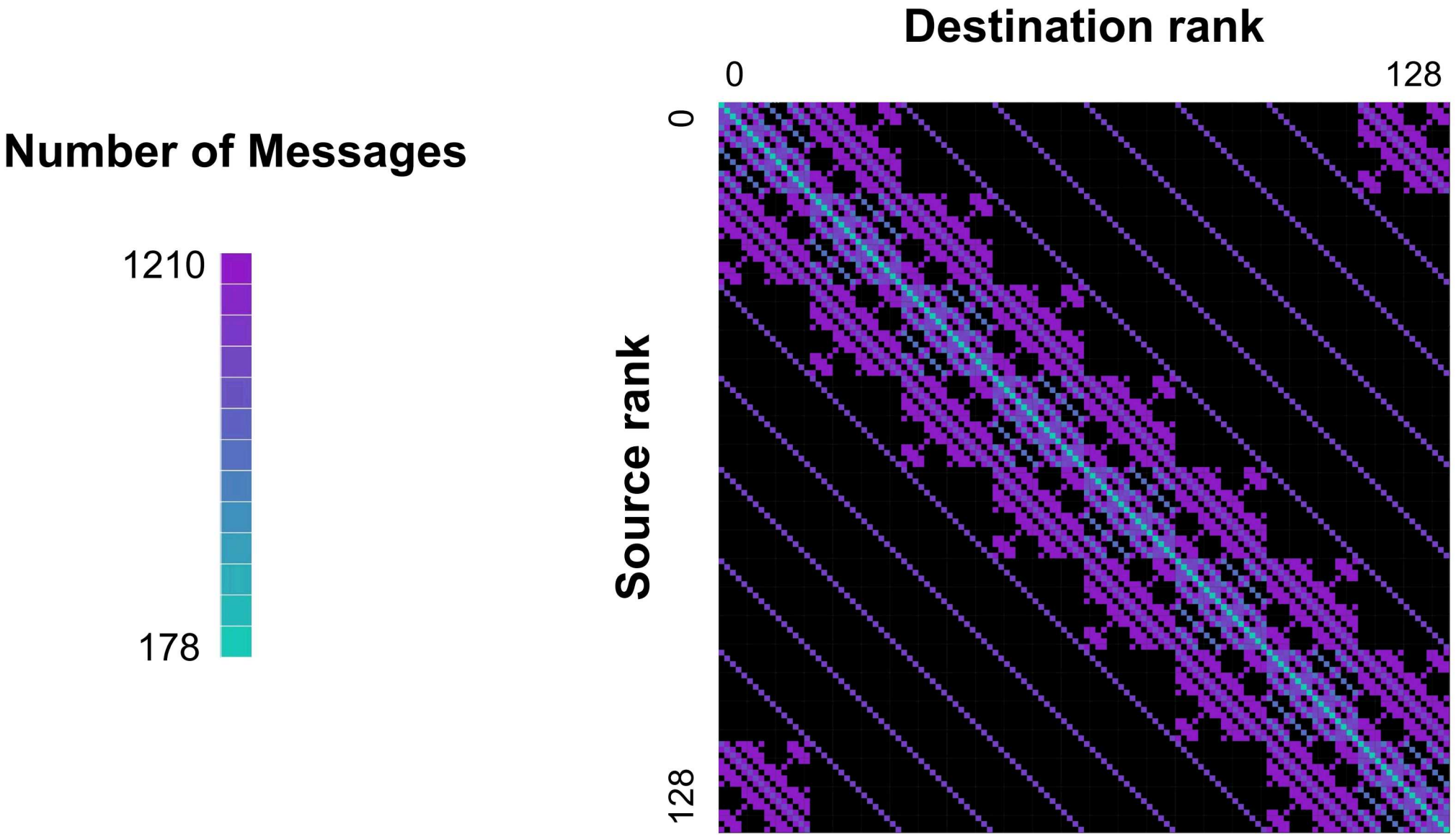


Parent		
Src	Dst	#Msg
0	1	35046
0	10	0
0	100	0
68	60	0
68	65	0
68	67	62
68	69	5887

Proxy		
Src	Dst	#Msg
0	1	152120
0	10	153422
0	100	1302
68	60	13020
68	65	13020
68	67	153422
68	69	1302



Communication Pairs



Parent/Proxy Pairs

List of the Proxy and real applications used in this work:

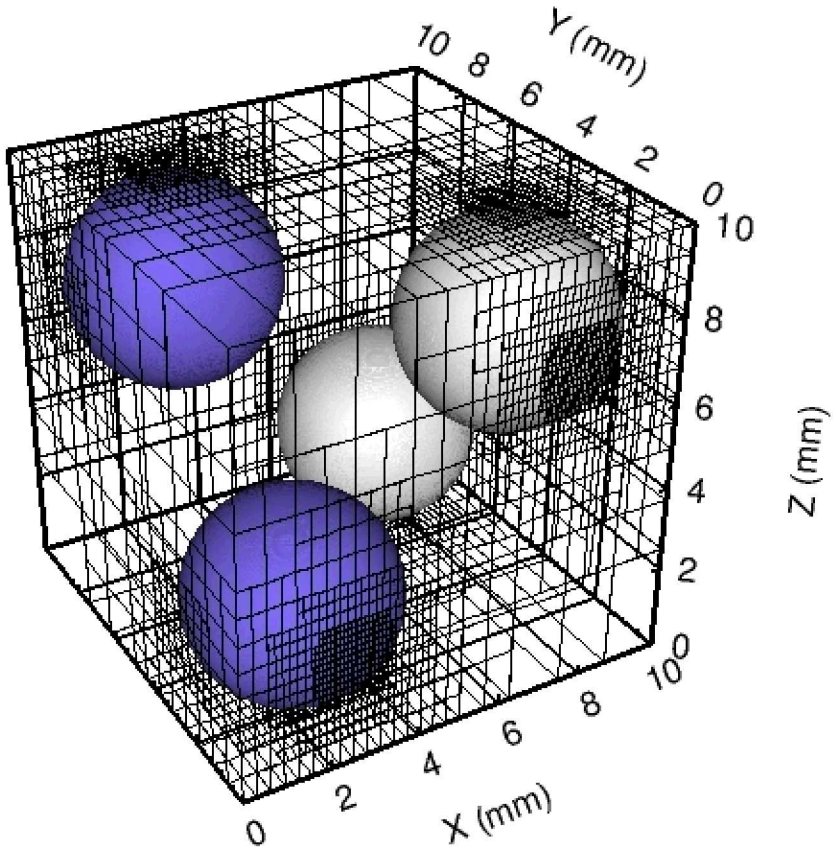
Proxy	Version	Parent	Version	
miniAMR	1.4.4	CTH	12.0	Shock Hydrodynamics
NEW miniAMR**Z**	TBD!	CTH	12.0	Shock Hydrodynamics
SWFFT	1.0	HACC	1.0	Cosmology/FFT
ExaMiniMD	1.0	LAMMPS	17 Aug 2017	Molecular dynamics

miniAMR vs miniAMRZ

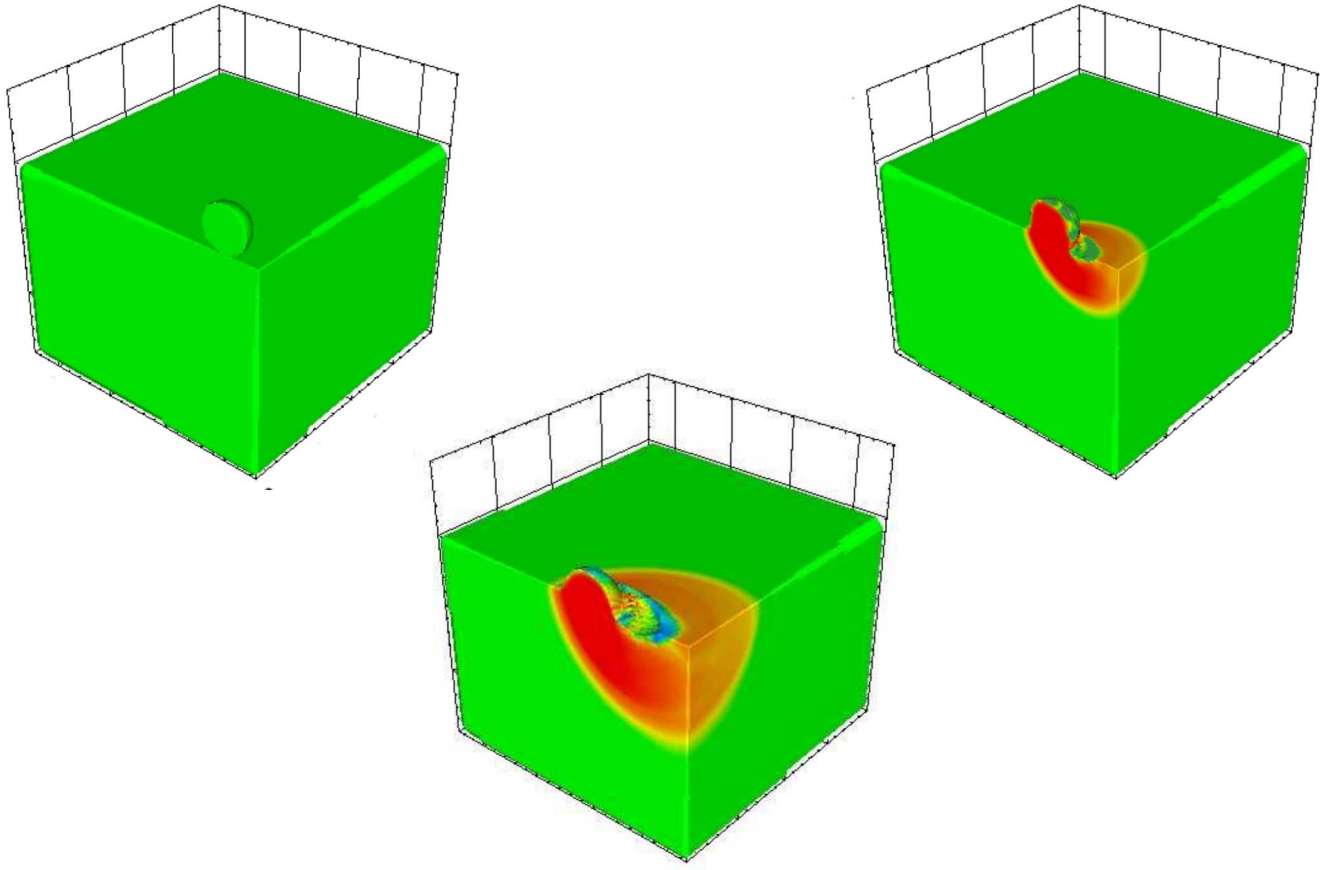
Need Courtenay help!

CTH/miniAMR(Z) problems

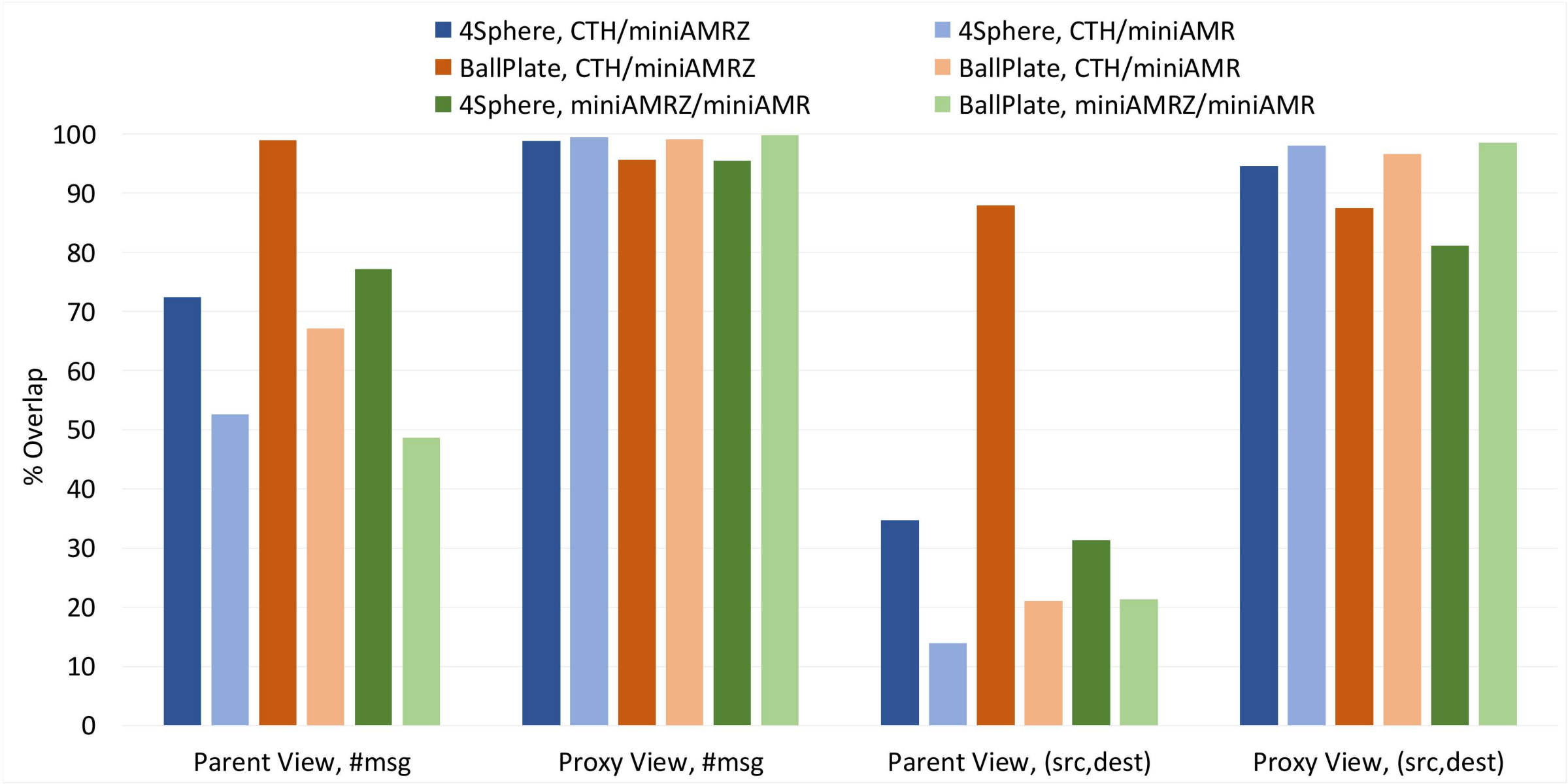
4 Sphere Problem



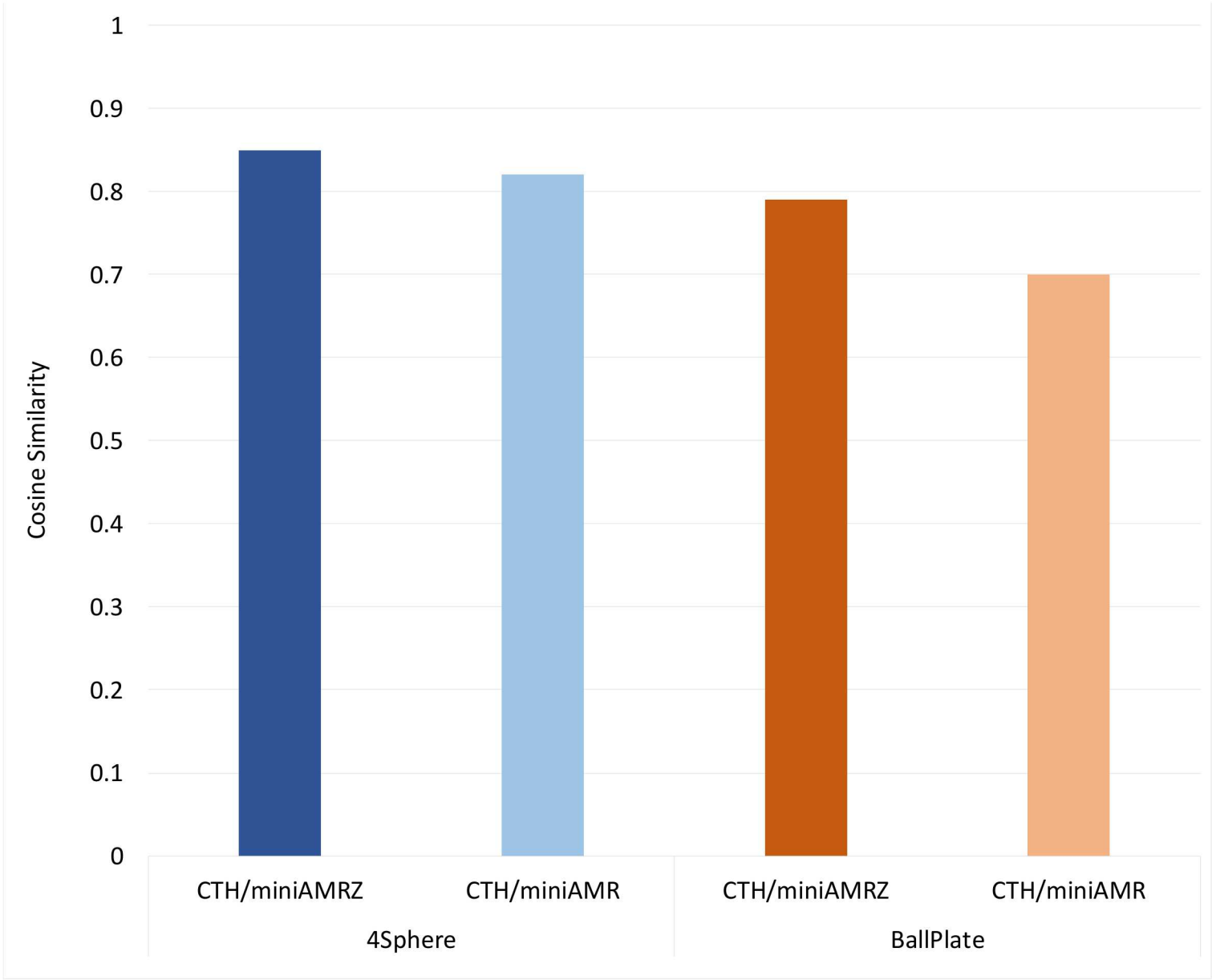
Ball & Plate Problem



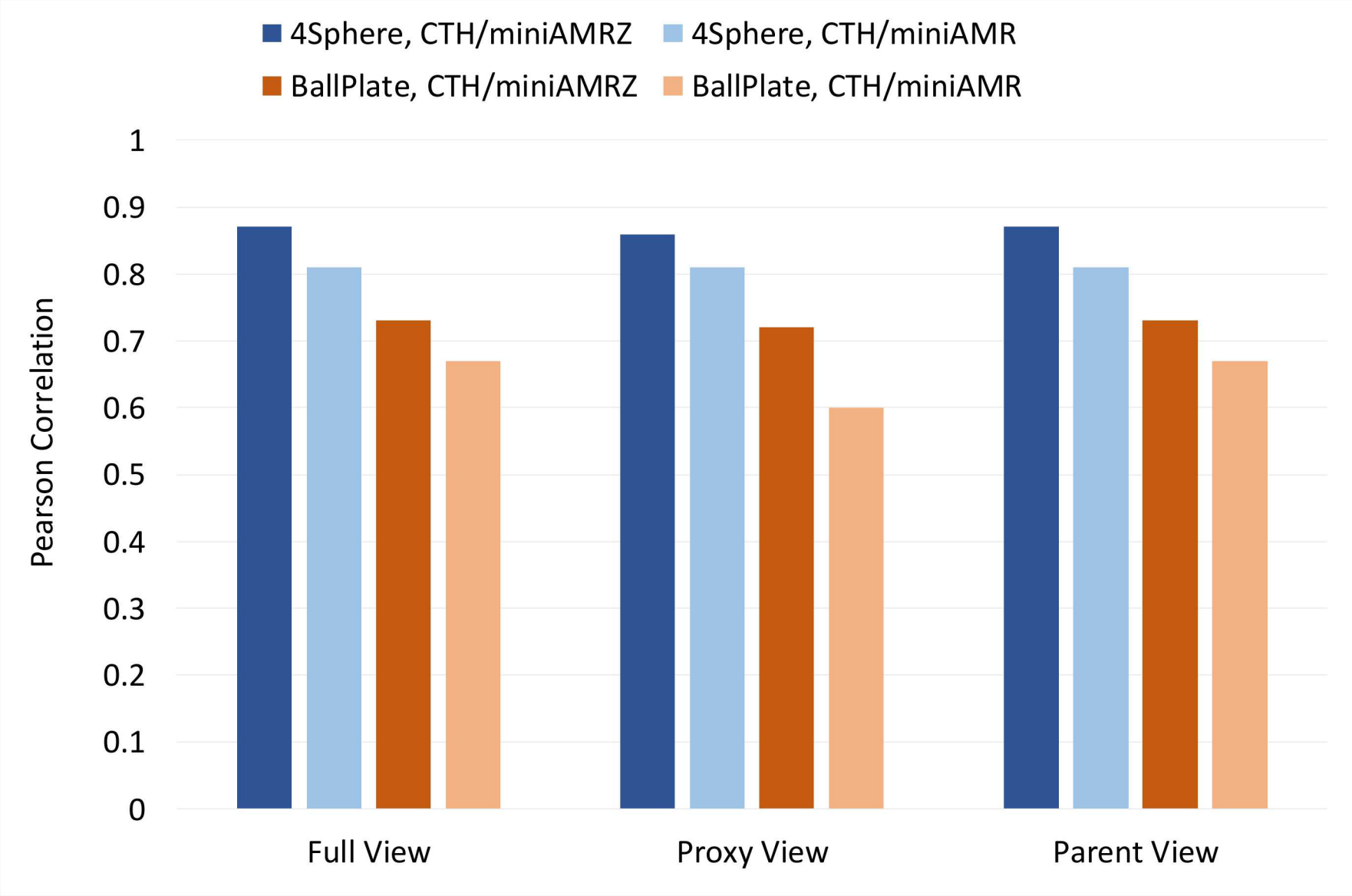
Result: Communication overlap relations



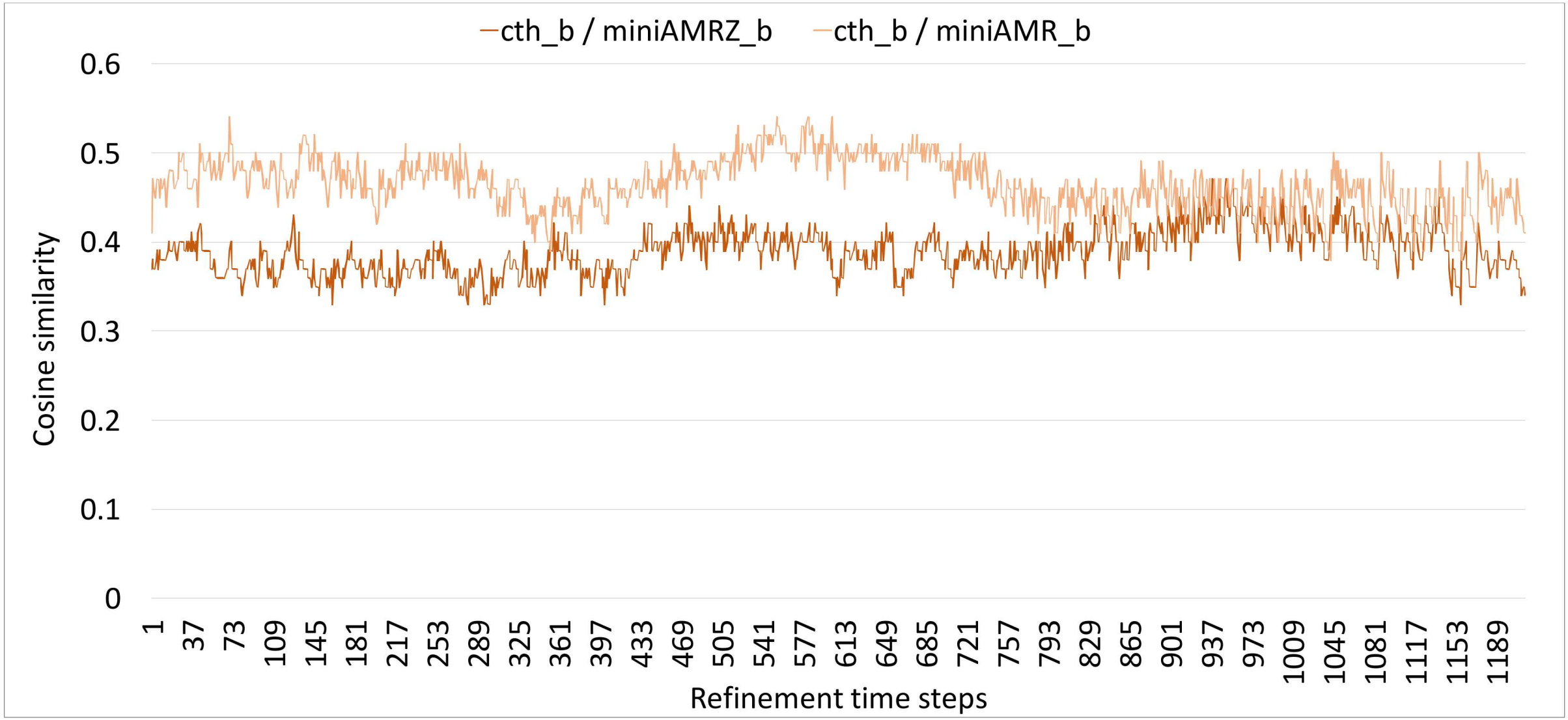
Result: Cosine Similarity



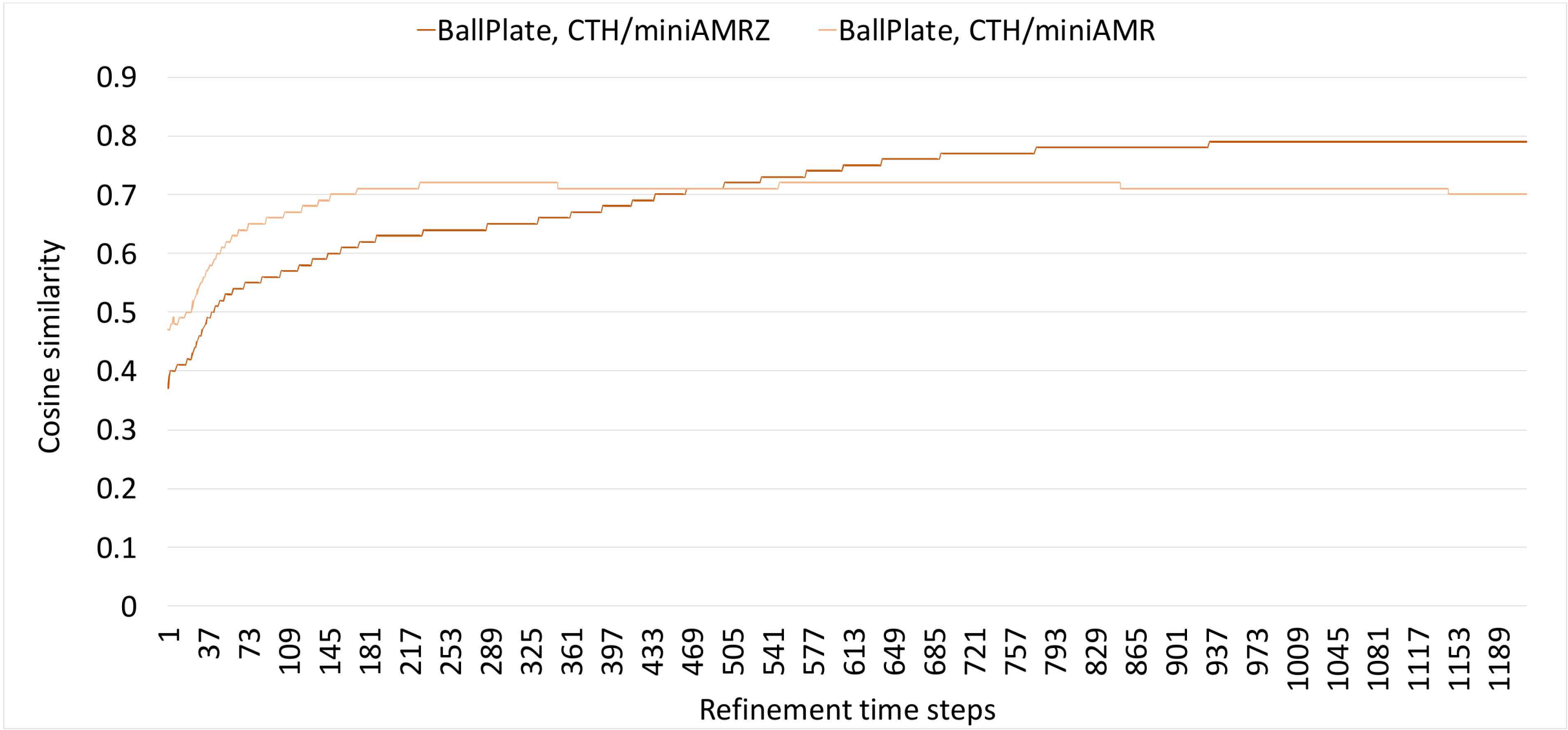
Result: Pearson Correlation



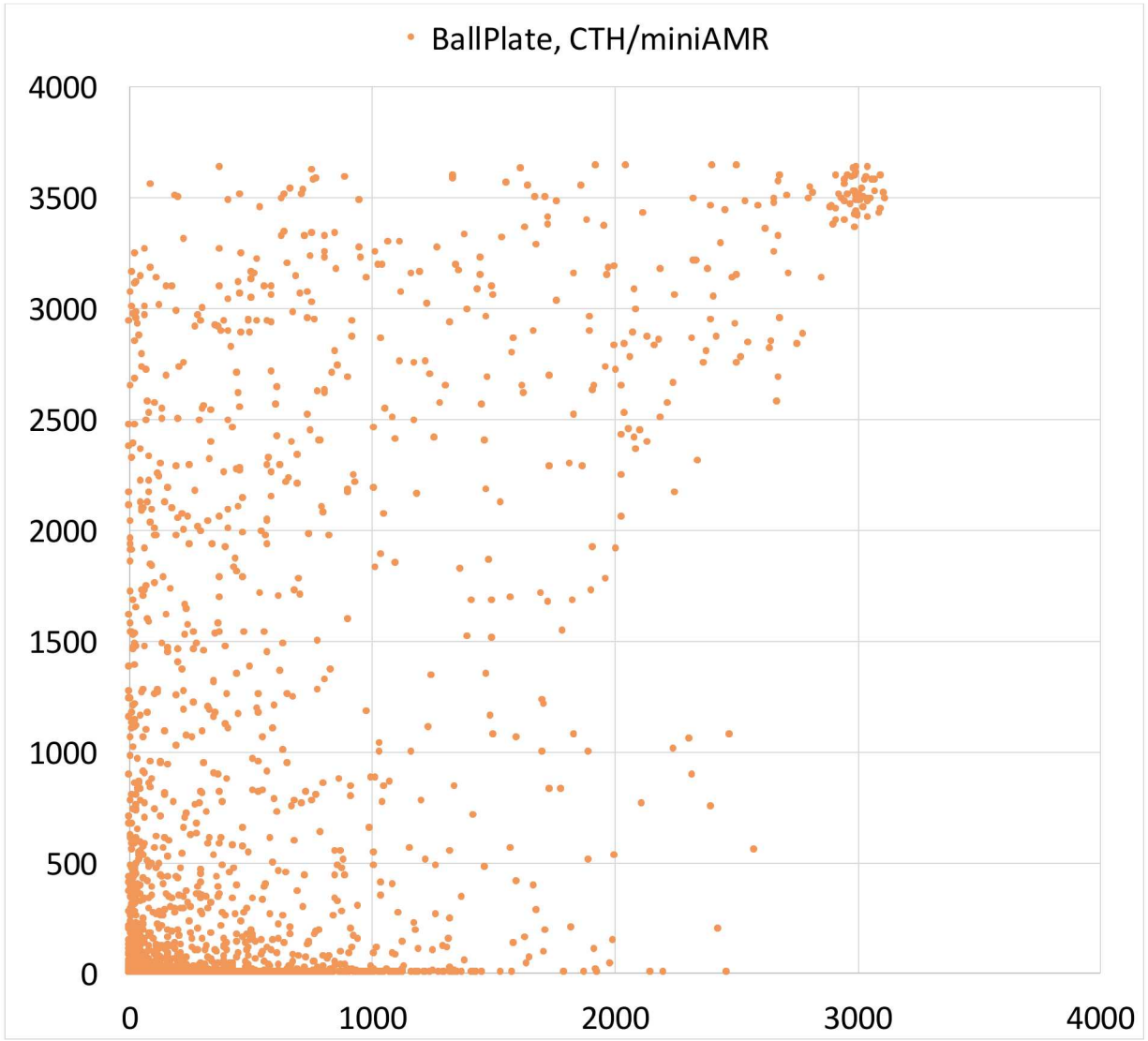
Similarity Metrics over Detailed Time Series Data, Ball & Plate

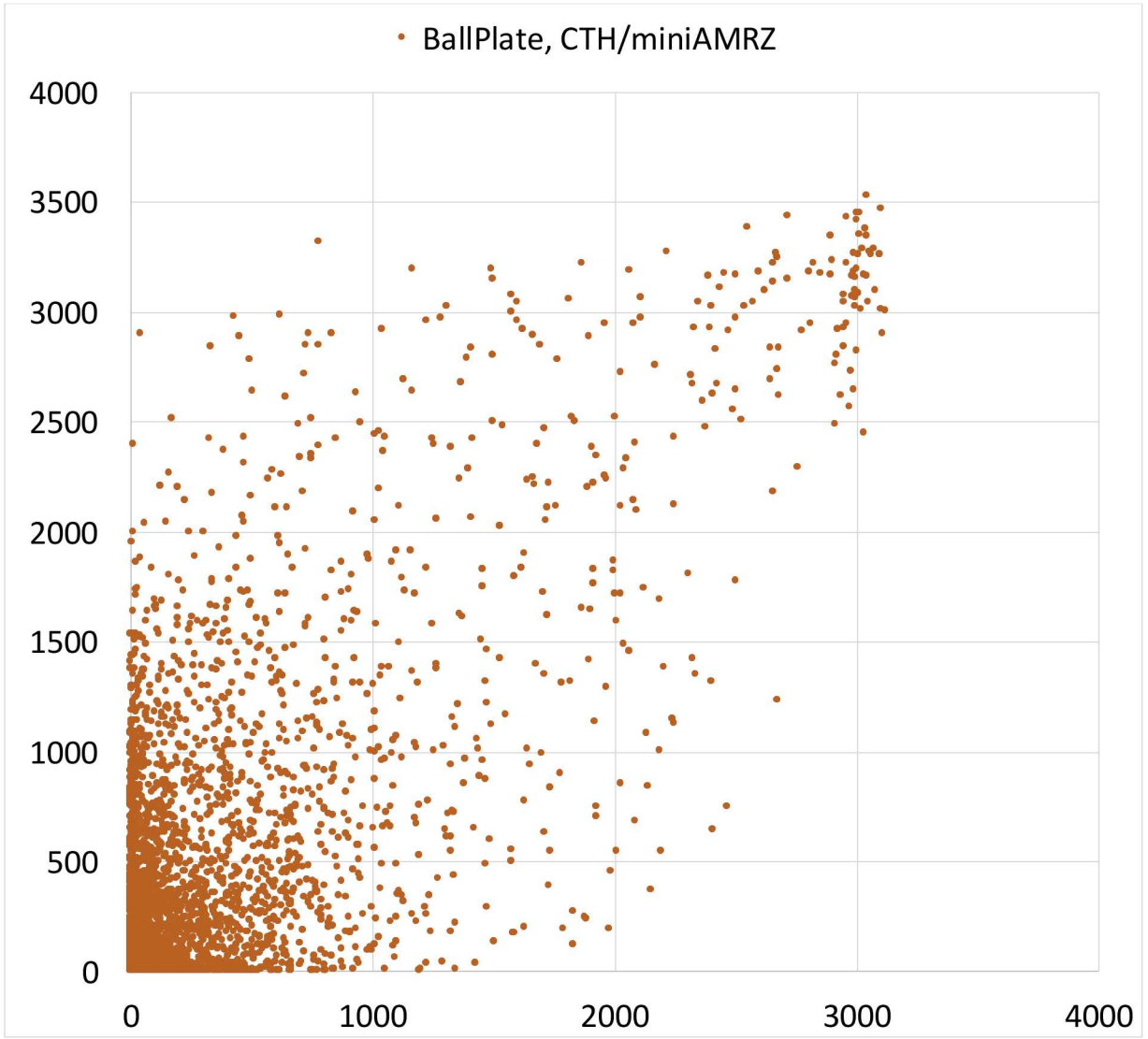


Cumulative Stepwise Cosine Similarity, Ball & Plate

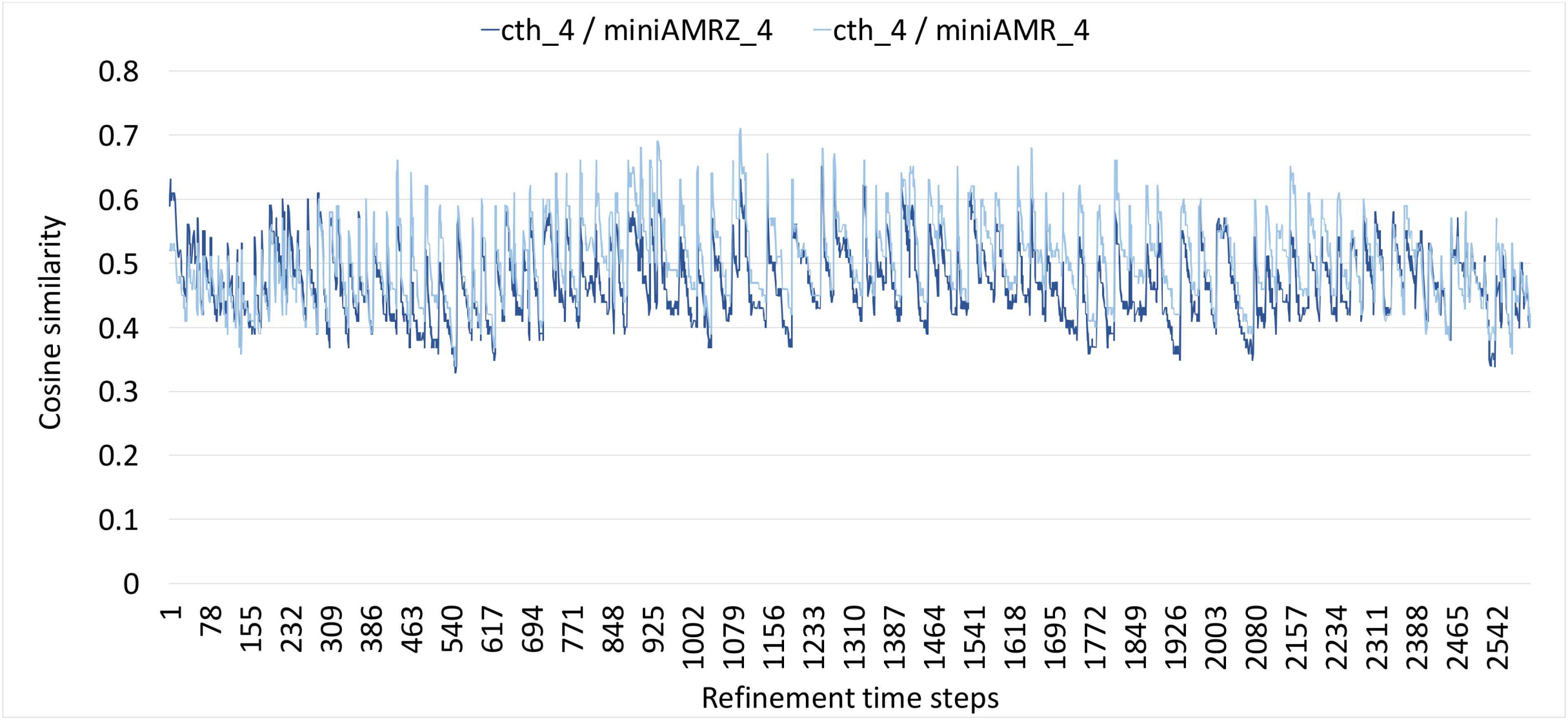


Messages Scatterplots, Ball & Plate

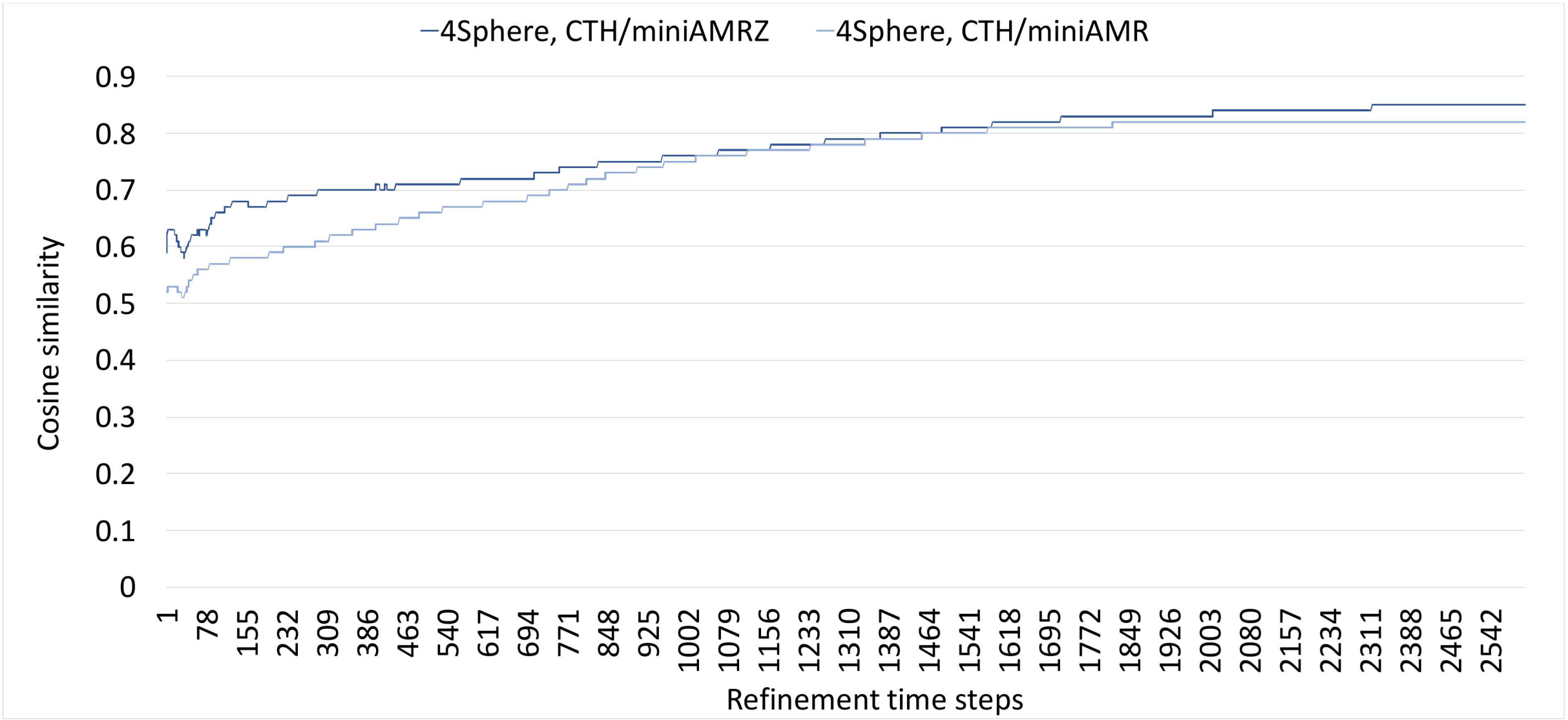




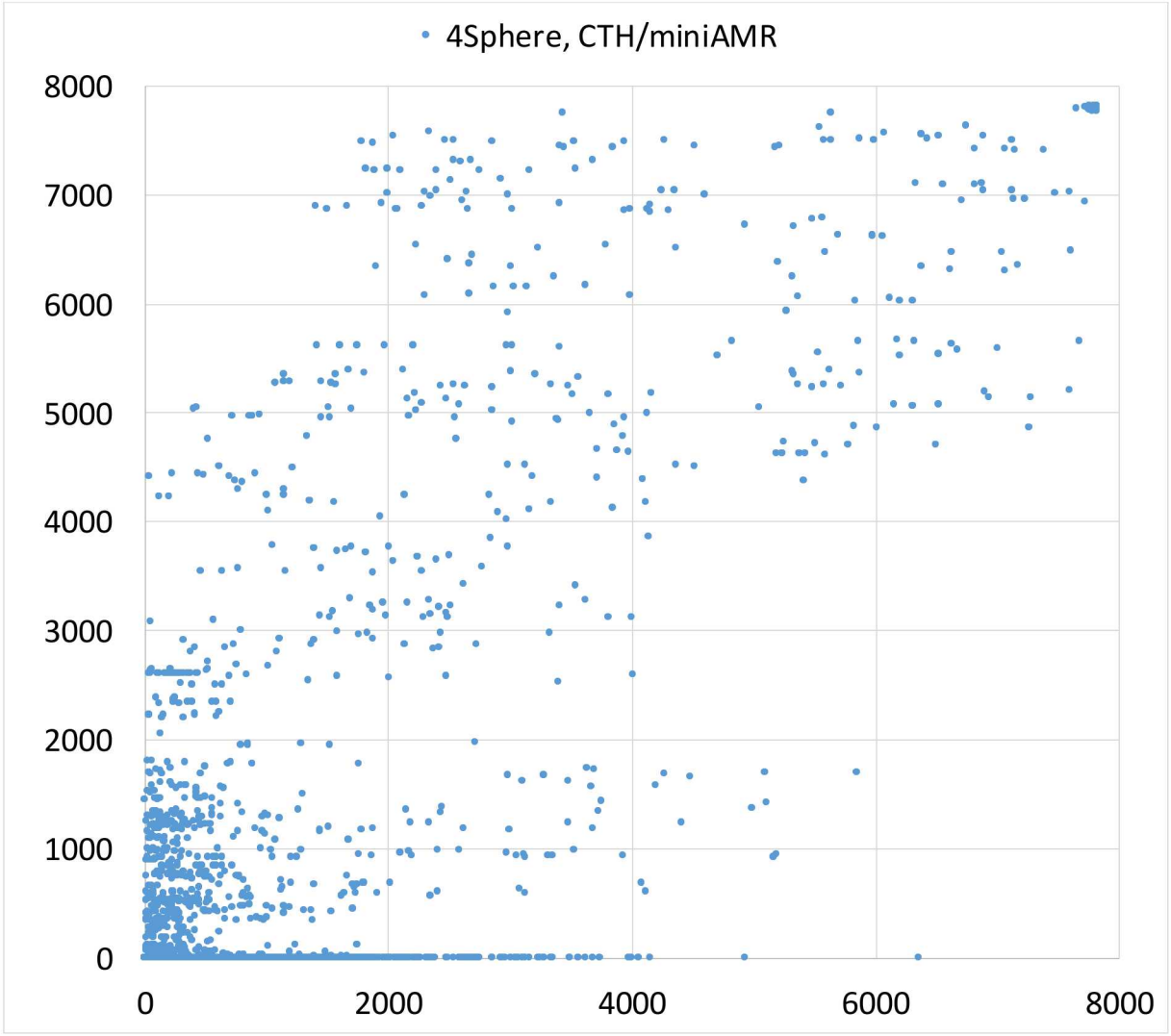
Similarity Metrics over Detailed Time Series Data, 4Sphere

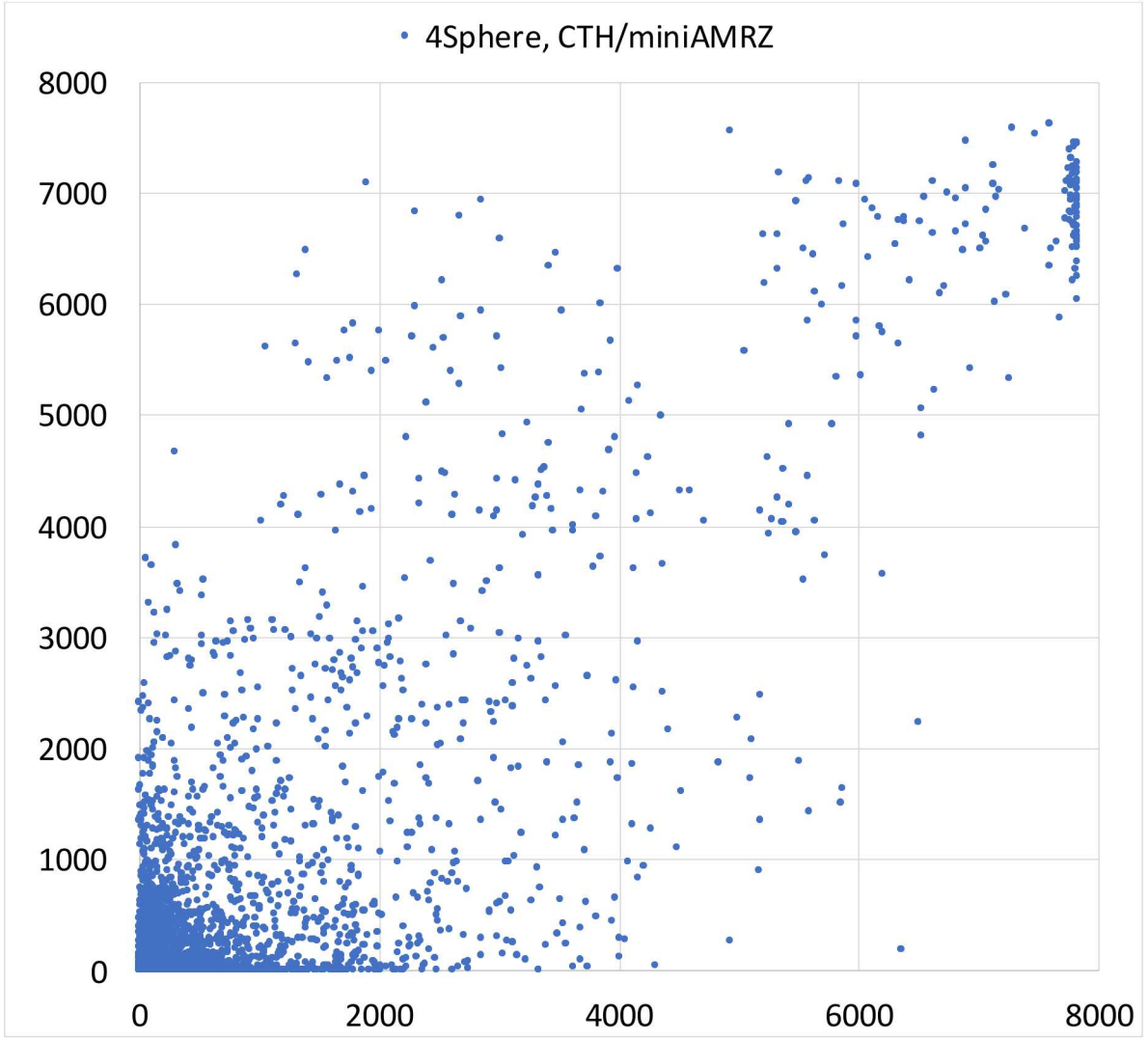


Cumulative Stepwise Cosine Similarity, 4Sphere



Messages Scatterplots, 4Sphere





Conclusion

Explored how the communication of a proxy application relates to its parent application

Improved version of miniAMR – with Z!

Presented a time-varying quantification methods:

- A time-varying pairwise communication quantification method that captures how much of one application matches the other
- How we should look at data cumulative vs time-varying

Future Work

Examine alternative methods

Incorporate the work on the communication patterns by exploring pairs with communication implementation

Explore more parent/proxy pairs

Acknowledgment

This research was supported by the **Exascale Computing Project (ECP)**, Project Number 17-SC-20-SC, a collaborative effort of two DOE organizations, the Office of Science and the National Nuclear Security Administration, responsible for the planning and preparation of a capable exascale ecosystem including software, applications, hardware, advanced system engineering, and early testbed platforms, to support the nation's exascale computing imperative.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE- NA0003525.

Questions ?

What's new?

- Dynamic profiling
- Roofline modeling
- Qualitative comparison of quantitative metrics
- Intel’s Top-Down Microarchitecture Analysis (TMA)