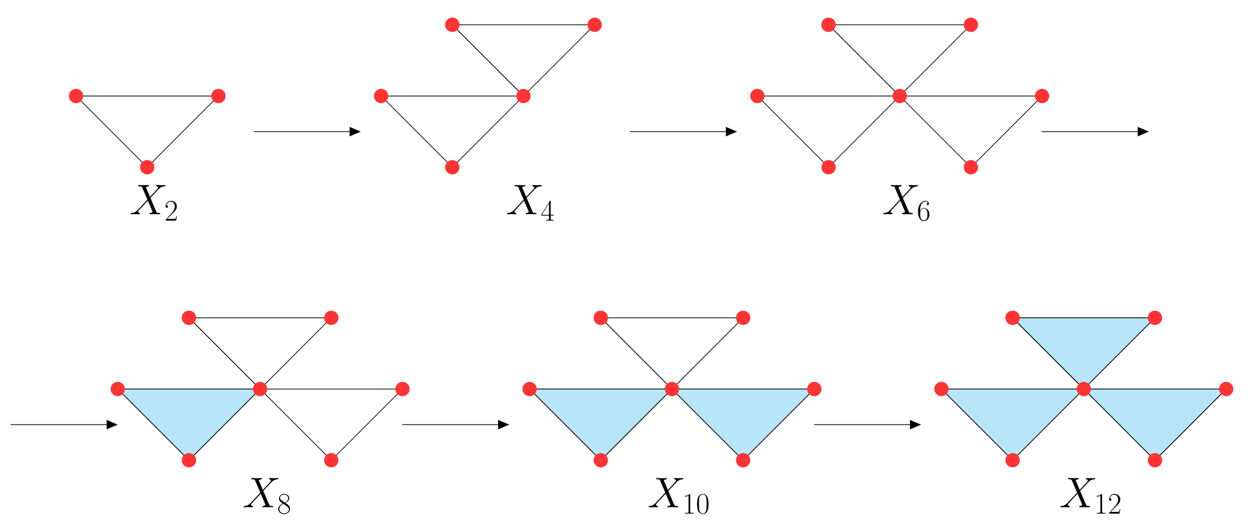
Graded Persistence Diagrams and Persistence Landscapes

Parker Edwards (people.clas.ufl.edu/pedwards), Leo Betthauser, Peter Bubenik

Department of Mathematics, University of Florida

Background

Consider applying the 1-dimensional simplicial homology functor $H_1(\bullet;\mathbb{F})$ to the following sequence $X_2 \subseteq X_4 \subseteq \cdots \subseteq X_{12}$ of simplicial complexes. $\mathbb{F}$ is a field.



The theory of persistent homology tracks the algebraic features which appear and disappear as homology changes through the sequence. The tracked information can be summarized via the rank function.

The rank function

If $[12]^2_<$ is the set of pairs (i,j) where $1 \leq i < j \leq 12$, the rank function $\mathrm{rank} : [12]^2_< \to \mathbb{Z}$ is defined by $\mathrm{rank}(i,j) = \mathrm{rank}(H_1(X_i;\mathbb{F}) \to H_1(X_j;\mathbb{F}))$.

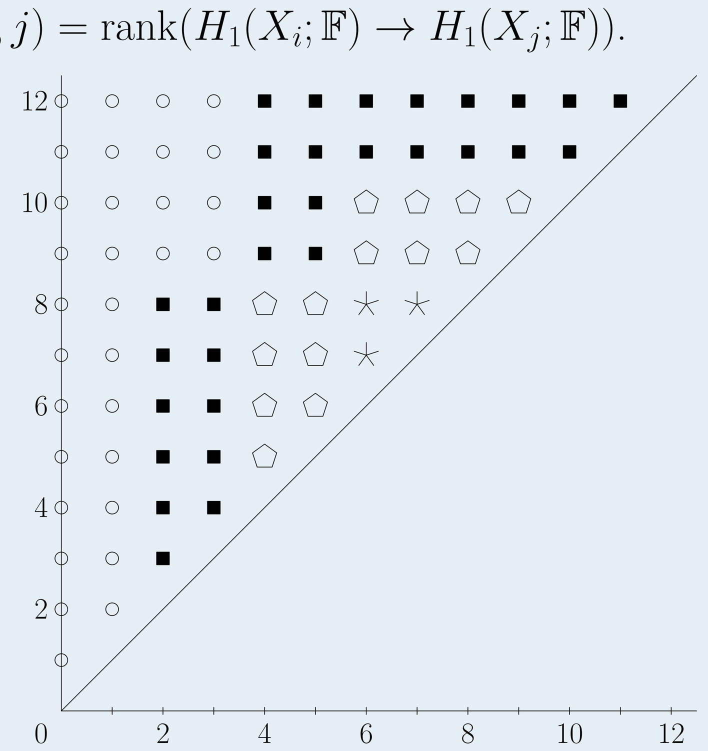


Figure: Visualization of the rank function for a persistence module. Circles indicate points that evaluate to 0, squares indicate points that evaluate to 1, pentagons indicate points that evaluate to 2, and stars indicate points that evaluate to 3.

Persistent homology summaries can serve as useful input to statistical machinery. In particular, the following persistence landscape encoding based on the rank function has desirable statistical properties for this purpose [1].

Persistence landscapes

The persistence landscape is a sequence of functions $\lambda_k : \mathbb{R} \to \mathbb{R}$ for $k \geq 1$ derived from the rank function.

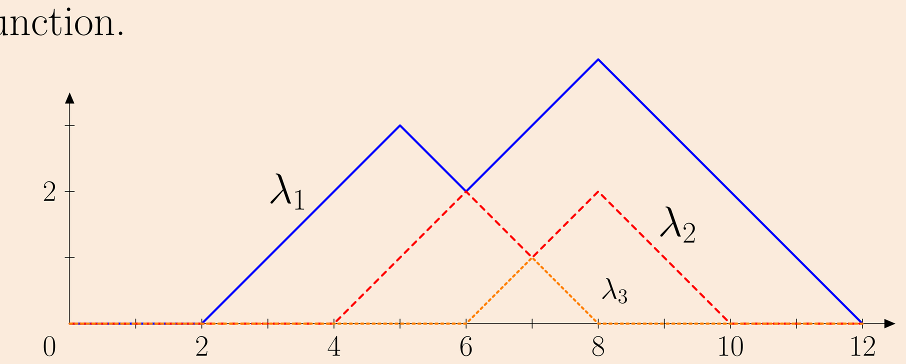


Figure: Graph of the persistence landscape for the rank function above. Each λ_k corresponds to the boundary of the region in $[12]^2_<$ with at least rank k.

Graded rank and graded persistence diagrams

Idea: We can grade the rank function by splitting it into simple component functions that track, for $k \geq 1$, points in $[12]^2_<$ where the rank is greater than or less than k.

The graded rank function

For $k \geq 1$, the k'th graded rank function $\mathrm{rank}_k : [12]^2_< \to \mathbb{Z}$ is given by

$$\mathrm{rank}_k(i,j) = \begin{cases} 1 & \mathrm{rank}(i,j) \geq k \\ 0 & \text{else} \end{cases}$$

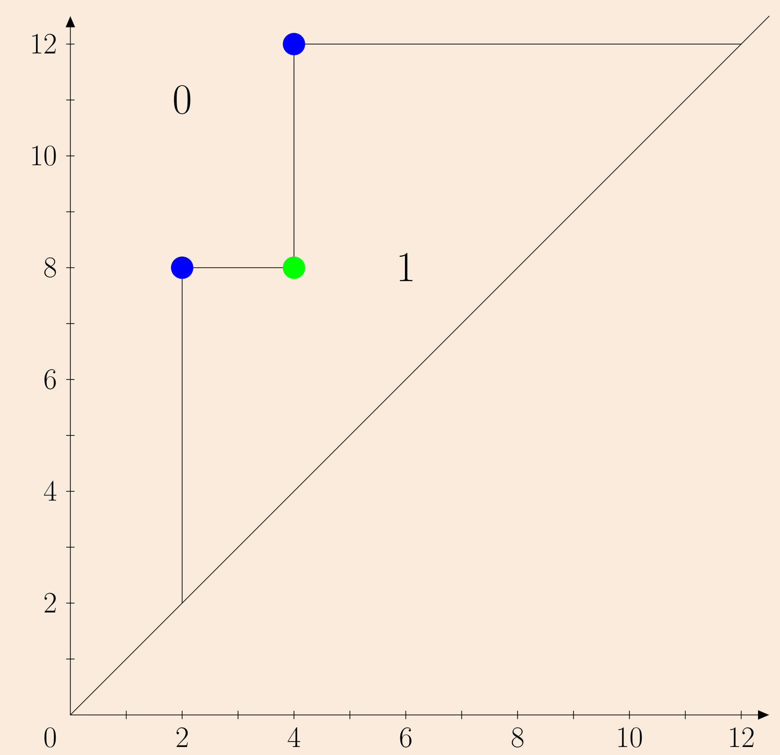


Figure: The 1st graded rank function of our example.

The graded rank function $(\mathrm{rank}_k)_{k\geq 1} : [12]^2_< \to \bigoplus_{k\geq 1}\mathbb{Z}$ is given by $(i,j) \mapsto (\mathrm{rank}_1(i,j), \mathrm{rank}_2(i,j), \dots)$.

Graded persistence diagrams

For $k \geq 1$, the k'th graded persistence diagram $\mathrm{PD}_k : [12]^2_< \to \mathbb{Z}$ is given by

$$\begin{aligned}\mathrm{PD}_k(i,j) = {} & \mathrm{rank}_k(i,j) - \mathrm{rank}_k(i-1,j) \\ & - \mathrm{rank}_k(i,j+1) + \mathrm{rank}_k(i-1,j+1)\end{aligned}$$

The graded persistence diagram $(\mathrm{PD}_k)_{k\geq 1} : [12]^2_< \to \bigoplus_{k\geq 1}\mathbb{Z}$ is given by the same Möbius inversion formula above with $(\mathrm{rank}_k)_{k\geq 1}$ replacing rank_k.

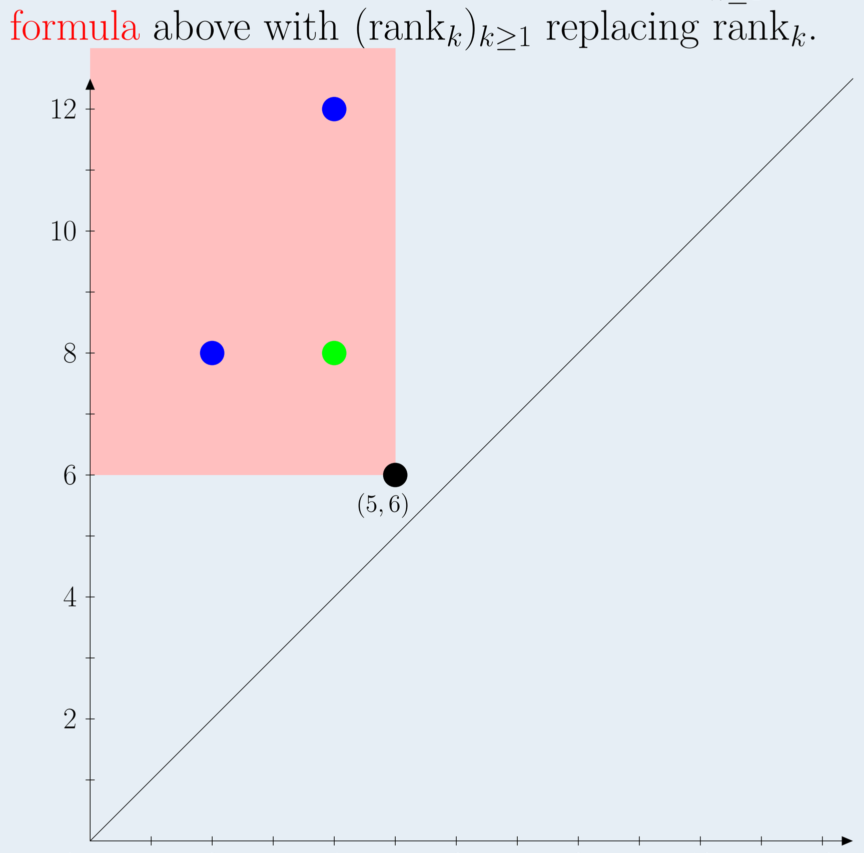


Figure: The 1st graded rank persistence diagram for our example. Blue points have Λ_1 value equal to 1, green points have value equal to -1, and all others have value 0. The rank_1 value of $(5,6)$ is the sum of PD_1 values in the shaded region.

Results

Theorem: Consistency of grading

Let the persistence diagram function $\mathrm{PD} : [12]^2_< \to \mathbb{Z}$ be defined analogously to PD_k replacing rank_k with rank, let "summation" denote summing all values of a function above and left of a point in $[12]^2_<$, and let $\Sigma : \bigoplus_{k\geq 1}\mathbb{Z} \to \mathbb{Z}$ be the function $(a_1, a_2, \dots) \mapsto a_1 + a_2 + \dots$.

The following diagram of function transformations commutes.

$$\begin{tikzcd}
\mathrm{rank} \arrow[r, shift left, "\text{Möbius inversion}"] \arrow[d, shift right, "\text{Grading}"'] & \mathrm{PD} \arrow[l, shift left, "\text{Summation}"] \\
(\mathrm{rank}_k)_{k\geq 1} \arrow[u, shift right, "\Sigma"'] \arrow[r, shift left, "\text{Möbius inversion}"] & (\mathrm{PD}_k)_{k\geq 1} \arrow[l, shift left, "\text{Summation}"] \arrow[u, "\Sigma"']
\end{tikzcd}$$

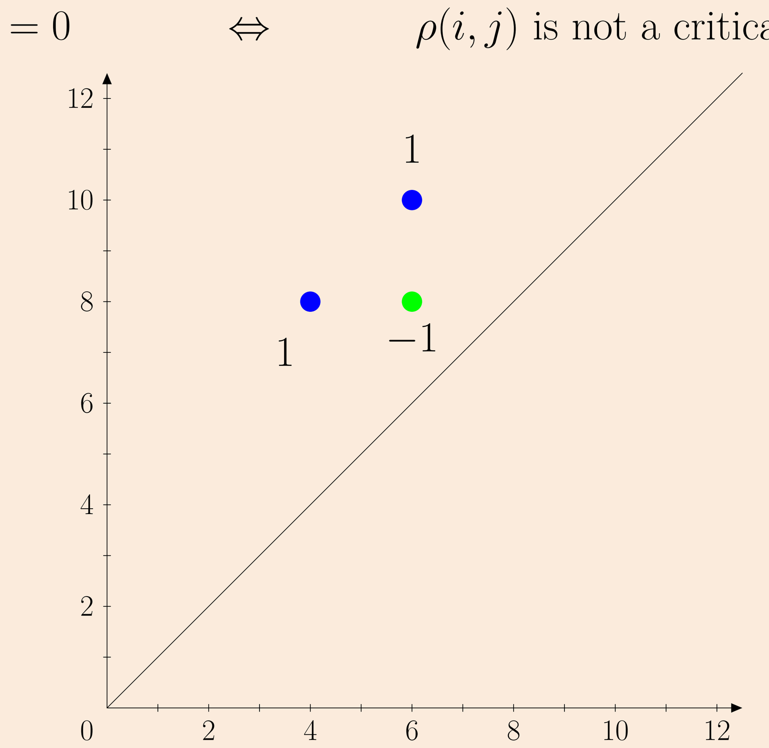
Remark: Persistence landscapes are piecewise linear functions, so are determined by their critical points.

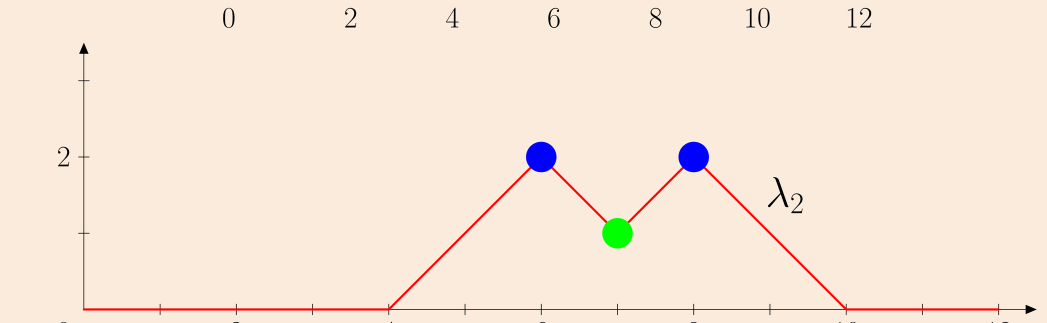
Theorem: Graded diagrams and landscapes are equivalent

The k'th graded persistence diagram PD_k is equivalent to the k'th persistence landscape λ_k.

More precisely: There is a bijection ρ between points $(i,j) \in [12]^2_<$ and points in $\mathbb{R}^2$ such that

$$\begin{aligned}
\mathrm{PD}_k(i,j) = 1 \quad & \Leftrightarrow \quad \rho(i,j) \text{ is a local maximum of } \lambda_k \\
\mathrm{PD}_k(i,j) = -1 \quad & \Leftrightarrow \quad \rho(i,j) \text{ is a local minimum of } \lambda_k \\
\mathrm{PD}_k(i,j) = 0 \quad & \Leftrightarrow \quad \rho(i,j) \text{ is not a critical point of } \lambda_k
\end{aligned}$$





References

[1] Betthauser, Leo, Peter Bubenik, and Parker B. Edwards. "Graded persistence diagrams and persistence landscapes." arXiv preprint https://arxiv.org/abs/1904.12807

[2] Bubenik, Peter. "Statistical Topological Data Analysis using Persistence Landscapes" J. of Machine Learning Research 16.1 (2015): 77-102