We thank the reviewers for their careful and thoughtful reviews.

Concerns about the time series experiments (R1, R2, R3)

The reviewers pointed out that we did not use standard baselines; we agree the paper would be strengthened by including them. Hence, we compared with Euclidean nearest neighbor (B1), dynamic time warping with optimized (B2) and constant window width (B3), persistence scale-space kernel on the Rips filtration (without density), (R), standard 1D persistence image (I) and landscapes (L) of a weighted average of Rips and density, and the multiparameter kernel (MP-K), landscape (MP-L), and image (MP-I). We show results for only a few datasets due to lack of space.

Dataset	B1	B2	B3	R	I	L	MP-K	MP-L	MP-I
GunPointOldVersusYoung	95.2	**96.5**	83.8	-	98.7	95.9	99.0	97.1	**100.0**
ProximalPhalanxOutlineCorrect	**80.8**	79.0	78.4	78.4	67.4	72.5	78.7	78.7	**81.8**
ECG200	**88.0**	**88.0**	77.0	67.0	68.0	68.0	77.0	74.0	**83.0**
Plane	96.2	**100.0**	**100.0**	82.9	64.8	82.9	92.4	84.8	**97.1**
SyntheticControl	88.0	98.3	**99.3**	50.0	45.7	44.0	50.7	**60.3**	56.3

We interpret this table to indicate that multiparameter persistence is a competitive option for this problem, and our method is the best of the persistence-based approaches. However, our view is that the real virtue of this method comes in situations such as the immunofluorescence images experiment.

Standard deviations (std) on the train set for these results when cross-validating were on the order of 0.05. There is no std for the test set since we used suggested train / test splits. **Computation times** for MP-I on these tasks averaged around 350 seconds; this was 3-4 times faster than MP-L and around 30-50 times faster than MP-K. The other summaries based on 1D persistence had running times similar to MP-I. **Size of feature space** It is correct that we worked with 10x10 and 50x50 images. We cross-validated all of the methods with up to 2,500 features, to ensure fair comparison; hence we do not think the size of the feature space accounts for performance improvements.

Stability with respect to matching (R1, R2, R3)

We agree with the reviewers that there is a disconnect between the theory and stability in practice. We will expand the discussion of this in the revision.

Perturbing the lines We experimented with random perturbations of the lines selected (both by randomly shifting the angles and making them closer and further apart). Accuracies degraded slowly in the size of the perturbations. For instance, when perturbing the line endpoints with noise whose amplitude was up to 10 times the distance between the lines, the mean accuracy of ECG200 and GunPoint went down by only 2-3%, with std around 0.02. **Matching instability** We never detected this in practice in any of the examples we ran, although it is easy to create synthetic examples exhibiting it. We will detail how to detect it (by comparing with bottleneck matchings) in the revision.

Extension to multiparameter persistence (R2)

We agree with R2 that extending the MP-I to multiparameter persistence is potentially subtle, and requires discussion of the choice of lines. For example, in R^3, one would sweep the 3D space with planes that all intersect on a given line. Since each plane would itself be swept with lines, 2D summands can be computed in each plane, and then connected through the common line to generate 3D summands. The stability result generalizes to this situation with a different geometric base case.

Clarity (R1, R2, R3)

All reviewers made very useful comments about how to improve clarity of the paper, which we will incorporate. We also thank R1 for the writing comments and reference works. We answer their questions here. **DTM**: we didn't study stability w.r.t. the DTM parameters but this is a very interesting question to look at. **Bounding rectangle** refers to the rectangle given by the minima and maxima of both filtrations. **Fig 2**: n should be replaced by N. **Lines**: The reason why N is not fixed for lines with same slope is because we implemented it that way. There is no theoretical obstruction for this though. $f|_\ell$ is indeed an abuse of notation. **Def 3.3**: n should be p and a, b, c, d should satisfy $a \leq b$, $c \leq d$. **Fig 4** was generated with lines of same slope. Sets of large bars indicate topological features that are persistent in both filtrations: in our case, circles with large diameters that are made of points with large density. Smaller loops would generate bars with smaller lengths for at least one filtration. Difference between a big loop and many small loops shows in the decomposition, but can be missed in MP-I since all the Gaussian functions (corresponding to the sets of bars) are summed.