Using iterated learning to reveal biases for well-structured meanings in language

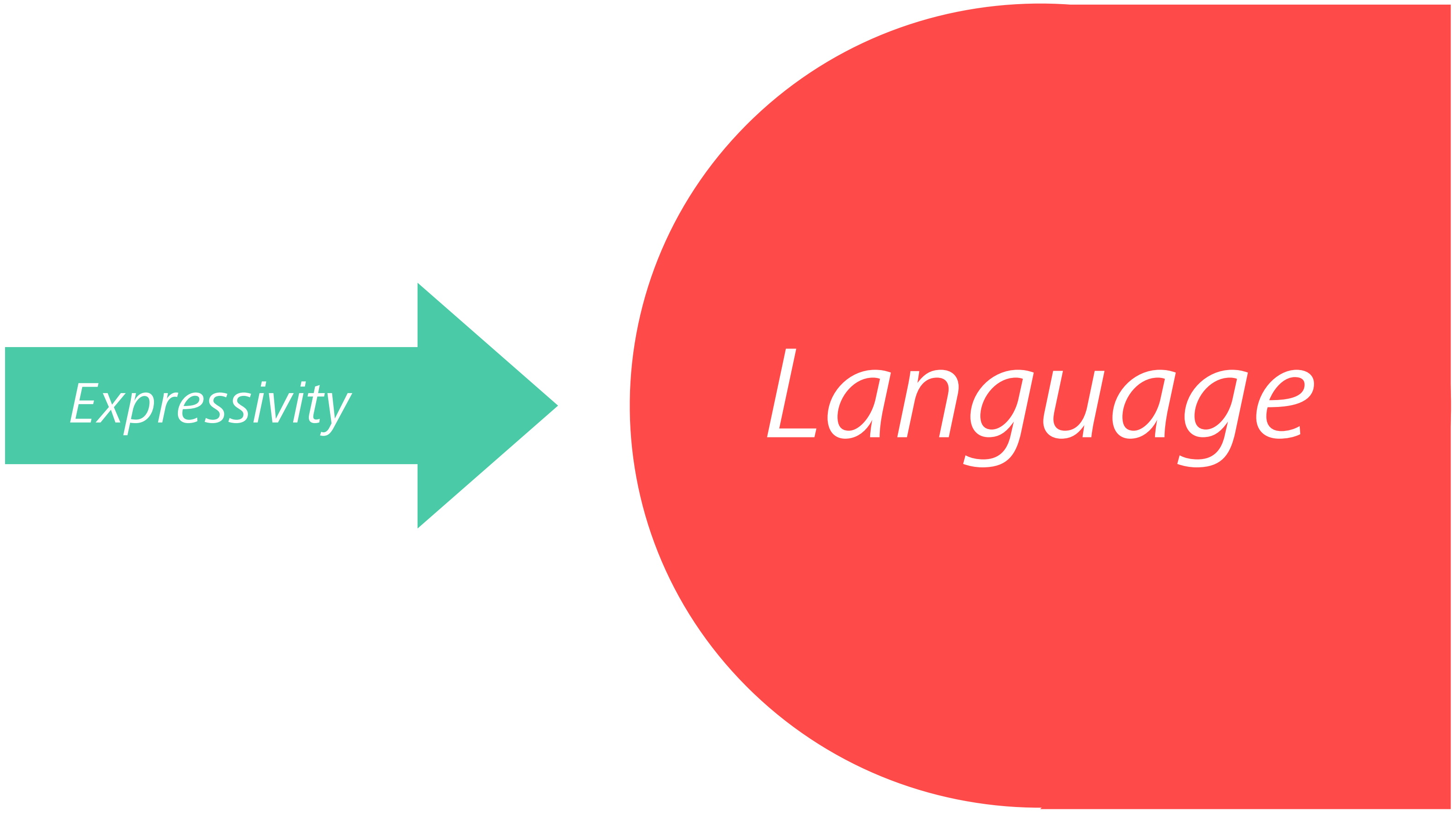
Jon W. Carr
Centre for Language Evolution
University of Edinburgh

Linguistics and English Language Postgraduate Conference 2016

What shapes language?

Language

What shapes language?



What shapes language?

Kirby, Tamariz, Cornish, & Smith, 2015, *Cognition*

What shapes language?

Kirby, Tamariz, Cornish, & Smith, 2015, *Cognition*

Kemp & Regier, 2012, *Science*

What shapes language?

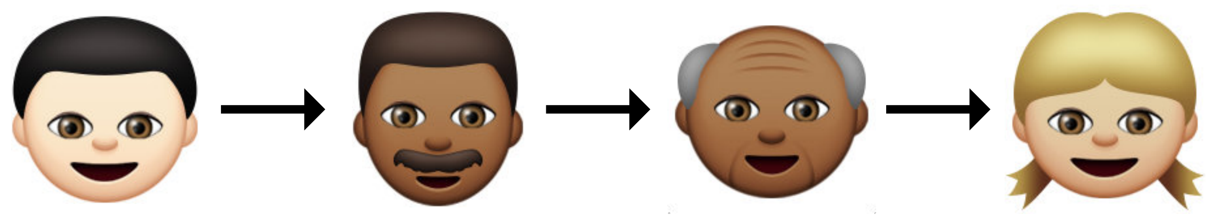
Kirby, Tamariz, Cornish, & Smith, 2015, *Cognition*

Kemp & Regier, 2012, *Science*

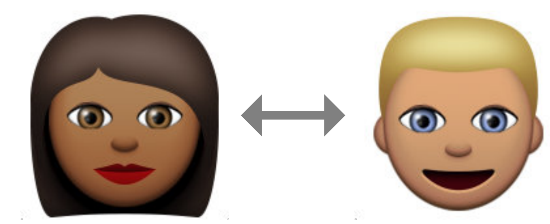
Gabelentz, 1901

Models of learning vs communication

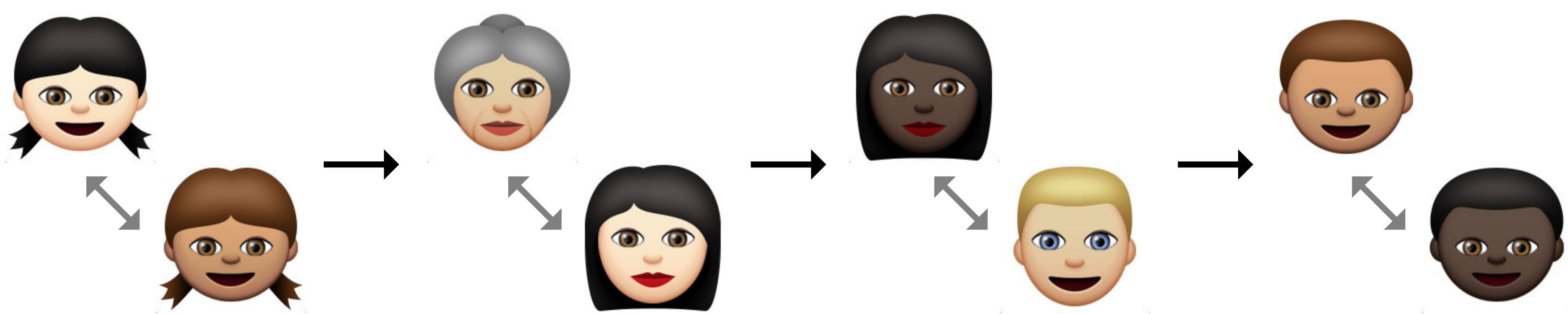
Transmission chain



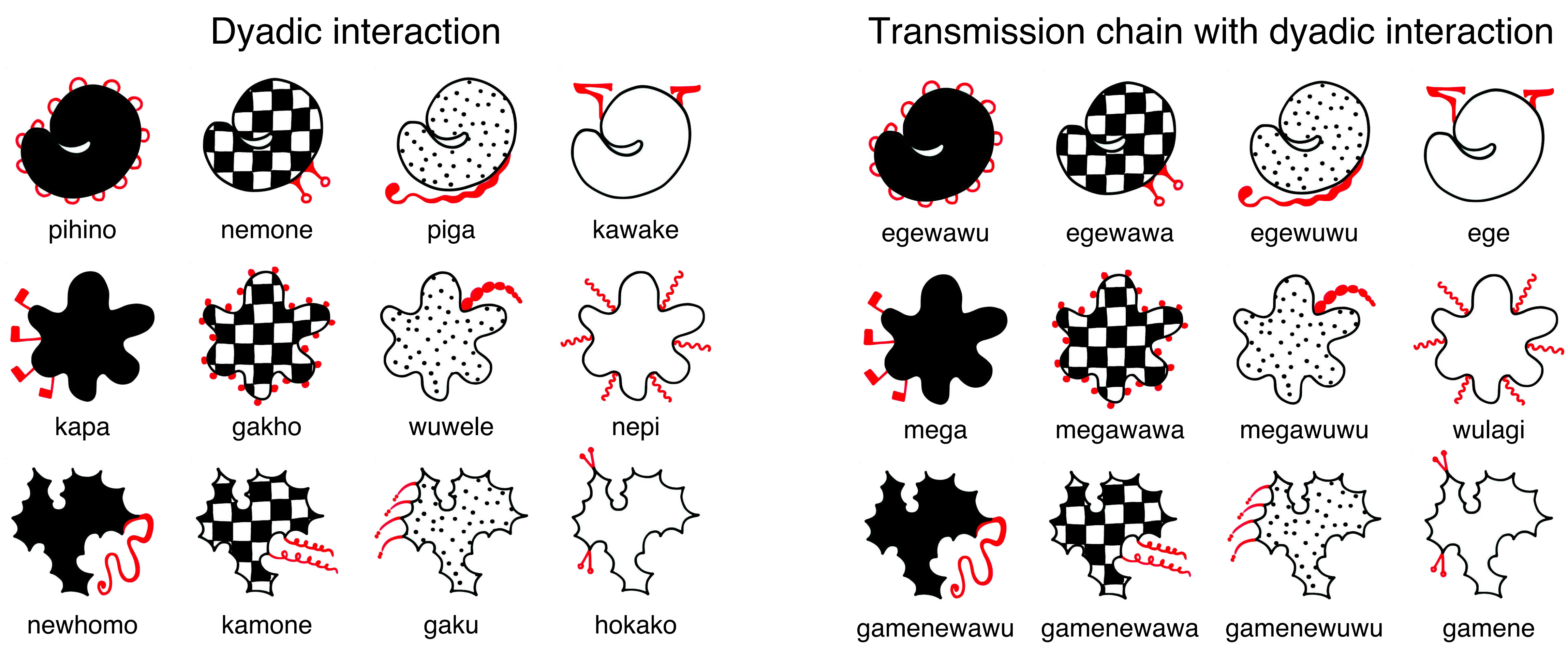
Dyadic interaction



Transmission chain with dyadic interaction



Learning vs communication



Kirby, Tamariz, Cornish, & Smith, 2015, *Cognition*

How do learning and communication shape the structure of semantic categories?

Previous work

Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2016). The cultural evolution of structured languages in an open-ended, continuous world. *Cognitive Science*. doi:10.1111/cogs.12371


COGNITIVE SCIENCE
A Multidisciplinary Journal
Cognitive Science (2016) 1–32

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12371

The Cultural Evolution of Structured Languages in an Open-Ended, Continuous World

Jon W. Carr,[a] Kenny Smith,[a] Hannah Cornish,[b] Simon Kirby[a]
[a]*School of Philosophy, Psychology and Language Sciences, University of Edinburgh*
[b]*Psychology, School of Natural Sciences, University of Stirling*

Received 5 October 2015; received in revised form 11 January 2016; accepted 15 January 2016

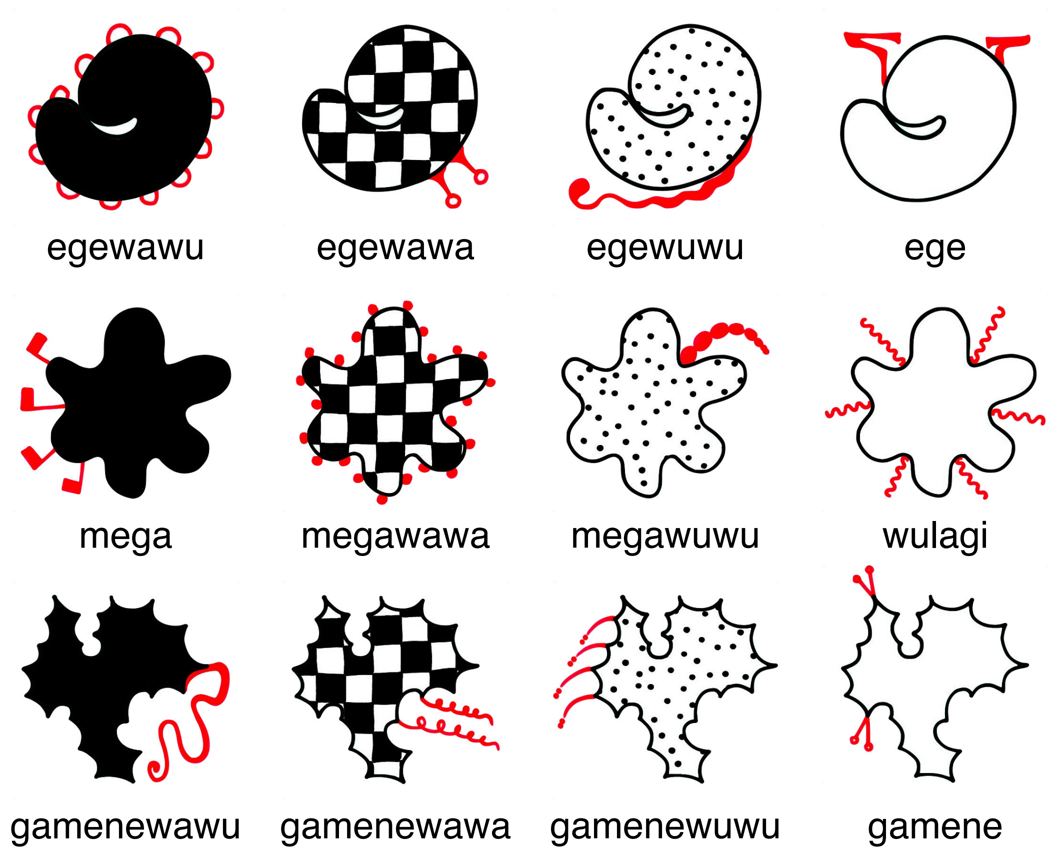
Abstract

Language maps signals onto meanings through the use of two distinct types of structure. First, the space of meanings is discretized into categories that are shared by all users of the language. Second, the signals employed by the language are compositional: The meaning of the whole is a function of its parts and the way in which those parts are combined. In three iterated learning experiments using a vast, continuous, open-ended meaning space, we explore the conditions under which both structured categories and structured signals emerge ex nihilo. While previous experiments have been limited to either categorical structure in meanings or compositional structure in signals, these experiments demonstrate that when the meaning space lacks clear preexisting boundaries, more subtle morphological structure that lacks straightforward compositionality—as found in natural languages—may evolve as a solution to joint pressures from learning and communication.

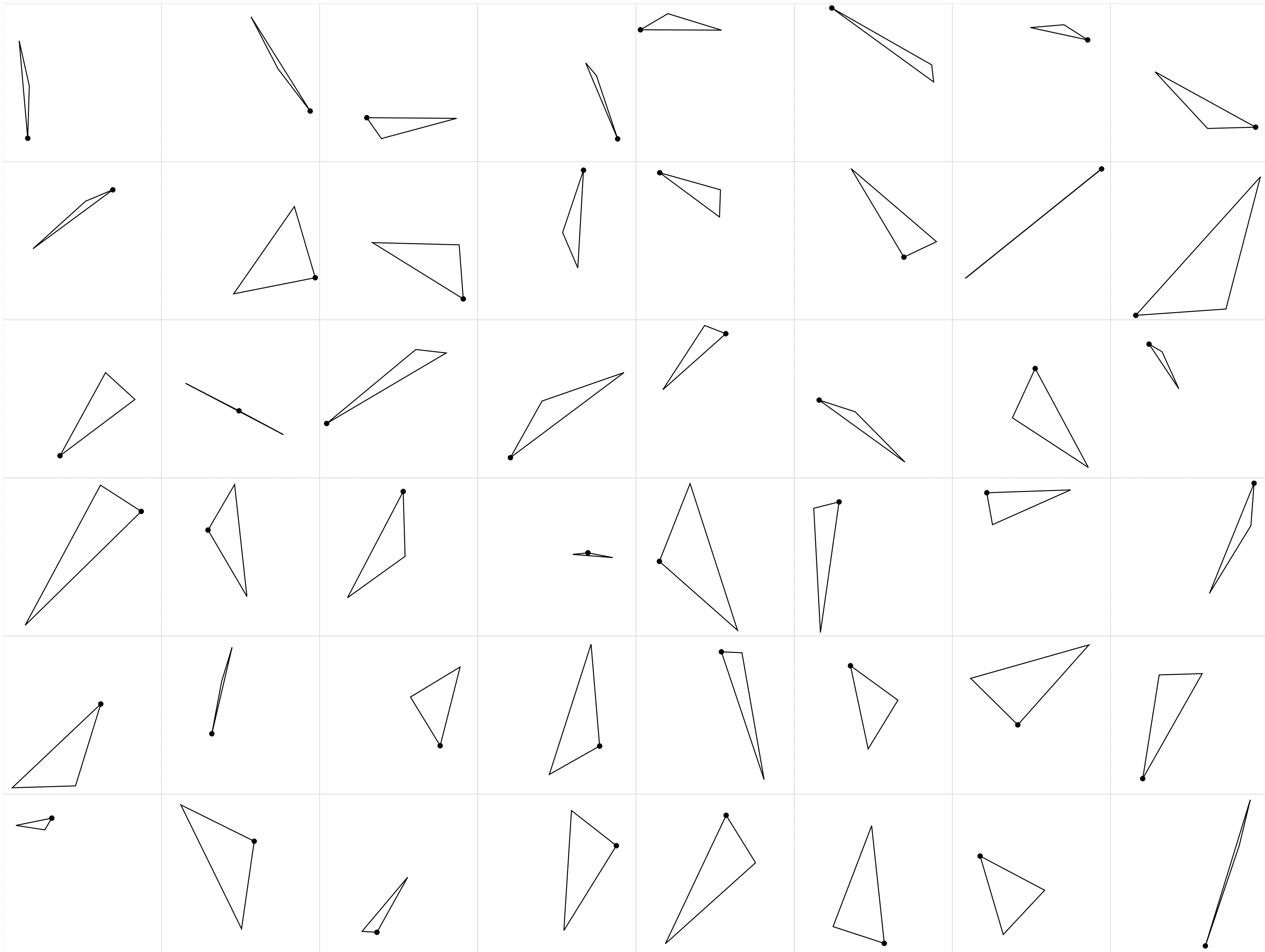
Keywords: Categorization; Communication; Compositionality; Cultural evolution; Iterated learning; Language evolution; Sound symbolism

1. Introduction

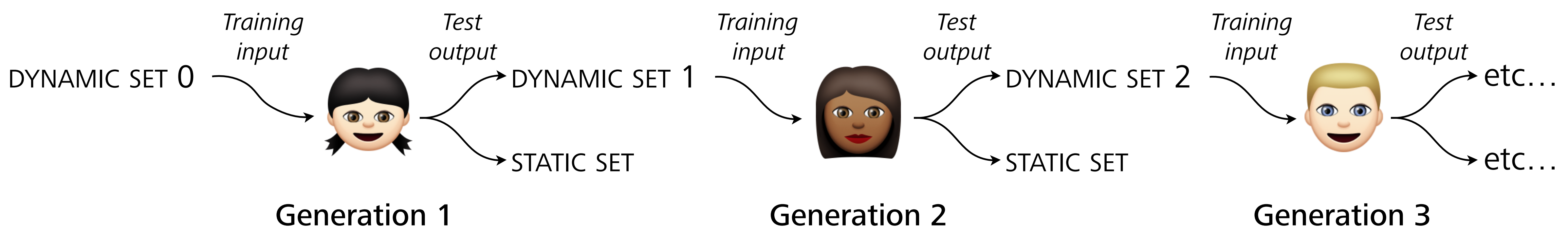
Discrete meaning space



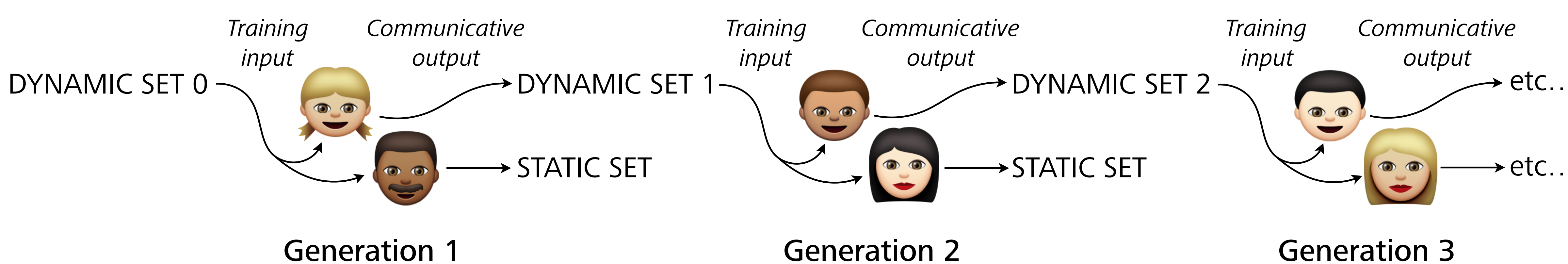
Open-ended meaning space

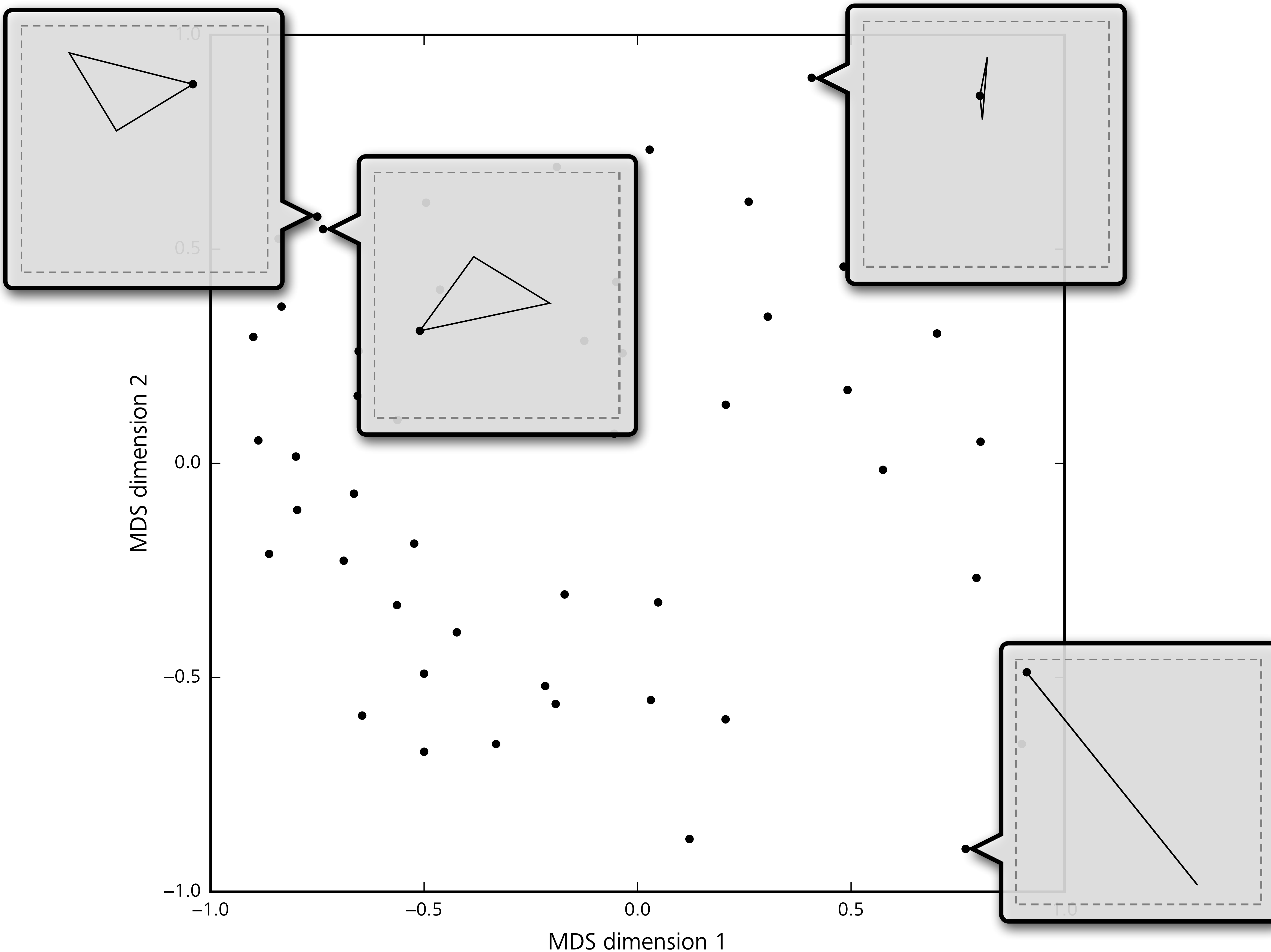


Experiment 1

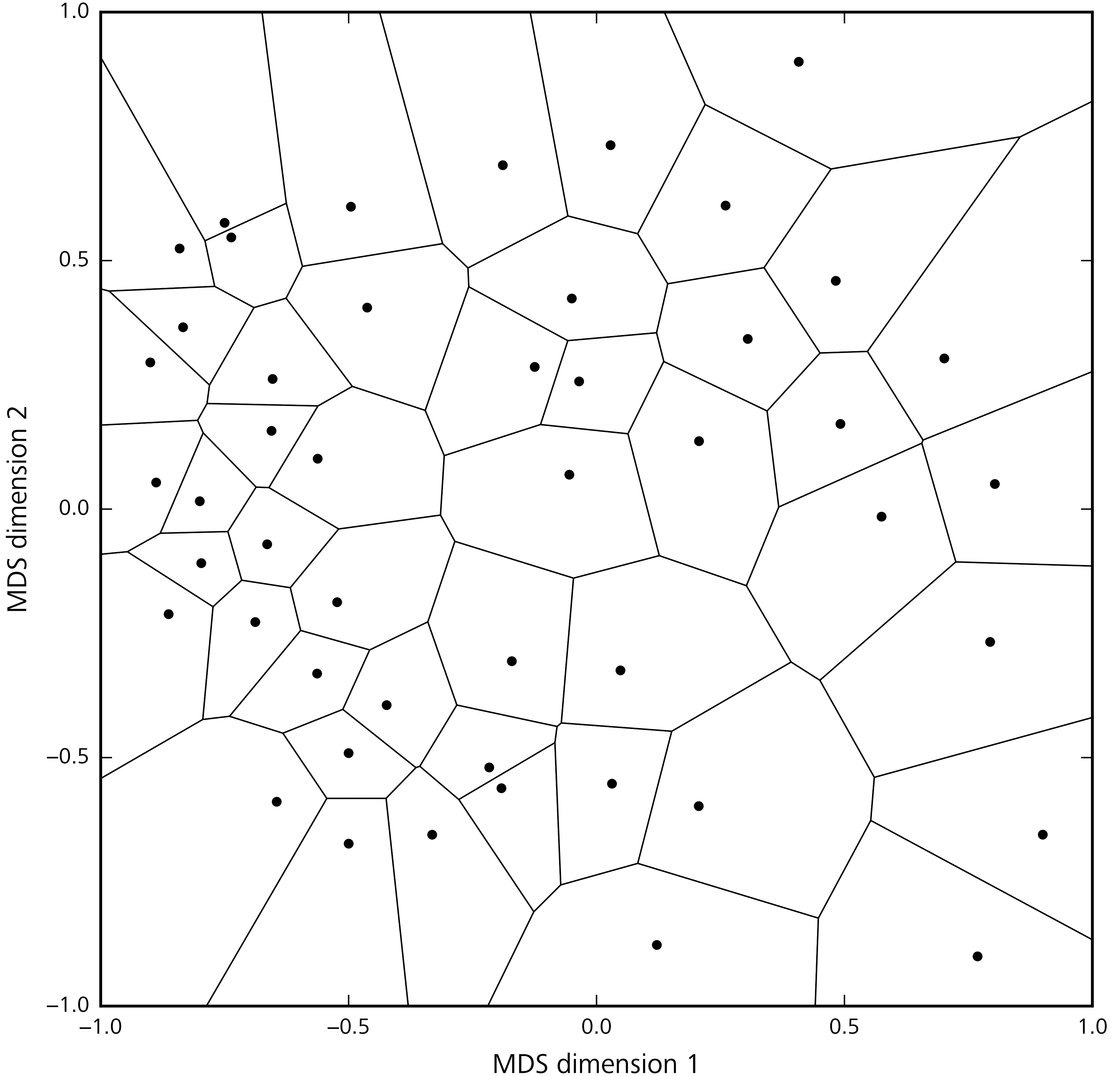


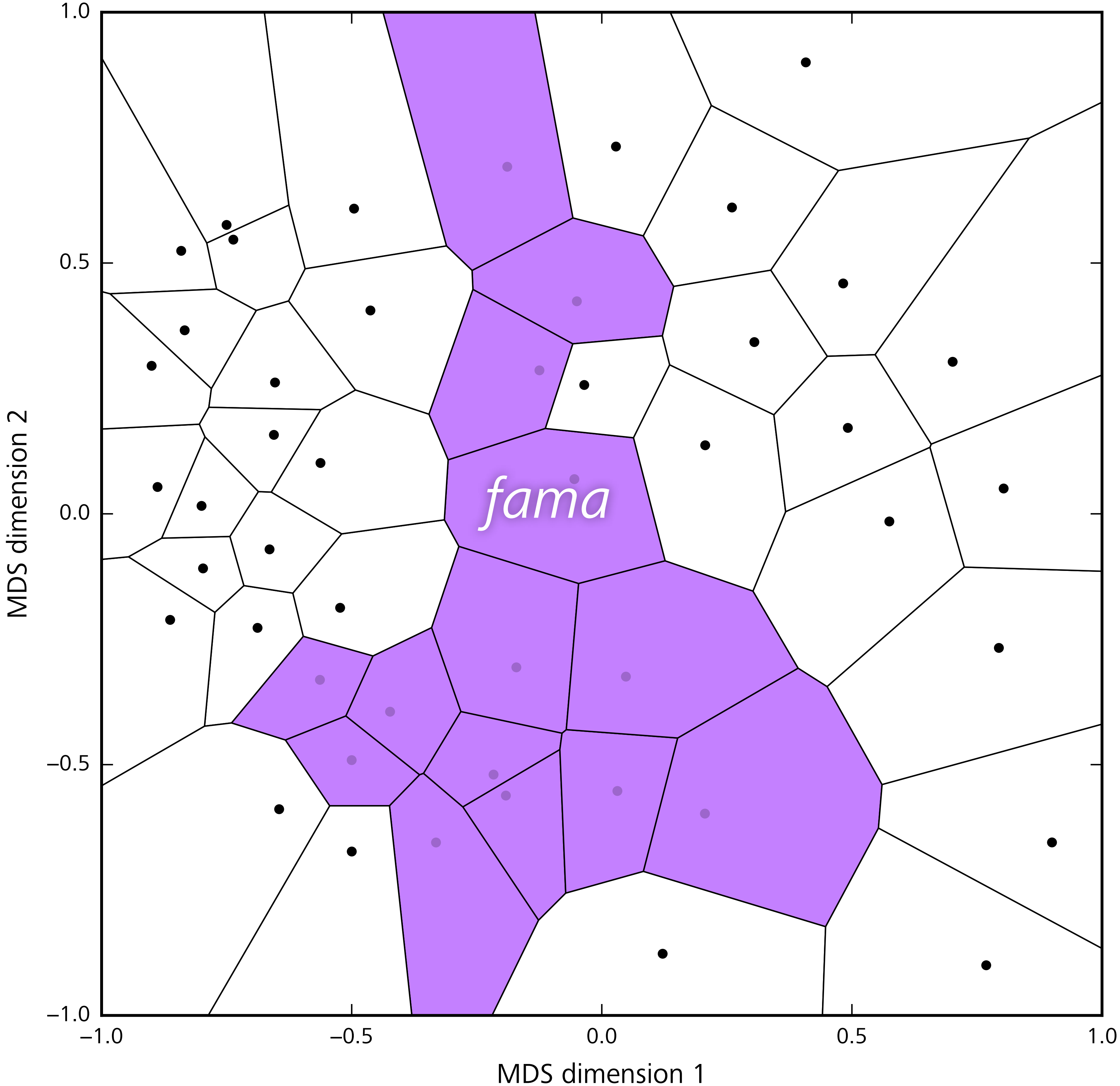
Experiment 2



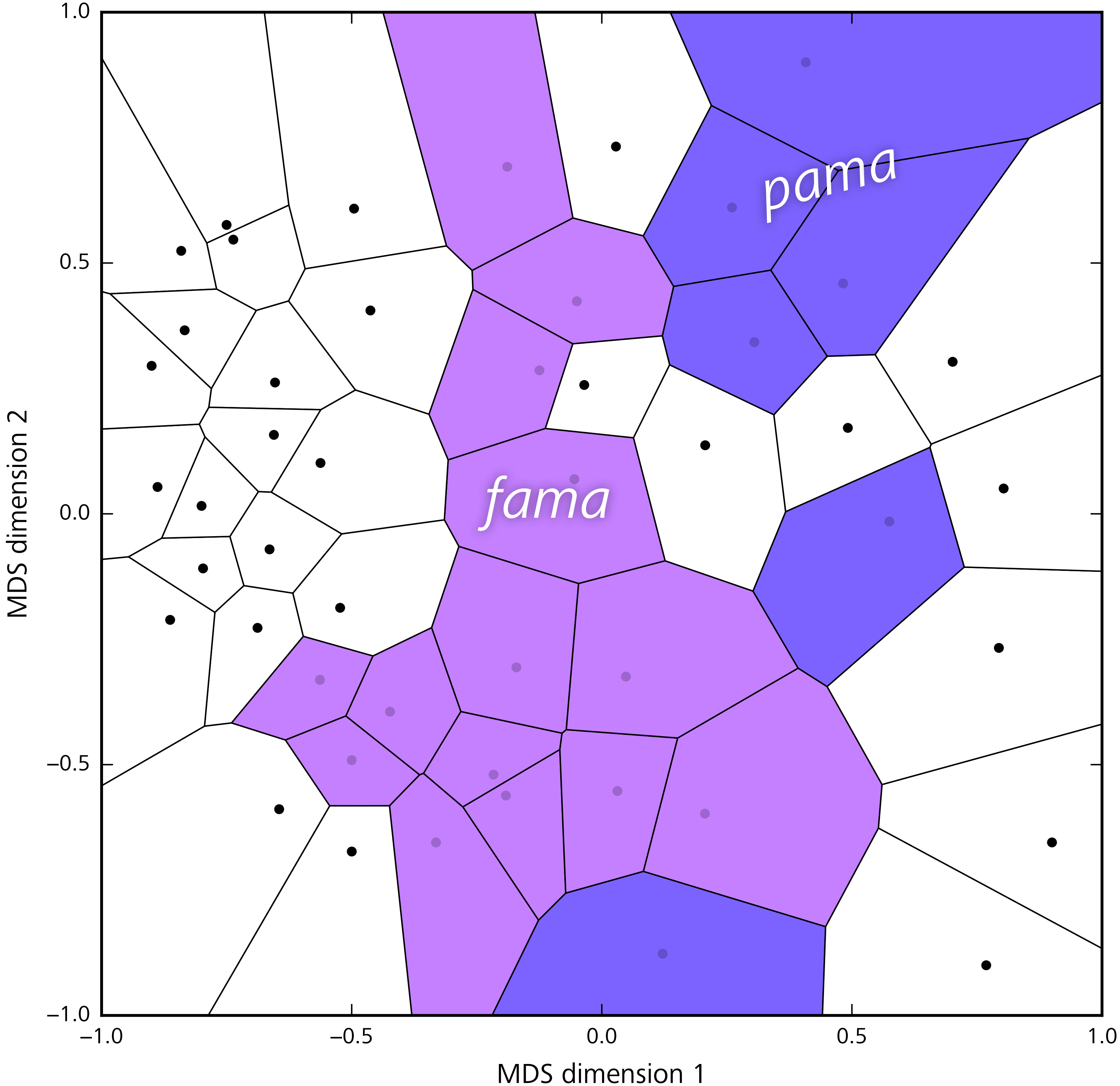


MDS dimension 2
MDS dimension 1
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

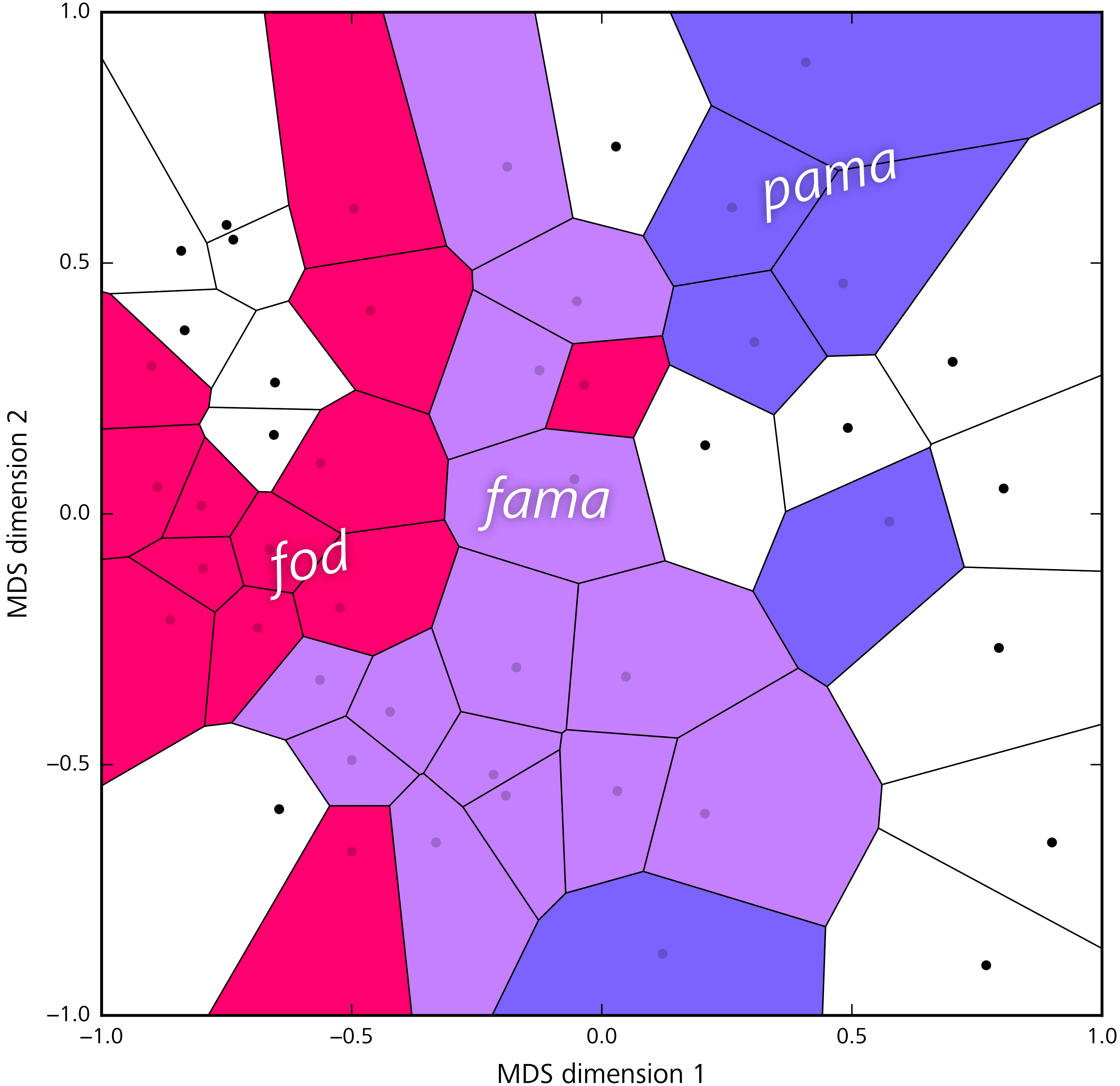

1.0
0.5
0.0
−0.5
−1.0
MDS dimension 2
−1.0
−0.5
0.0
0.5
1.0
MDS dimension 1



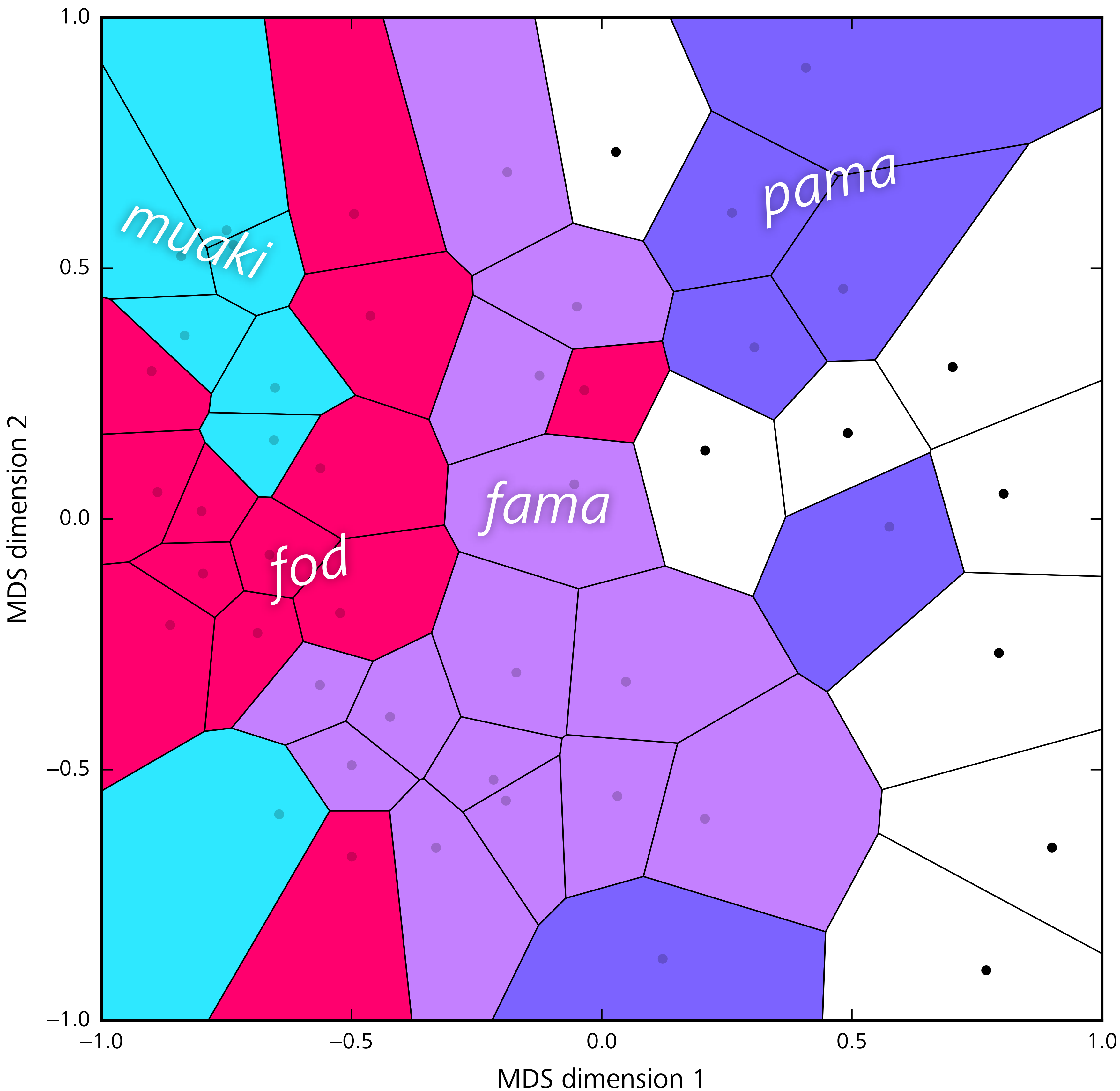
fama
MDS dimension 1
MDS dimension 2
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0



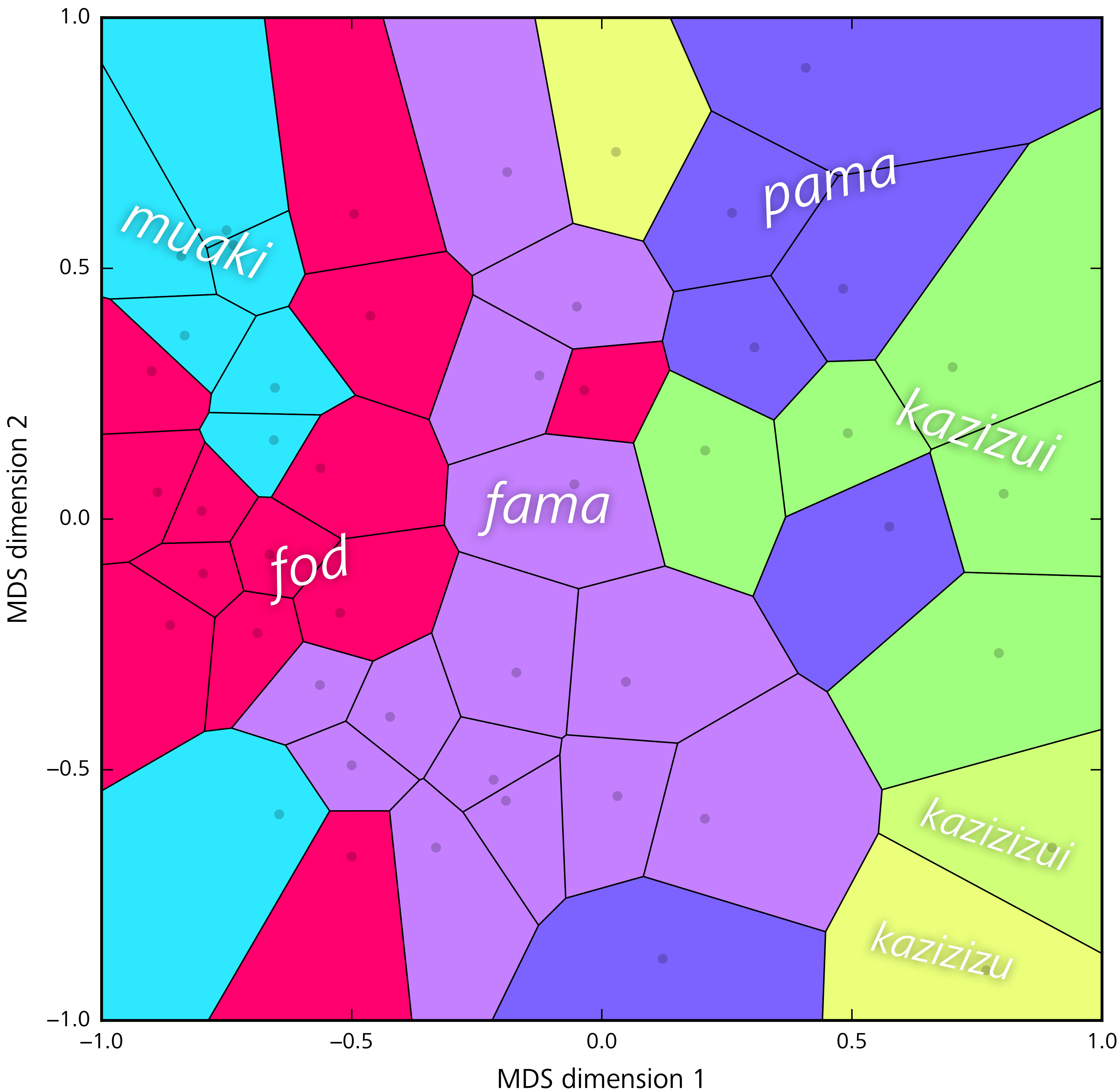
pama
fama
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
MDS dimension 1
MDS dimension 2

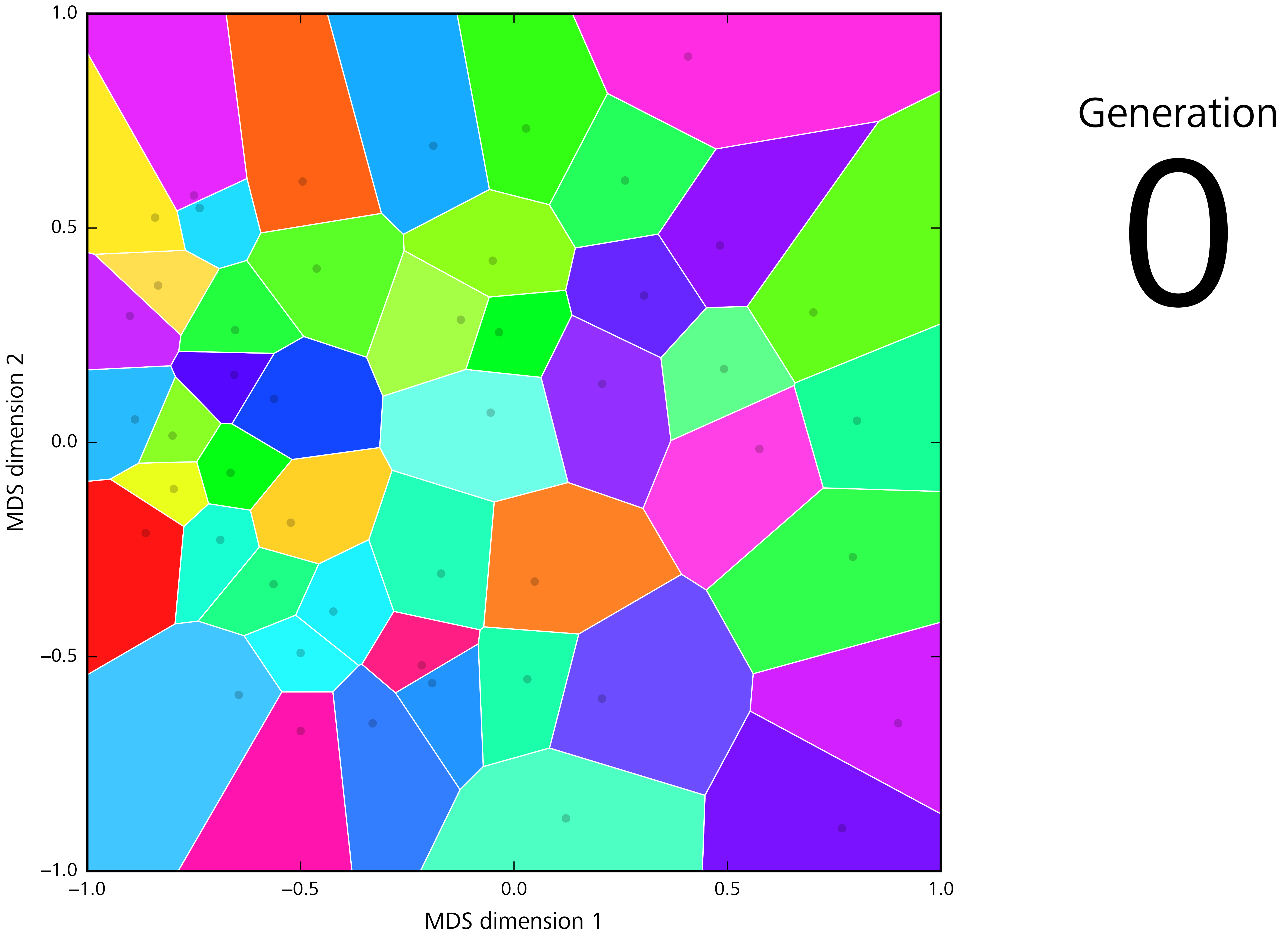


pama
fama
fod
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
MDS dimension 1
MDS dimension 2

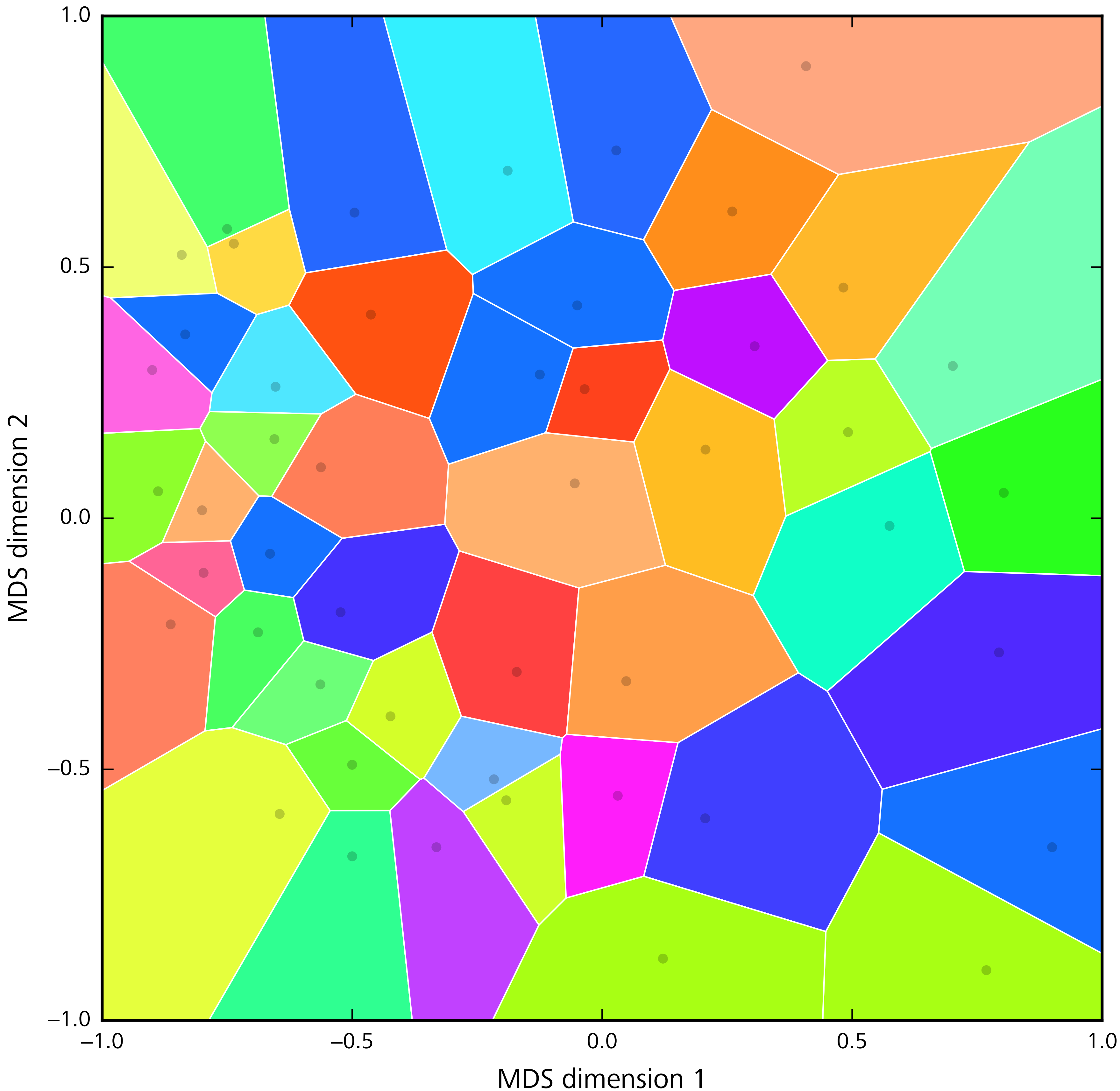


muaki
pama
fama
fod
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
MDS dimension 1
MDS dimension 2


muaki
pama
kazizui
fama
fod
kazizizui
kazizizu
MDS dimension 1
MDS dimension 2
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

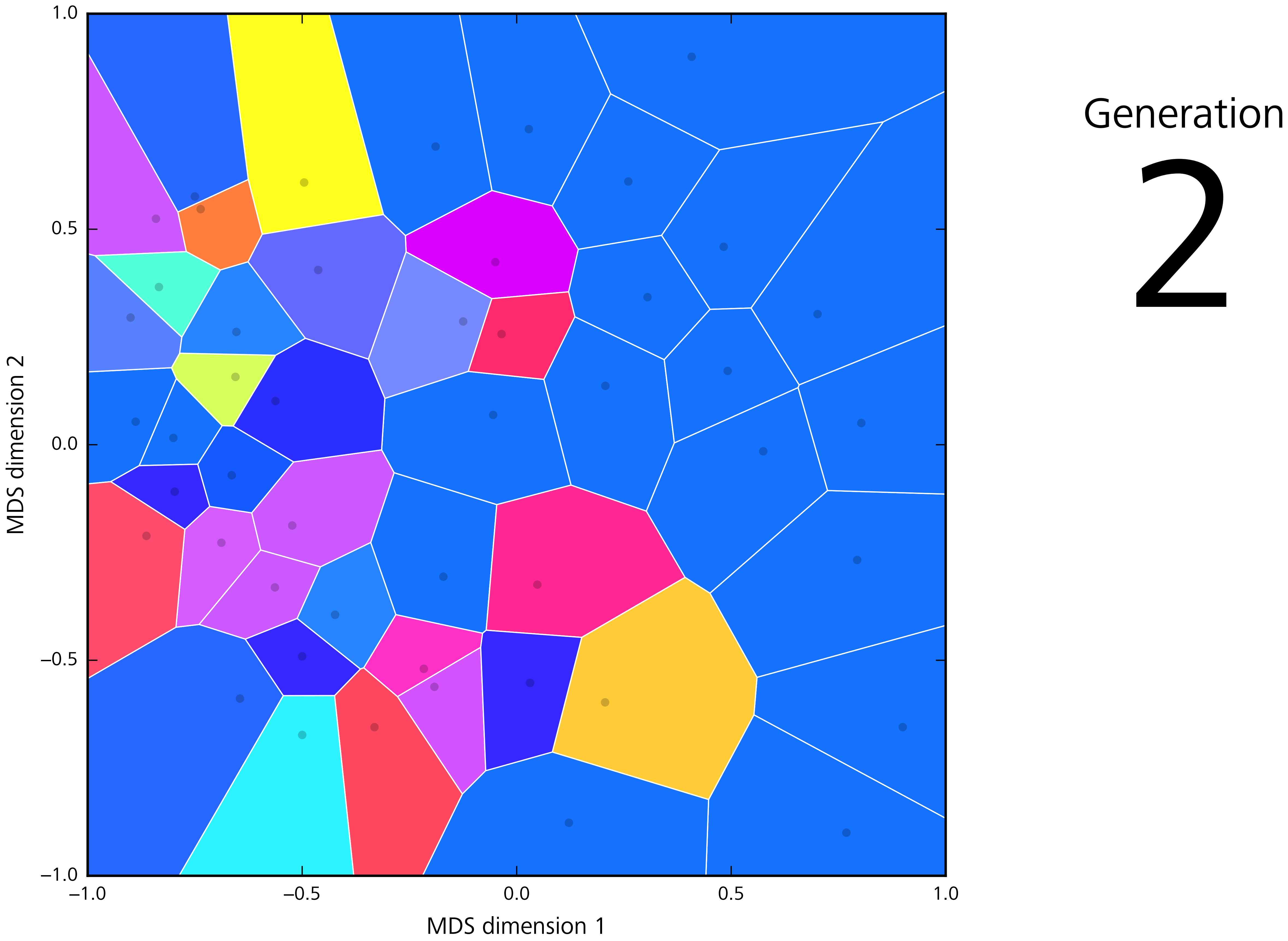


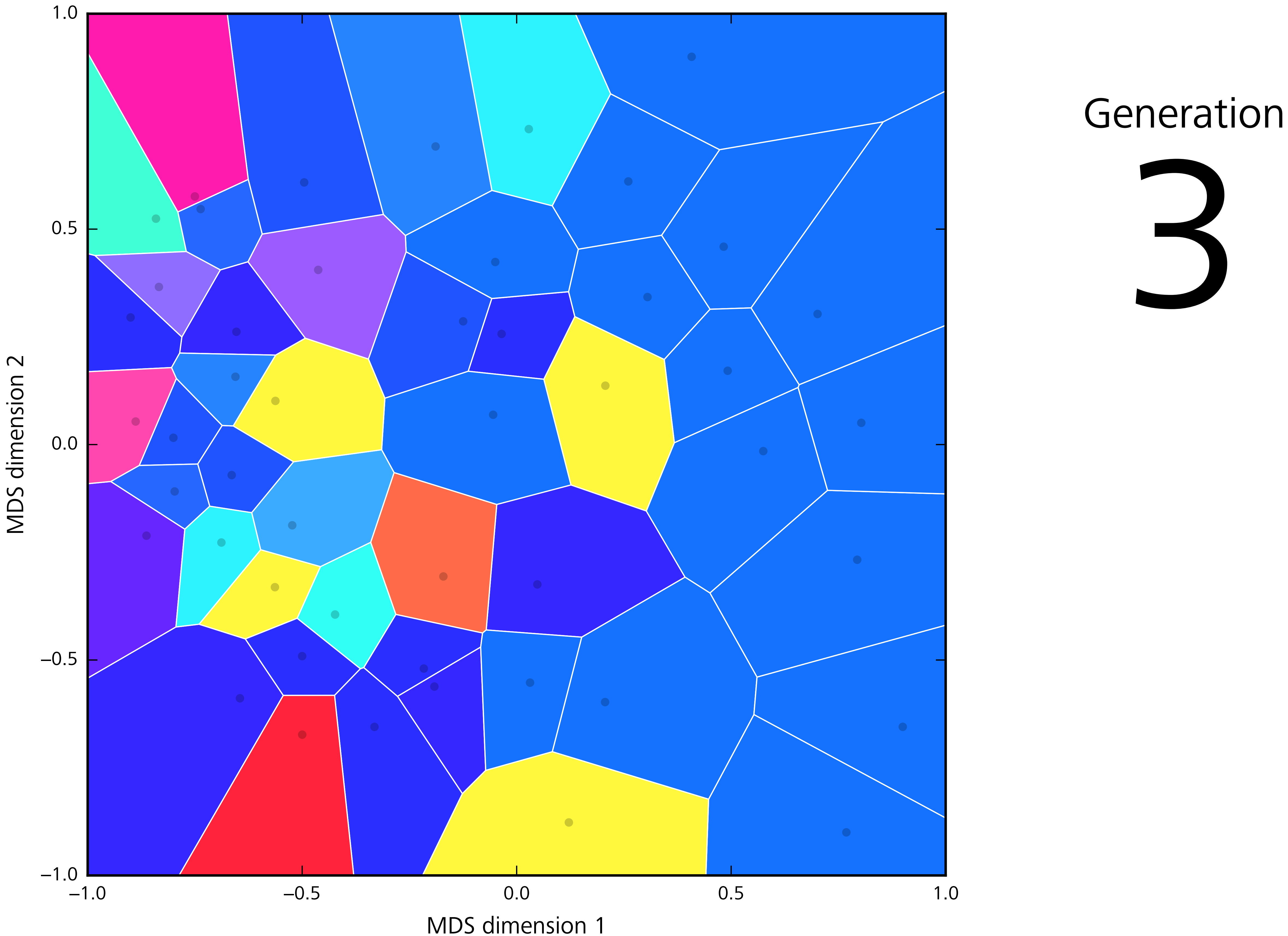
1.0
0.5
0.0
−0.5
−1.0
MDS dimension 2
−1.0
−0.5
0.0
0.5
1.0
MDS dimension 1
Generation
0

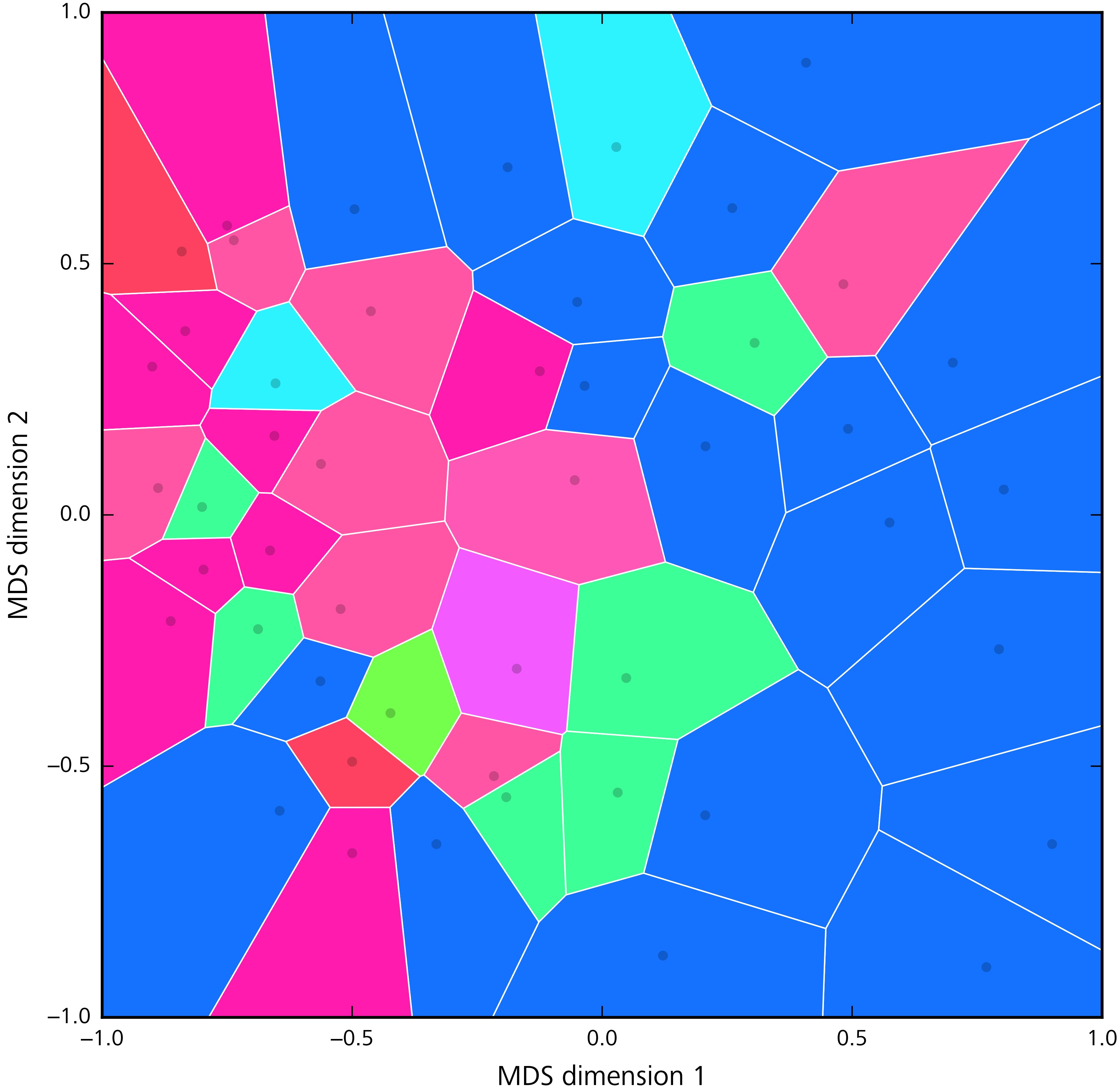


Generation

1

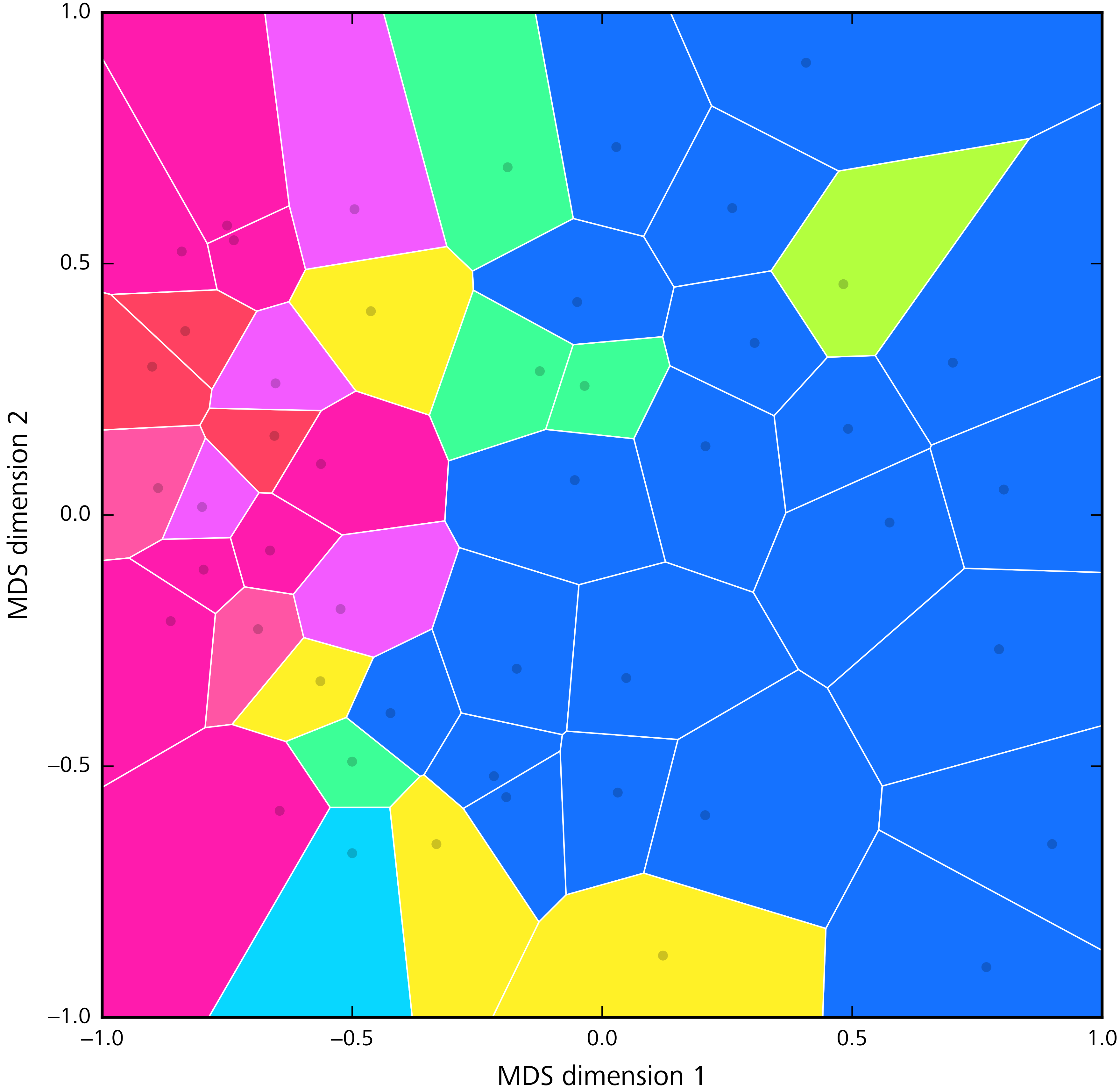

Generation
2
MDS dimension 1
MDS dimension 2
−1.0
−0.5
0.0
0.5
1.0


MDS dimension 2
MDS dimension 1
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
Generation
3



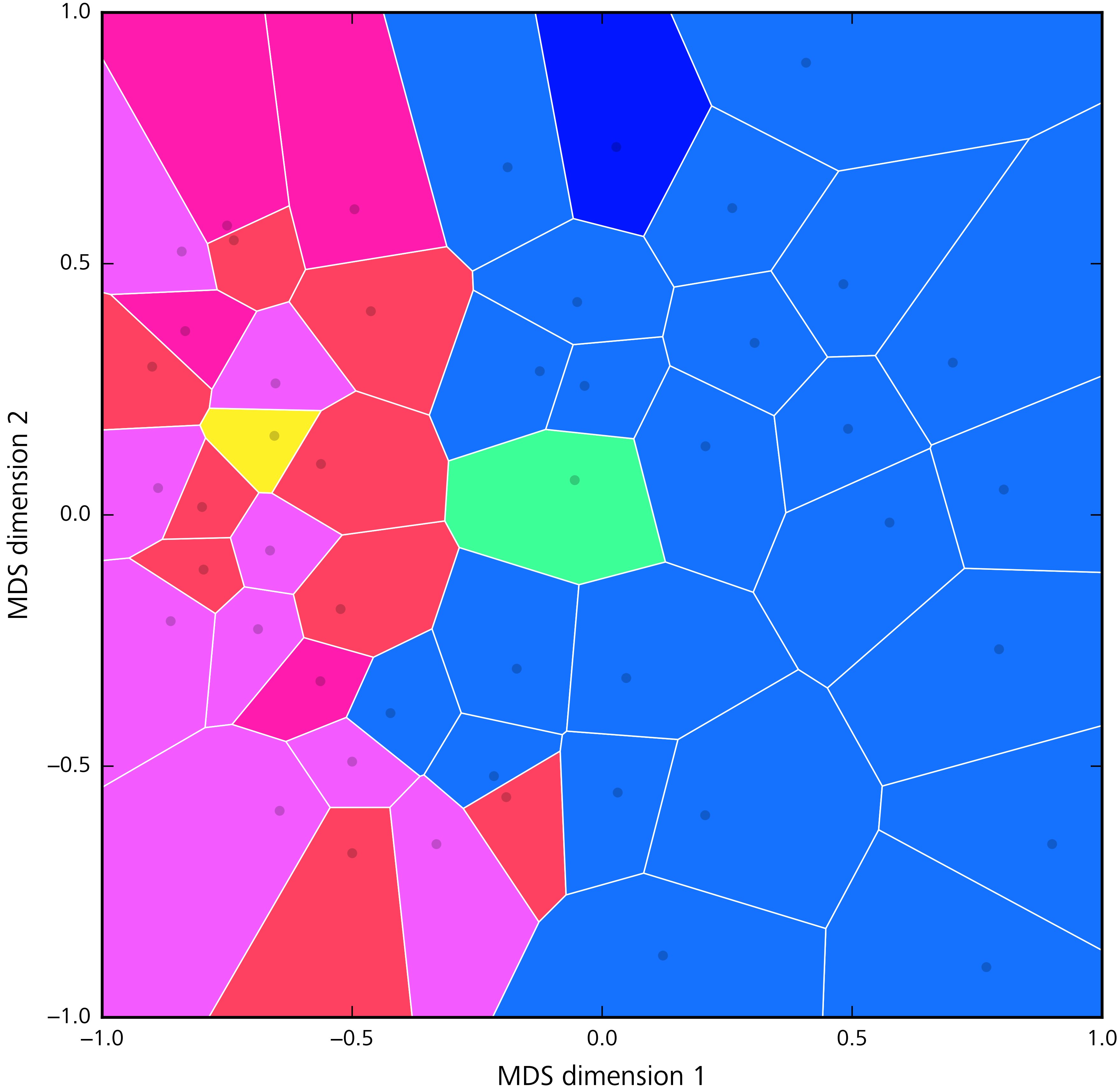
Generation

4



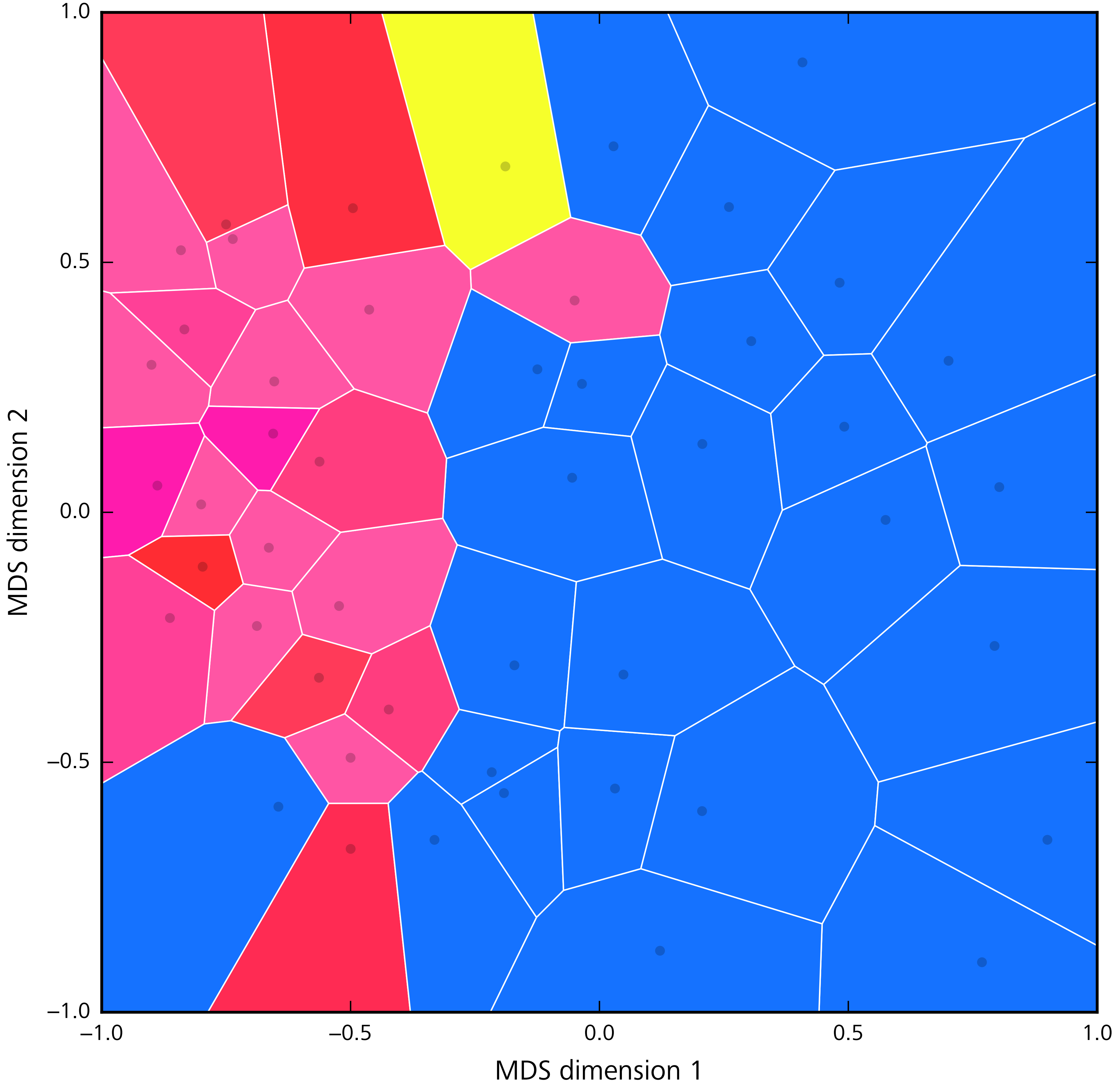
Generation

5



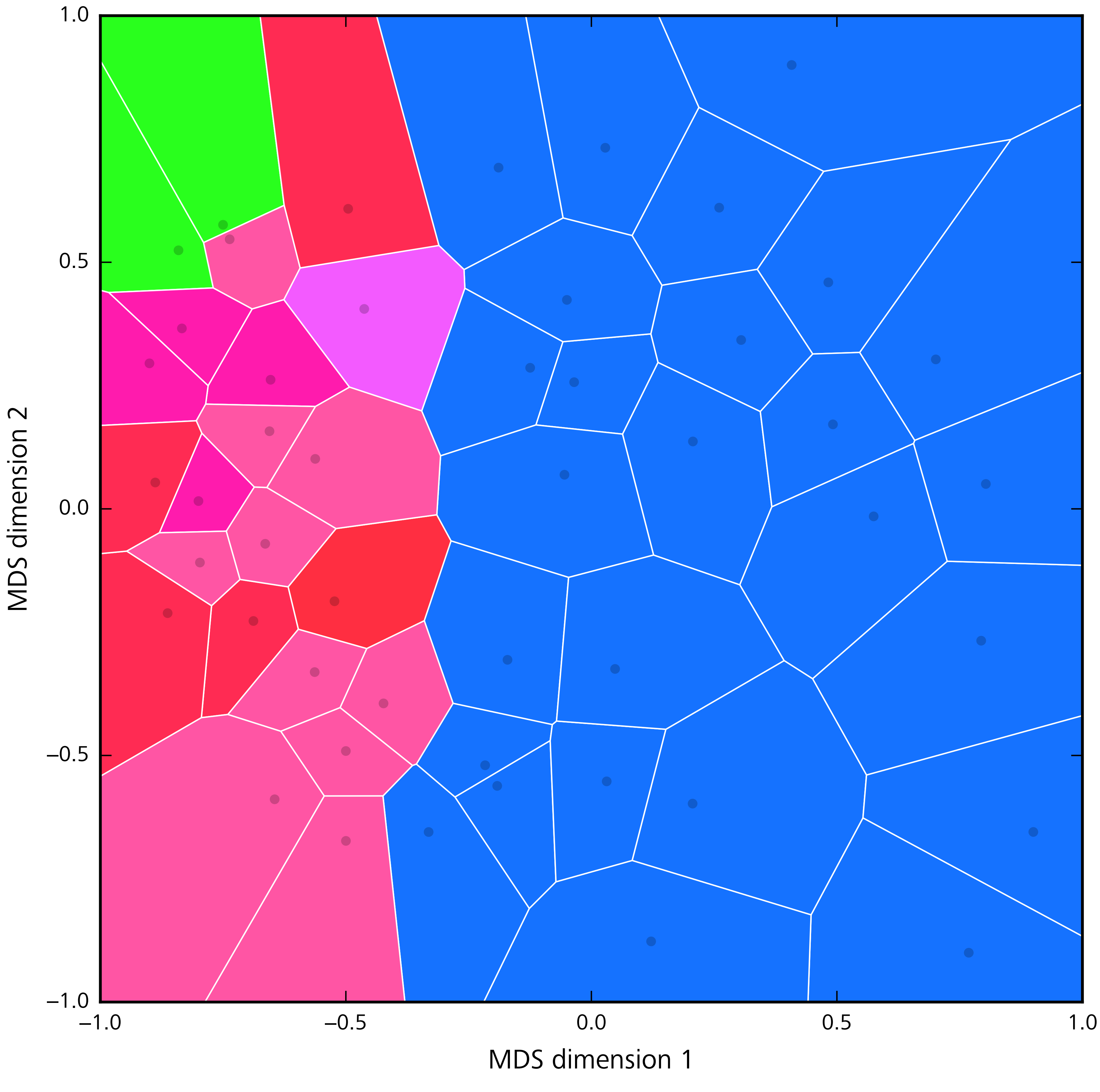
Generation

6



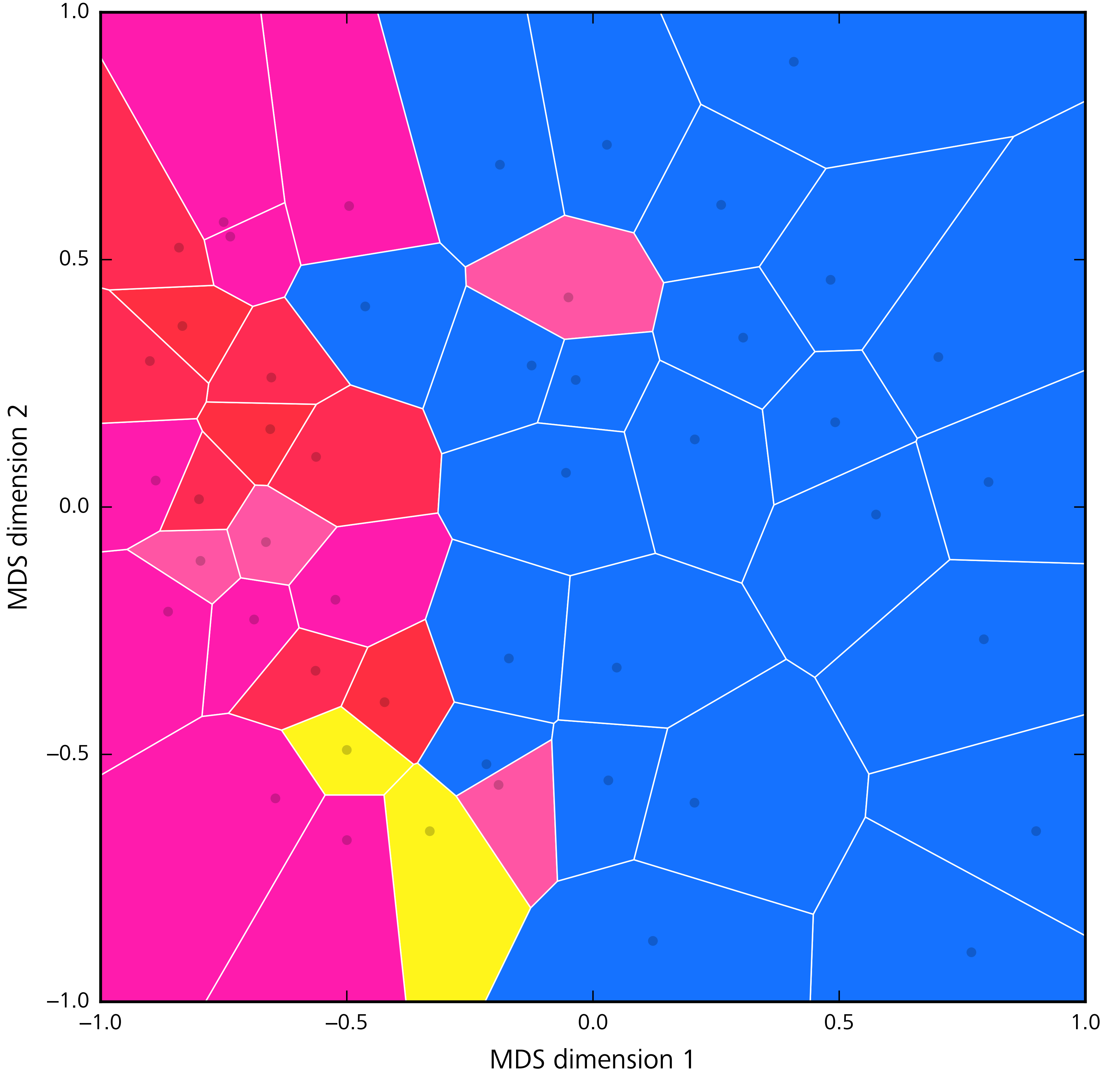
Generation

7



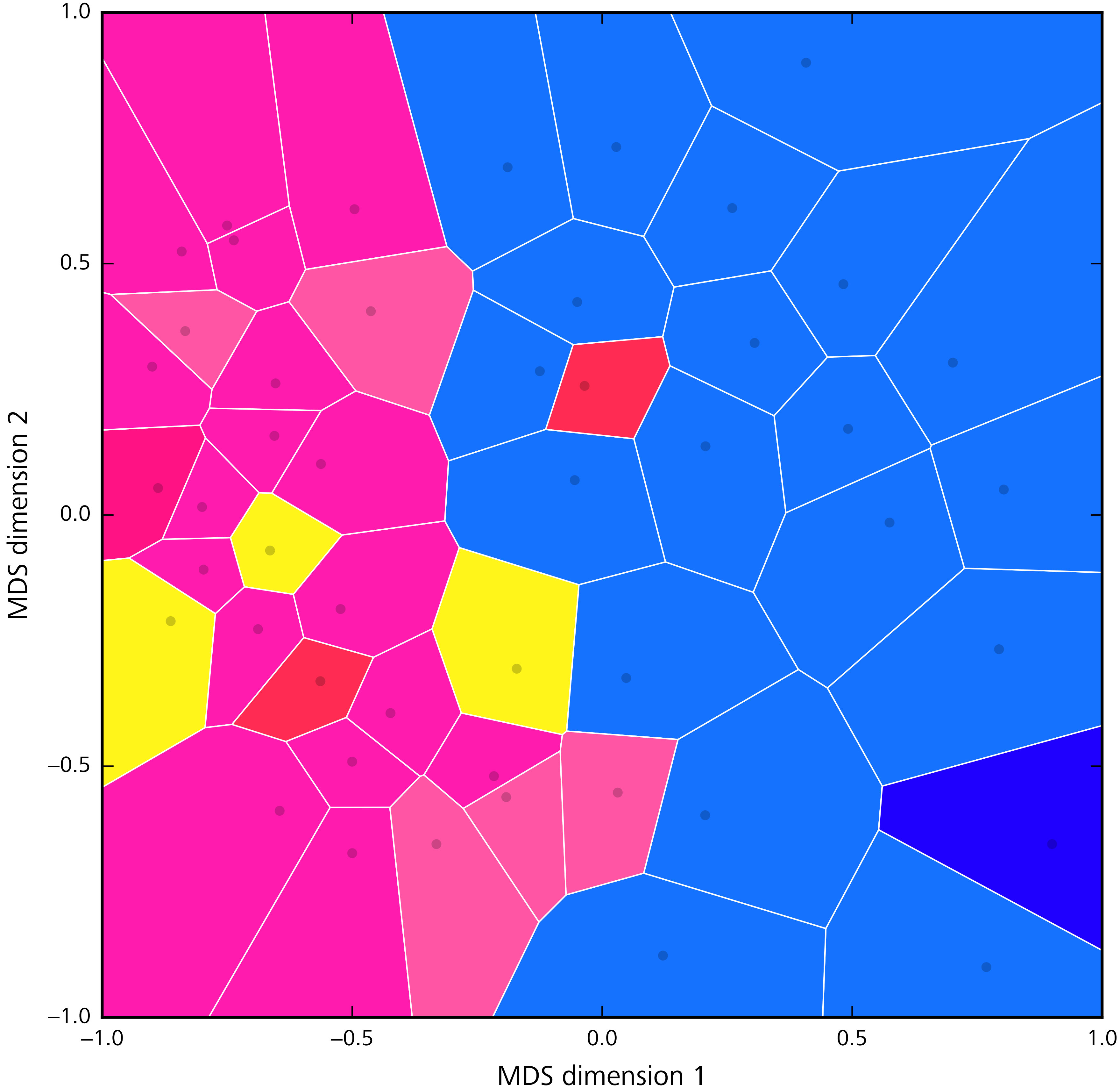
Generation

8



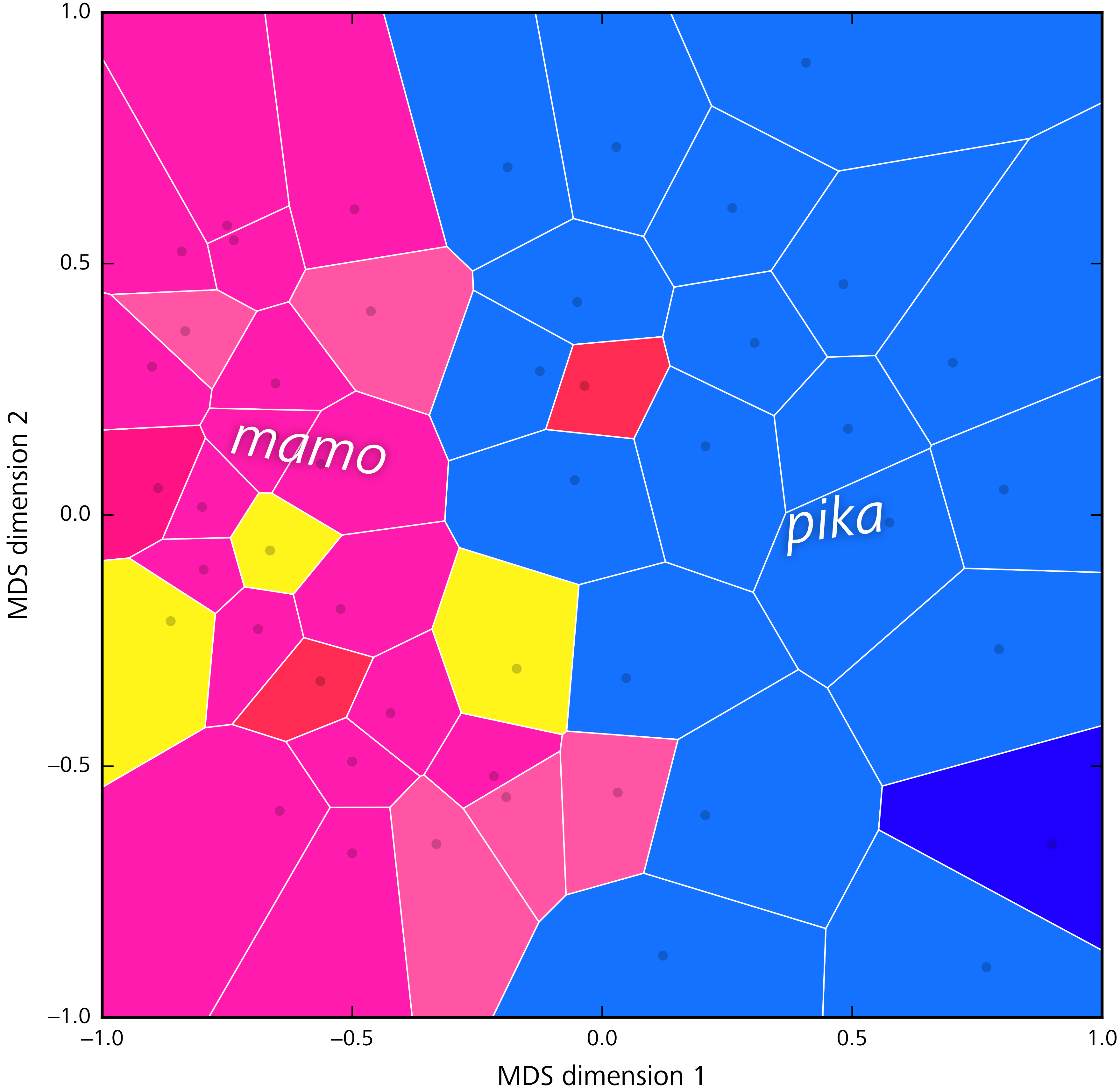
Generation

9



Generation

10



Generation

10

Conclusions

Experiment 1 showed that cultural evolution can deliver languages that categorize the meaning space under pressure from learnability.

This happens by losing categories and structuring the space in such a way that is easy to learn.

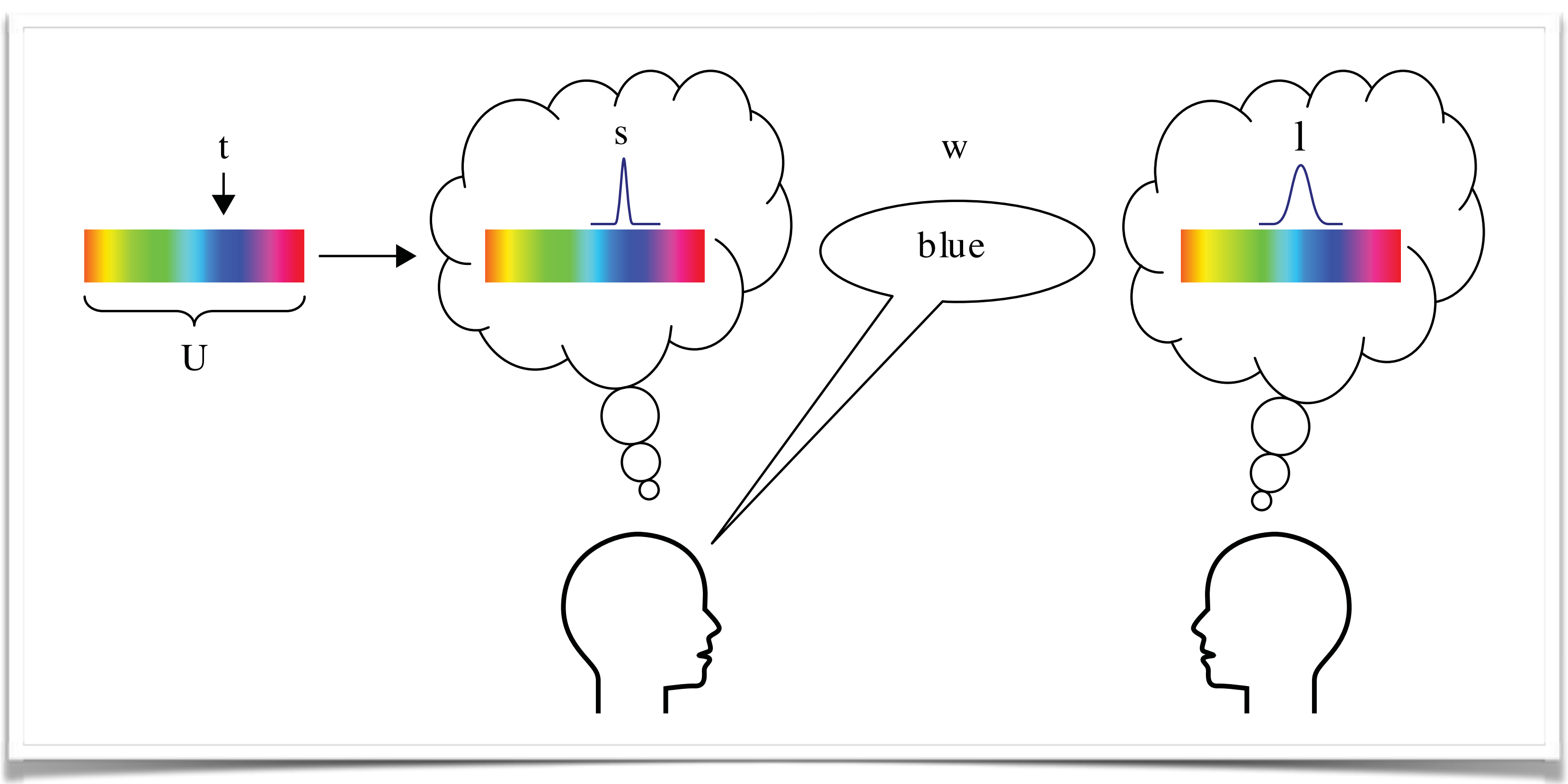
Experiment 2 combined a pressure for learnability **and** a pressure for expressivity derived from a genuine communicative task. This gave rise to languages that use both categorization and string-internal structure to be both learnable and expressive.

Ongoing work

Informativeness

Natural category systems "provide maximum information with the least cognitive effort" (Rosch, 1999).

Regier et al. formalize informativeness as "communicative cost". The most informative category system is one that **minimizes** communicative cost.



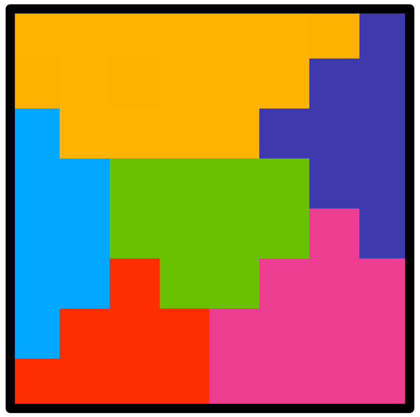
Informativeness

Natural category systems "provide maximum information with the least cognitive effort" (Rosch, 1999).

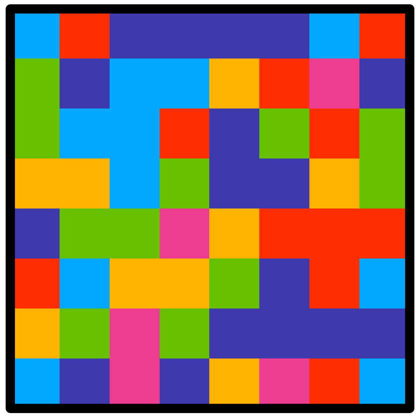
Regier et al. formalize informativeness as "communicative cost". The most informative category system is one that **minimizes** communicative cost.

Negative logarithm of average within-category similarity, summed for all possible targets.

$$\sum_{t \in T} -\log_2\left(\frac{1}{|C_t|}\sum_{c \in C_t} e^{-s \cdot d_{t,c}^2}\right)$$

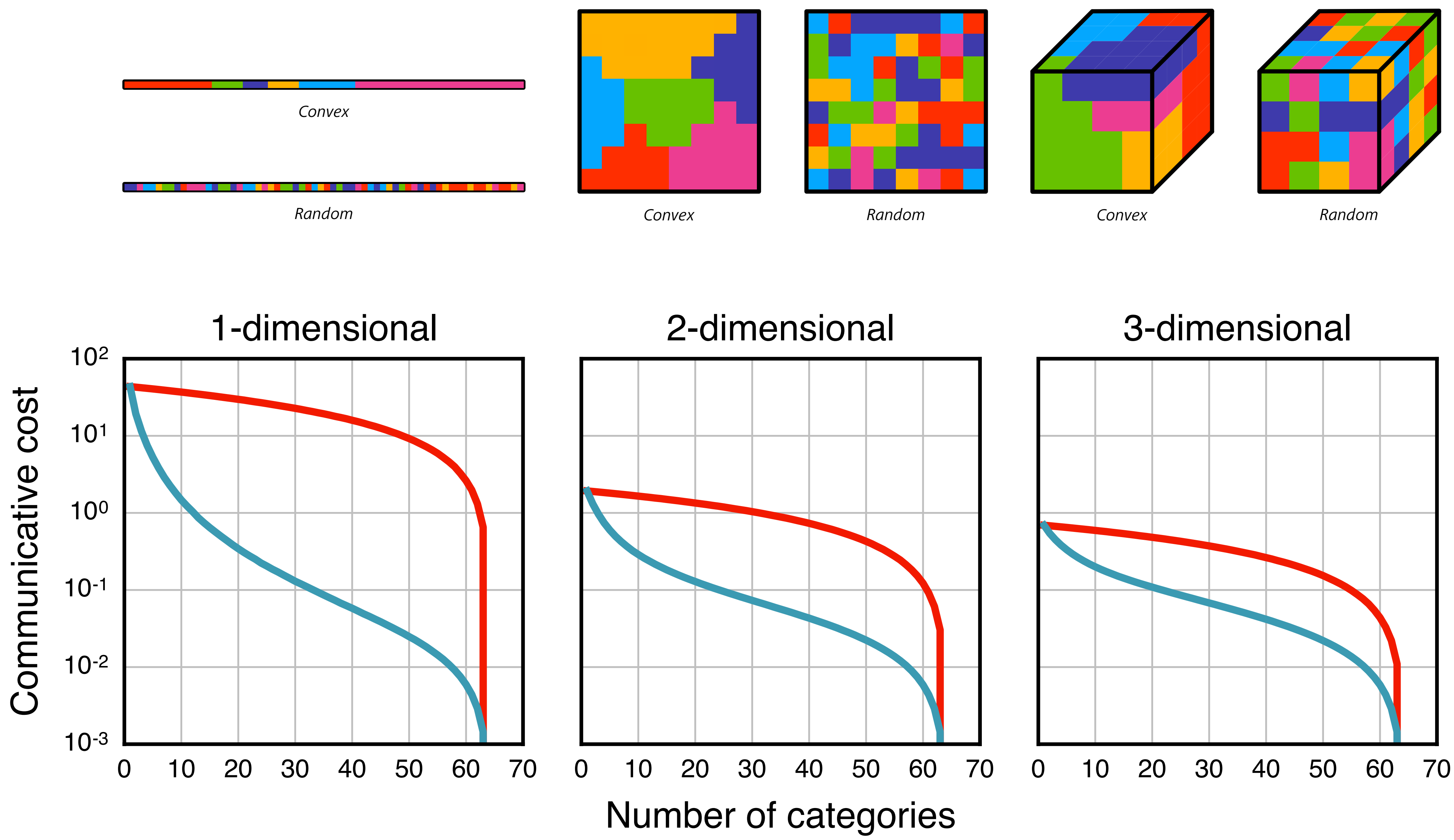


Convex



Random

Informativeness

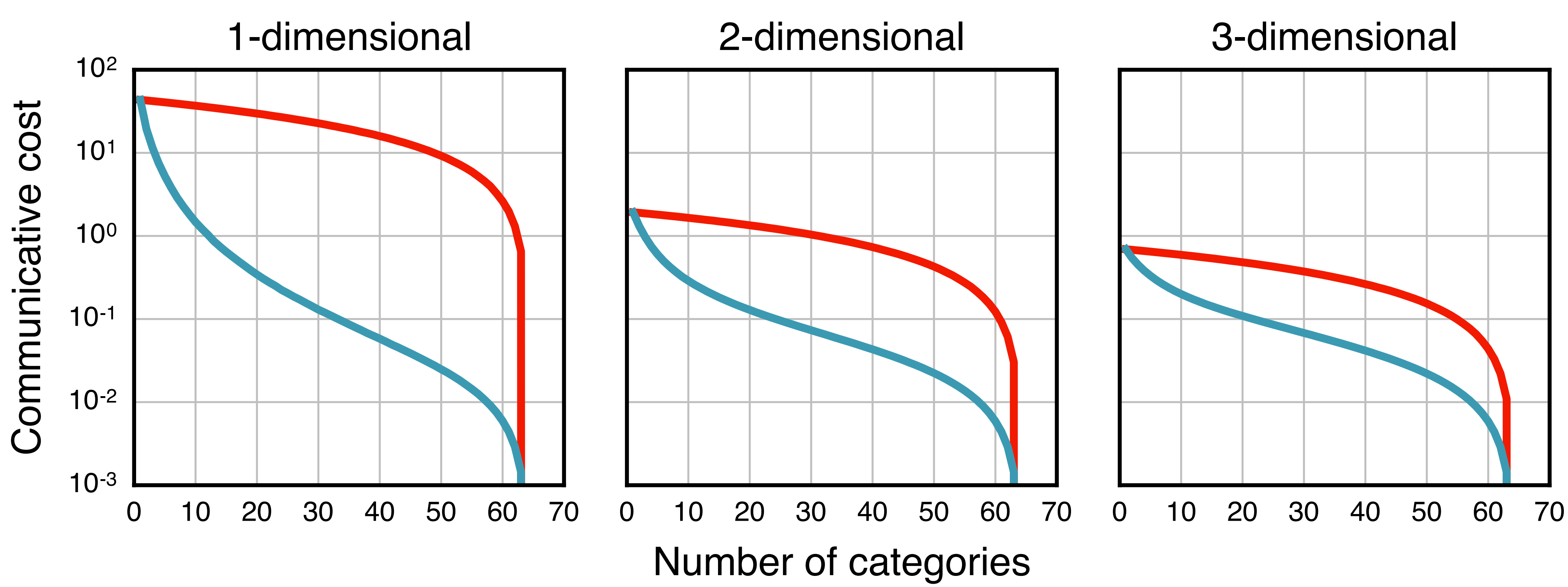


Three predictions

Maximize number of categories The use of a single category has the highest cost and is therefore the least informative system; placing every item in its own category reduces the cost to 0 (maximum informativeness).

Maximize dimensionality Representing 64 items using three dimensions is more informative than representing 64 items using one dimension (for a given number of categories).

Maximize convexity A convex category system is always better than or equal to a non-convex system in terms of minimizing communicative cost. Convex category structures are optimal.



Except: The learnability tradeoff

Maximize number of categories But: Learning an infinite number of categories is not possible given finite time and cognitive resources.

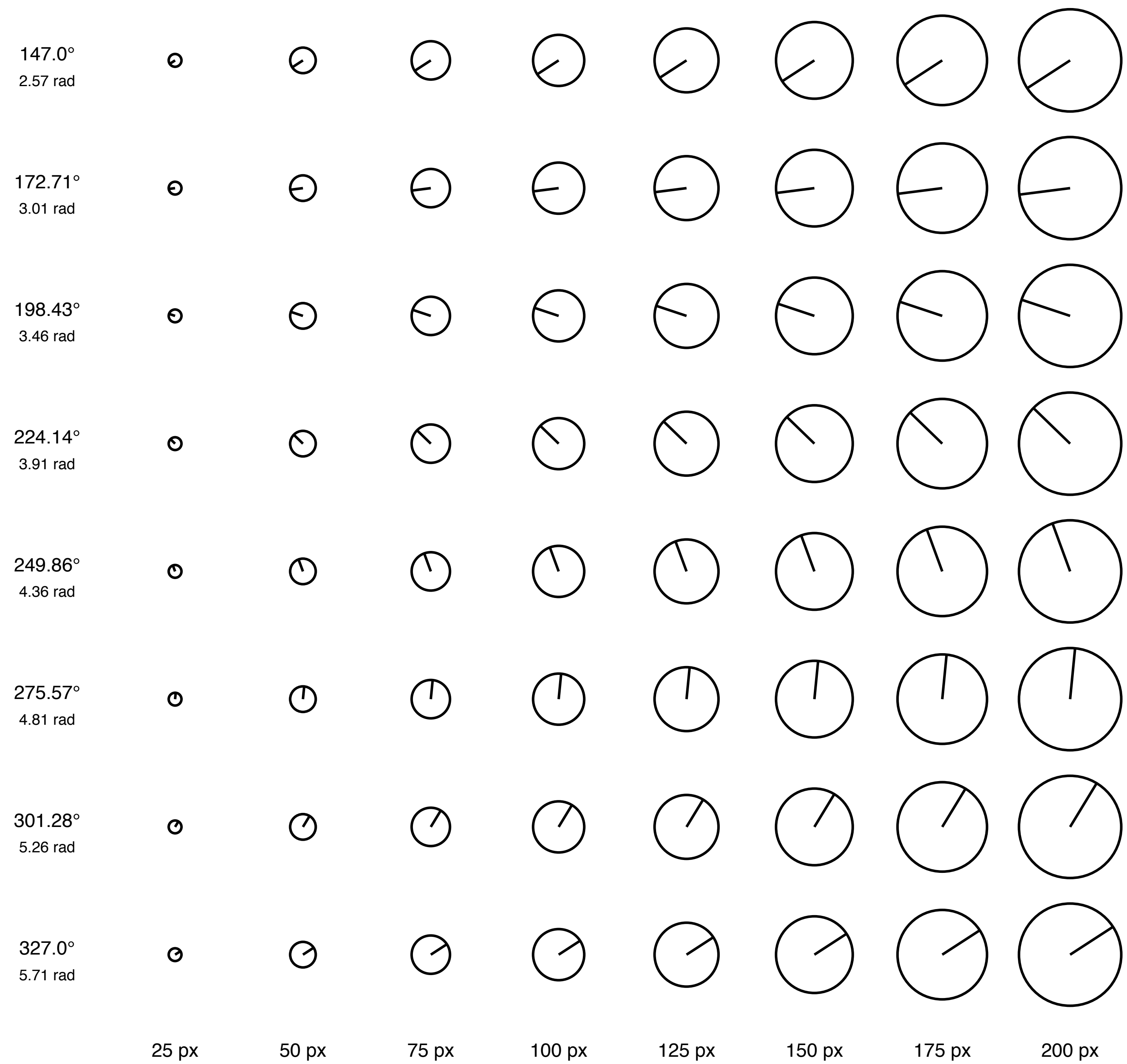
Maximize dimensionality But: Representing categories using infinite feature dimensions would be impossible to process.

However:

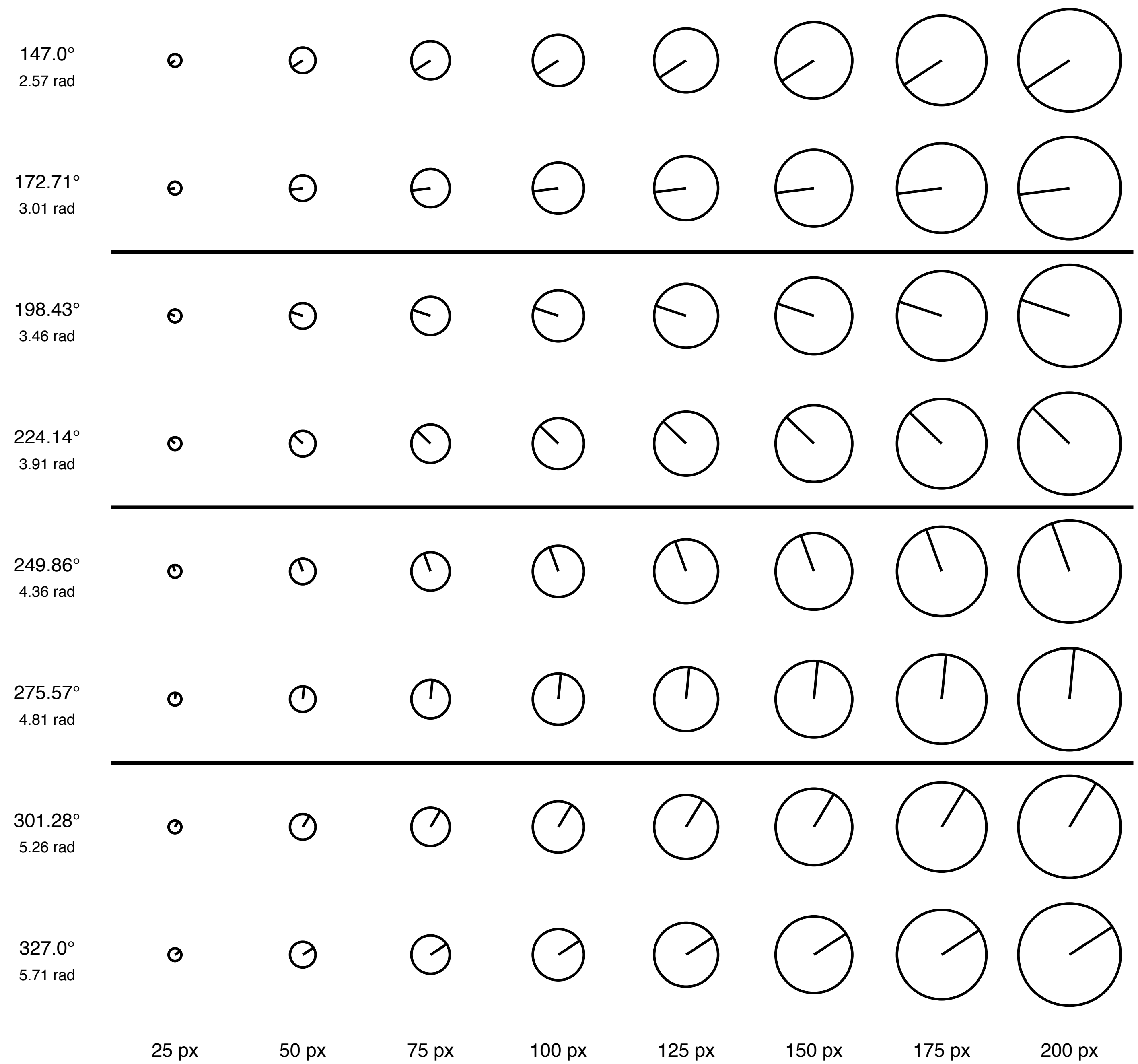
Maximize convexity Convexity leads to systems that are both more informative and potentially easier to learn.

Thus, the property of convexity seems to be particularly interesting (Gärdenfors, 2000, 2014).

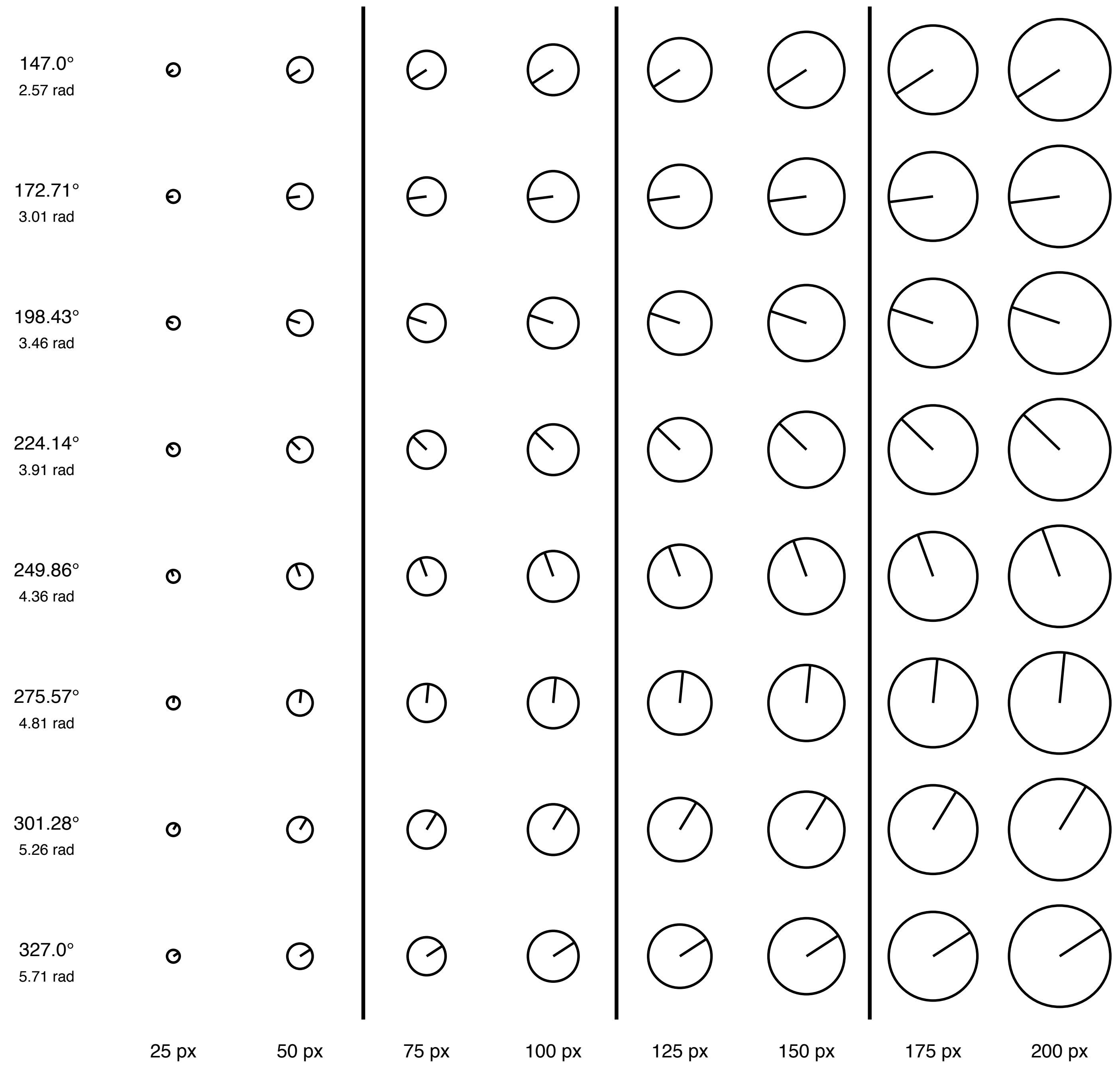
Stims: Shepard circles



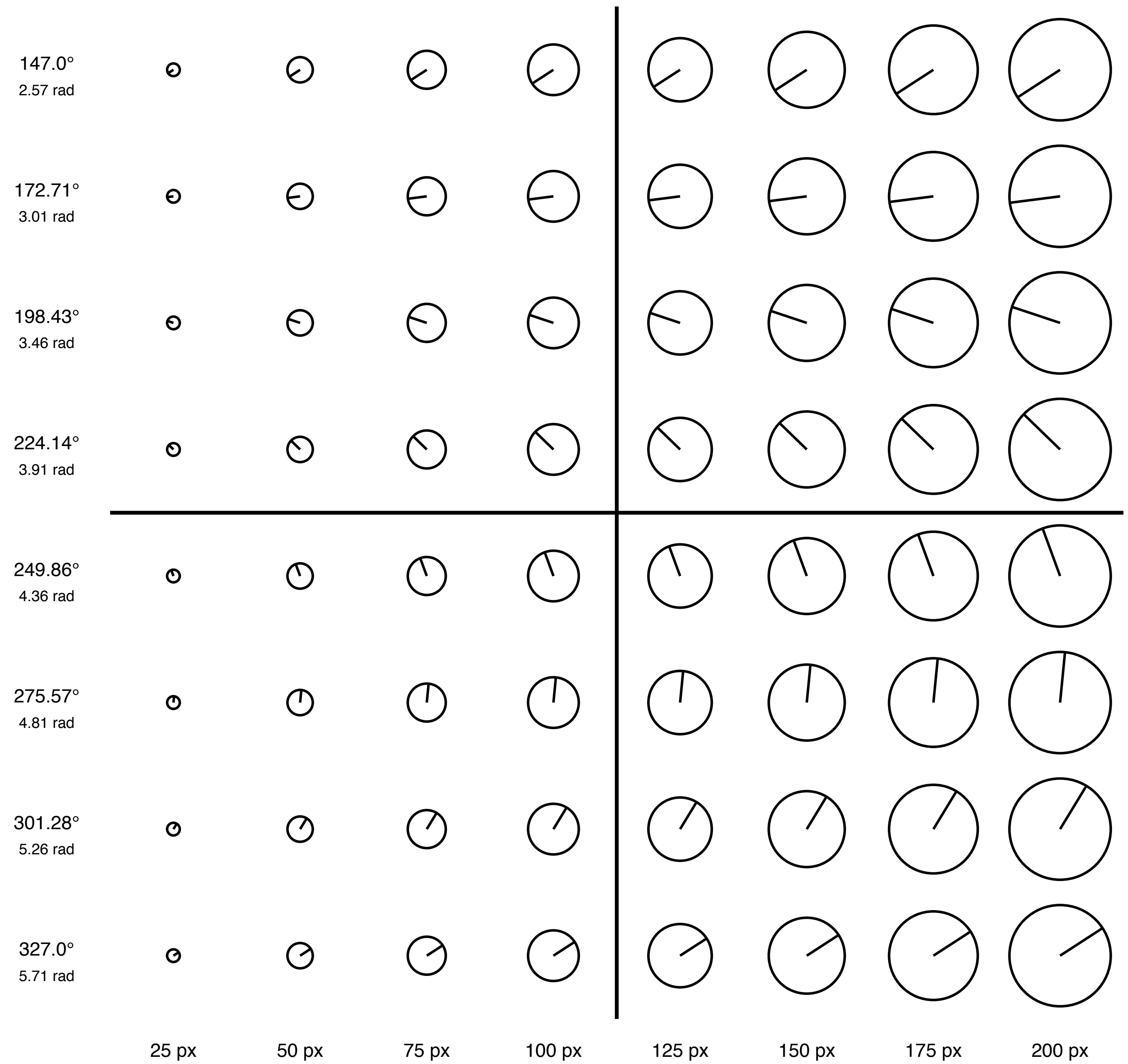
Stims: Shepard circles



Stims: Shepard circles

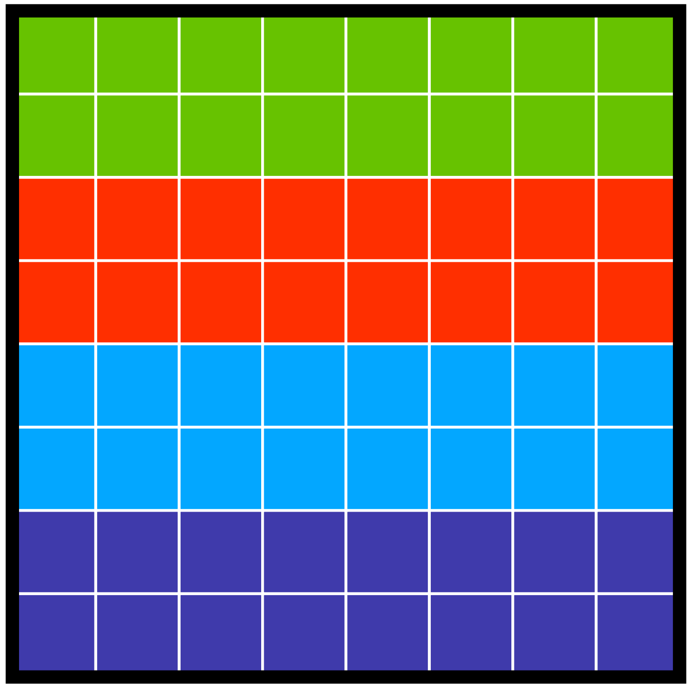


Stims: Shepard circles

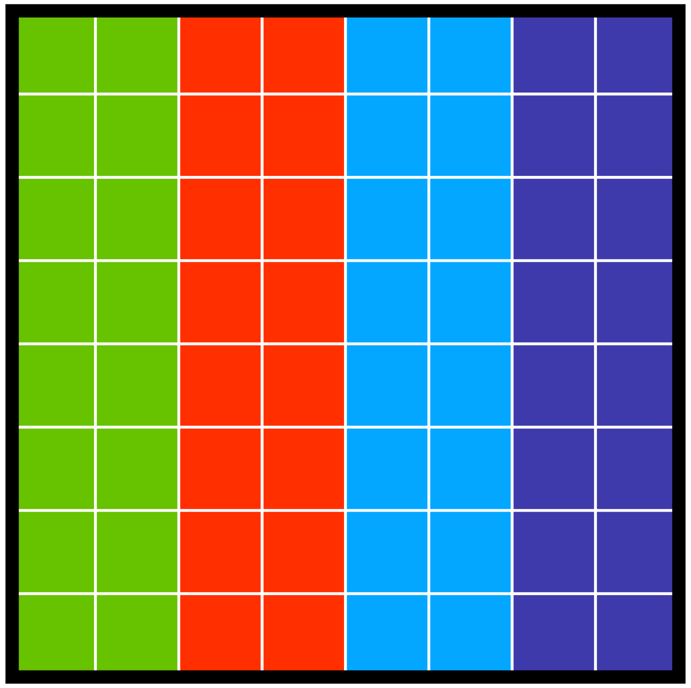


Squares and Stripes: Three category systems

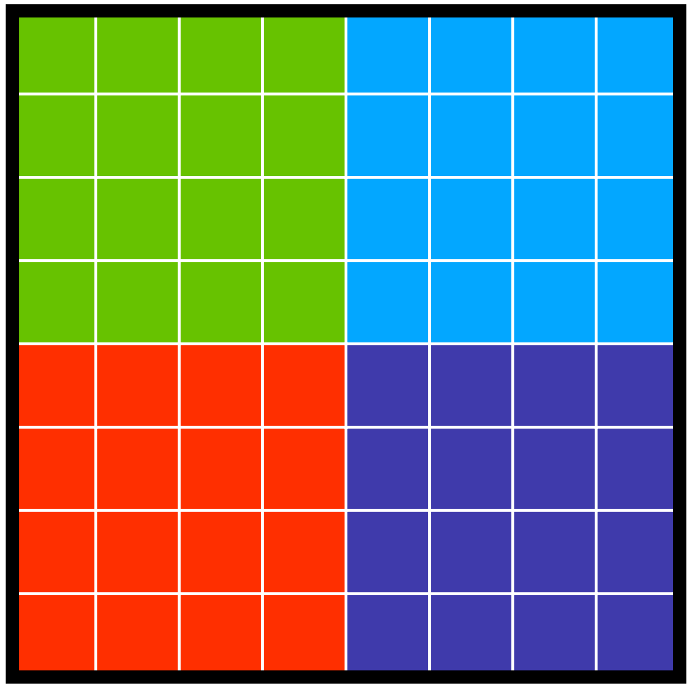
Angle-only



Size-only



Angle & Size

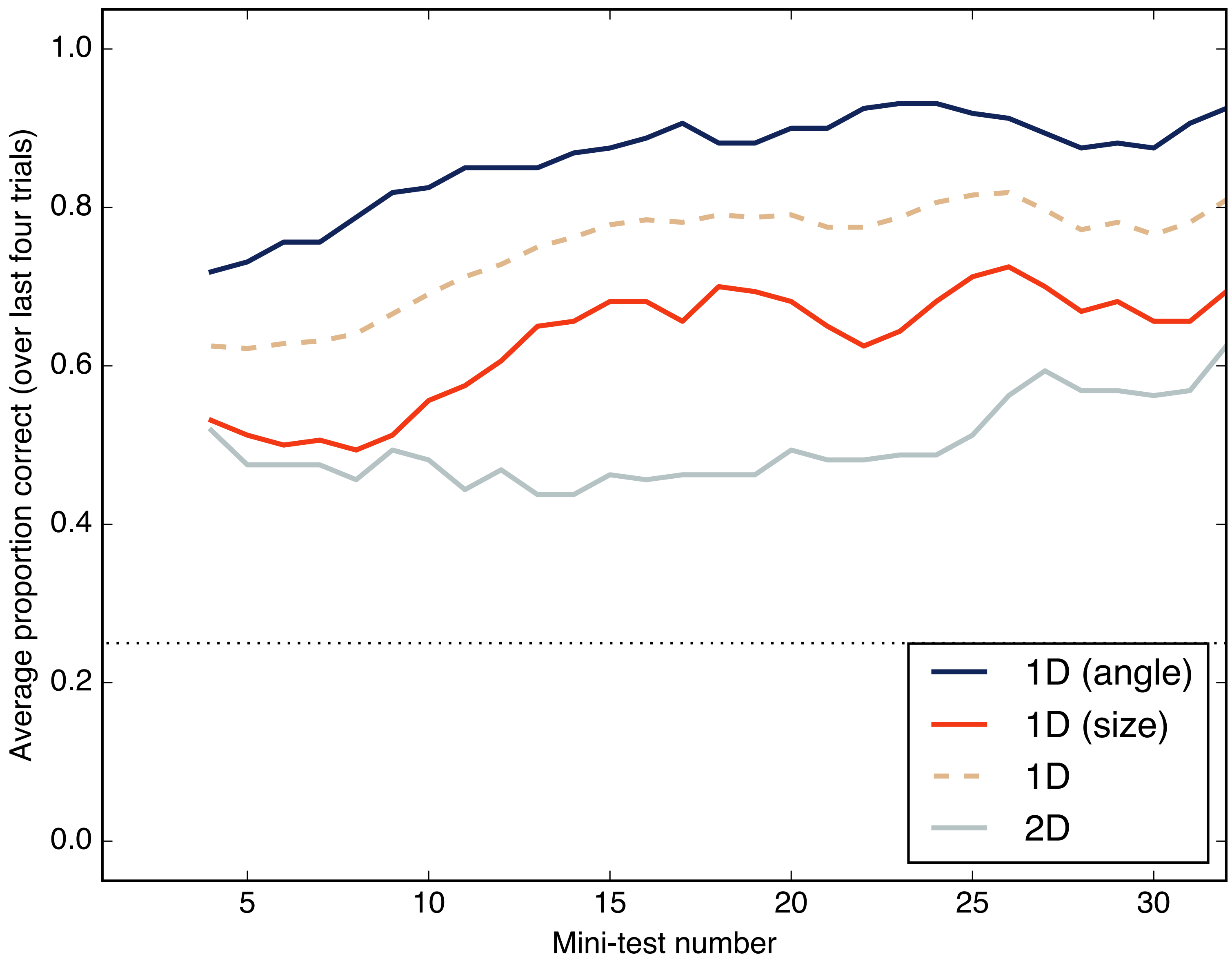


Easy to learn but low informativeness

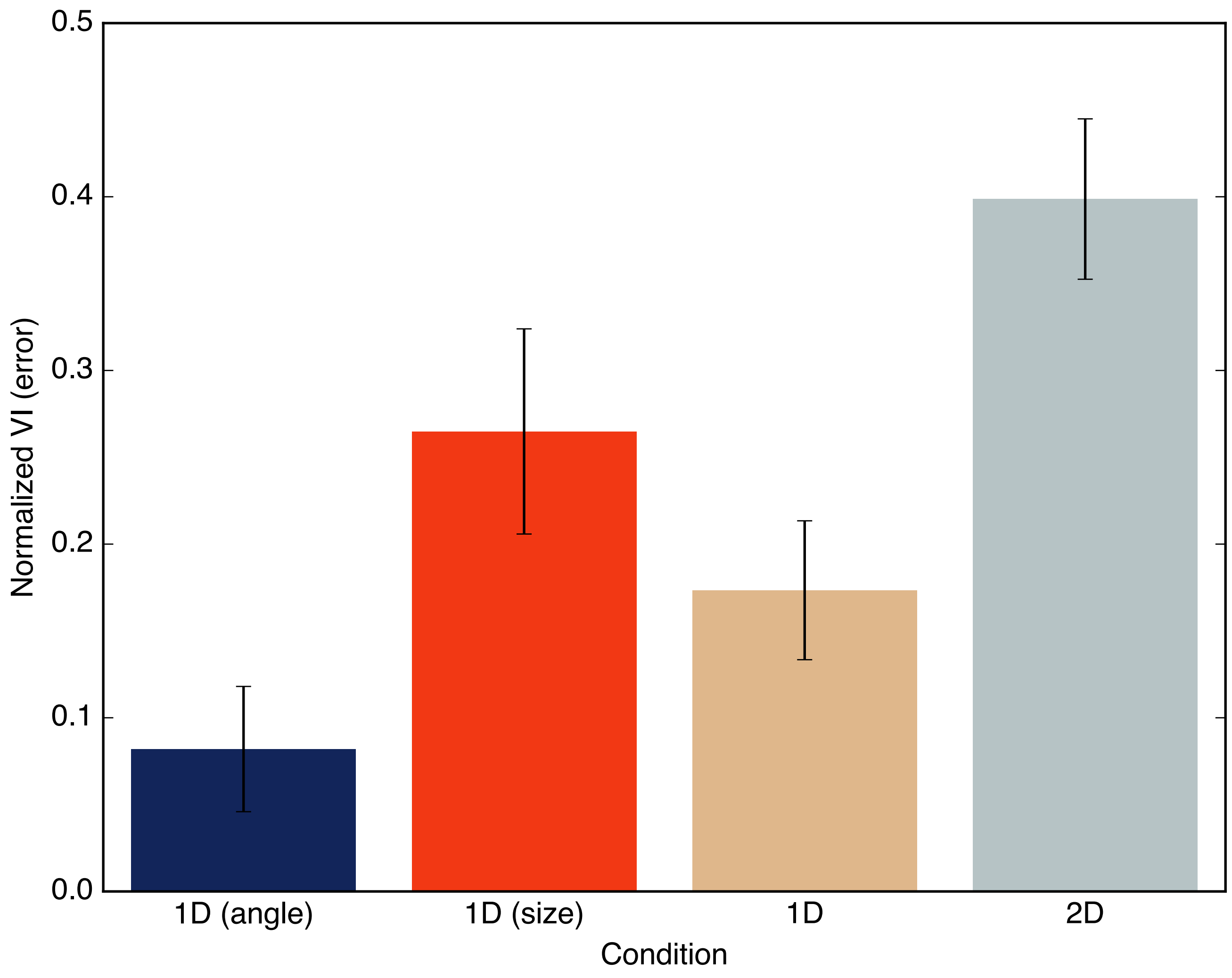
Informative but hard to learn

Results (so far. . .)

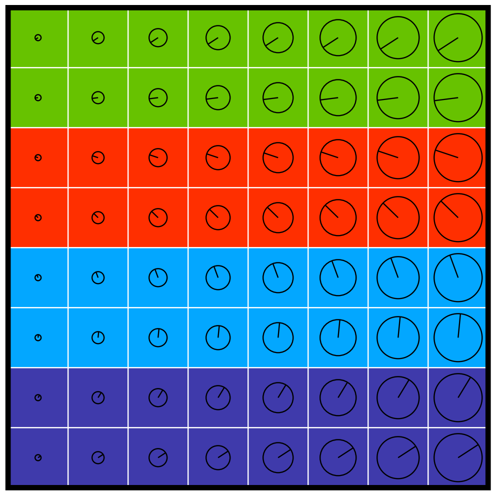
Results: Training trajectory



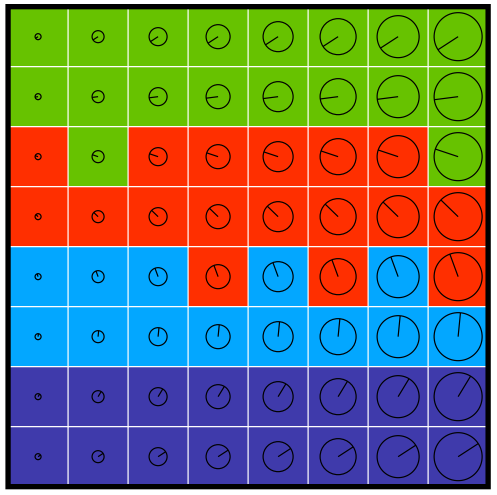
Results: Test performance



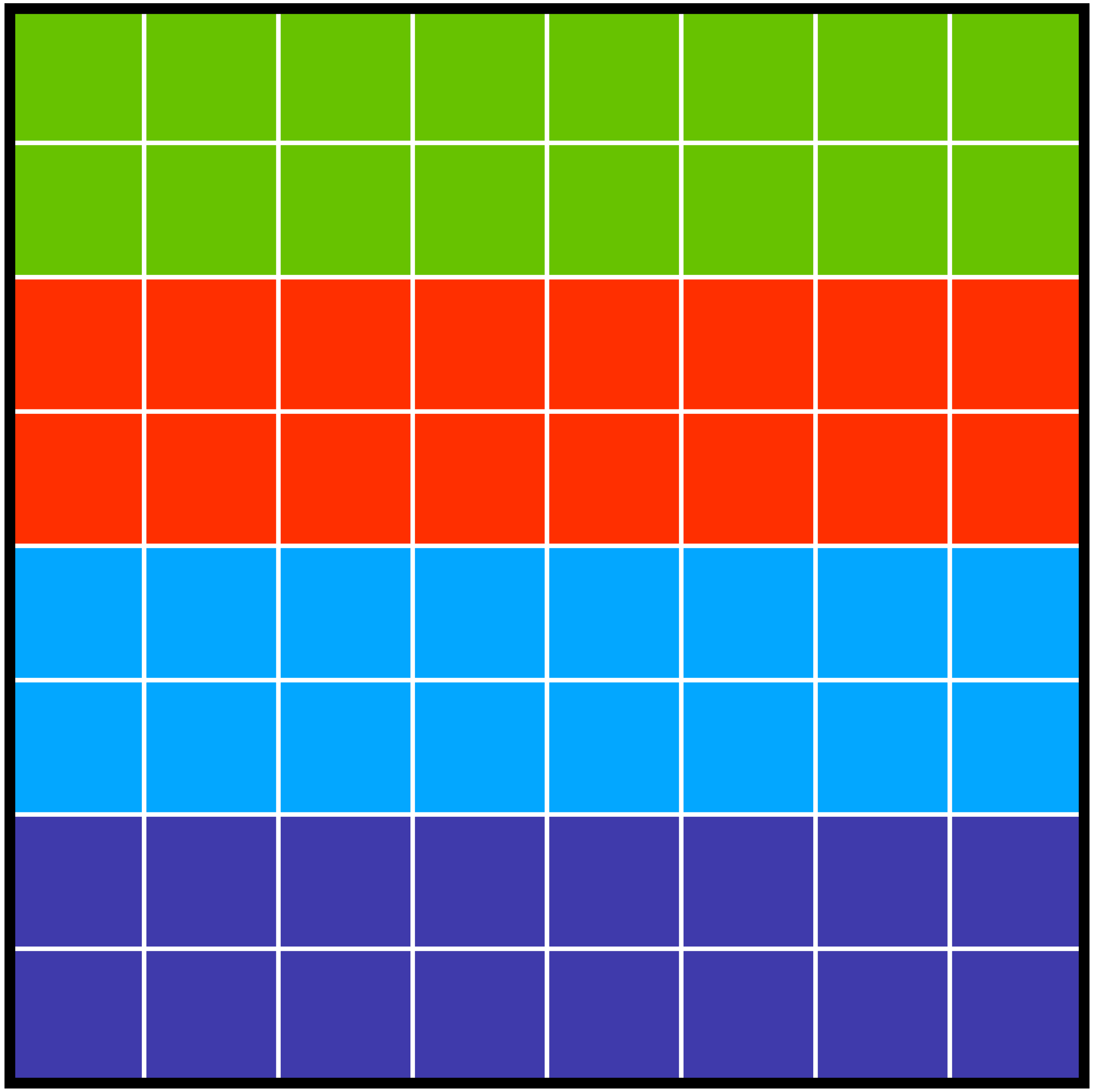
Training material



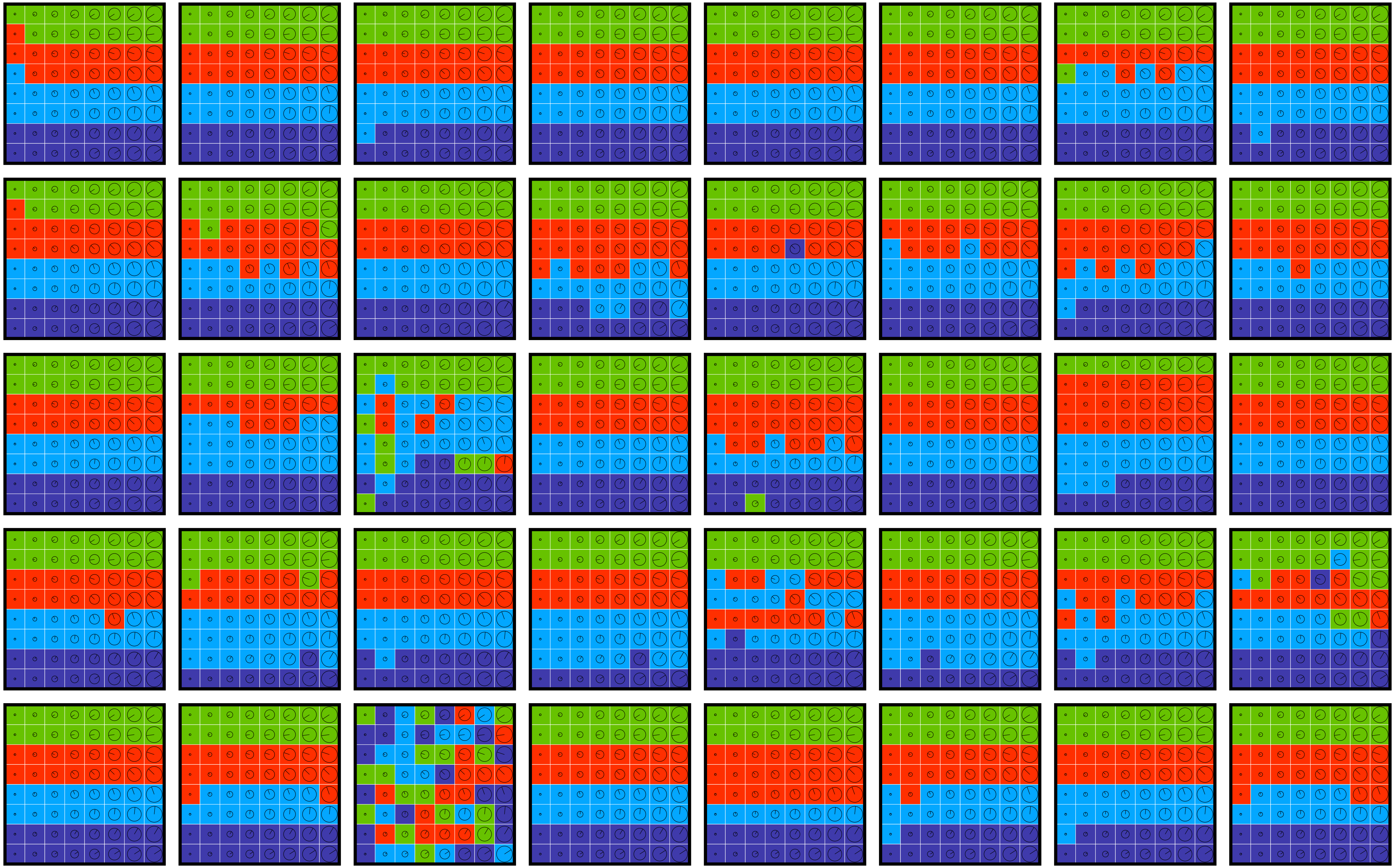
Participant's test outcome



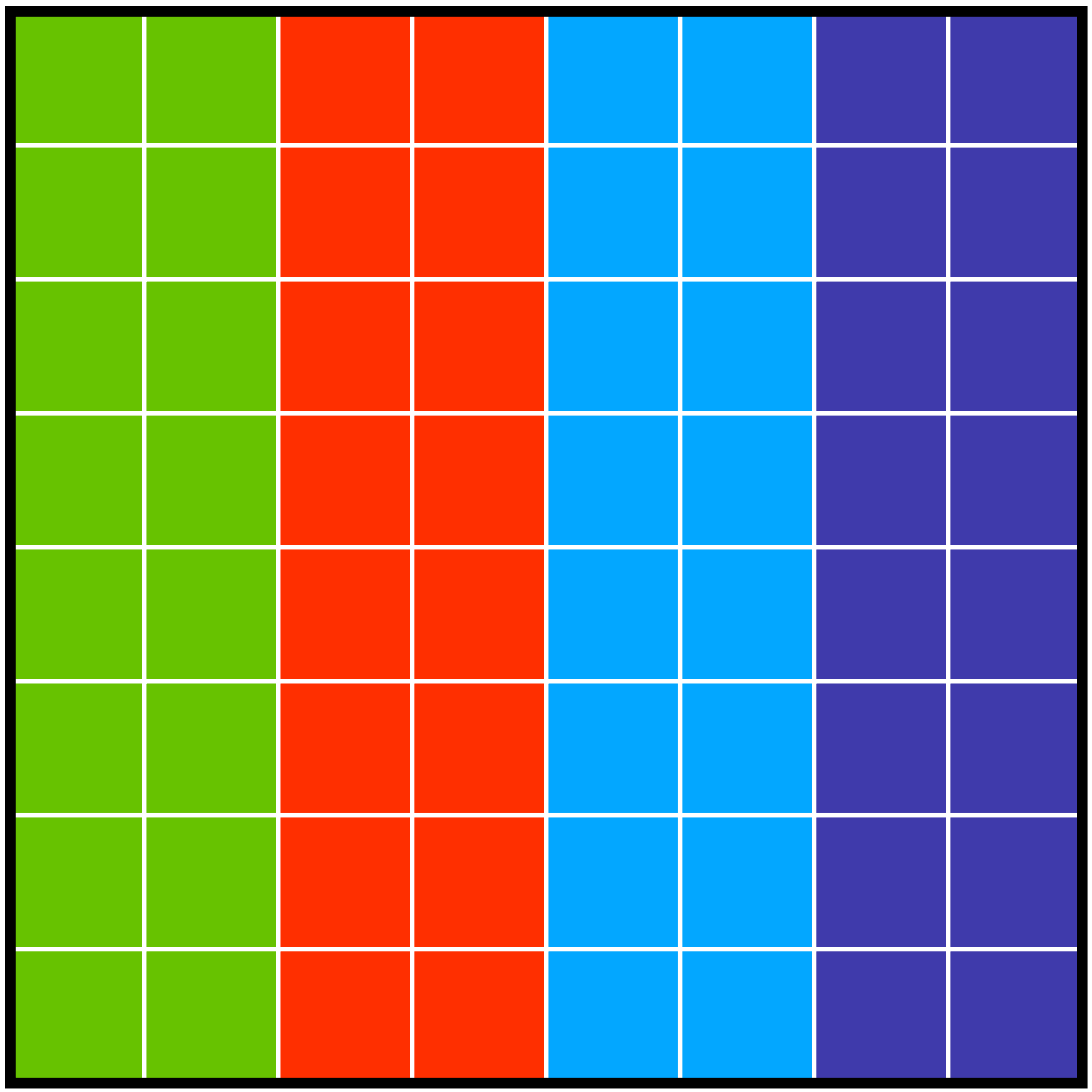
Results: Angle only



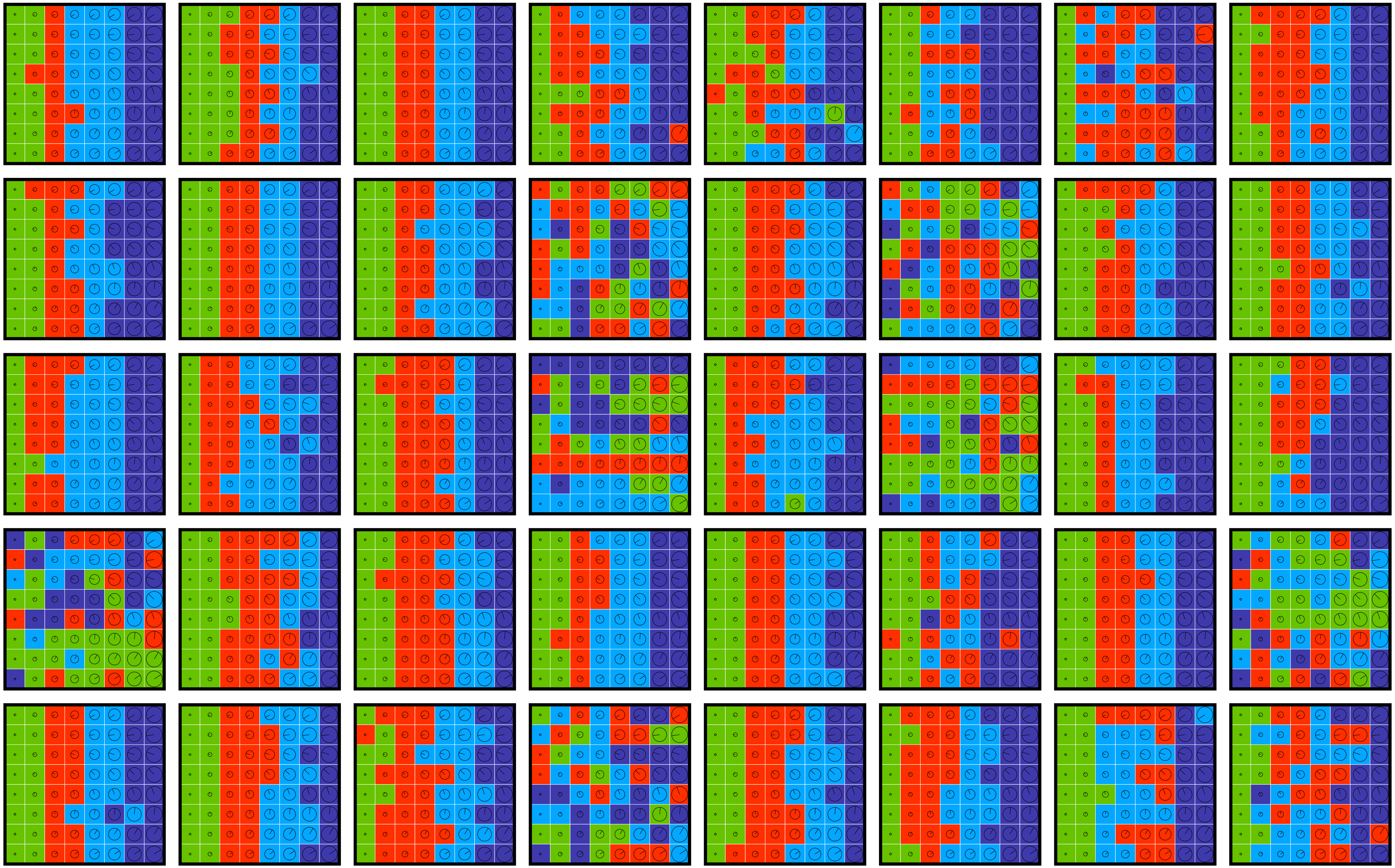
Results: Angle only



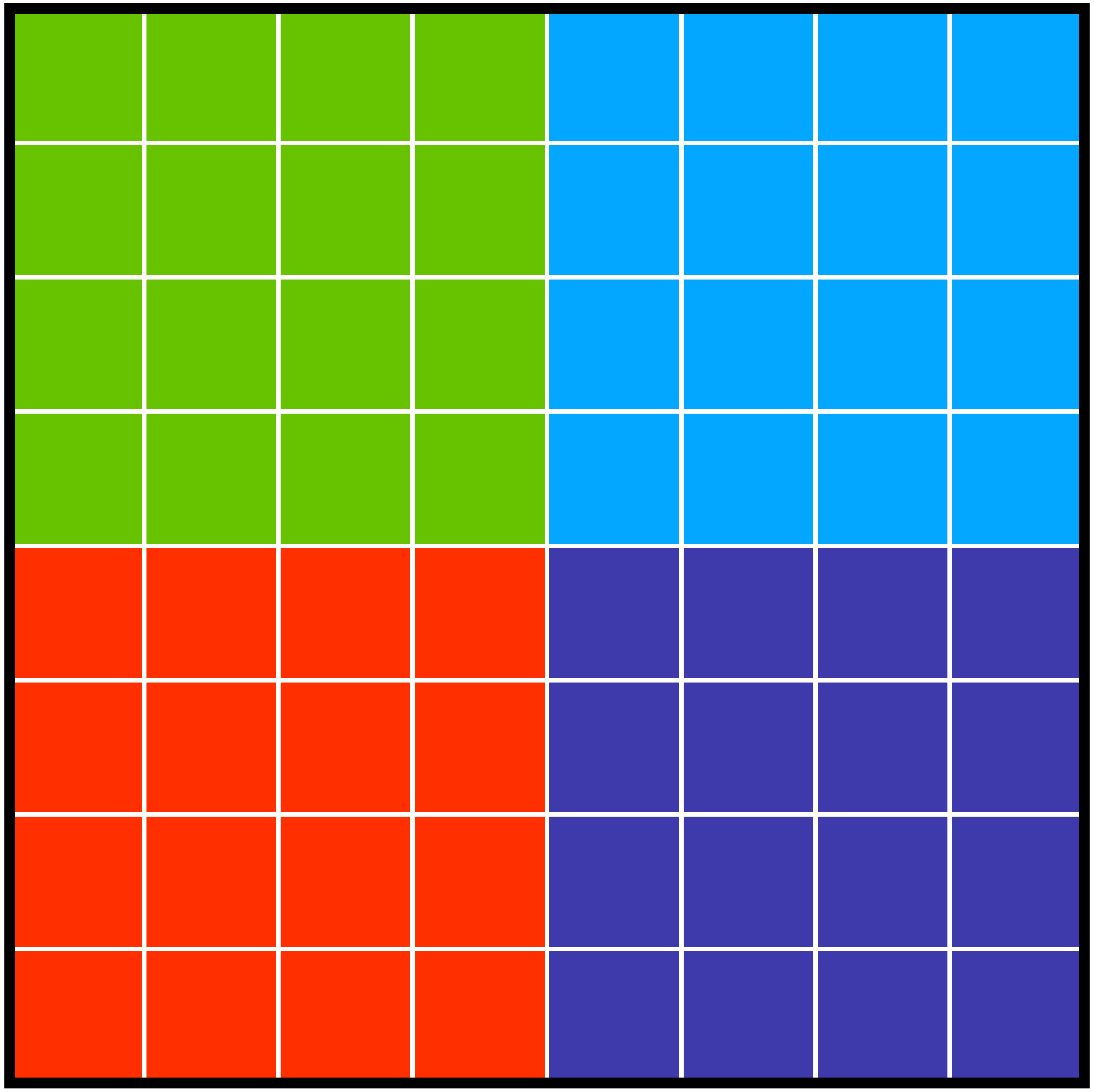
Results: Size only



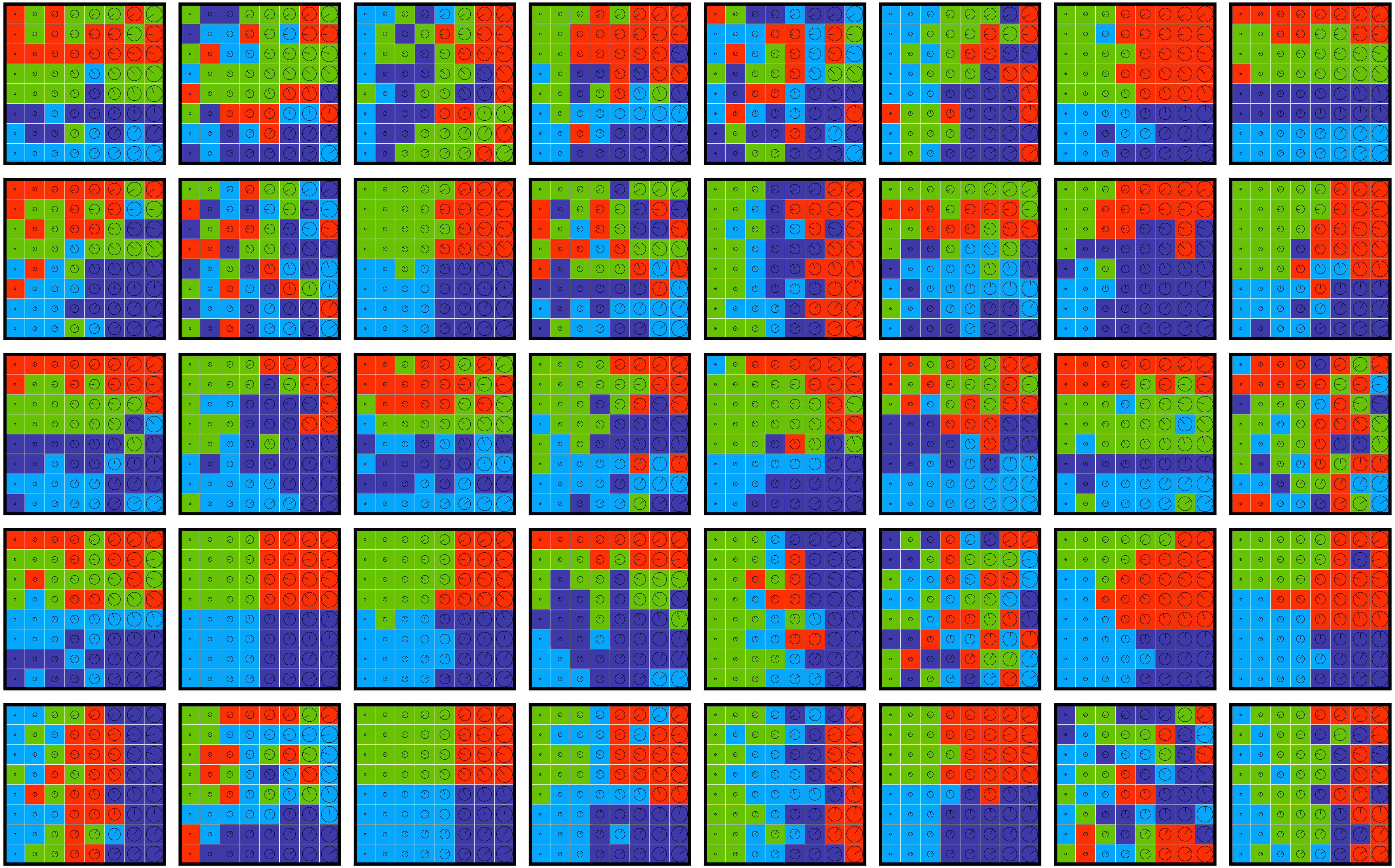
Results: Size only



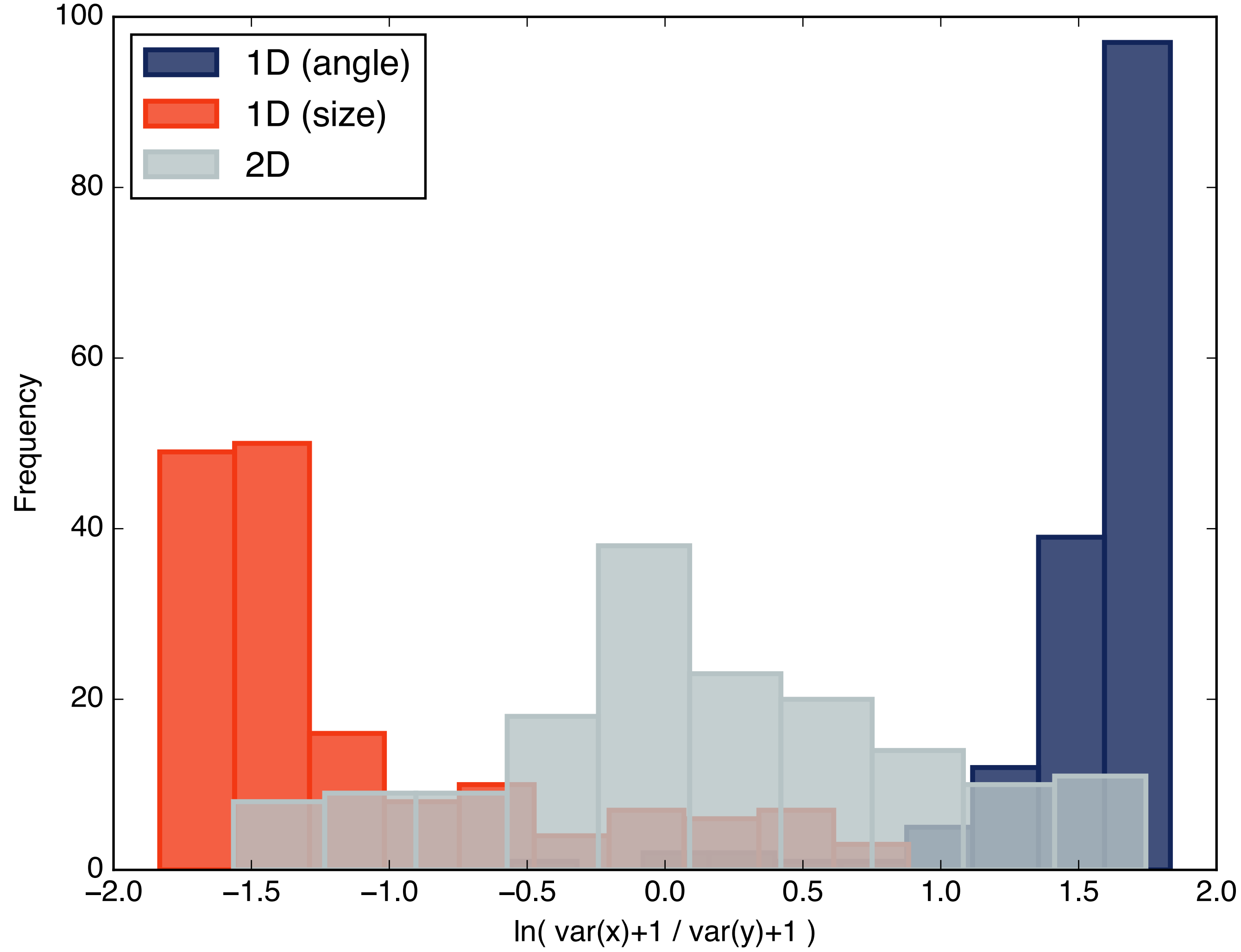
Results: Angle & Size



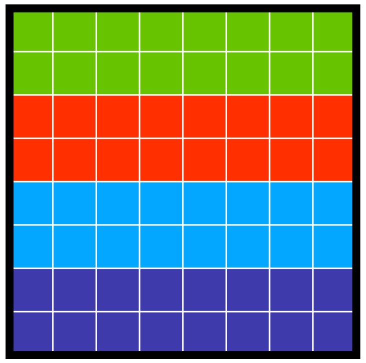
Results: Angle & Size



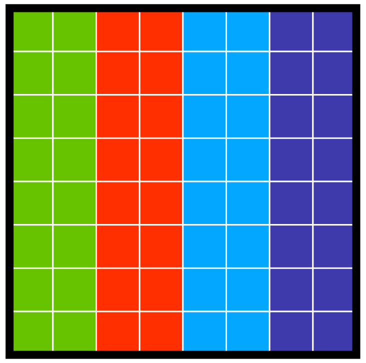
Results: Dimension preference



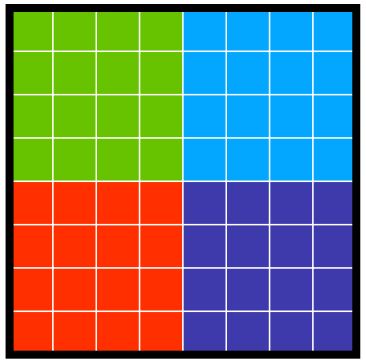
Angle-only system



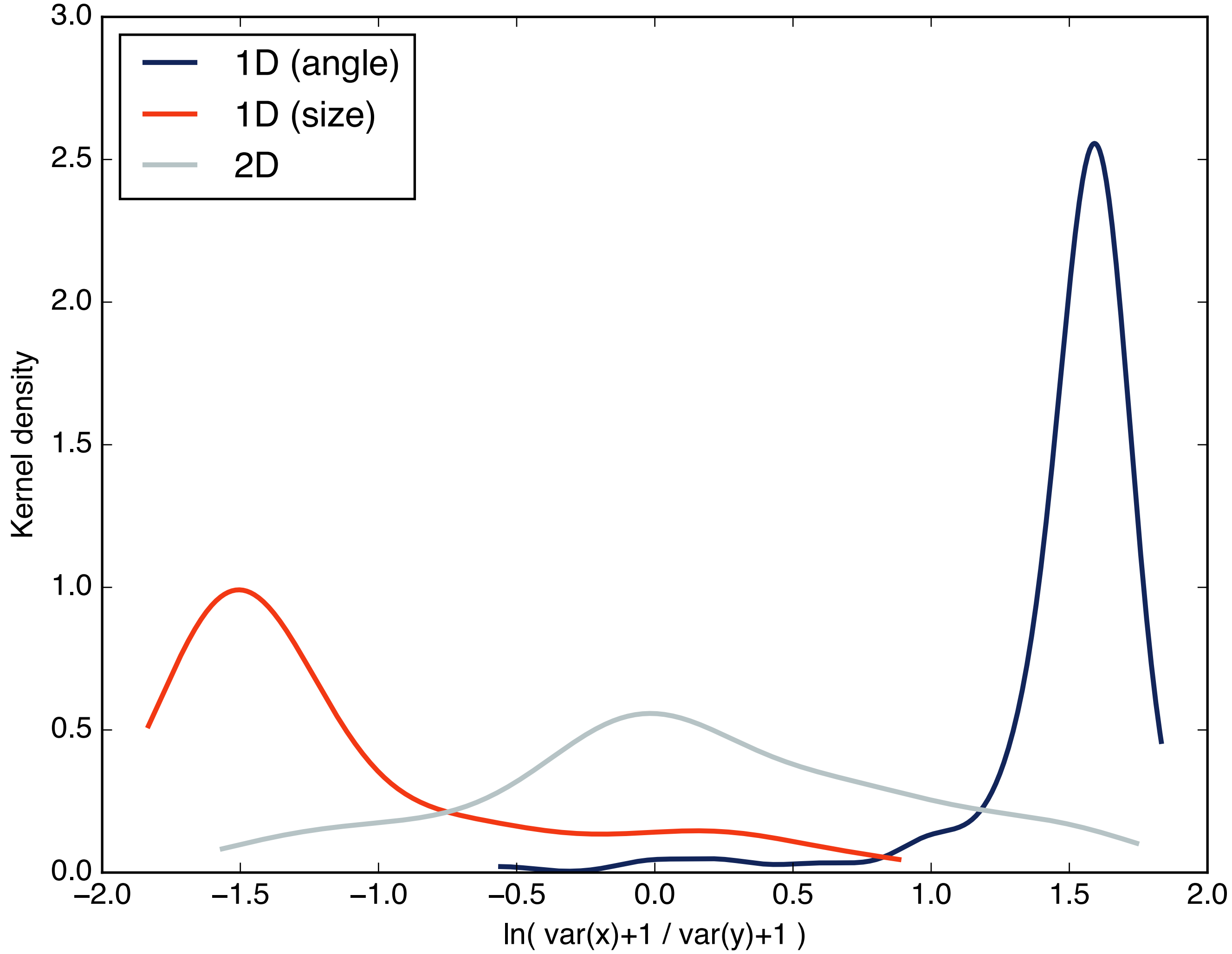
Size-only system



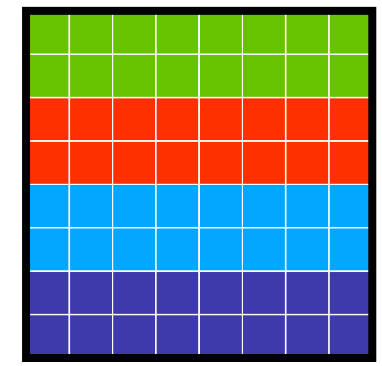
Angle & Size system



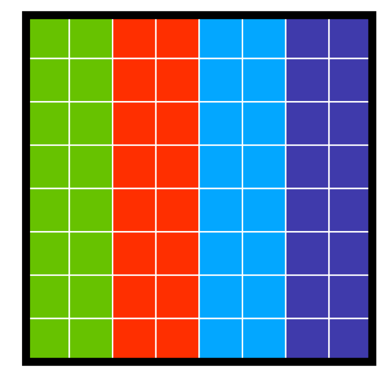
Results: Dimension preference



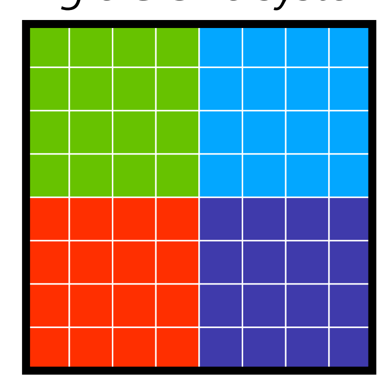
Angle-only system



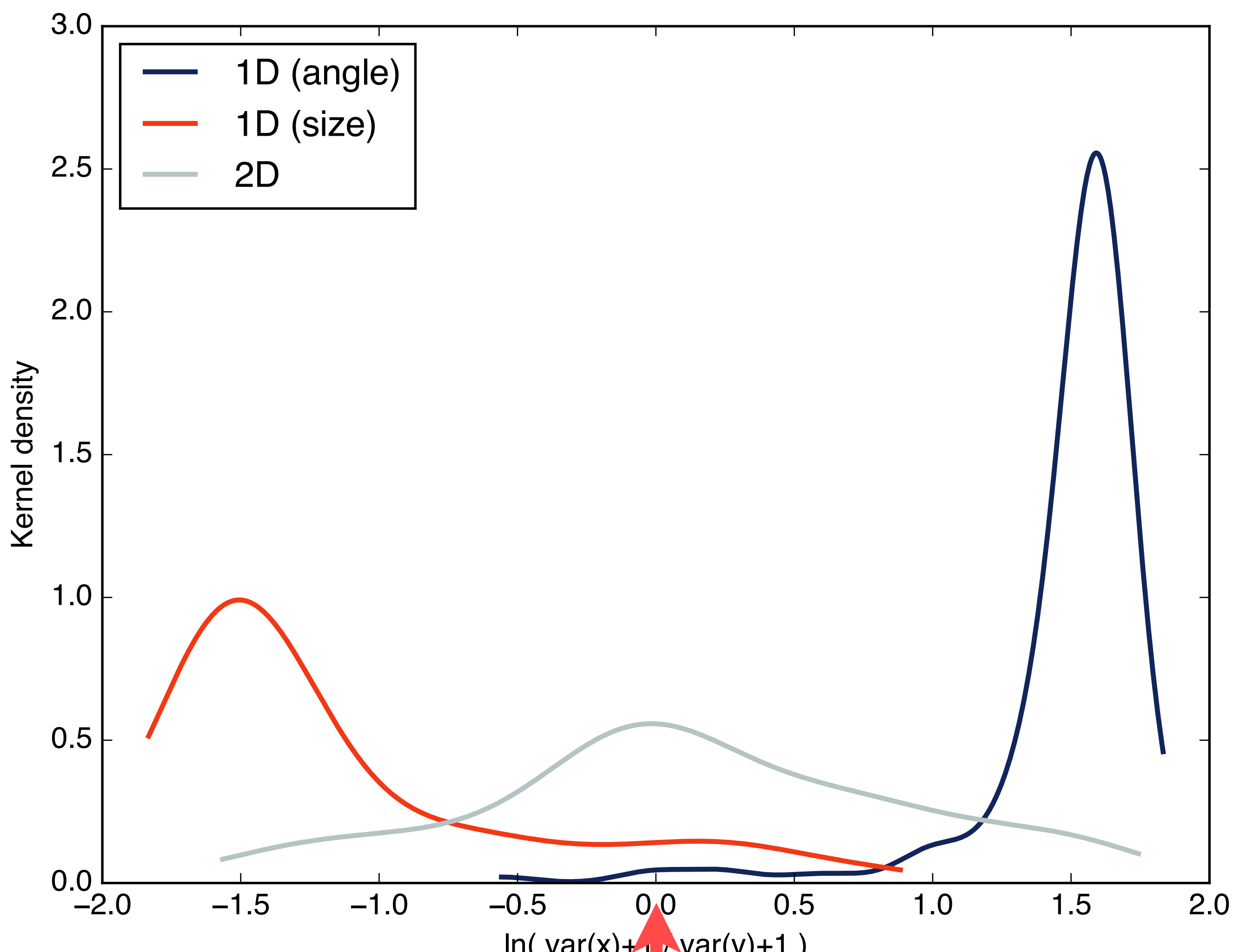
Size-only system



Angle & Size system

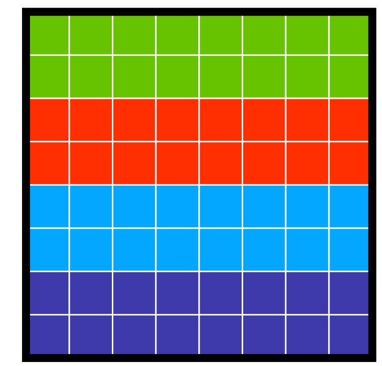


Results: Dimension preference

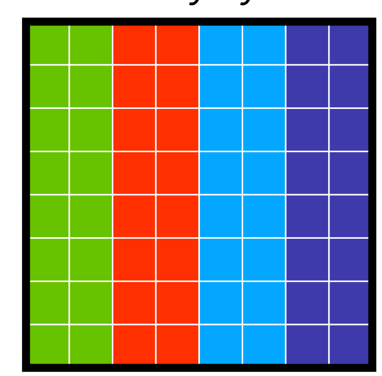


Angle and size equally important

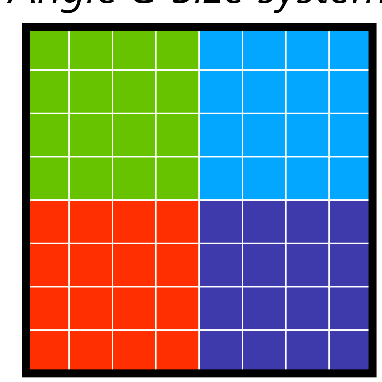
Angle-only system



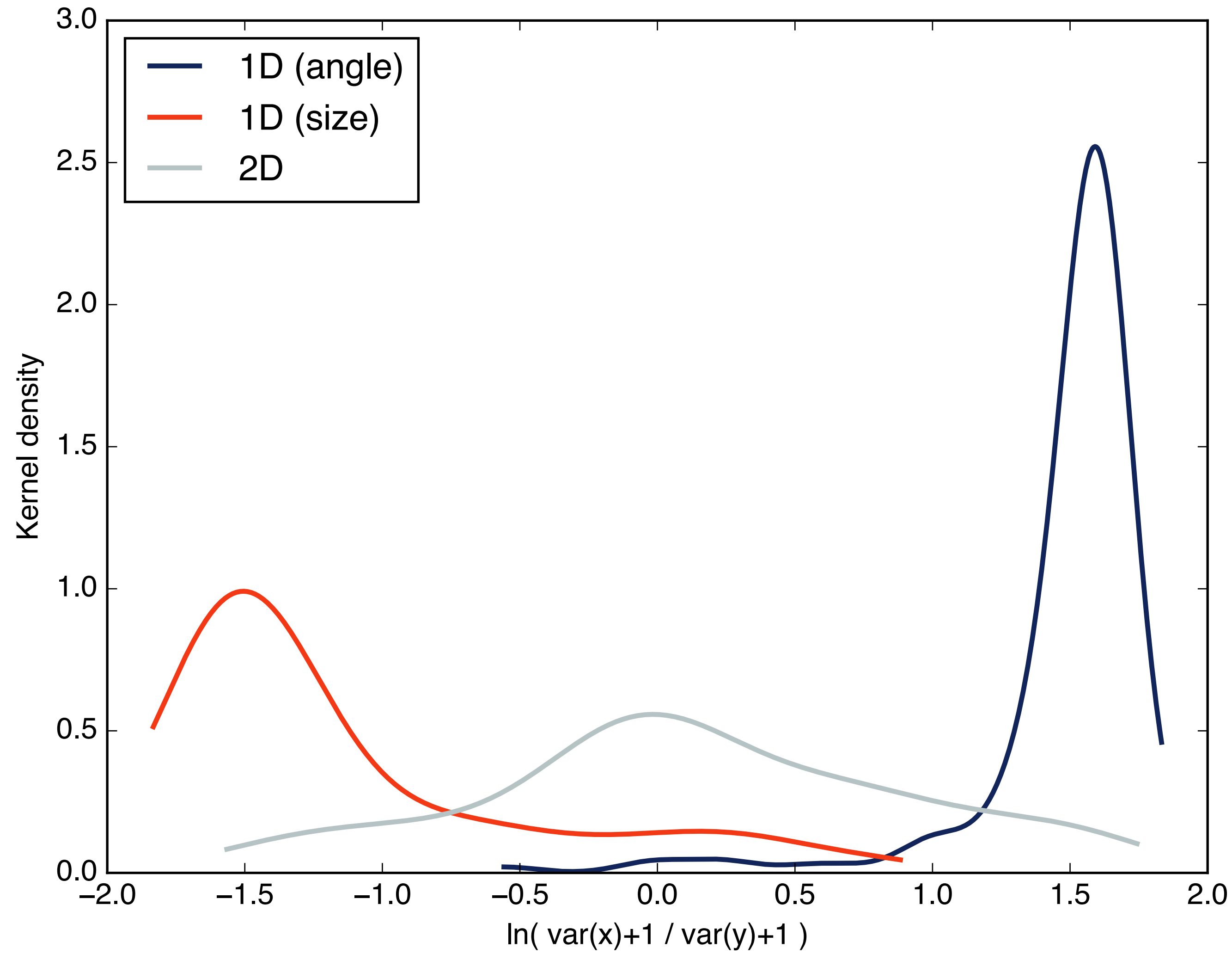
Size-only system



Angle & Size system

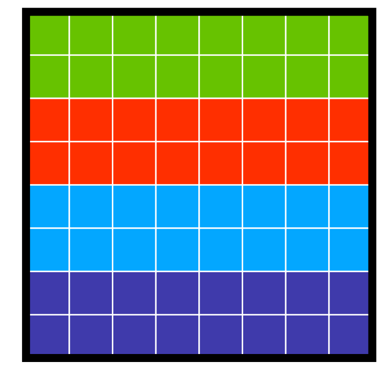


Results: Dimension preference

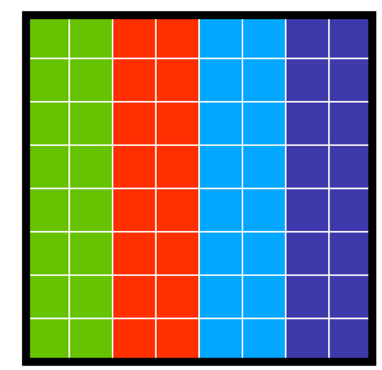


Angle more important than size

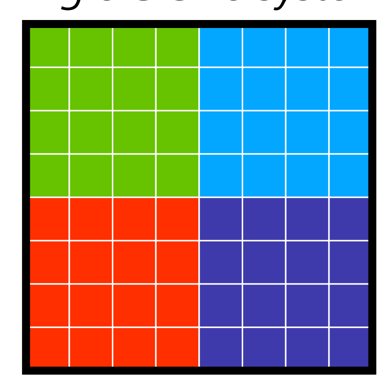
Angle-only system



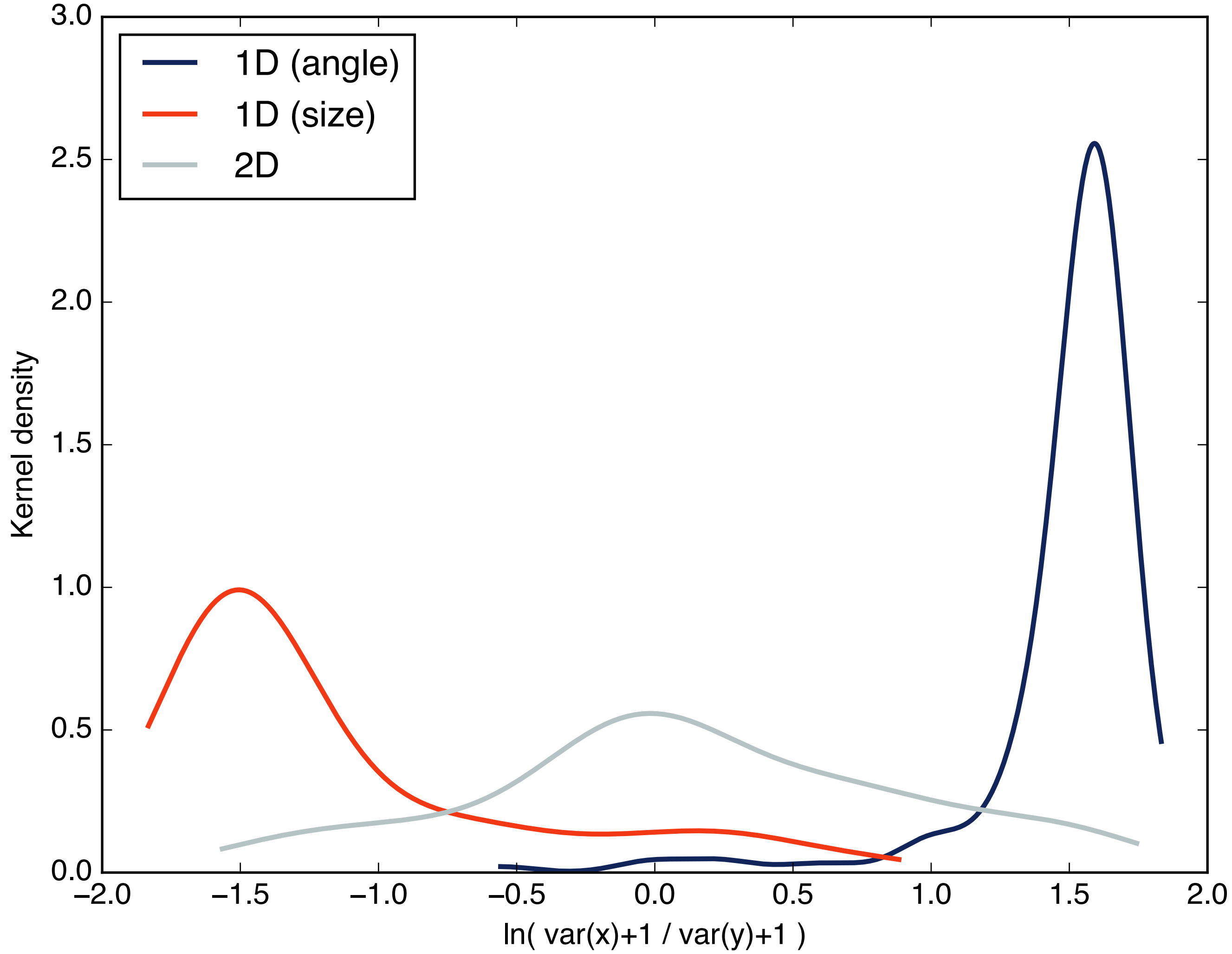
Size-only system



Angle & Size system

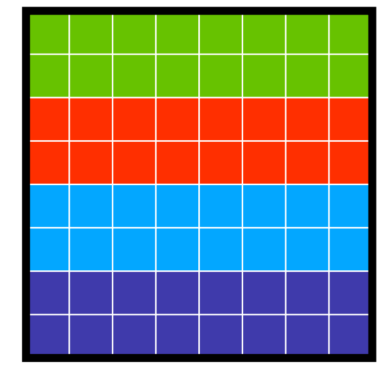


Results: Dimension preference

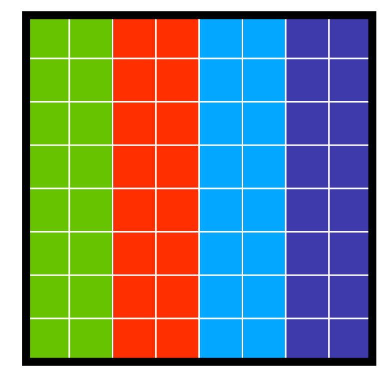


Size more important than angle

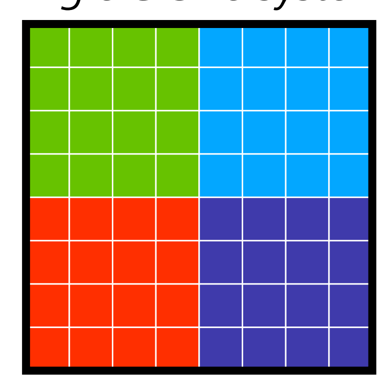
Angle-only system



Size-only system



Angle & Size system



Next steps

Why do people find the size-only condition so hard – is it just something weird with these stims?

What happens when the task is iterated in a transmission chain?

> Prediction: Everyone shifts to the angle-only system because it’s easiest
>
> Prediction: lots of noise
>
> Prediction: loss of categories

What happens when you introduce a communicative task?

> Prediction: Everyone shifts to the angle & size system because it’s the most informative.


Simon Kirby


Kenny Smith


Jennifer Culbertson

Thanks!

References

Canini, K. R., Griffiths, T. L., Vanpaemel, W., & Kalish, M. L. (2014). Revealing human inductive biases for category learning by simulating cultural transmission. *Psychonomic Bulletin & Review*, *21*, 785–793. doi: 10.3758/s13423-013-0556-3

Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2016). The cultural evolution of structured languages in an open-ended, continuous world. *Cognitive Science*, 1–32. doi:10.1111/cogs.12371

Carstensen, A., & Regier, T. (2013). Individuals recapitulate the proposed evolutionary development of spatial lexicons. In *Proceedings of the 35th annual conference of the Cognitive Science Society* (pp. 293–298). Austin, TX: Cognitive Science Society.

von der Gabelentz, G. (1901). *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherige Ergebnisse*. Leipzig, Germany: C. H. Tauchnitz.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*, 1049–1054. doi:10.1126/science.1218811

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the USA*, *105*, 10681–10686. doi:10.1073/pnas.0707835105

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102. doi:10.1016/j.cognition.2015.03.016

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge, MA: MIT Press.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge, MA: MIT Press.

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney & W. O'Grady, *The handbook of language emergence* (pp. 237–263). Hoboken, NJ: John Wiley & Sons, Inc. doi:10.1002/9781118346136.ch11

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*, 54–87. doi:10.1016/0022-2496(64)90017-3