# New Foundations for Topological Data Analysis – The Power of Cohomotopy –

Urs Schreiber on joint work with Hisham Sati

جامعـة نيويورك أبـوظـي NYU ABU DHABI NYU AD Science Division, Program of Mathematics & Center for Quantum and Topological Systems

New York University, Abu Dhabi



brief presentation

CQTS @ NYU Abu Dhabi, 26 May 2022

slides and further pointers at: ncatlab.org/schreiber/show/New+Foundations+for+TDA+--+Cohomotopy

#### Abstract.

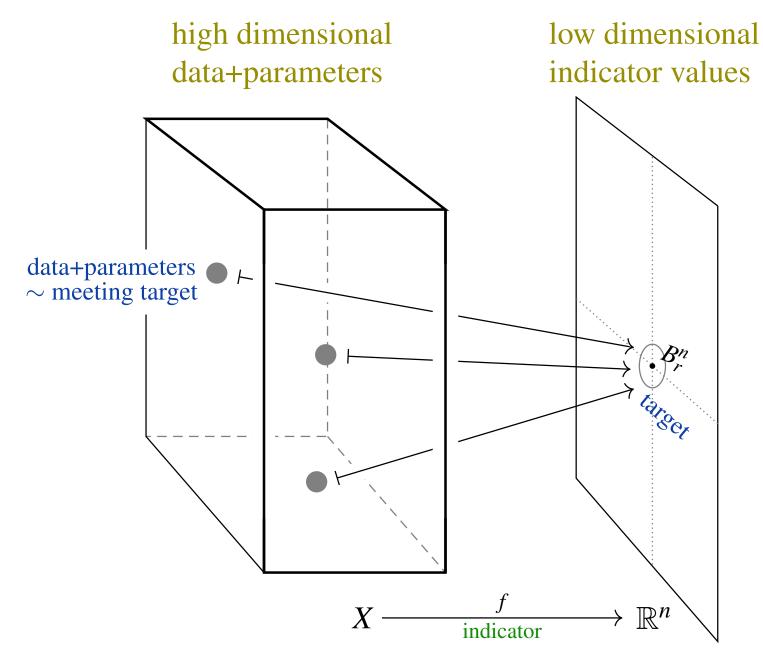
The aim of *topological data analysis* (TDA) is to provide qualitative analysis of large data/parameter sets in a way which is robust against uncertainties and noise. This is accomplished using tools and theorems from the mathematical field of *algebraic topology*. While a tool called *persistent homology* has become the signature method of TDA, it tends to produce answers that are either hard to interpret or impossible to compute.

Both problems are solved by a variant method [FK17] which we may call *persistent cohomotopy*: A first result shows [FKW18] that this new method provides computable answers to the concrete question of detecting whether there exist data+parameters that meet a prescribed target indicator precisely, even in the presence of uncertainty and noise.

More generally, efficient data analysis will require further refining persistent cohomotopy to *twisted equivariant cohomotopy* [SS-Orb, §5]. Curiously, this has profound relations (<u>Hypothesis H</u>) to formal high energy physics and quantum materials, connecting to which might serve to further enhance the power of topological data analysis.

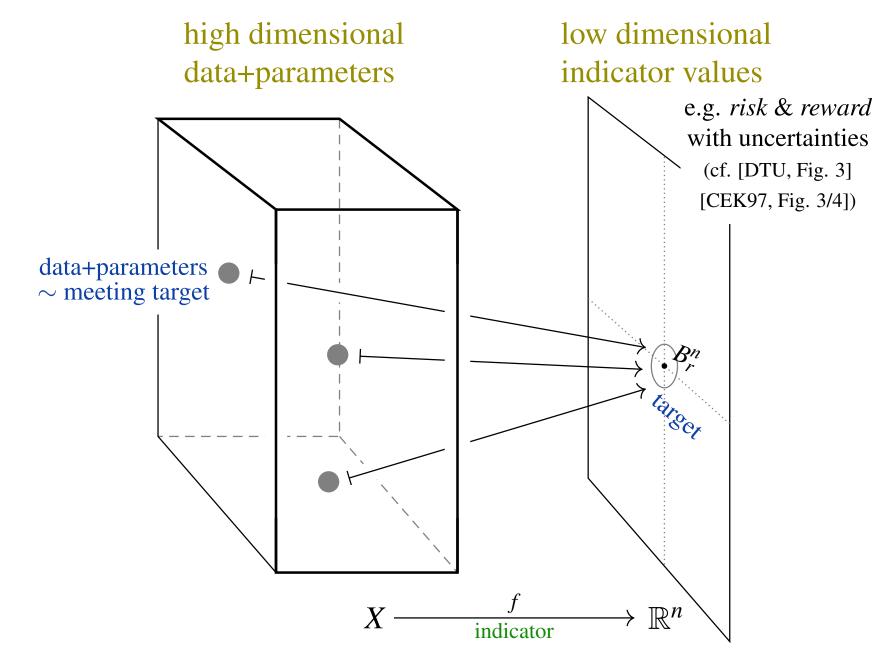
# Find data meeting prescribed target with uncertainties – The problem.

Given high-dimensional data+parameters and a handful of indicators subject to uncertainty & noise. *Can a given target be met?* 



# Find data meeting prescribed target with uncertainties – The problem.

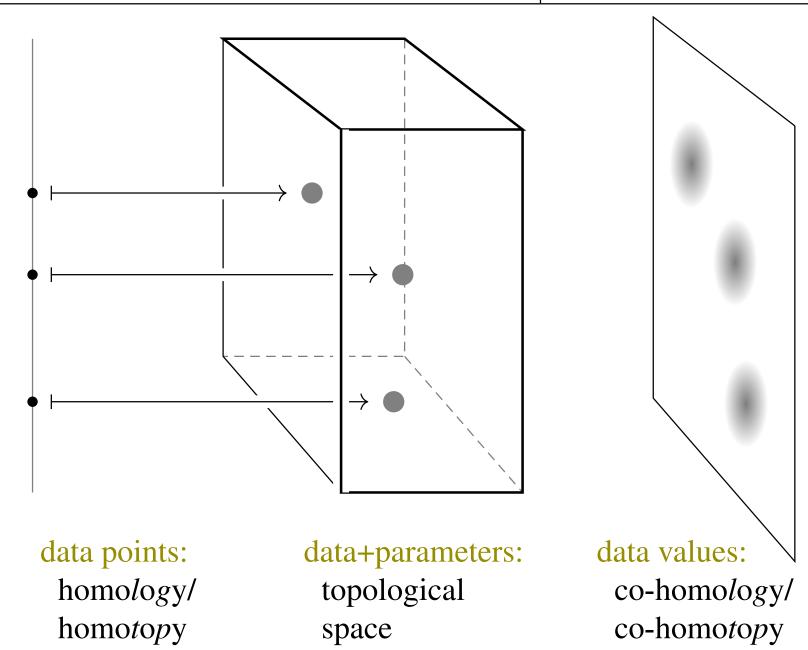
Given high-dimensional data+parameters and a handful of indicators subject to uncertainty & noise. *Can a given target be met?* 



# Find data meeting prescribed target with uncertainties – The strategy.

Use mathematical tools from *algebraic topology* (e.g. [Ca09][Ou15]):

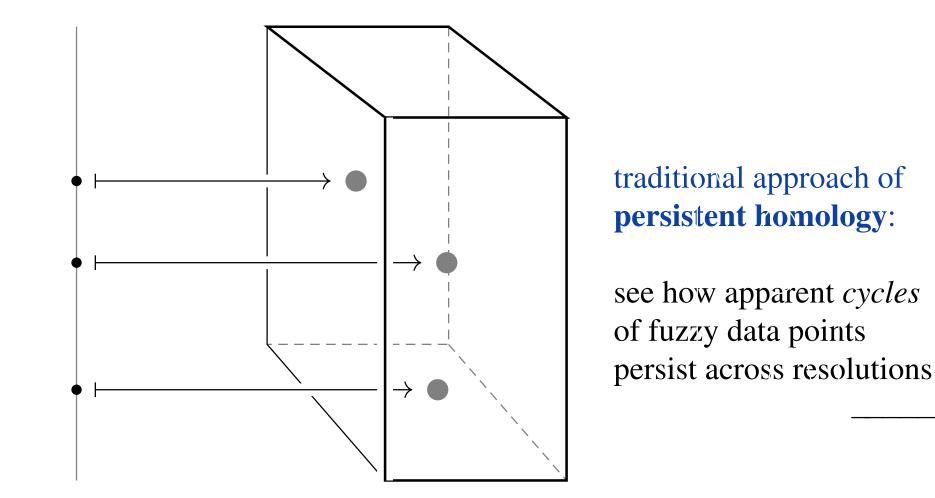
**topology:** robustness under mild deformations: **algebraic:** tractable invariants



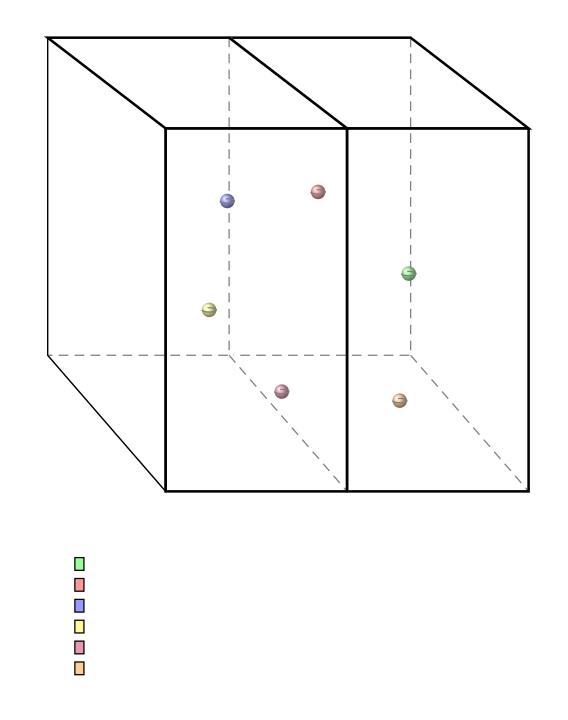
### Find data meeting prescribed target with uncertainties – The strategy.

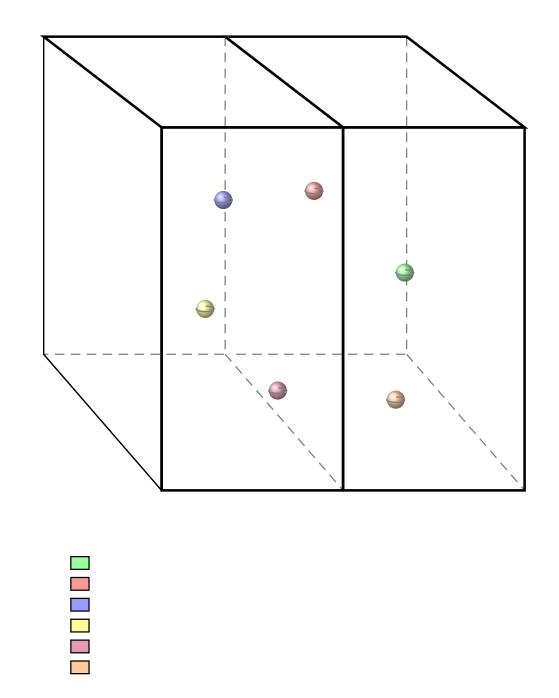
Use mathematical tools from *algebraic topology* (e.g. [Ca09][Ou15]):

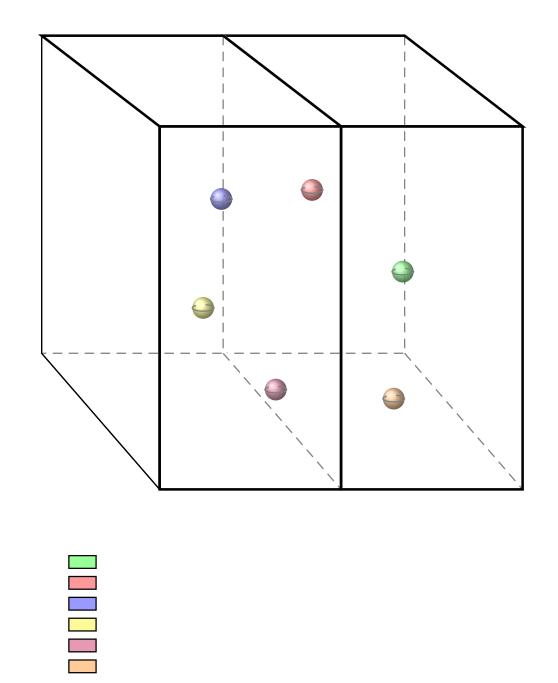
**topology:** robustness under mild deformations: **algebraic:** tractable invariants

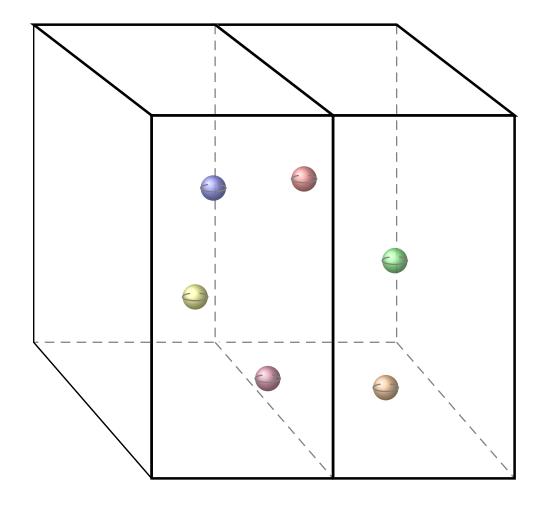


data points: homology/ homotopy data+parameters: topological space

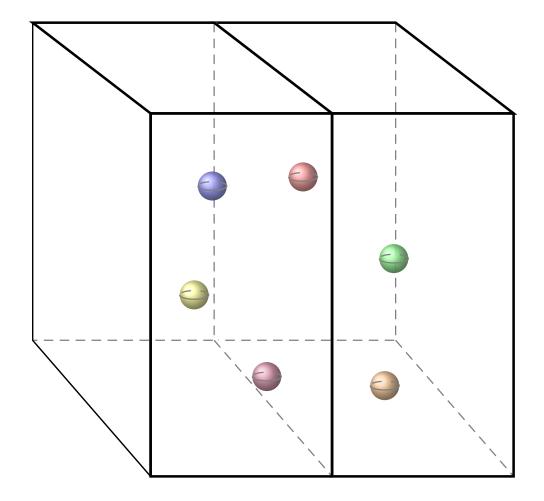




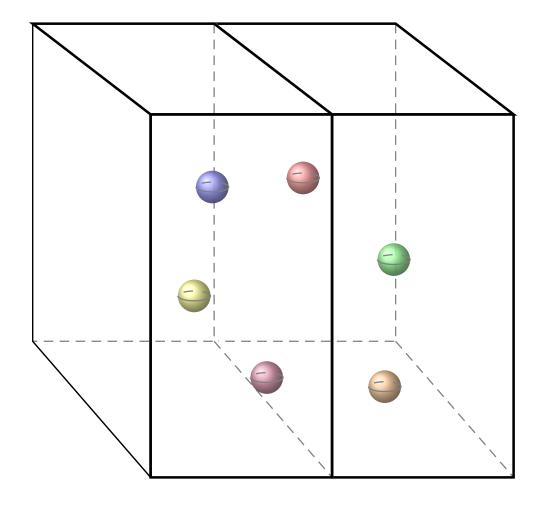




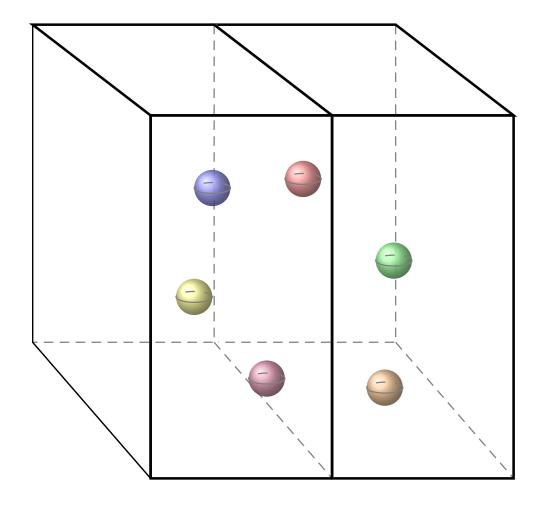




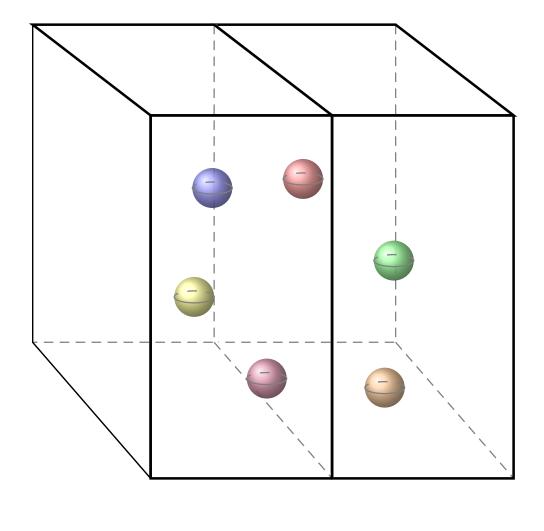




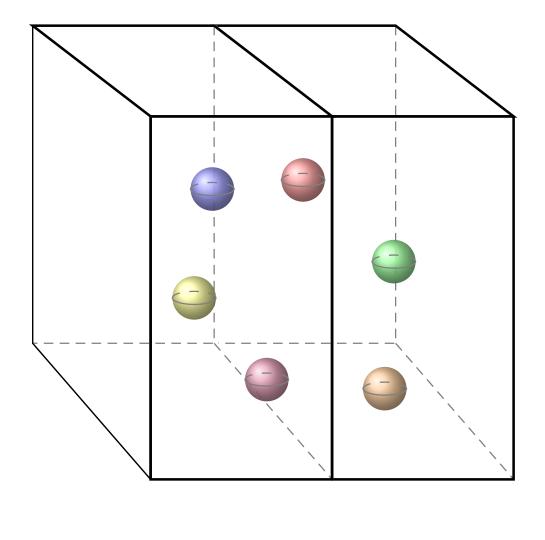




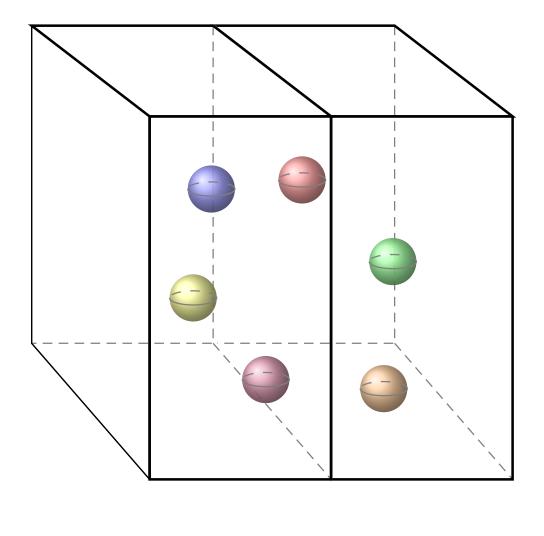




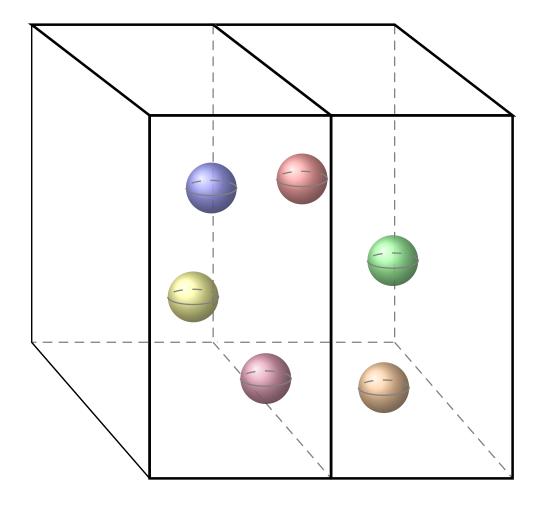


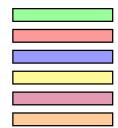


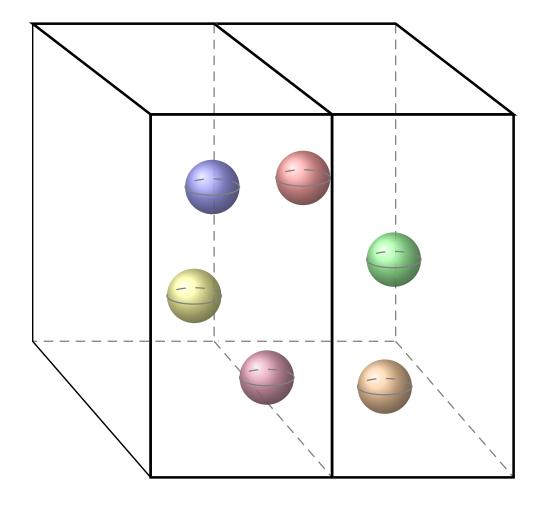




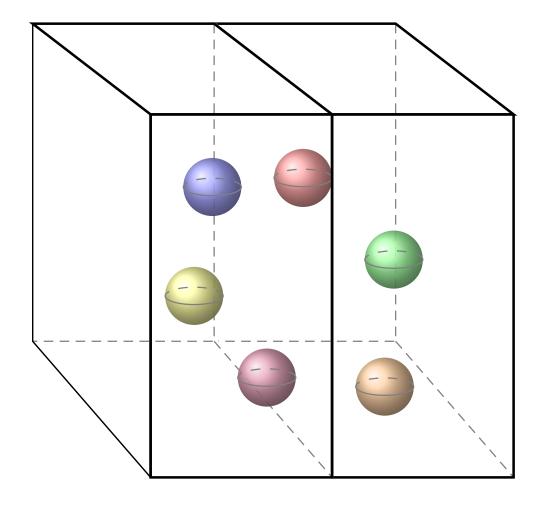




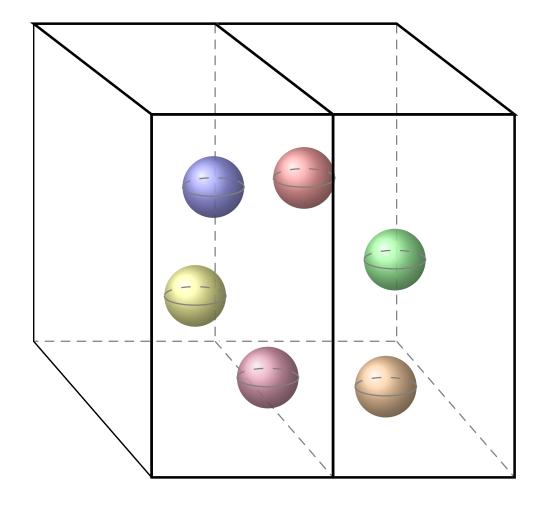


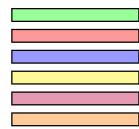


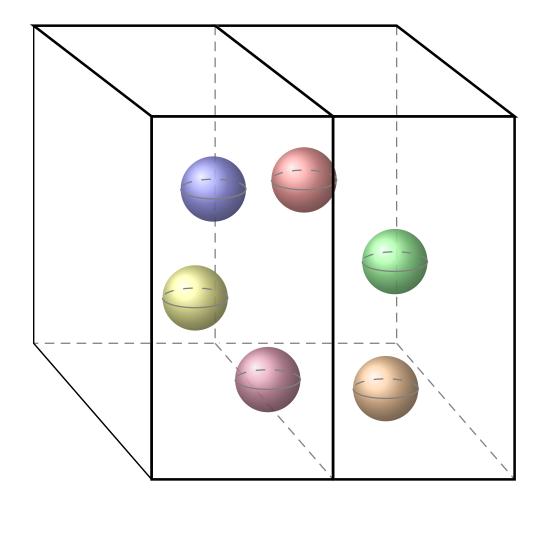


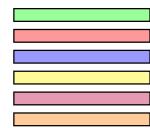


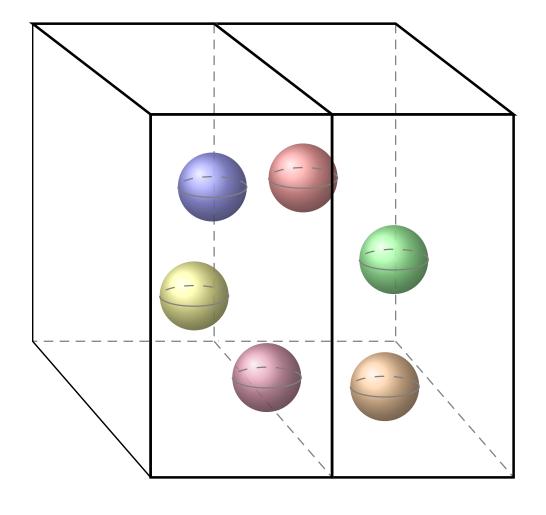


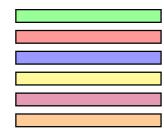


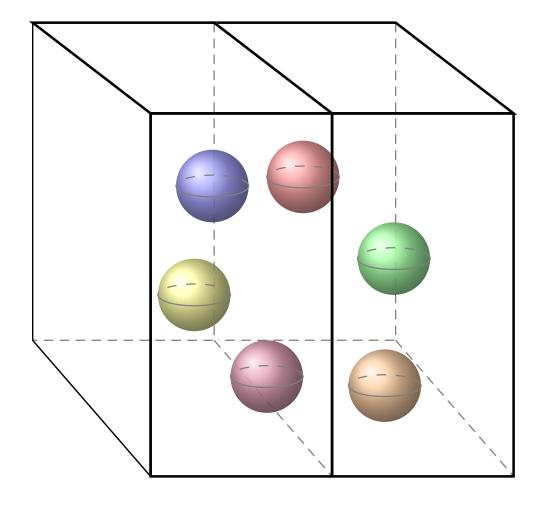


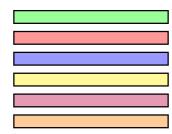


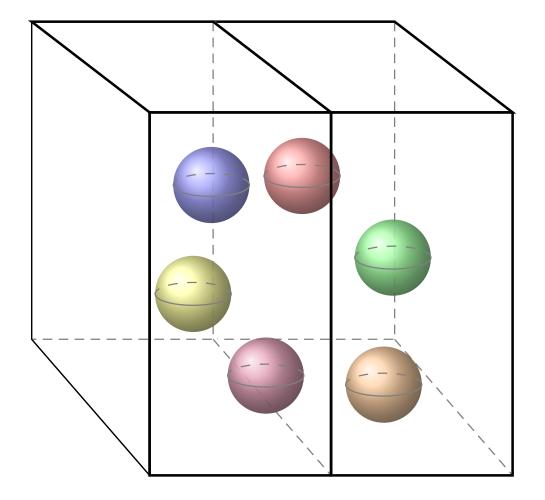


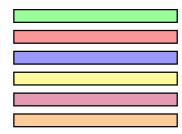


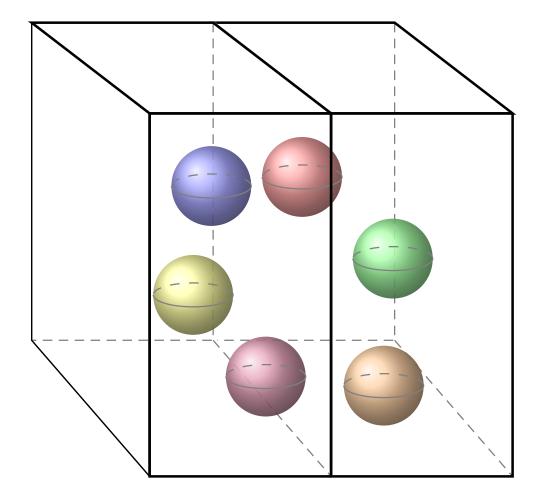


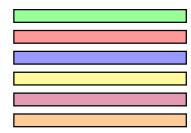


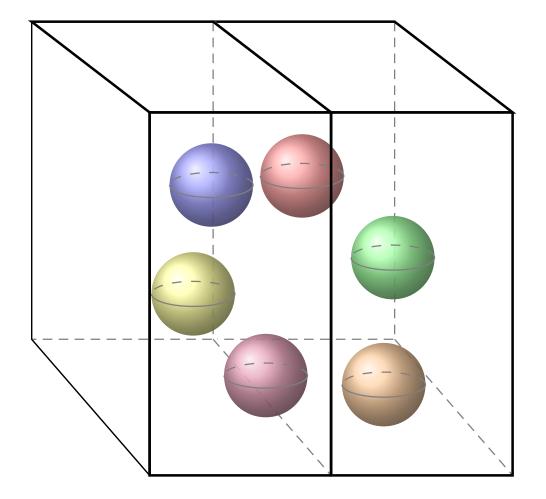


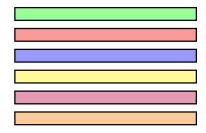


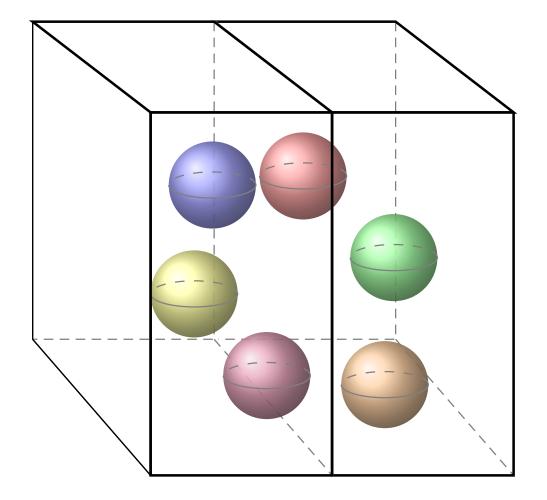


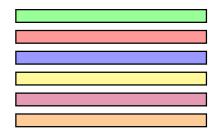


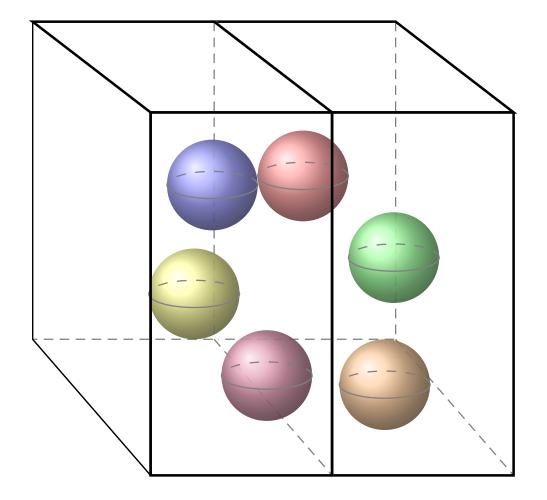




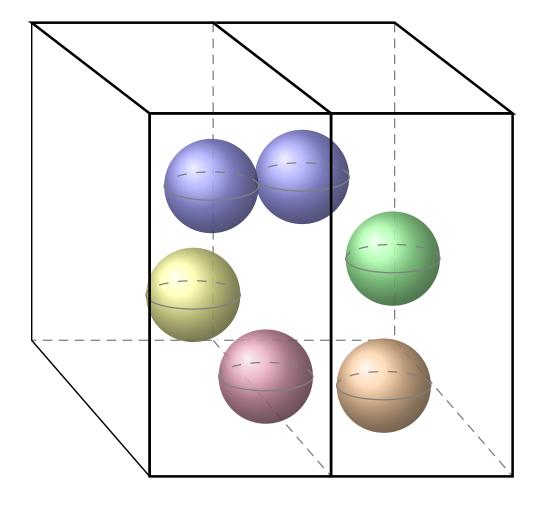




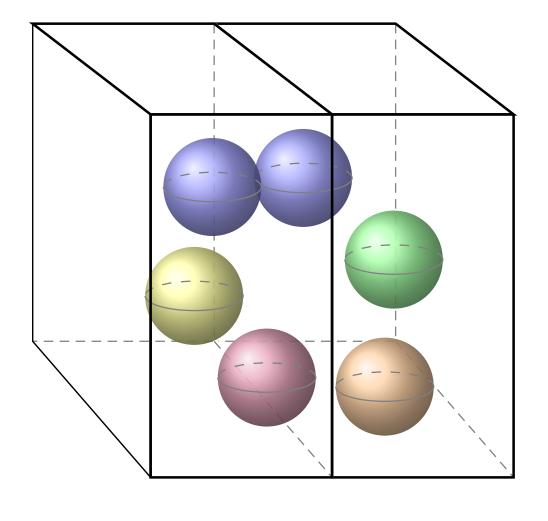




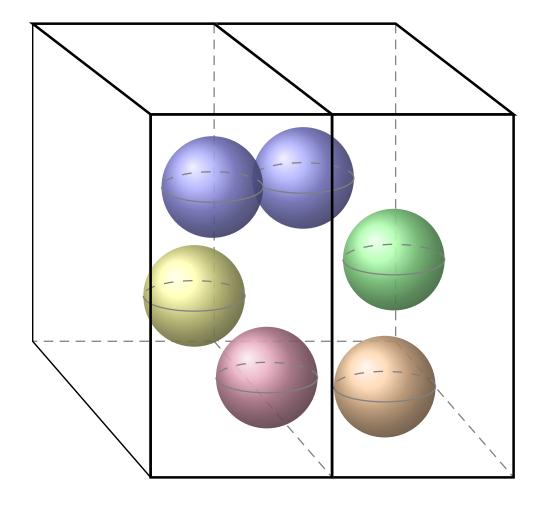


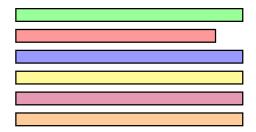


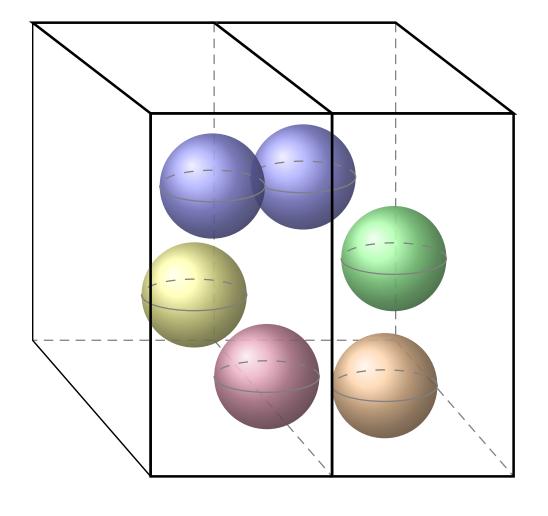




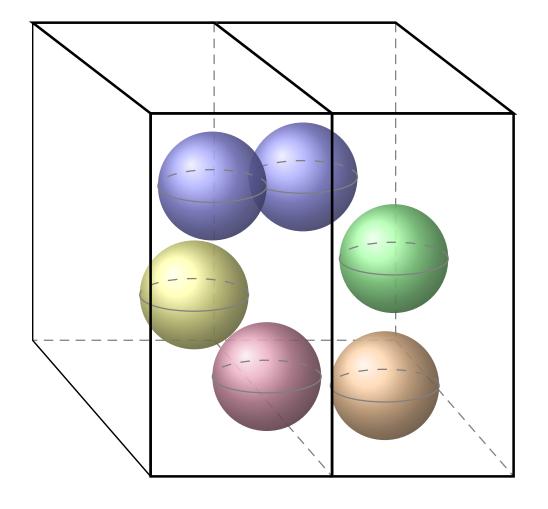


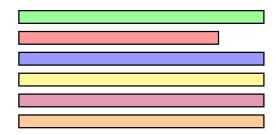


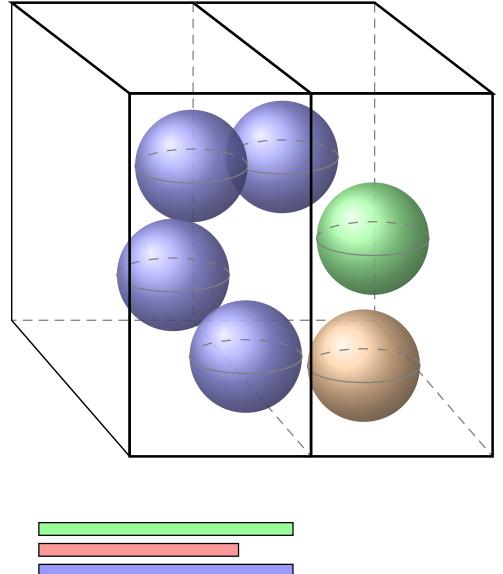




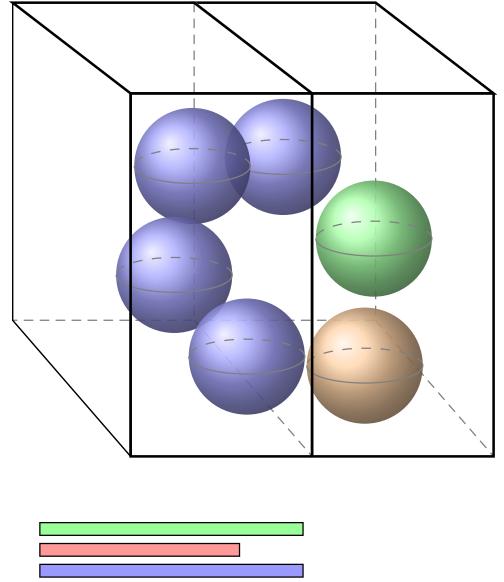


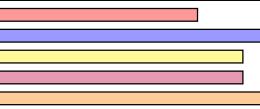


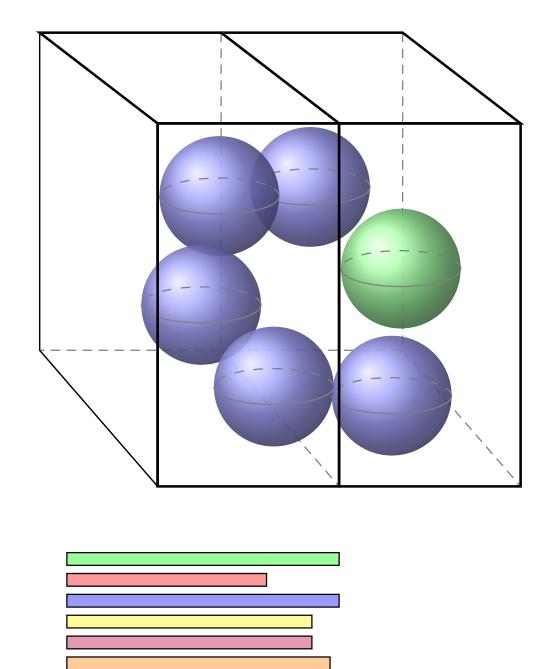


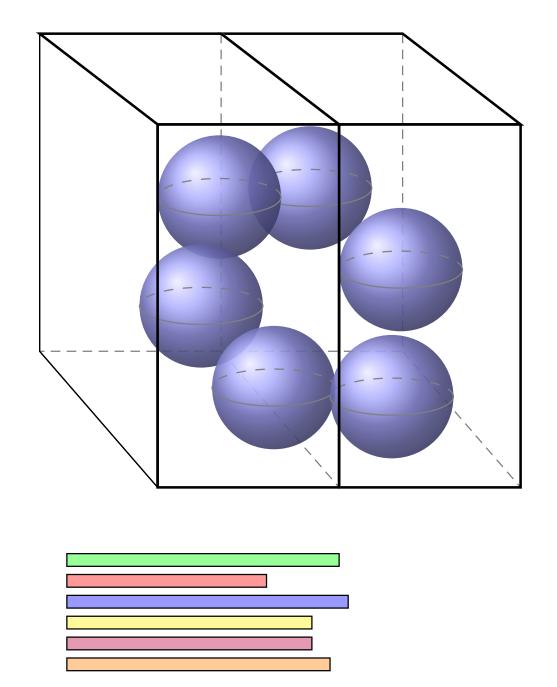


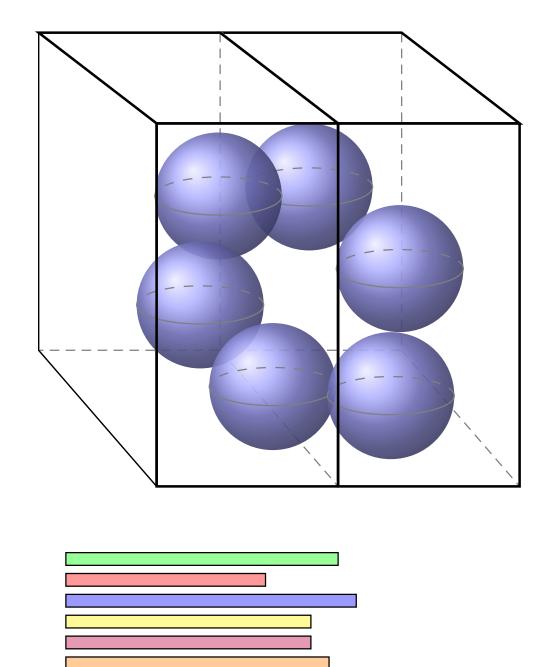


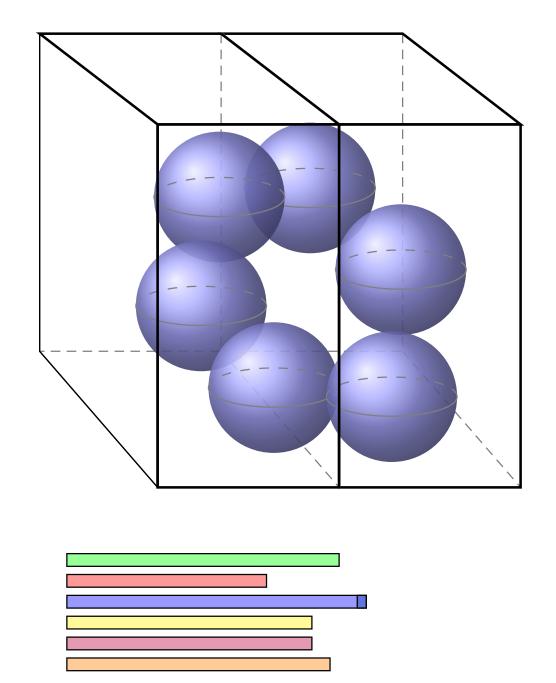


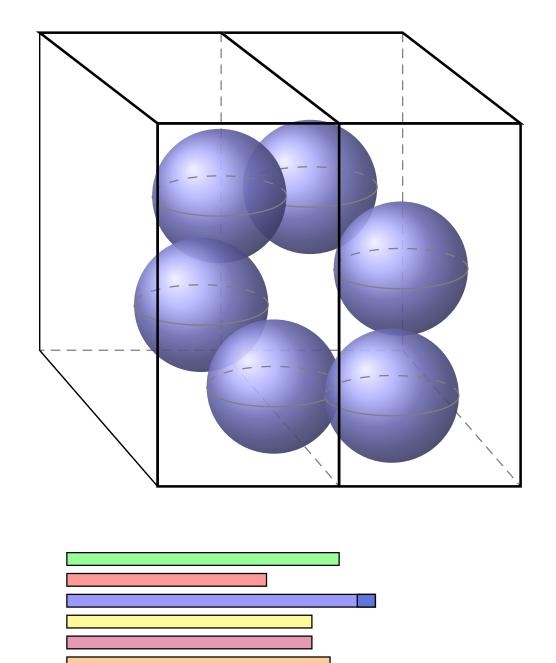


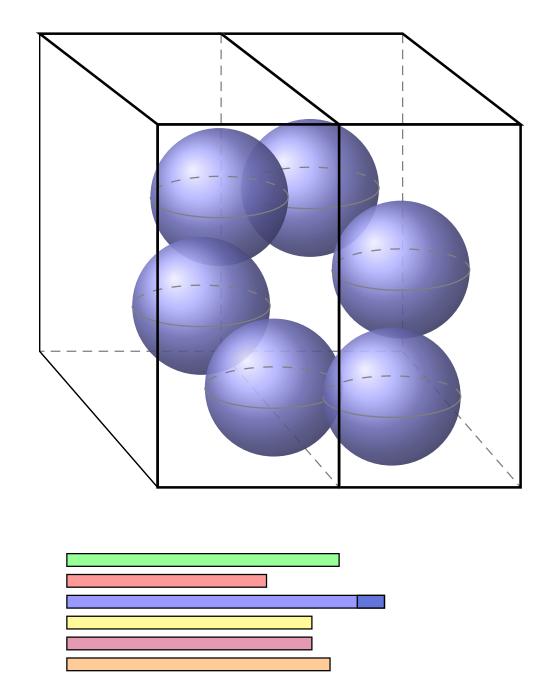


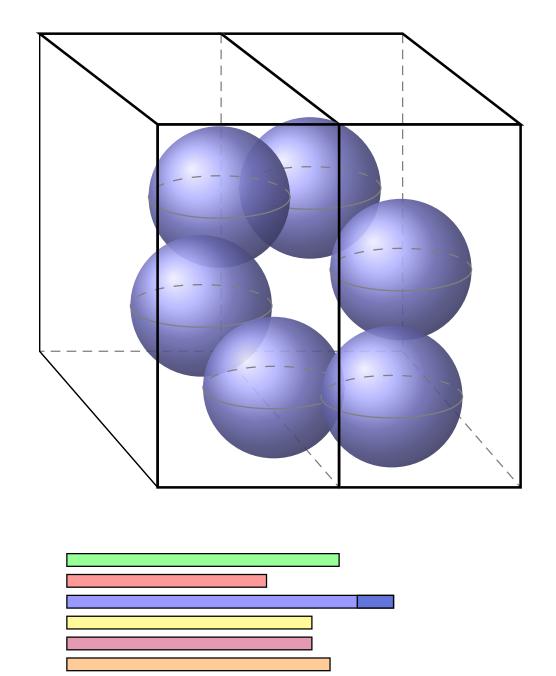


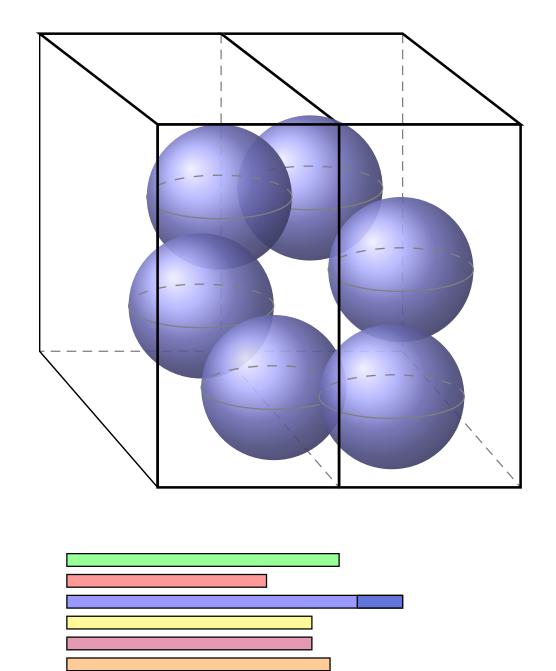


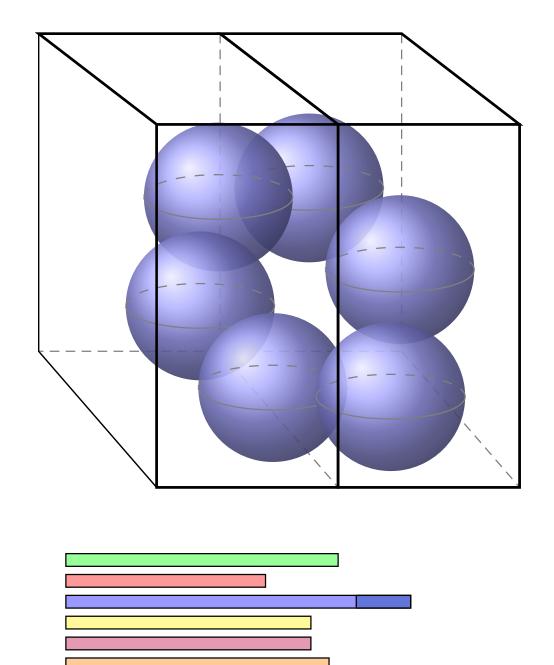


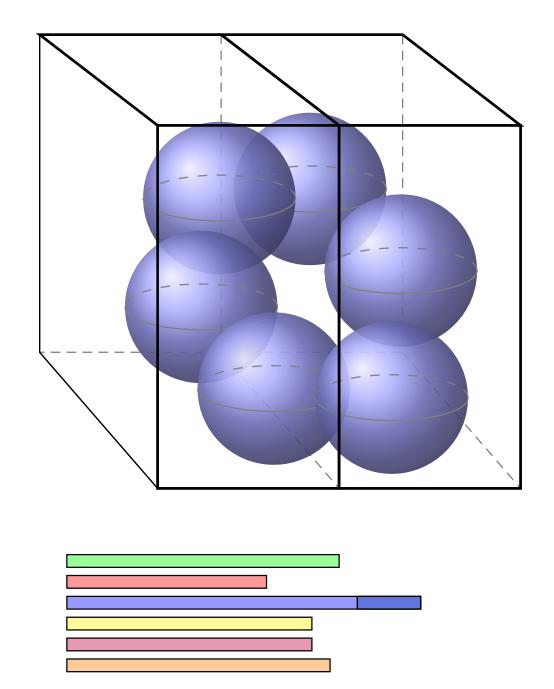


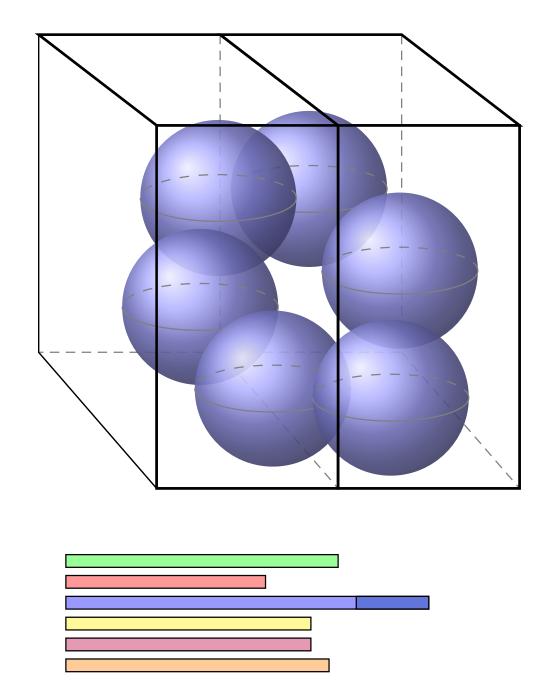


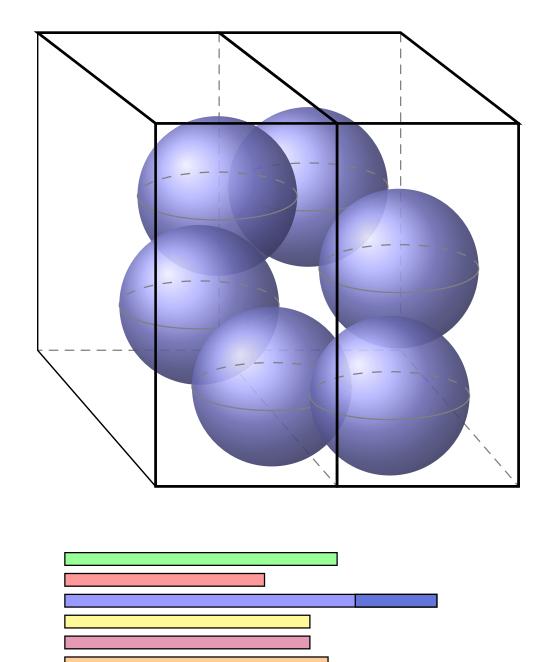


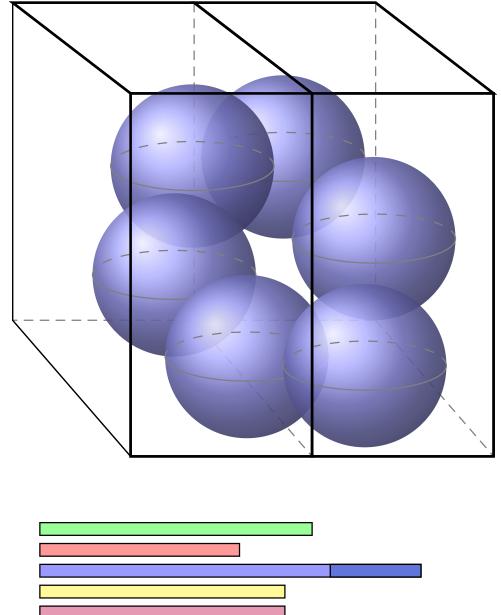


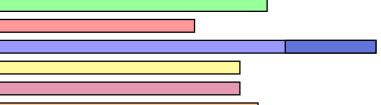


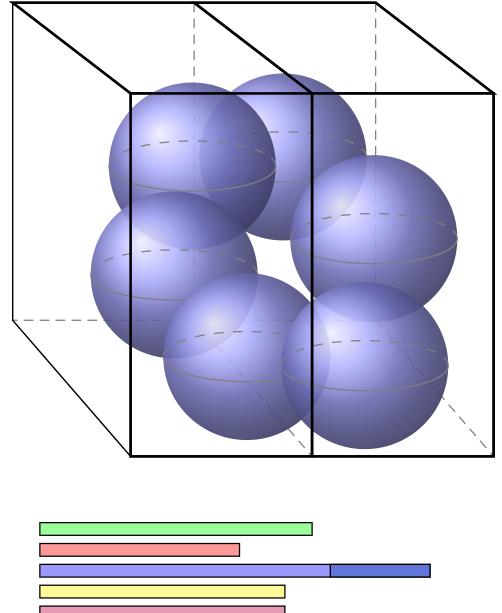




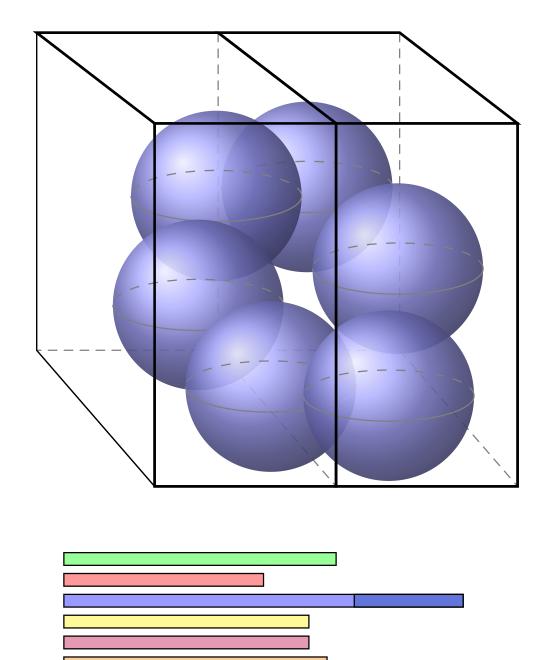


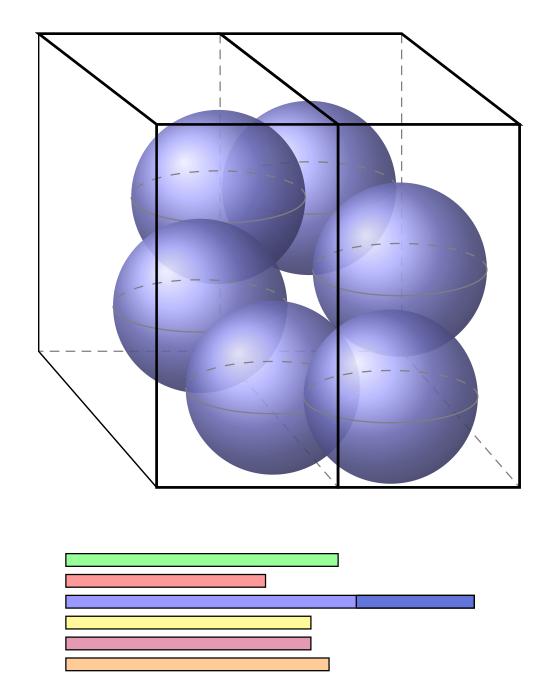


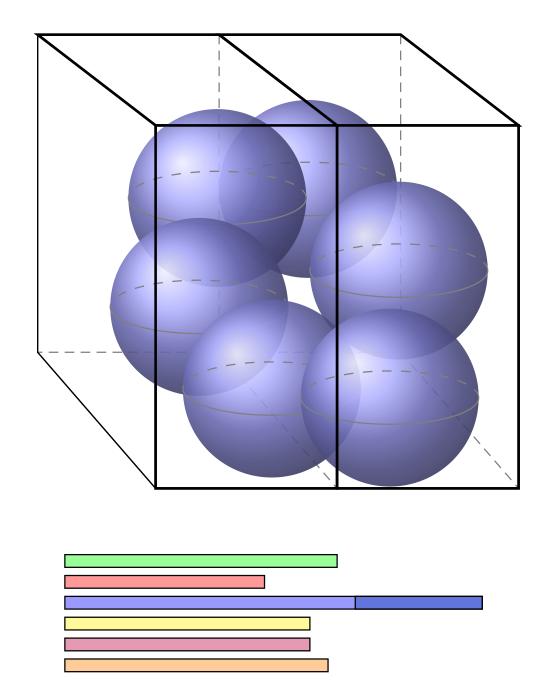


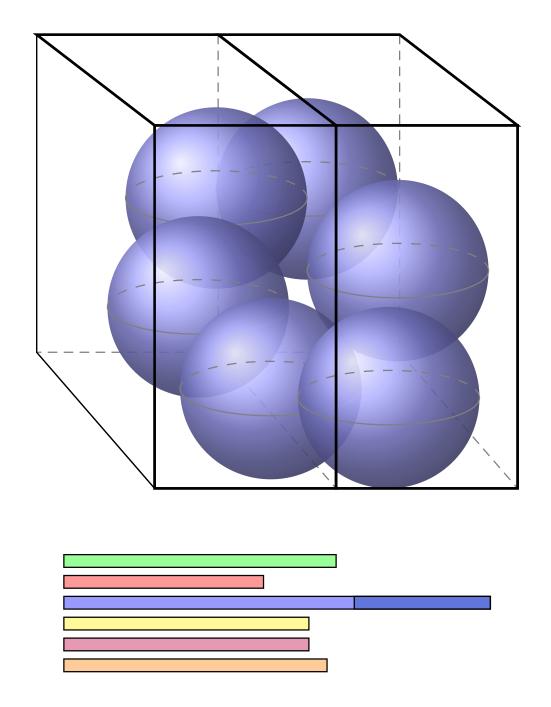


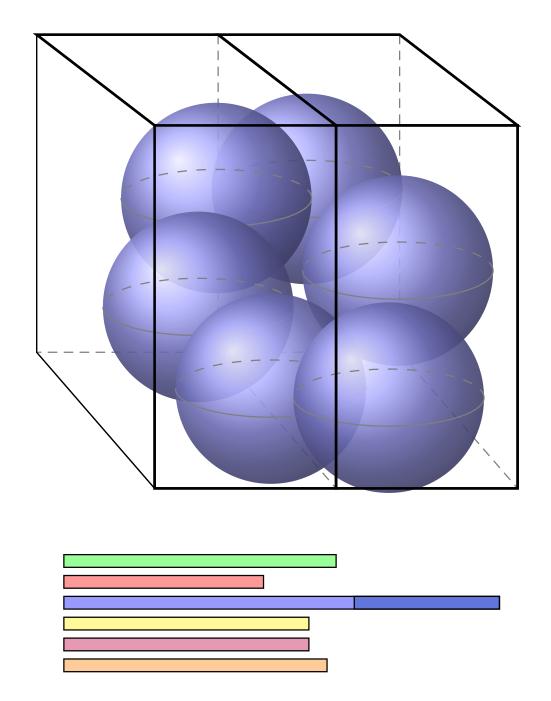


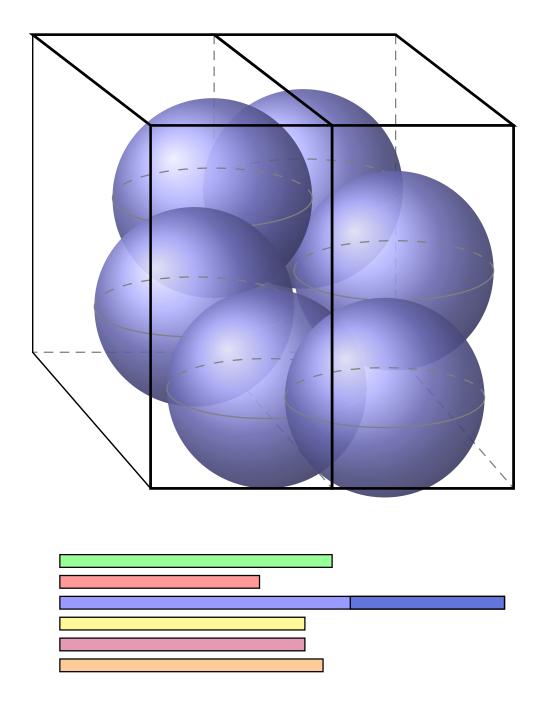


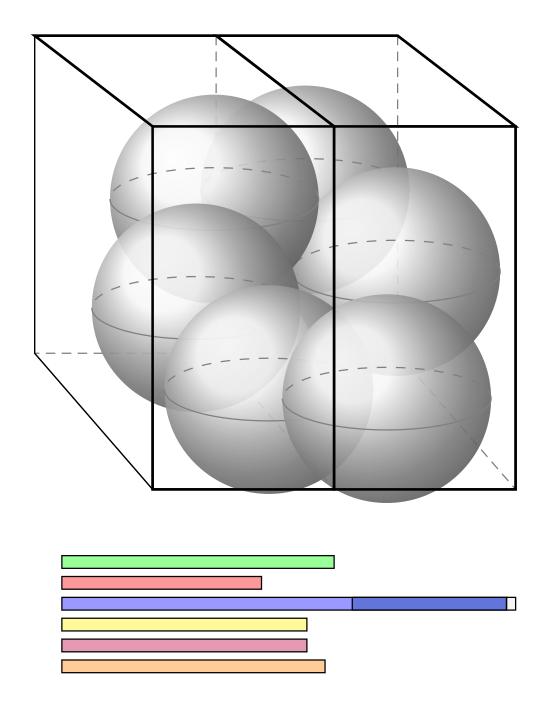


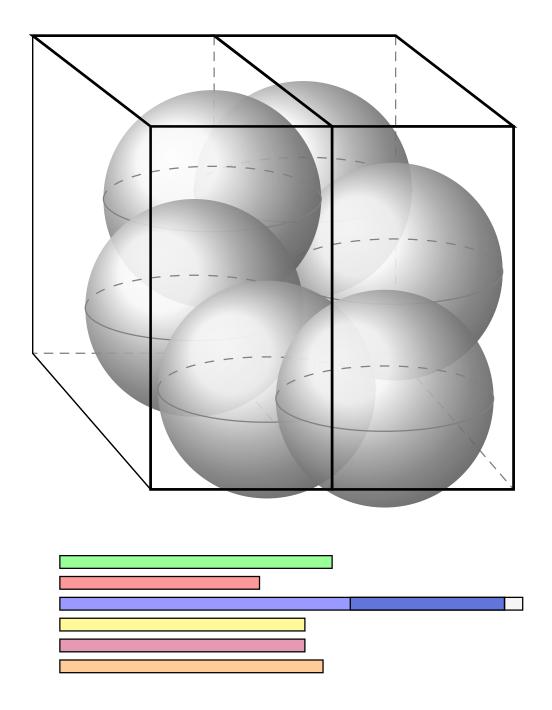


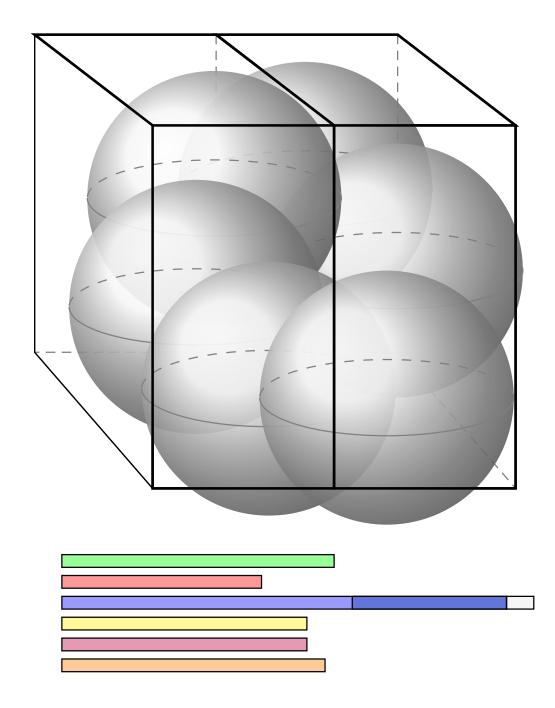


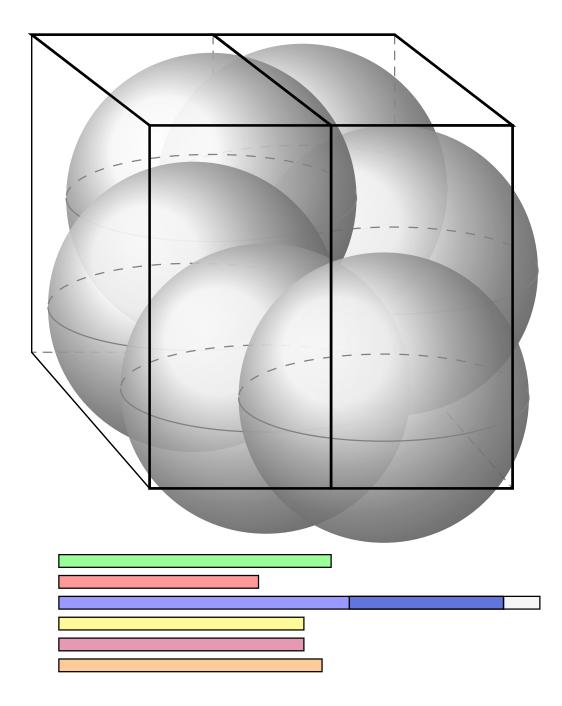


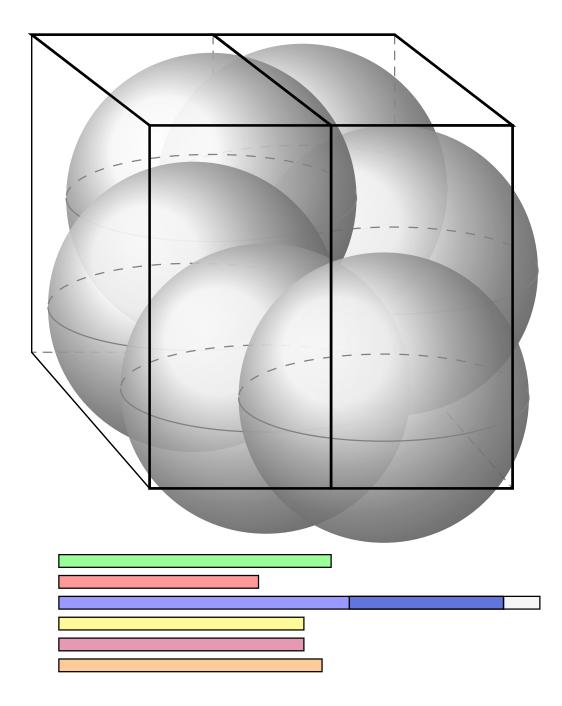








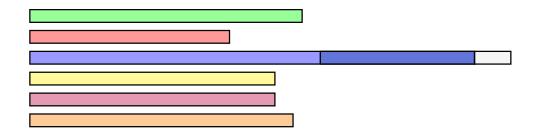




so persistent homology allows to detect apparent cycles in the data

fascinating – but

implication for practical data analysis needs to be figured out by other means and often remains mysterious.

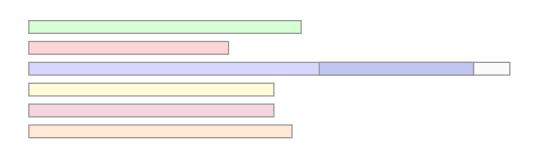


so persistent homology allows todetectapparent cyclesin the data

fascinating – but

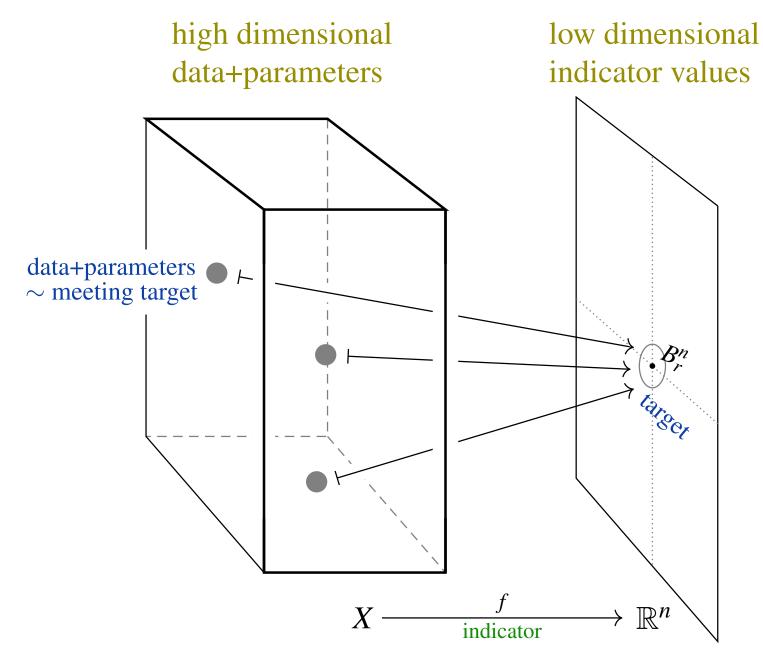
implication for practical data analysis needs to be figured out by other means and often remains mysterious.

let's recall the practically relevant question



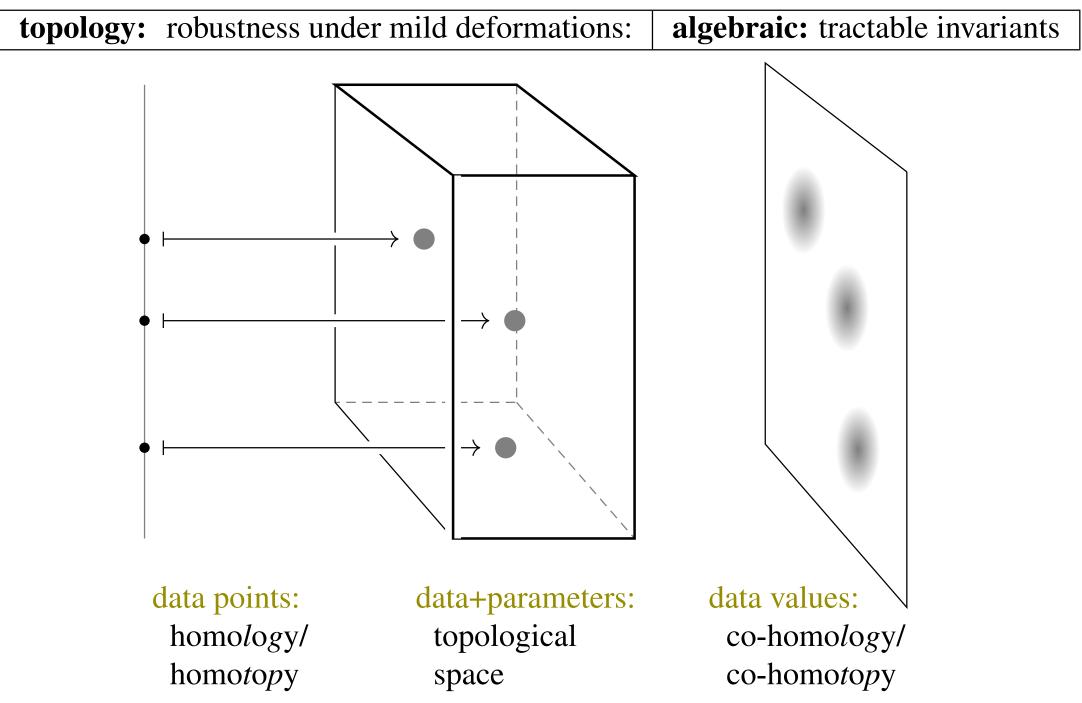
#### Find data meeting prescribed target with uncertainties – The problem.

Given high-dimensional data+parameters and a handful of indicators subject to uncertainty & noise. *Can a given target be met?* 



#### Find data meeting prescribed target with uncertainties – The strategy.

Use mathematical tools from *algebraic topology*:

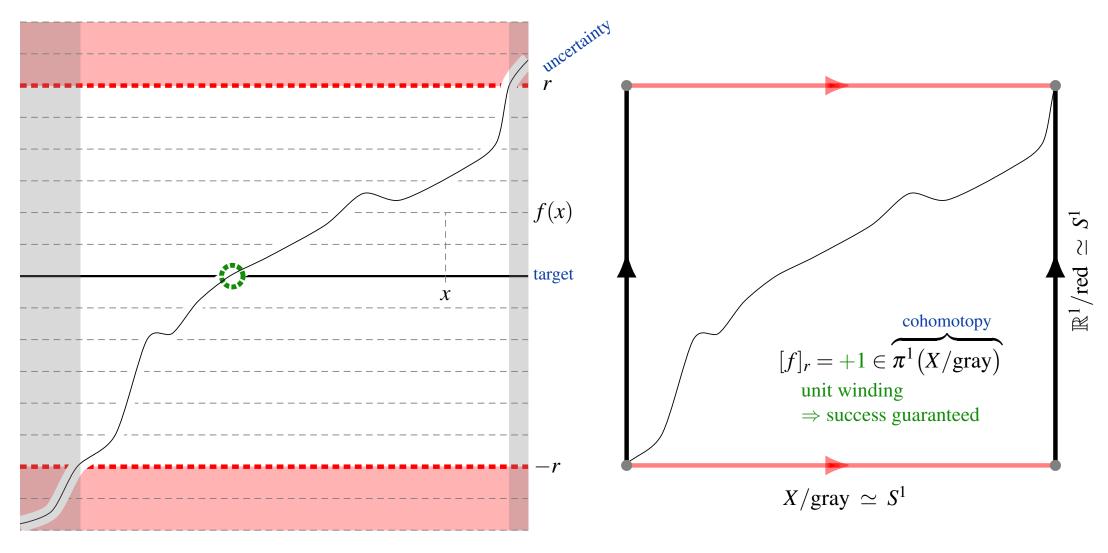


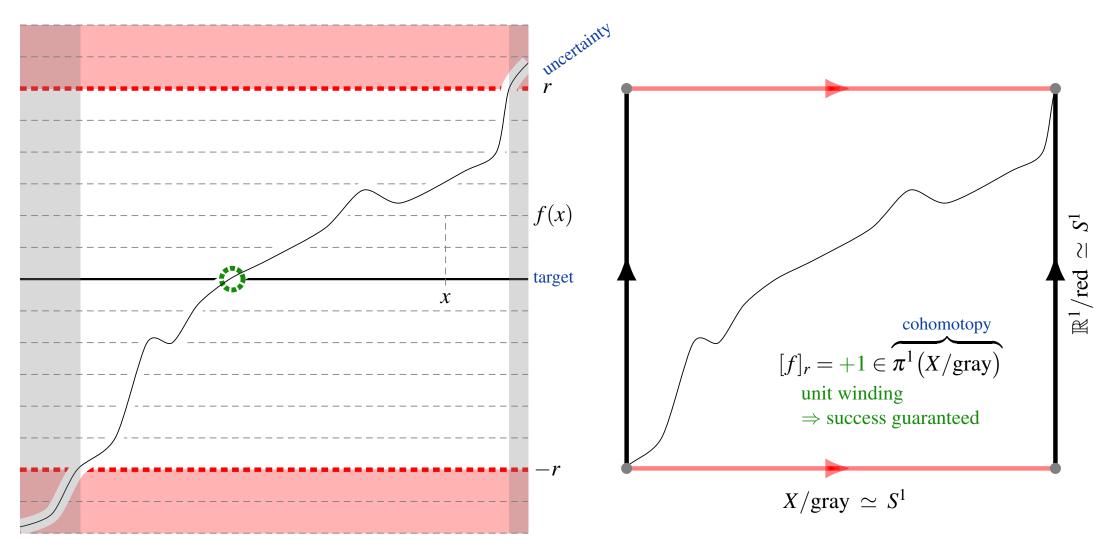
#### Find data meeting prescribed target with uncertainties – The strategy.

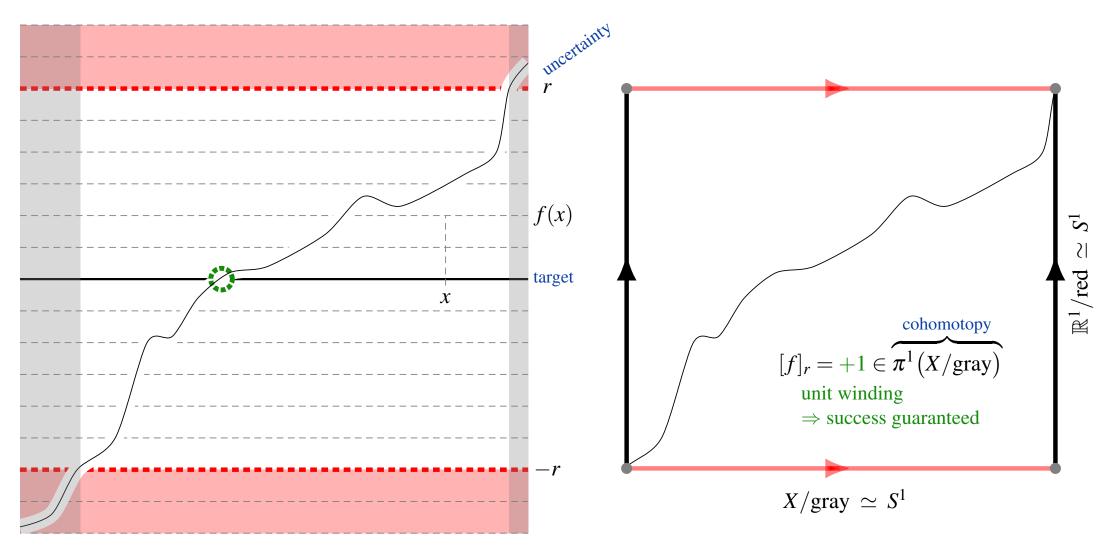
Use mathematical tools from *algebraic topology*: algebraic: tractable invariants **topology:** robustness under mild deformations: novel approach of persistent cohomotopy: find shape of data with fixed indicator value under given uncertainties data values: data+parameters: topological co-homology/

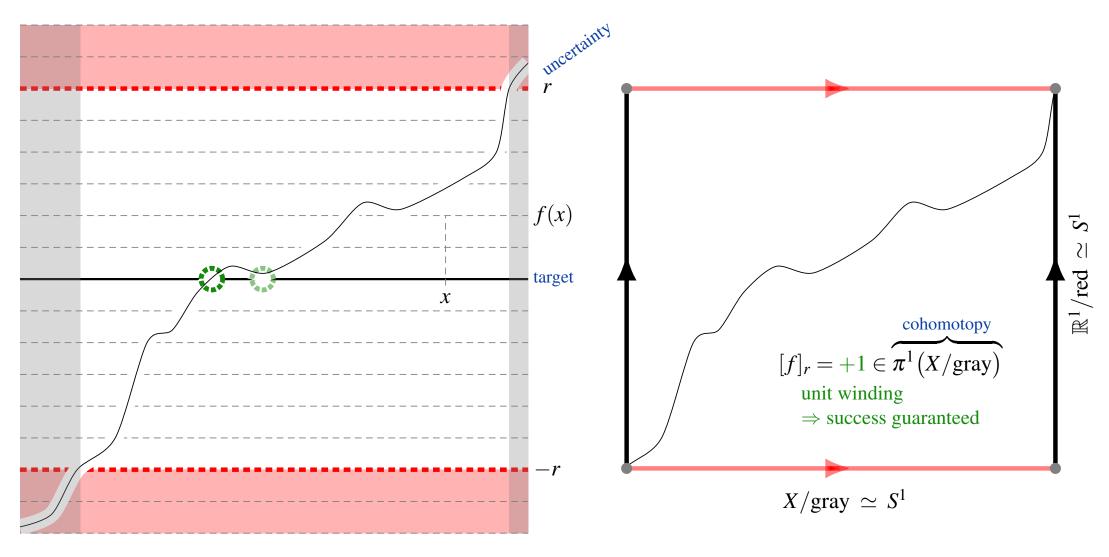
space

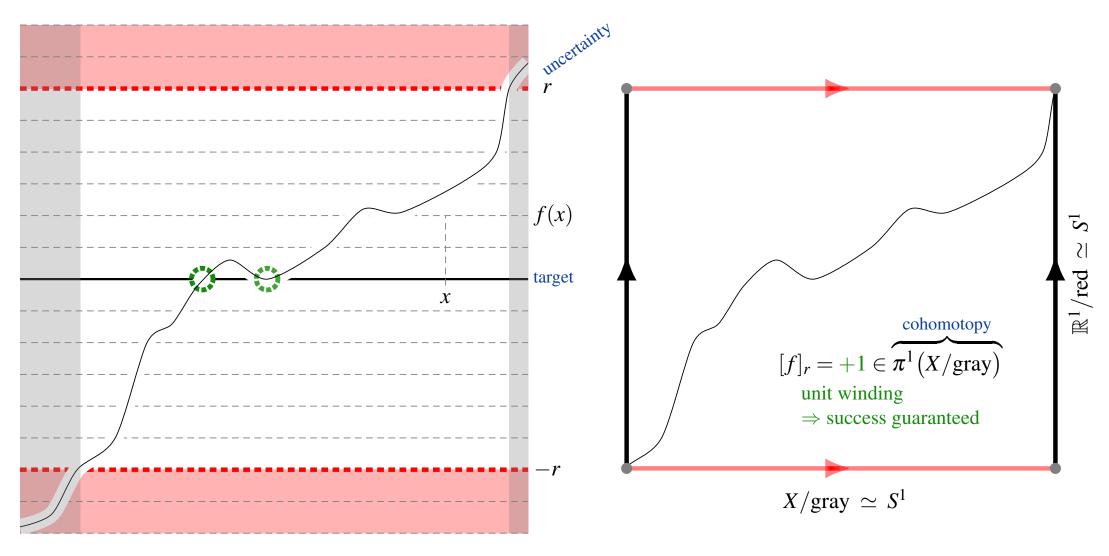
co-homotopy



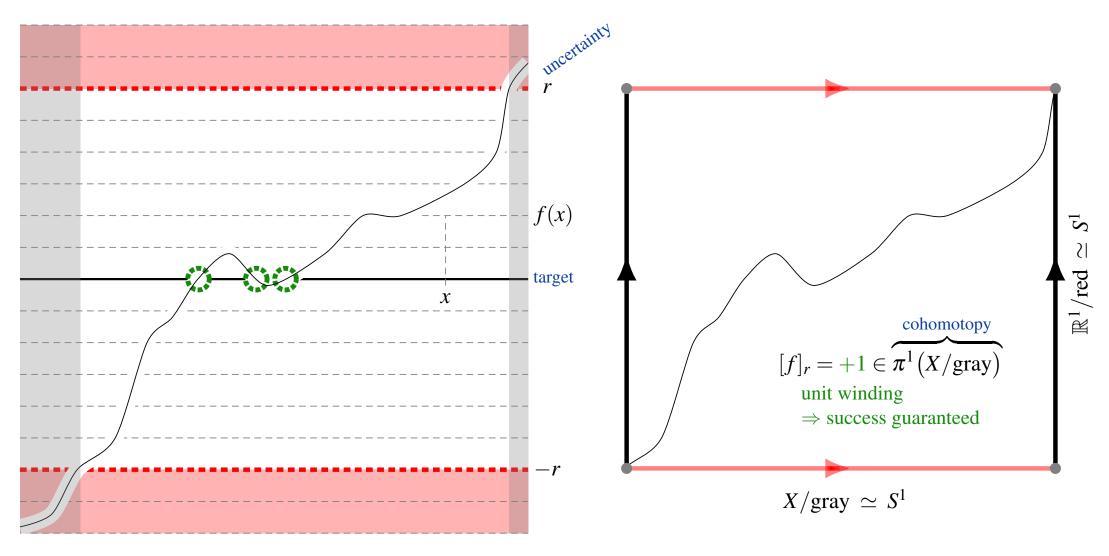




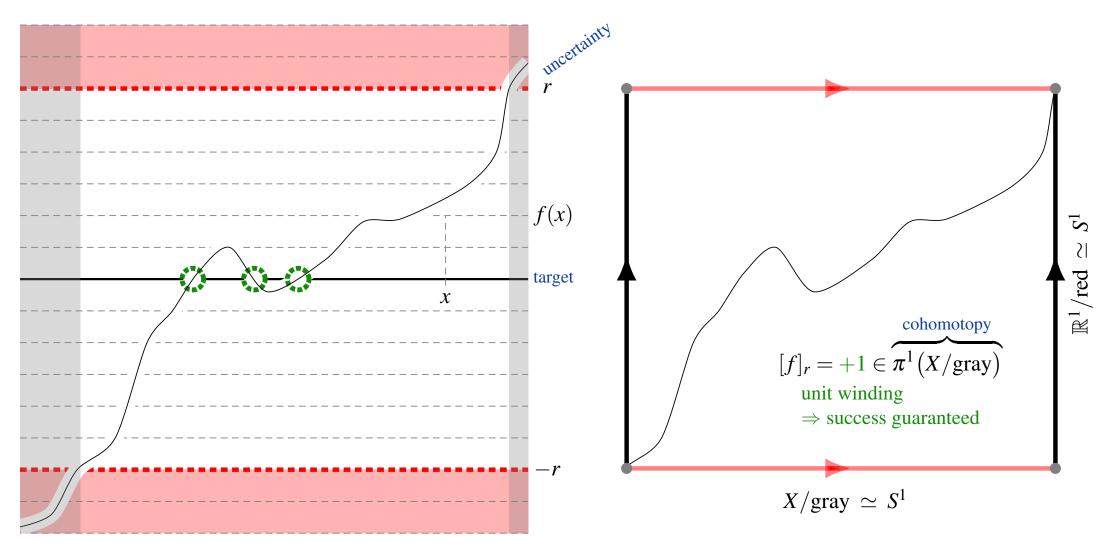


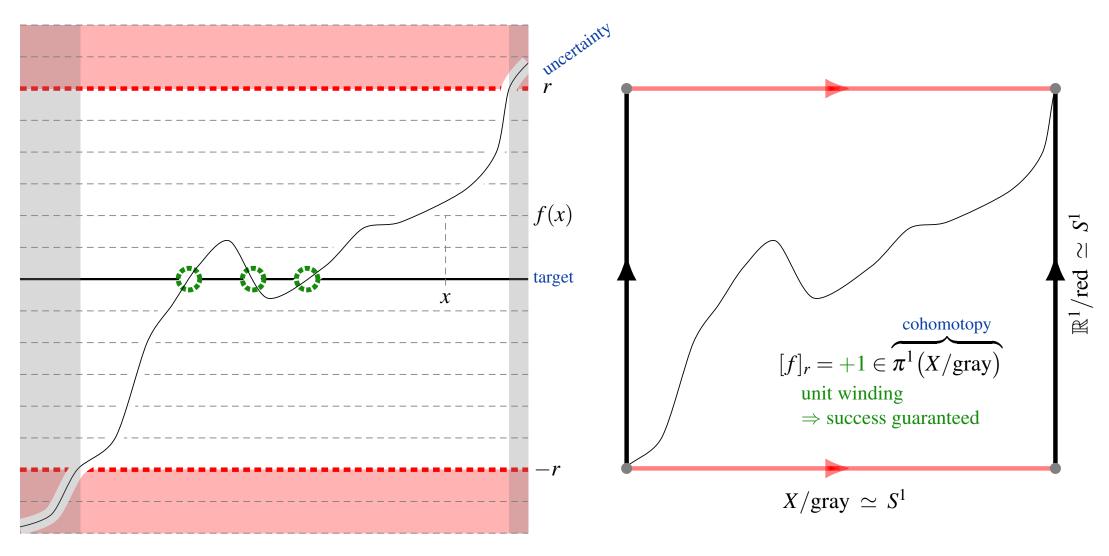


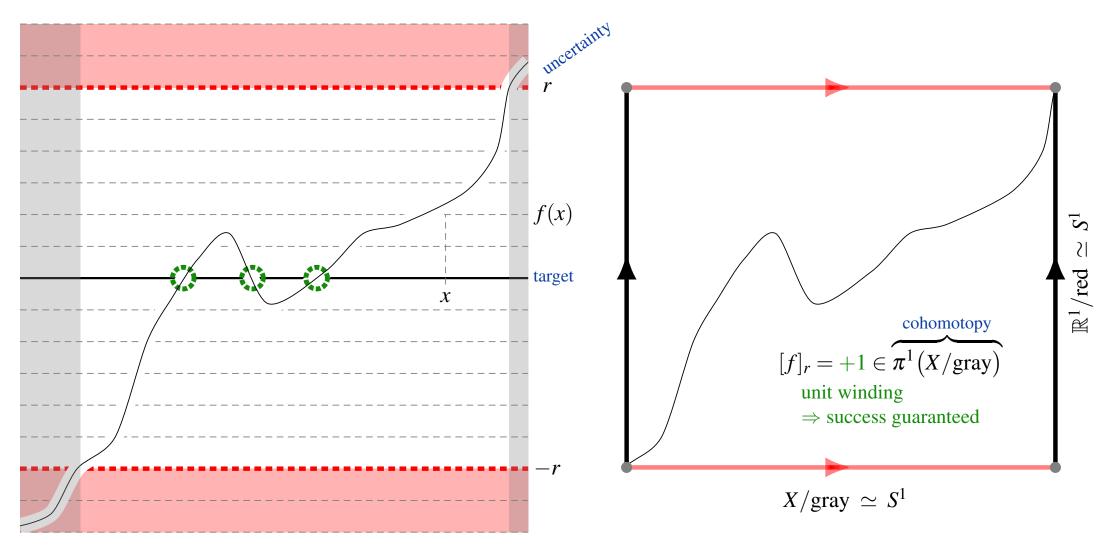
Narrow in on potentially successful data points within given uncertainty.
*Indicator winding* guarantees that at least one is successful *with certainty*.

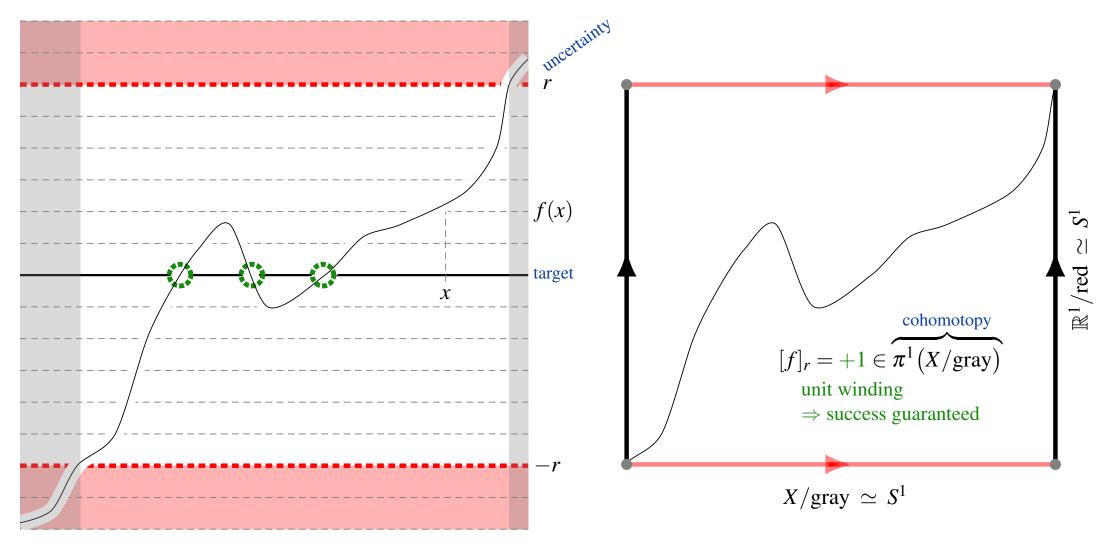


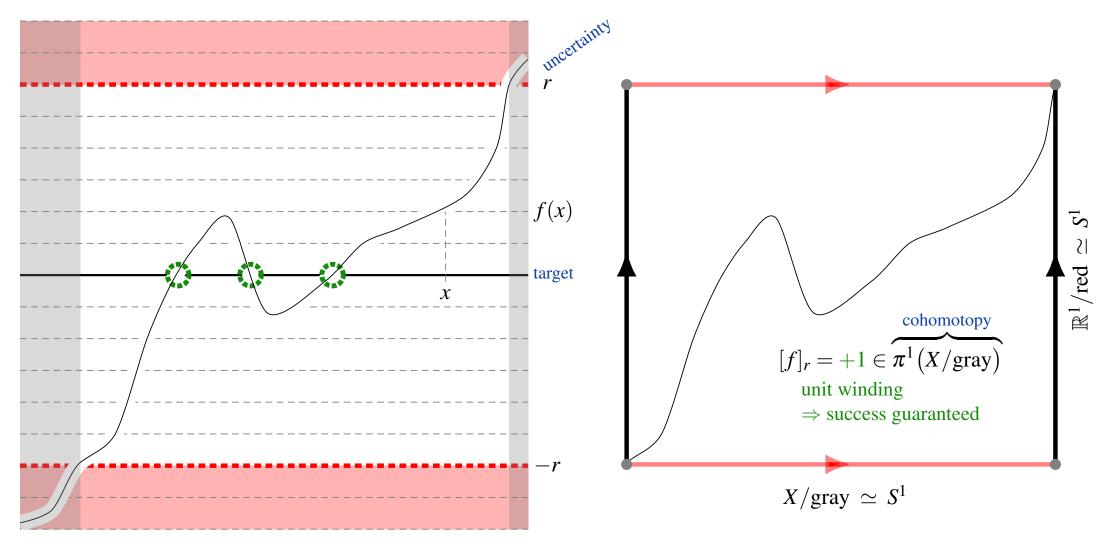
#### = data+parameters meet target

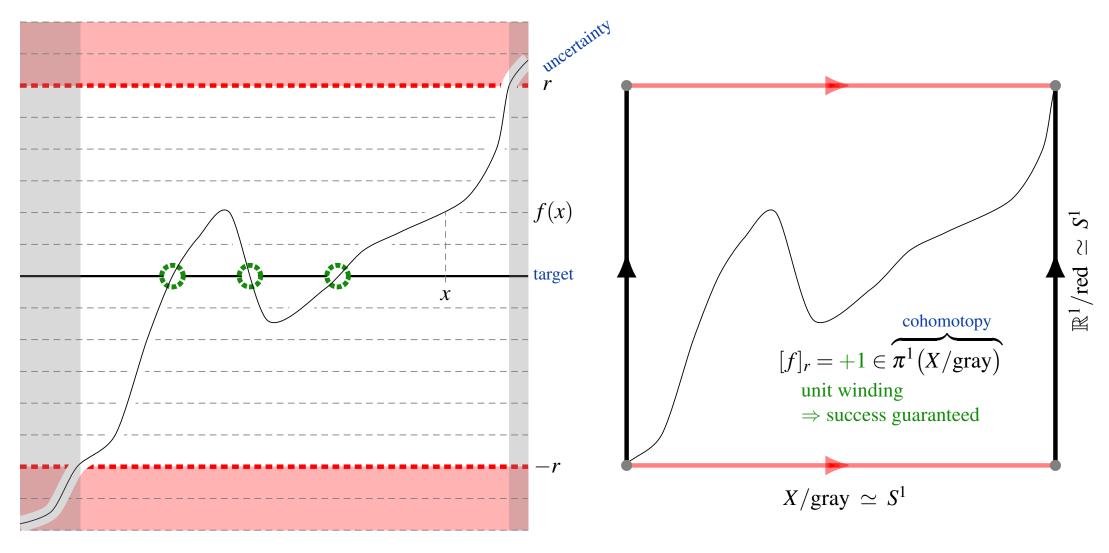




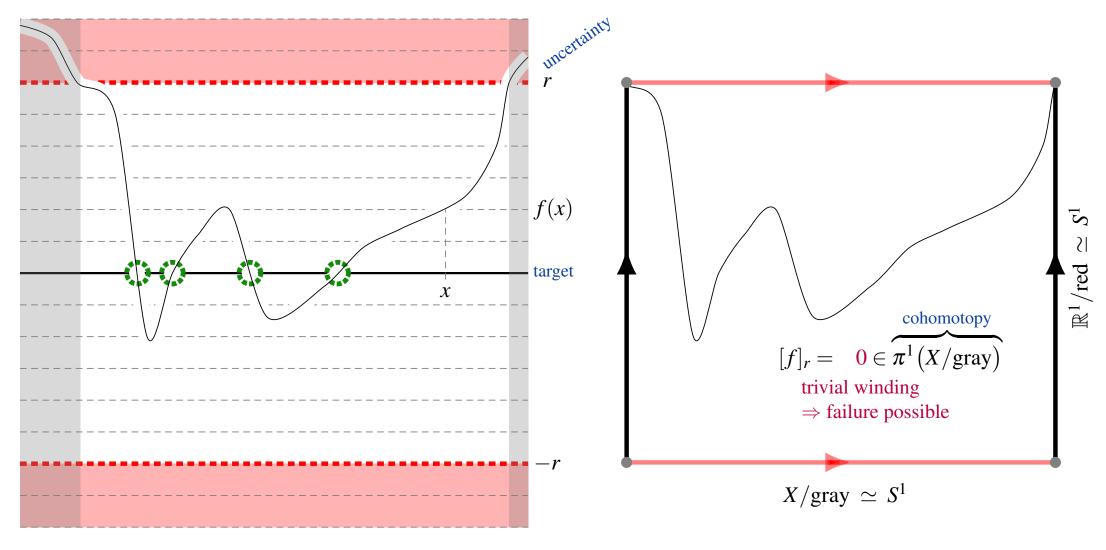




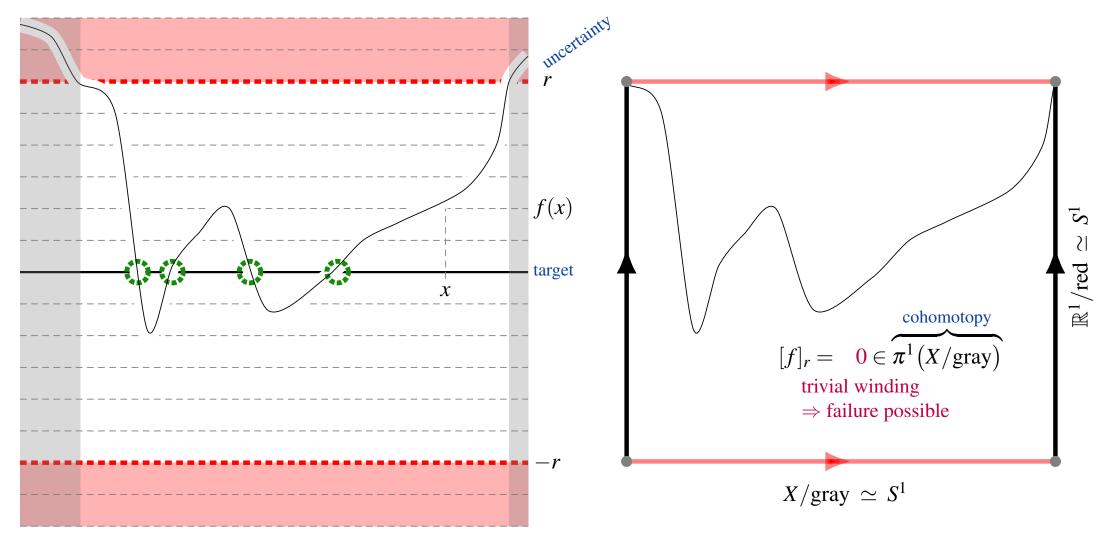




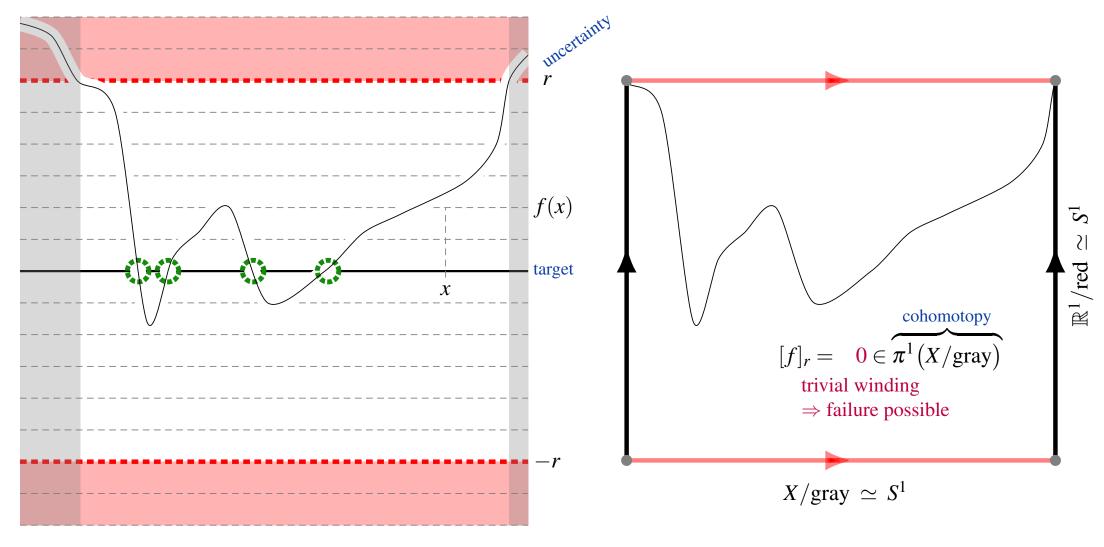
1. Narrow in on potentially successful data points within given uncertainty.



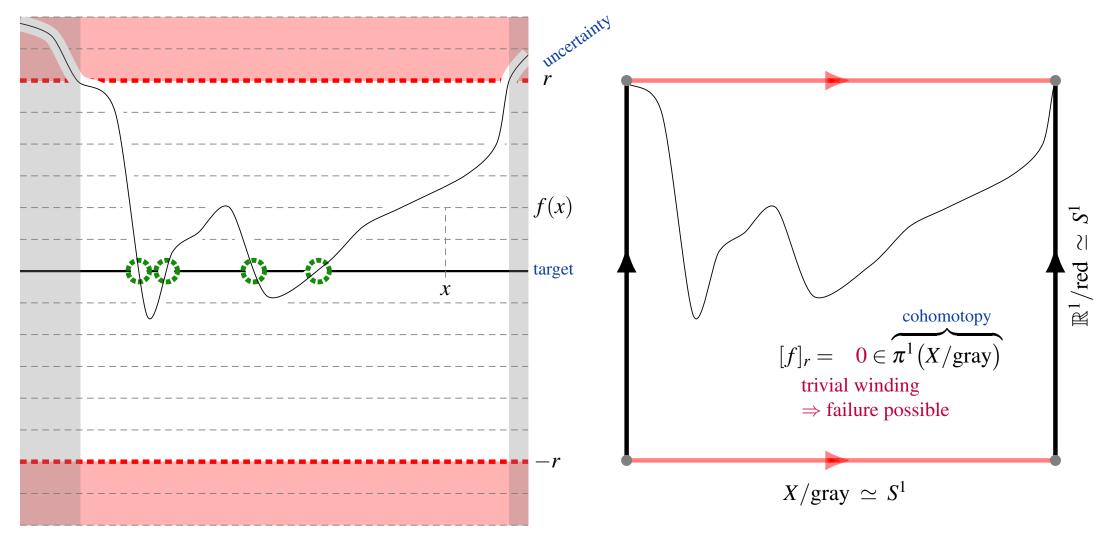
1. Narrow in on potentially successful data points within given uncertainty.



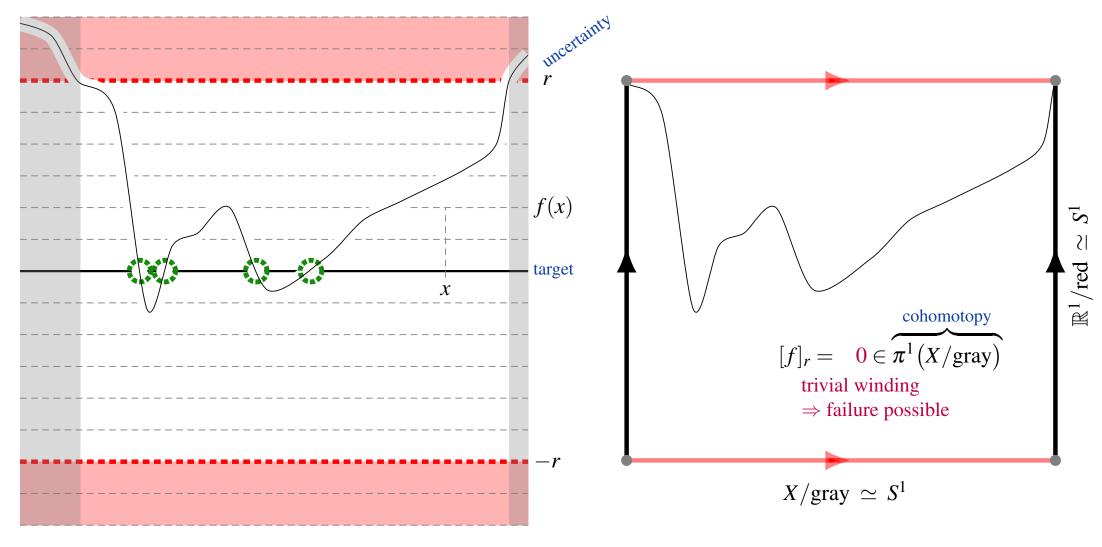
1. Narrow in on potentially successful data points within given uncertainty.



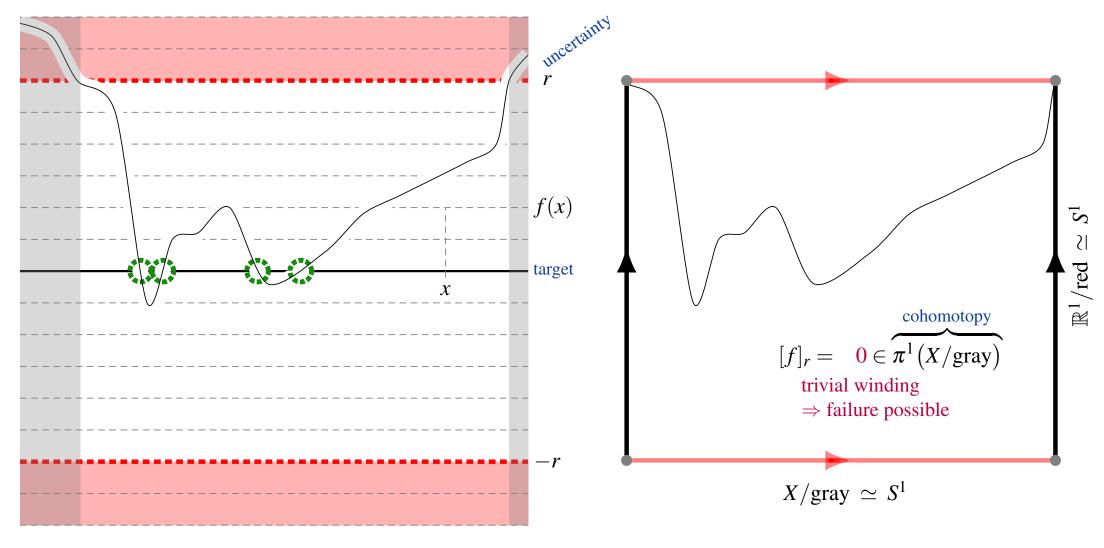
1. Narrow in on potentially successful data points within given uncertainty.



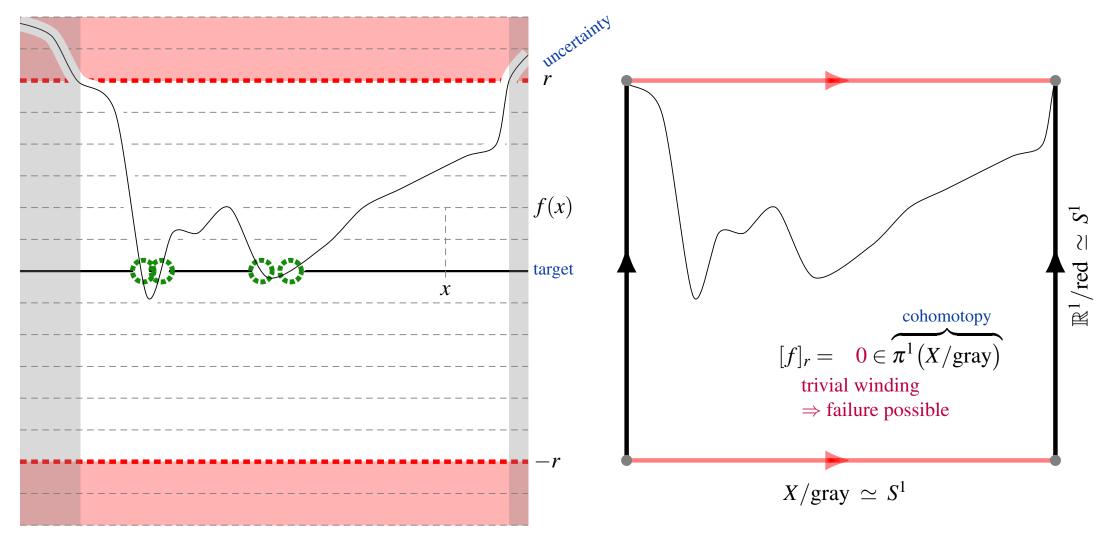
1. Narrow in on potentially successful data points within given uncertainty.



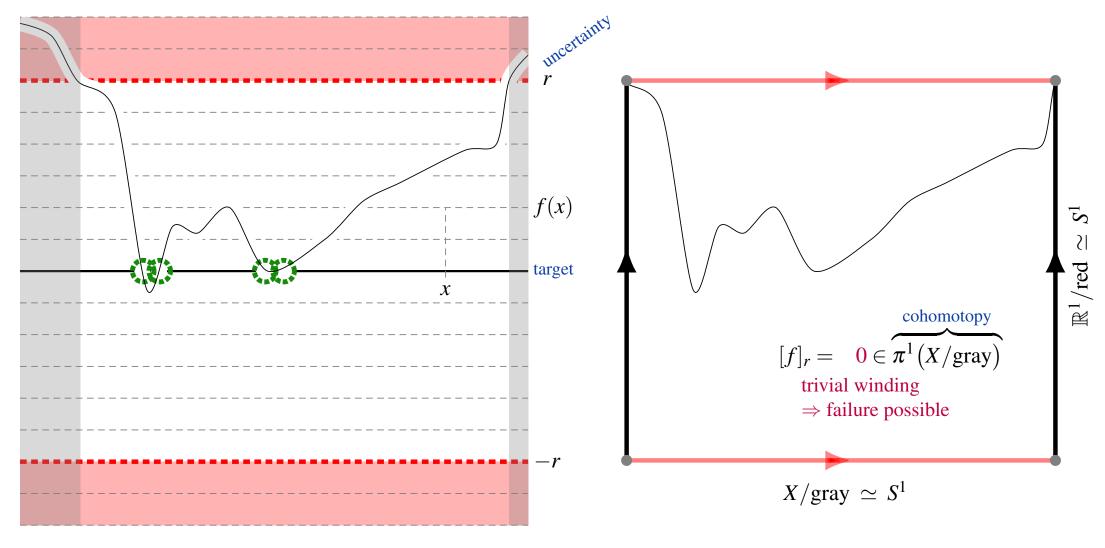
1. Narrow in on potentially successful data points within given uncertainty.



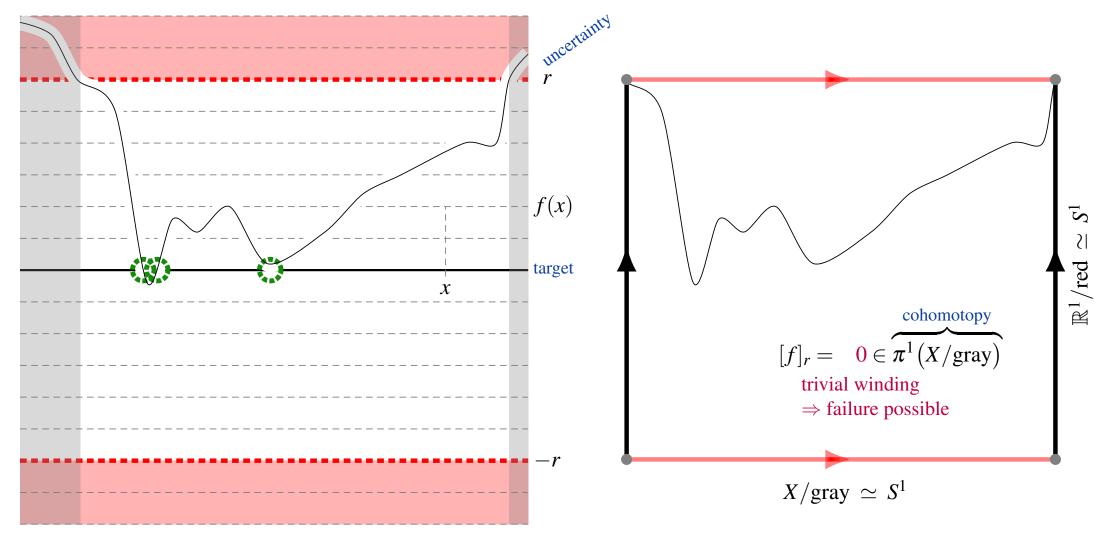
1. Narrow in on potentially successful data points within given uncertainty.



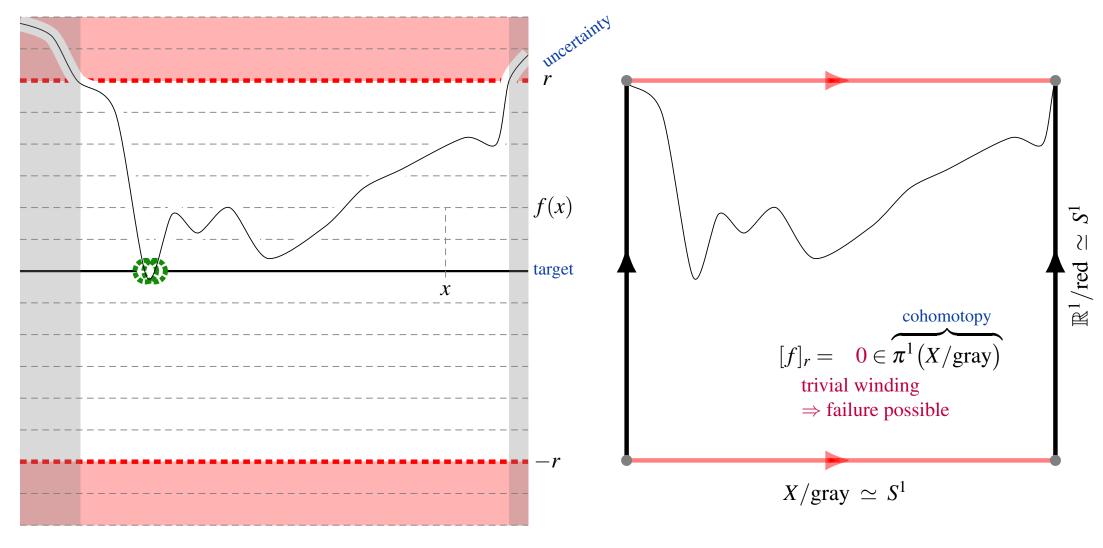
1. Narrow in on potentially successful data points within given uncertainty.



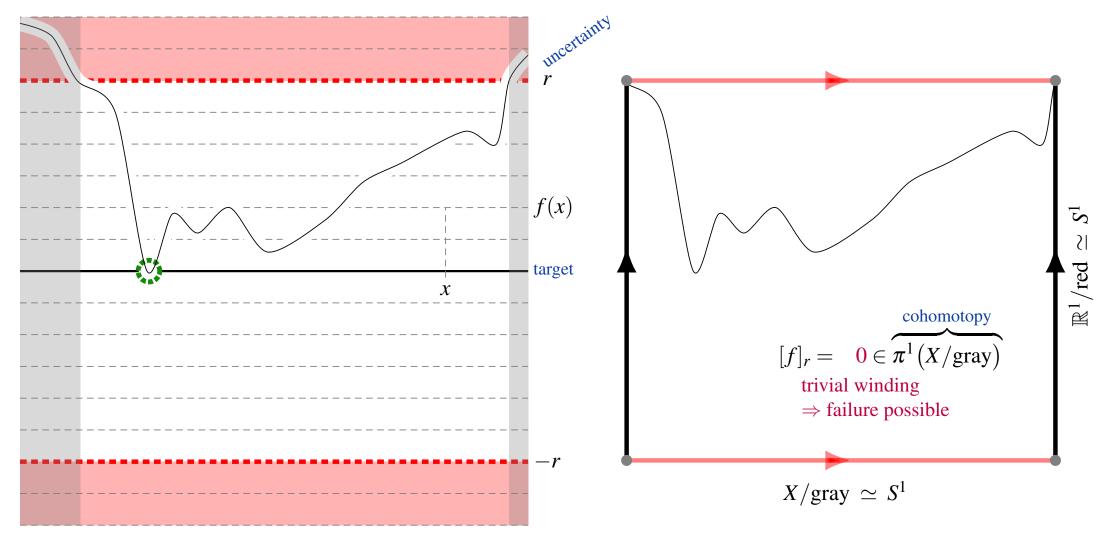
1. Narrow in on potentially successful data points within given uncertainty.



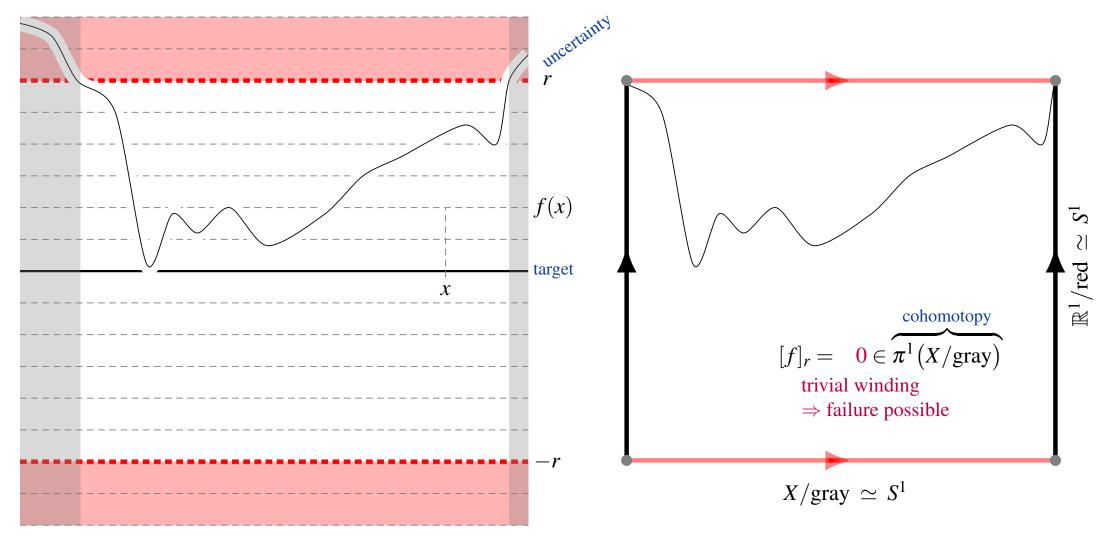
1. Narrow in on potentially successful data points within given uncertainty.



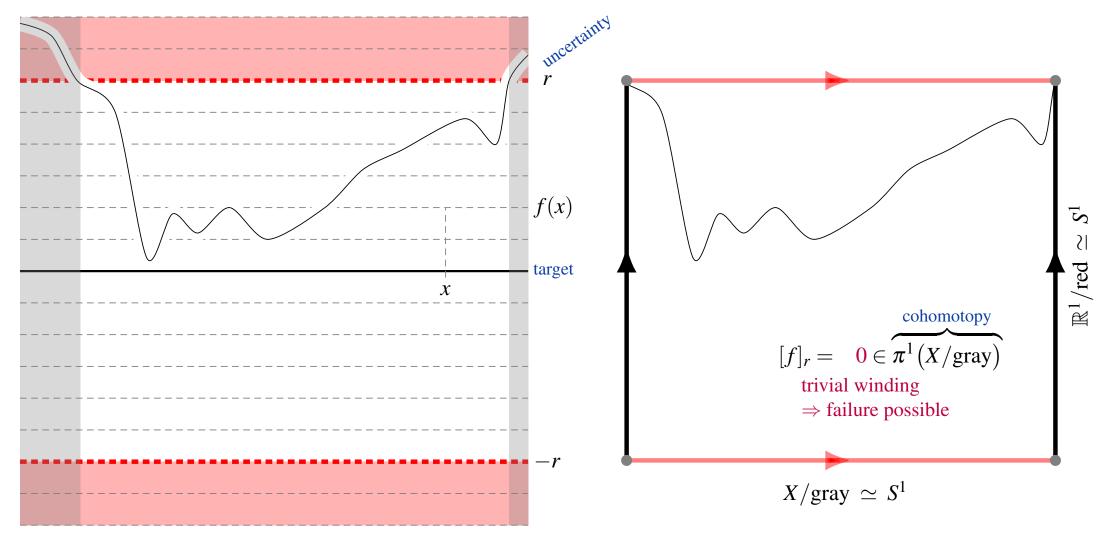
1. Narrow in on potentially successful data points within given uncertainty.



1. Narrow in on potentially successful data points within given uncertainty.

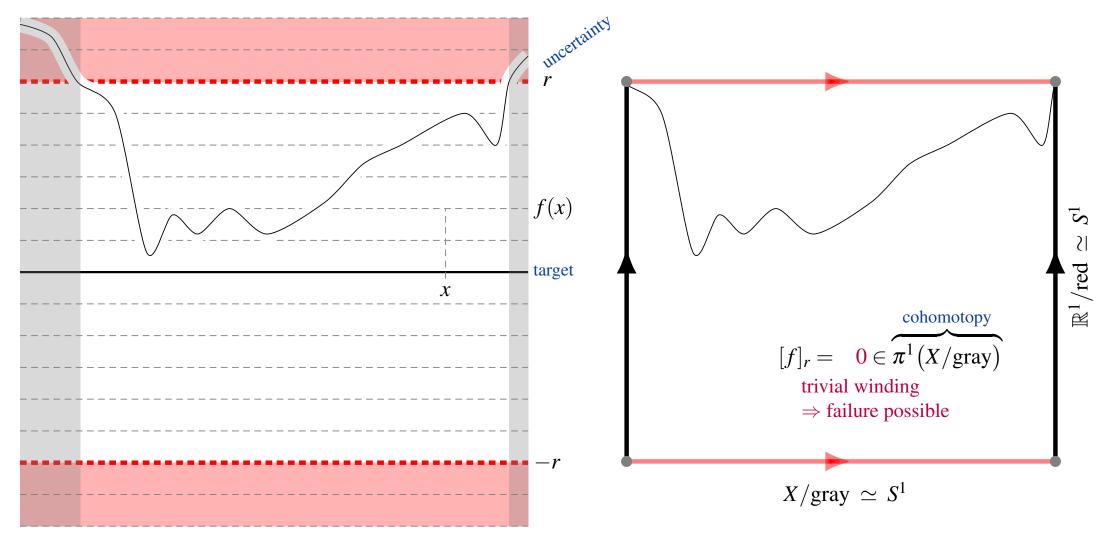


1. Narrow in on potentially successful data points within given uncertainty.



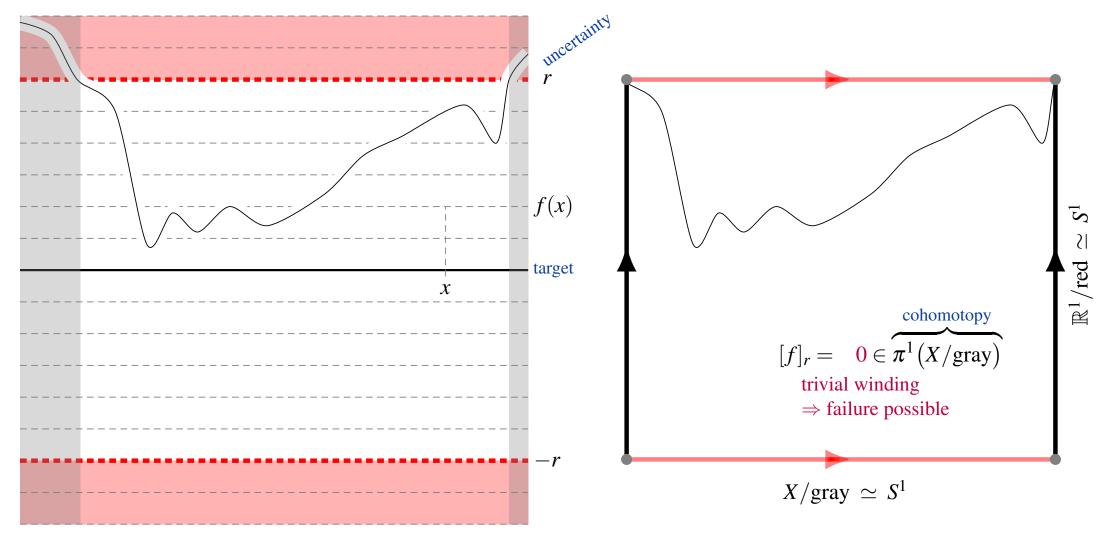
1. Narrow in on potentially successful data points within given uncertainty.

2. Indicator winding guarantees that at least one is successful with certainty.

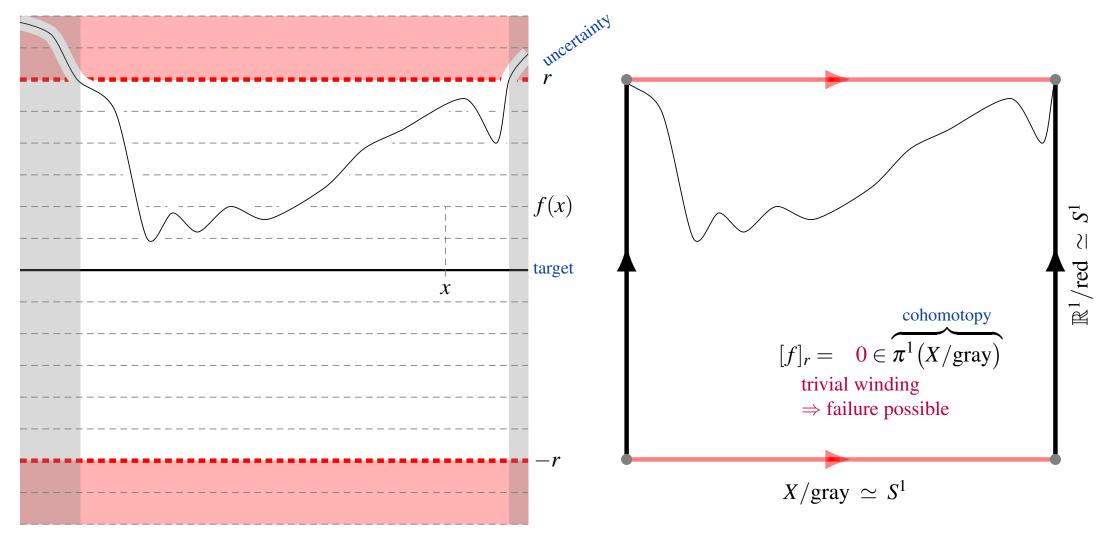


= data+parameters meet target

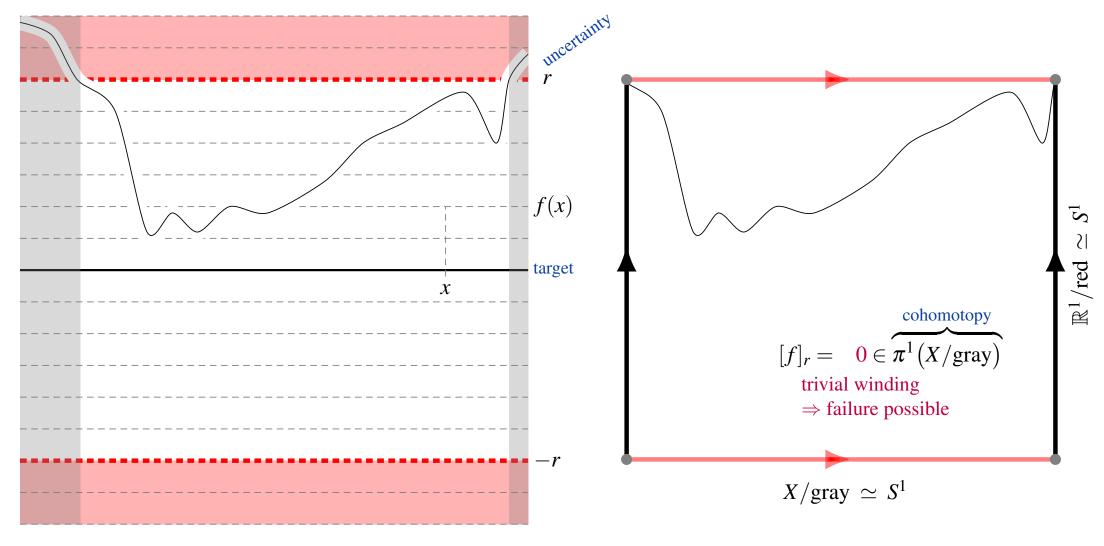
1. Narrow in on potentially successful data points within given uncertainty.

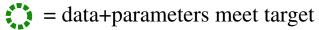


1. Narrow in on potentially successful data points within given uncertainty.

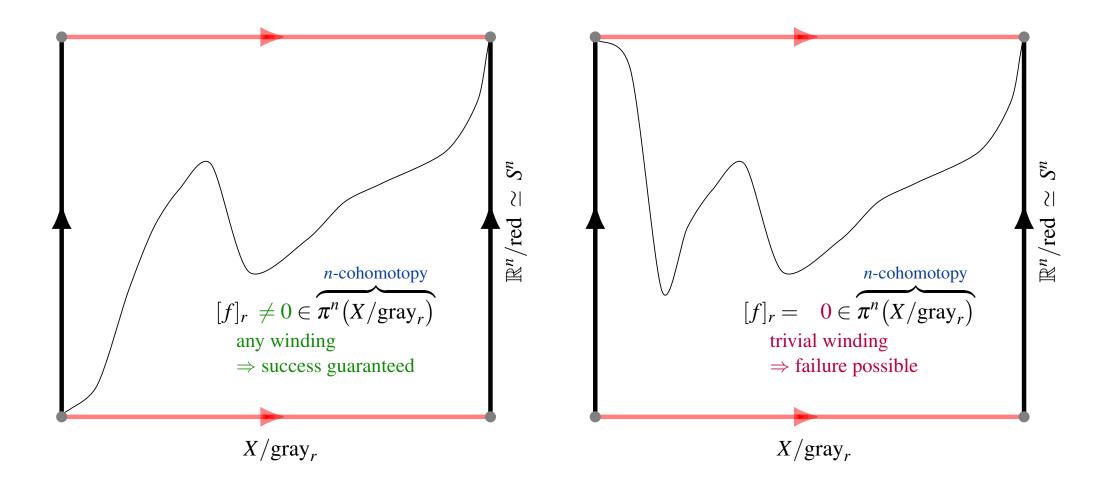


1. Narrow in on potentially successful data points within given uncertainty.



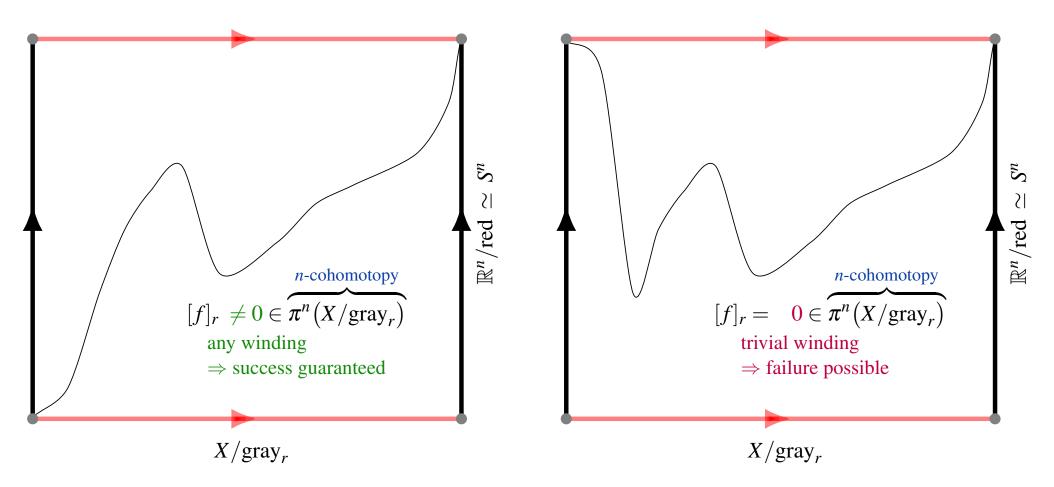


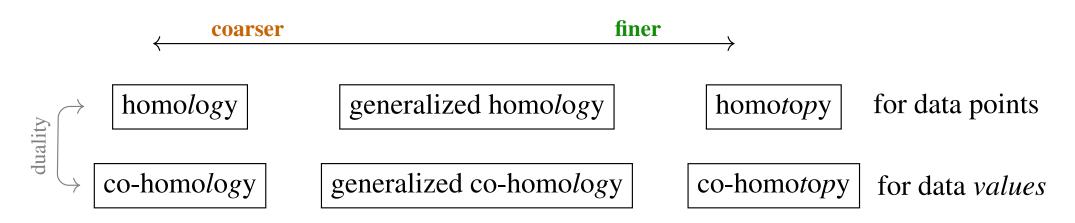
**Theorem** [FK17][FKW18]: For dim(X)  $\leq 2n - 4$ the persistent cohomotopy  $[f]_r$  is *computable*, hence the success guarantee is *decidable*.

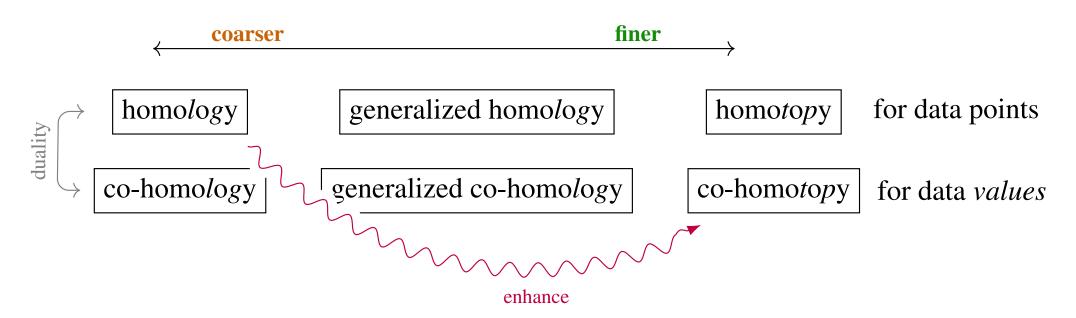


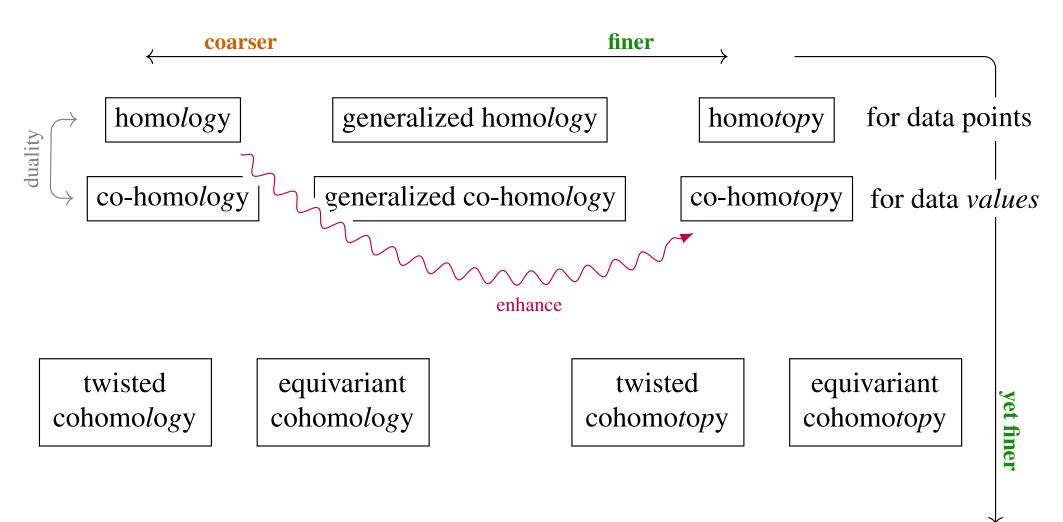
**Theorem** [FK17][FKW18]: For dim(X)  $\leq 2n - 4$ the persistent cohomotopy  $[f]_r$  is *computable*, hence the success guarantee is *decidable*.

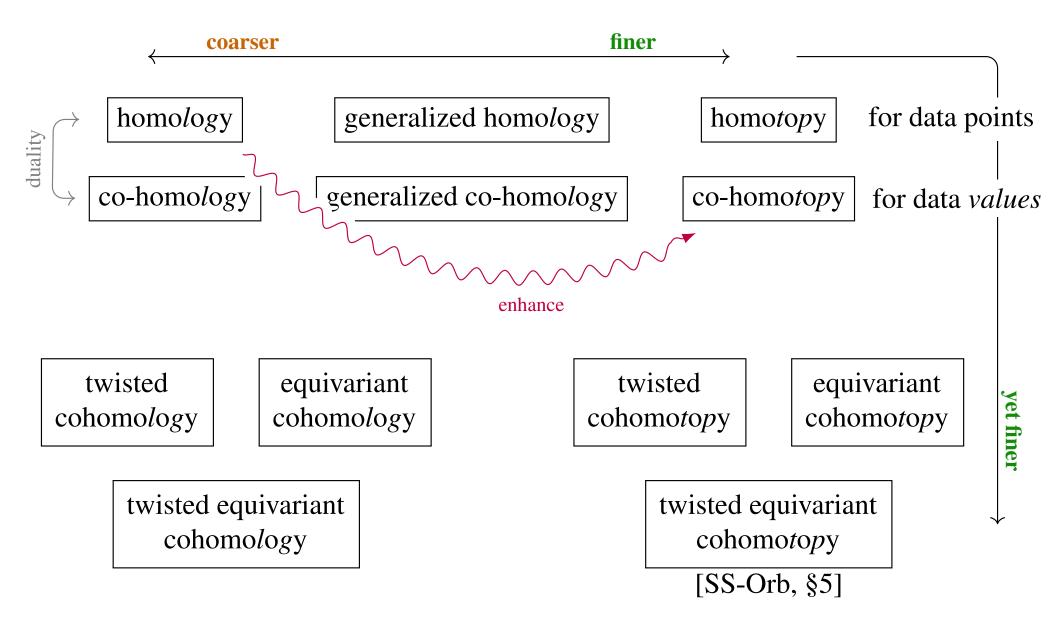
**Bonus:** Persistence of  $[f]_{\bullet}$  yields the max. tolerable uncertainties.

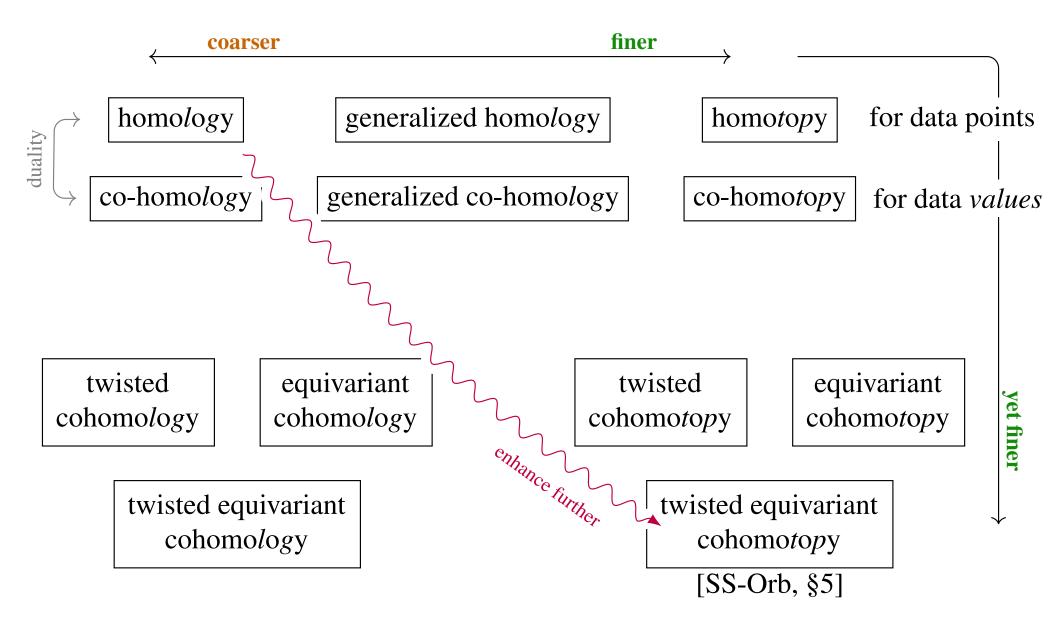






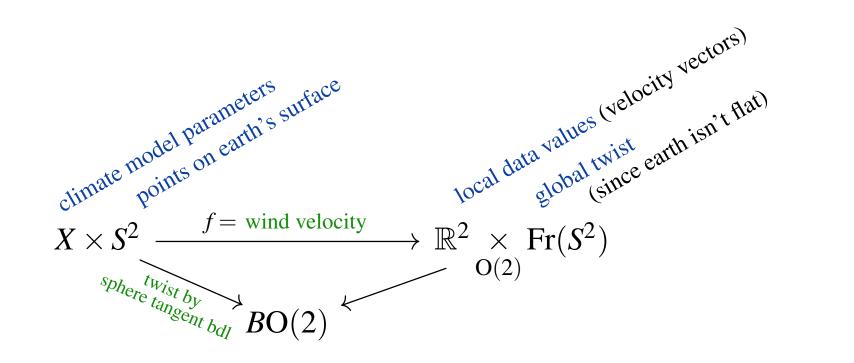






### **Cohomotopy – further enhancements: twisting.**

Often indicator values include *tangent vectors* to a manifold for example: *global wind velocity* 

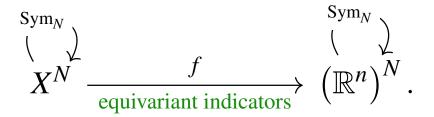


In such case indicator winding is in **tangentially twisted cohomotopy** ([FSS-Char][FSS19][FSS21a][FSS21b]).

#### **Cohomotopy – further enhancements: equivariance.**

Often data arises in *multiple copies*  $X^N = X \times \cdots \times X$ , where the order of the copies must not matter  $X \times X \times X \cdots \times X \times X$ 

this means that indicator values must by *equivariant* under the action of the permutation group  $Sym_N$ :



In this situation indicator winding is in **equivariant cohomotopy** 

([SS20a][BSS21]).

#### **Cohomotopy – further enhancements: equivariance.**

Often data arises in *multiple copies*  $X^N = X \times \cdots \times X$ , where the order of the copies must not matter  $X \times X \times X \cdots \times X \times X$ 

this means that indicator values must by *equivariant* under the action of the permutation group  $Sym_N$ :



In this situation indicator winding is in **equivariant cohomotopy** 

([SS20a][BSS21]).

In general, indicator values are both: equivariant *and* twisted. In this general case indicator winding is in **twisted equivariant cohomotopy** ([SS-Orb, §5]).

• Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.

- Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.
- Popular *persistent homology* is only the coarsest of these tools.

- Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.
- Popular *persistent homology* is only the coarsest of these tools.
- Novel *persistent homotopy* is a much finer tool, but its output still hard to interpret in practice of data analysis.

- Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.
- Popular *persistent homology* is only the coarsest of these tools.
- Novel *persistent homotopy* is a much finer tool, but its output still hard to interpret in practice of data analysis.
- Practically relevant question of data meeting target indicators is answered instead by *co*-homology and the finer *co-homotopy*.

- Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.
- Popular *persistent homology* is only the coarsest of these tools.
- Novel *persistent homotopy* is a much finer tool, but its output still hard to interpret in practice of data analysis.
- Practically relevant question of data meeting target indicators is answered instead by *co*-homology and the finer *co-homotopy*.
- Basics of *persistent cohomotopy* have been developed [FK17] and *proven* to be practically useful (computable) [FKW18].

- Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.
- Popular *persistent homology* is only the coarsest of these tools.
- Novel *persistent homotopy* is a much finer tool, but its output still hard to interpret in practice of data analysis.
- Practically relevant question of data meeting target indicators is answered instead by *co*-homology and the finer *co-homotopy*.
- Basics of *persistent cohomotopy* have been developed [FK17] and *proven* to be practically useful (computable) [FKW18].
- General applications will require twisted & equivariant cohomotopy [SS-Orb].

- Topological data analysis brings power tools of algebraic topology to robustly detect global structure in large data.
- Popular *persistent homology* is only the coarsest of these tools.
- Novel *persistent homotopy* is a much finer tool, but its output still hard to interpret in practice of data analysis.
- Practically relevant question of data meeting target indicators is answered instead by *co*-homology and the finer *co-homotopy*.
- Basics of *persistent cohomotopy* have been developed [FK17] and *proven* to be practically useful (computable) [FKW18].
- General applications will require twisted & equivariant cohomotopy [SS-Orb].
- At <u>CQTS</u> we plan to develop the refined tool of persistent twisted equivariant cohomotopy for practical TDA.

[Ca09] G. Carlsson: *Topology and data* Bull. Amer. Math. Soc. 46 2 (2009) 255-308 [doi:10.1090/S0273-0979-09-01249-X]

[Ou15] S. Y. Oudot:

*Persistence Theory: From Quiver Representations to Data Analysis* Math Surv Monogr **209** AMS (2015) [pdf]

[DTU] Technical University of Denmark, ProjectLab: Bubble diagram in portfolio management [apppm.man.dtu.dk/index.php/Bubble\_diagram\_in\_portfolio\_management]

[CEK97] R. G. Cooper, S. J. Edgett and E. J. Kleinschmidt: *Portfolio Management in New Product Development* Research Technology Management **40** 5 (1997) 16-28, [www.stage-gate.la/wp-content/uploads/2018/06/wp12.pdf] [jstor:24131605]

[FK17] P. Franek, M. Krčál: *Persistence of Zero Sets* Homology, Homotopy Appl **19** 2 (2017) [arXiv:1507.04310] [doi:10.4310/HHA.2017.v19.n2.a16]

[FKW18] P. Franek, M. Krčál, H. Wagner: Solving equations and optimization problems with uncertainty J Appl. and Comput. Topol. 1 (2018) 29 [arXiv:1607.06344] [doi:10.1007/s41468-017-0009-6]

[SS-Orb] H. Sati, U. Schreiber: *Proper Orbifold Cohomology* [arXiv:2008.01101]

[FSS-Char] D. Fiorenza, H. Sati, U. Schreiber: The character map in (twisted differential) non-abelian cohomology [arXiv:2009.11909]

[FSS19] D. Fiorenza, H. Sati, U. Schreiber: *Twisted Cohomotopy implies M-theory anomaly cancellation* Comm Math Phys **377** (2020) 1961-2025 [doi:10.1007/s00220-020-03707-2] [arXiv:1904.10207]

[FSS21a] D. Fiorenza, H. Sati, U. Schreiber: *Twisted Cohomotopy implies M5 level quantization* Comm Math Phys **384** (2021) 403–432 [doi:10.1007/s00220-021-03951-0] [arXiv:1906.07417]

[FSS21b] D. Fiorenza, H. Sati, U. Schreiber: *Twisted cohomotopy implies twisted String structure* J Math Phys **62** (2021) 042301 [doi:10.1063/5.0037786] [arXiv:2002.11093]

[FSS22] D. Fiorenza, H. Sati, U. Schreiber: *Twistorial Cohomotopy implies GS anomaly cancellation* Rev Math Phys (2022) [doi:10.1142/S0129055X22500131] [arXiv:2008.08544]

[SS20a] H. Sati, U. Schreiber:

*Equivariant Cohomotopy implies orientifold tadpole cancellation* J Geom Phys **156** (2020) 103775 *Equivariant Cohomotopy implies tadpole cancellation* 

[doi:10.1016/j.geomphys.2020.103775] [arXiv:1909.12277]

[SS20b] H. Sati, U. Schreiber:

The Character Map in Equivariant Twistorial Cohomotopy [arXiv:2011.06533]

[BSS21] S. Burton, H. Sati, U. Schreiber, Lift of fractional D-brane charge to equivariant Cohomotopy theory, J Geom Phys 161 (2021) 104034 [doi:10.1016/j.geomphys.2020.104034] [arXiv:1812.09679]

[SS22] H. Sati, U. Schreiber:

Differential Cohomotopy implies brane observables Adv Theor Math Phys **26** 4 (2022) [ISSN:1095-0753] [arXiv:1912.10425]

