Appendix

A.1 TDA terminology and definitions.

Before introducing barcodes, we need to first know about homology groups, at least in terms of their operational definitions. For more mathematically-oriented readers, we recommend alternative references [11, 14-17, 62]. For this paper, it is enough to know that after obtaining a simplicial complex from a filtration process, we can get the homology groups with a few more operations. To understand these additional steps, some terms need to be defined. The first term is called the n-*chain*, the largest subset of n-dimensional simplices. For example, suppose we are given a simplicial complex $\Sigma = \{\{i\},\{j\},\{k\},\{l\},\{i,j\},\{j,k\},\{k,l\},\{k,i\},\{l,j\},\{i,j,k\}\}$. We refer to sets containing just one entity as nodes, those comprising two entities as links, and those with three entities as faces, etc. In this example, the 2-chain is $\{i,j,k\}$, whereas the 1-chain is $\{\{i,j\},\{j,k\},\{k,l\},\{k,i\},\{l,j\}\}$. Thus, a n-chain can be understood as a way to sort out sets consisting of n-dimensional simplices only. Next, we can convert the n-chain to its n-*chain group* $C_n(\Sigma)$, which is $C_n(\Sigma) = \sum \alpha_i \sigma_i$, where α_i is the ith coefficient that can be defined in terms of groups, e.g. a cyclic group $\mathbb{Z}_2$ will has $\alpha_i = 0$ or 1 only. σ_i is the ith simplex in $C_n(\Sigma)$.

Next, we introduce a few more terms, that are *chain complexes*, *boundary* (or boundary operator), *kernel*, and *image*. The boundary operator ∂ can operate on an n-simplex X, or $\partial(X) = \sum_{k=1}^{n}(-1)^k [N_1, N_2, \dots, \widehat{N_k}, \dots, N_n]$, where $\widehat{N_k}$ is a node removed from X. Hereafter, we shall acronym the n-chain group $C_n(\Sigma)$ as C_n, and acronym $\partial(C_n(\Sigma))$ as $\partial_n(\Sigma)$. The kernel of $\partial_n(\Sigma)$ is defined as finding an n-chain $Z_n \subseteq C_n$, which satisfy $\partial(Z_n) = 0$. The image of a boundary $\partial_n(\Sigma)$ can then be written as $\operatorname{Im}\partial_n(\Sigma)$, which gives the sets of boundaries. Finally, after introducing these four terms, we are now able to give a formal definition to the nth homology group of Σ:

$$H_n(\Sigma) \coloneqq \operatorname{Ker}\partial_n(\Sigma) \ / \operatorname{Im}\ \partial_{n+1}(\Sigma),$$

where ∂_n, and ∂_{n+1} operating on Σ can be translated into the boundary operator operating on the n-chain and the $n+1$-chain group of Σ. $H_n(\Sigma)$ thus satisfy all the properties that can be defined as a group in group theory, i.e. associativity, the identity element, and the inverse element, which is very convenient in understanding topology using group theory's language as well as performing group operations. Also, Betti numbers can be defined as the dimension of $H_n(\Sigma)$ or

$$\beta_n(\Sigma) = \dim\big(H_n(\Sigma)\big).$$

In this sense, if the homology group is a reducible representation of the data, then Betti numbers are the dimensions of the irreducible representation of the same homology group.

A2. *p*th Wasserstein distance

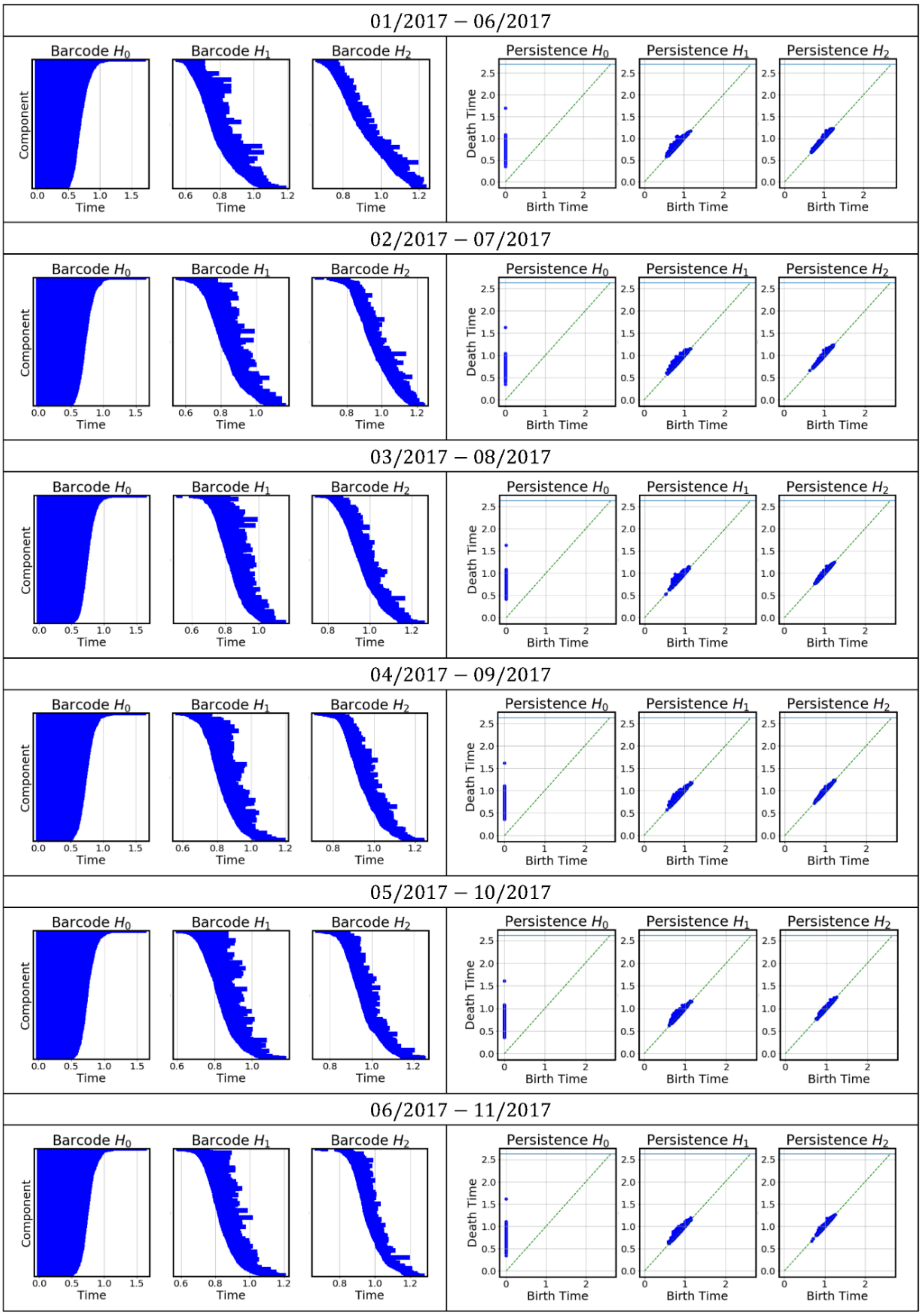
Suppose we have two persistence landscapes constructed using different data set in different periods, and we would like to scrutinize their differences. One of the effective approaches is resorting to the Wasserstein distances. Suppose now we have two simplicial complexes Σ_1, and Σ_2, where their Wasserstein distances can be defined as:

$$W_p(P_k^1, P_k^2) \coloneqq \inf_{\phi} \left[\sum_{x \in P_k^1} \|x - \phi(x)\|_\infty^{p} \right]^{\frac{1}{p}},$$

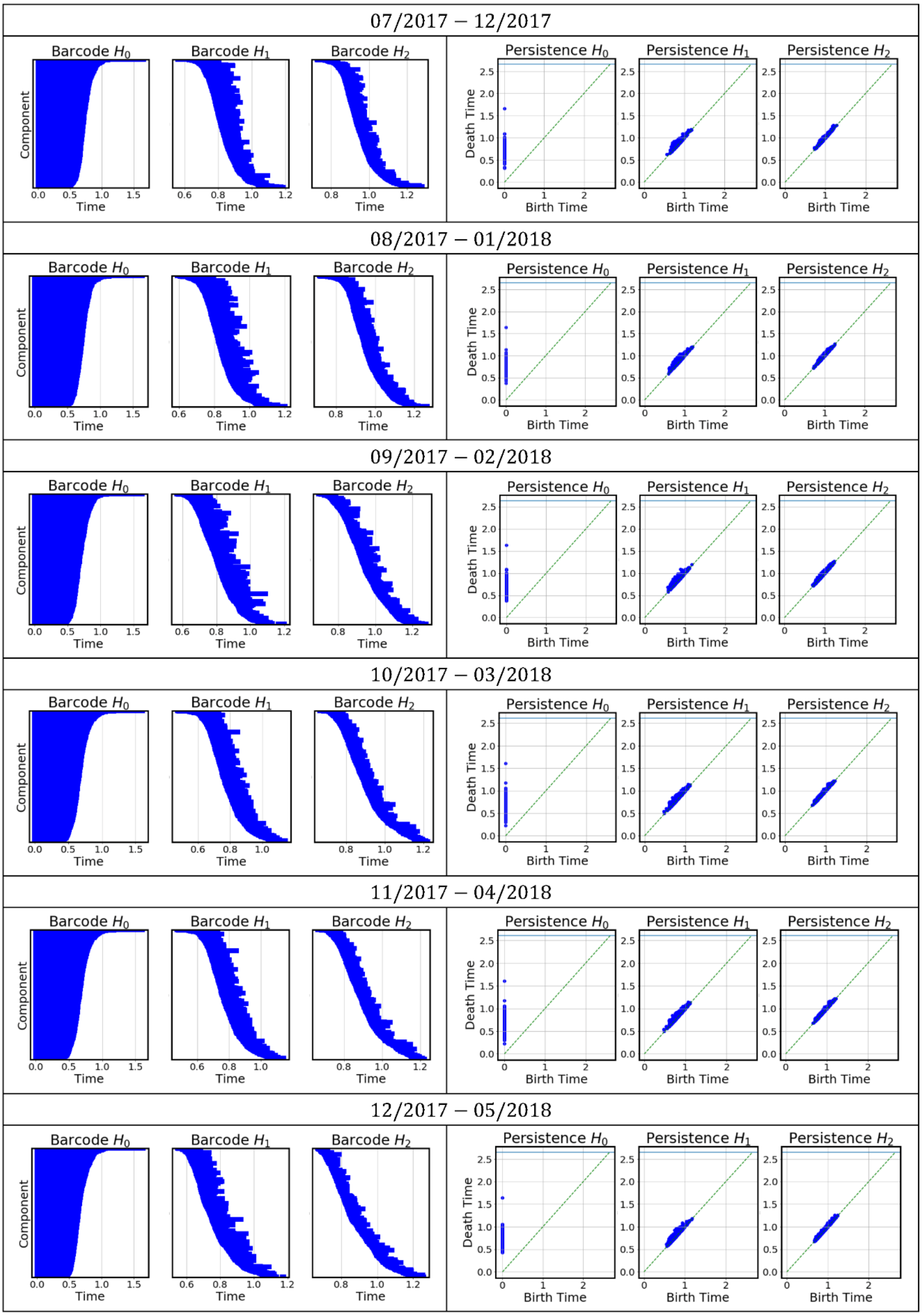
where W_p is the *p*th Wasserstein distance, P_k^1, P_k^2 are two of the persistence diagrams corresponding to Σ_1, and Σ_2 in the *k*th dimension; $\phi(x)$ is a bijective function that generates the points in P_k^2, and $\|\cdot\|_\infty$ denotes the sup norm. For $p = 1$, and 2, we have the *Manhattan distance*, and *Euclidean distance*, respectively by using the ℓ^{p}-norm definition, whereas if $p = \infty$, we use the ℓ^∞-norm, which is the supremum over the points x instead to arrive at a *bottleneck distance*. By choosing appropriate *p* orders, we can differentiate the metrical differences of the two persistence diagrams.

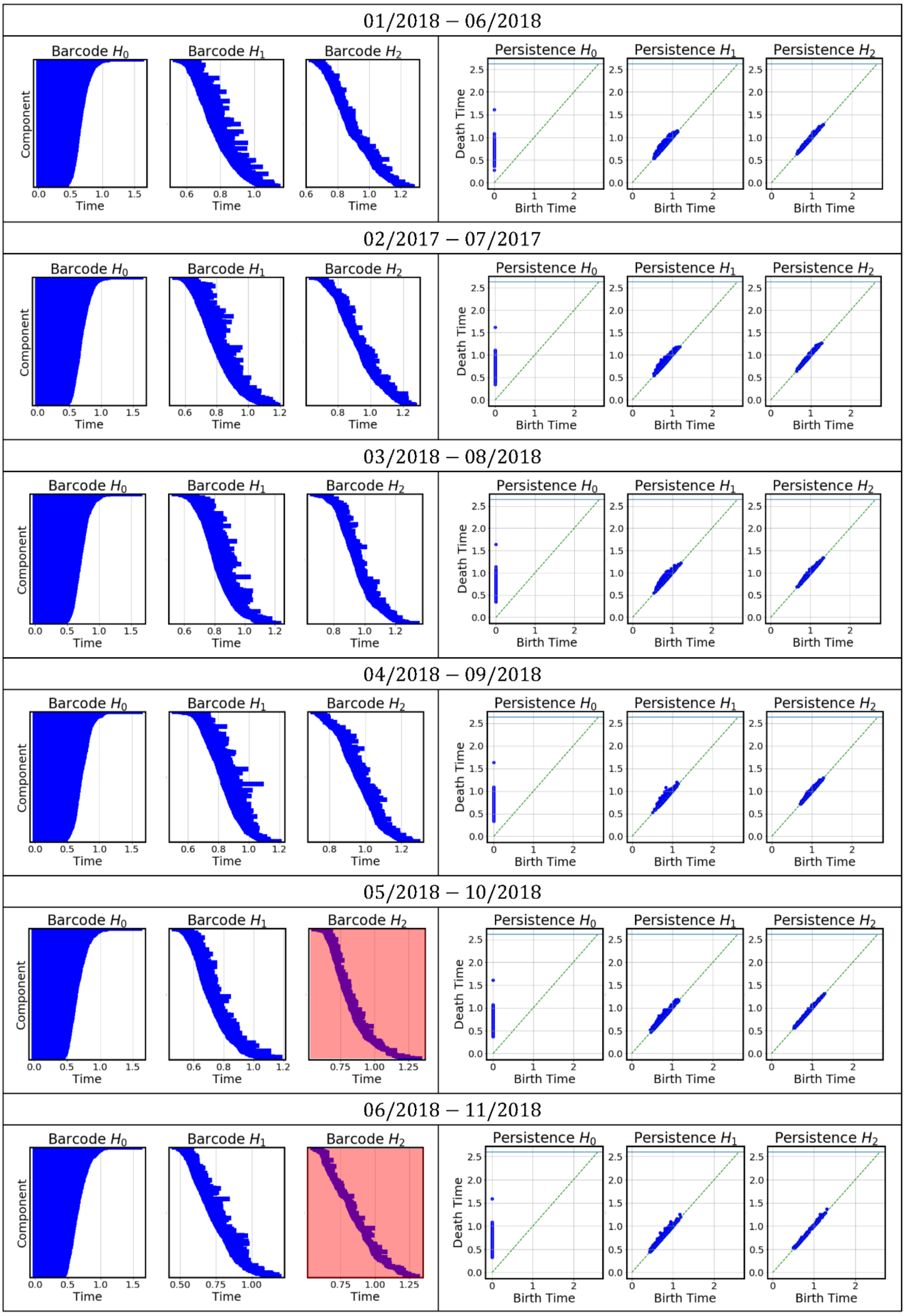
A3.

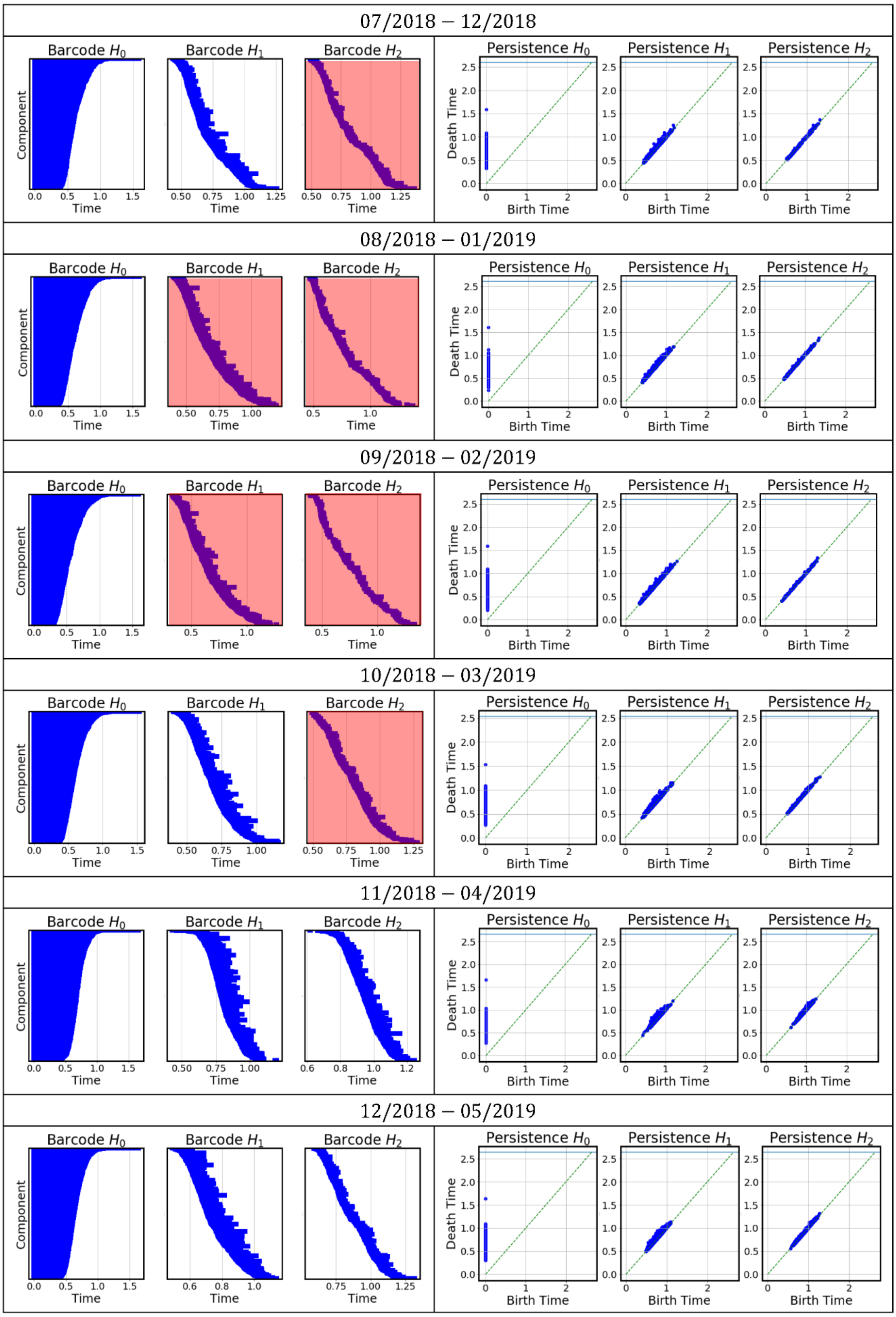
Figure A1, and Figure A2 and are too large, therefore we put them here for the reader’s convenience.

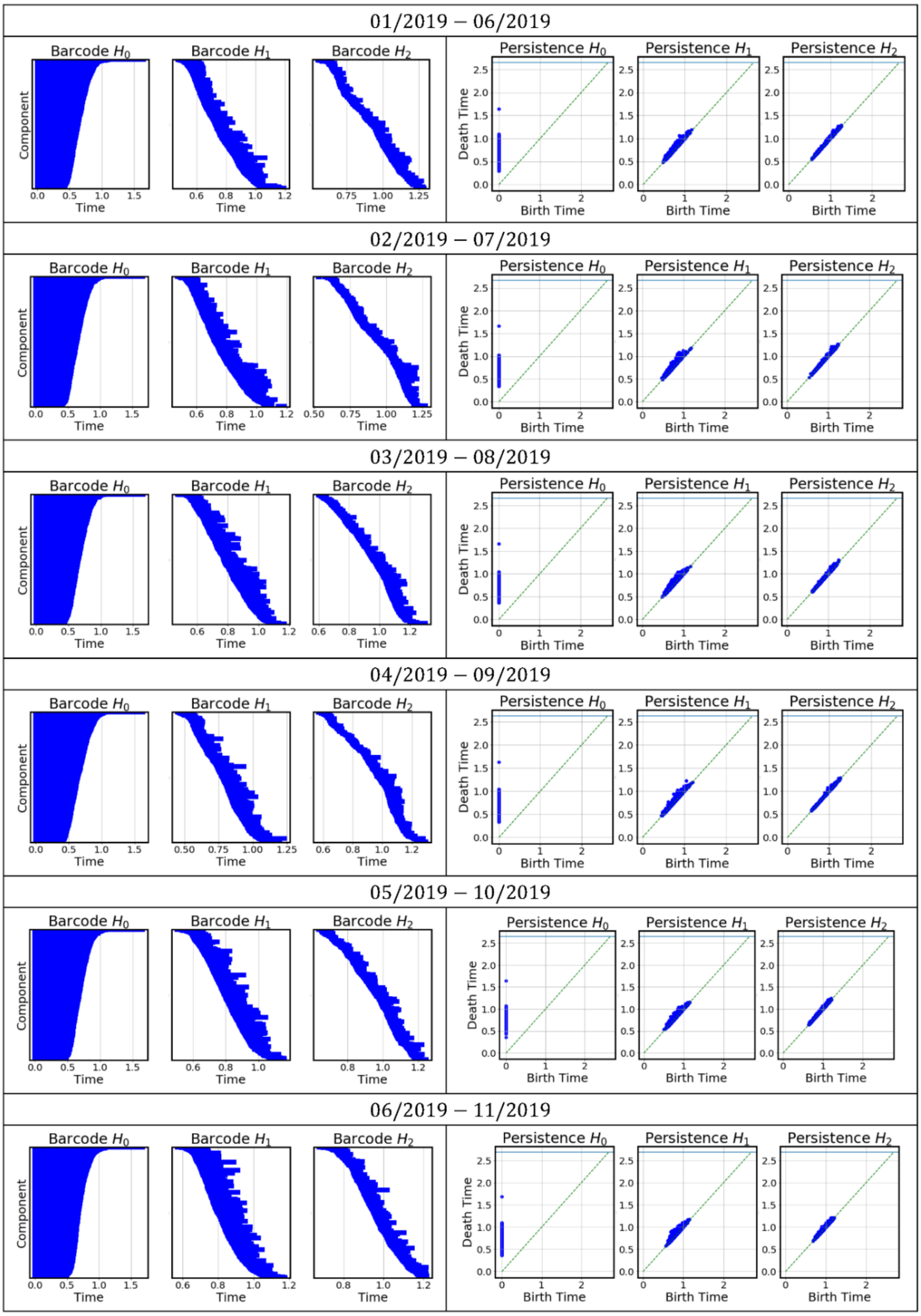


01/2017 – 06/2017
02/2017 – 07/2017
03/2017 – 08/2017
04/2017 – 09/2017
05/2017 – 10/2017
06/2017 – 11/2017
Barcode H_0
Barcode H_1
Barcode H_2
Persistence H_0
Persistence H_1
Persistence H_2
Component
Time
Death Time
Birth Time


07/2017 – 12/2017
08/2017 – 01/2018
09/2017 – 02/2018
10/2017 – 03/2018
11/2017 – 04/2018
12/2017 – 05/2018
Barcode H_0
Barcode H_1
Barcode H_2
Persistence H_0
Persistence H_1
Persistence H_2
Component
Time
Death Time
Birth Time


01/2018 – 06/2018
02/2017 – 07/2017
03/2018 – 08/2018
04/2018 – 09/2018
05/2018 – 10/2018
06/2018 – 11/2018
Barcode H_0
Barcode H_1
Barcode H_2
Component
Time
Persistence H_0
Persistence H_1
Persistence H_2
Death Time
Birth Time


07/2018 – 12/2018
08/2018 – 01/2019
09/2018 – 02/2019
10/2018 – 03/2019
11/2018 – 04/2019
12/2018 – 05/2019
Barcode H_0
Barcode H_1
Barcode H_2
Persistence H_0
Persistence H_1
Persistence H_2
Component
Time
Death Time
Birth Time



01/2019 – 06/2019
02/2019 – 07/2019
03/2019 – 08/2019
04/2019 – 09/2019
05/2019 – 10/2019
06/2019 – 11/2019
Barcode H_0
Barcode H_1
Barcode H_2
Persistence H_0
Persistence H_1
Persistence H_2
Component
Time
Death Time
Birth Time

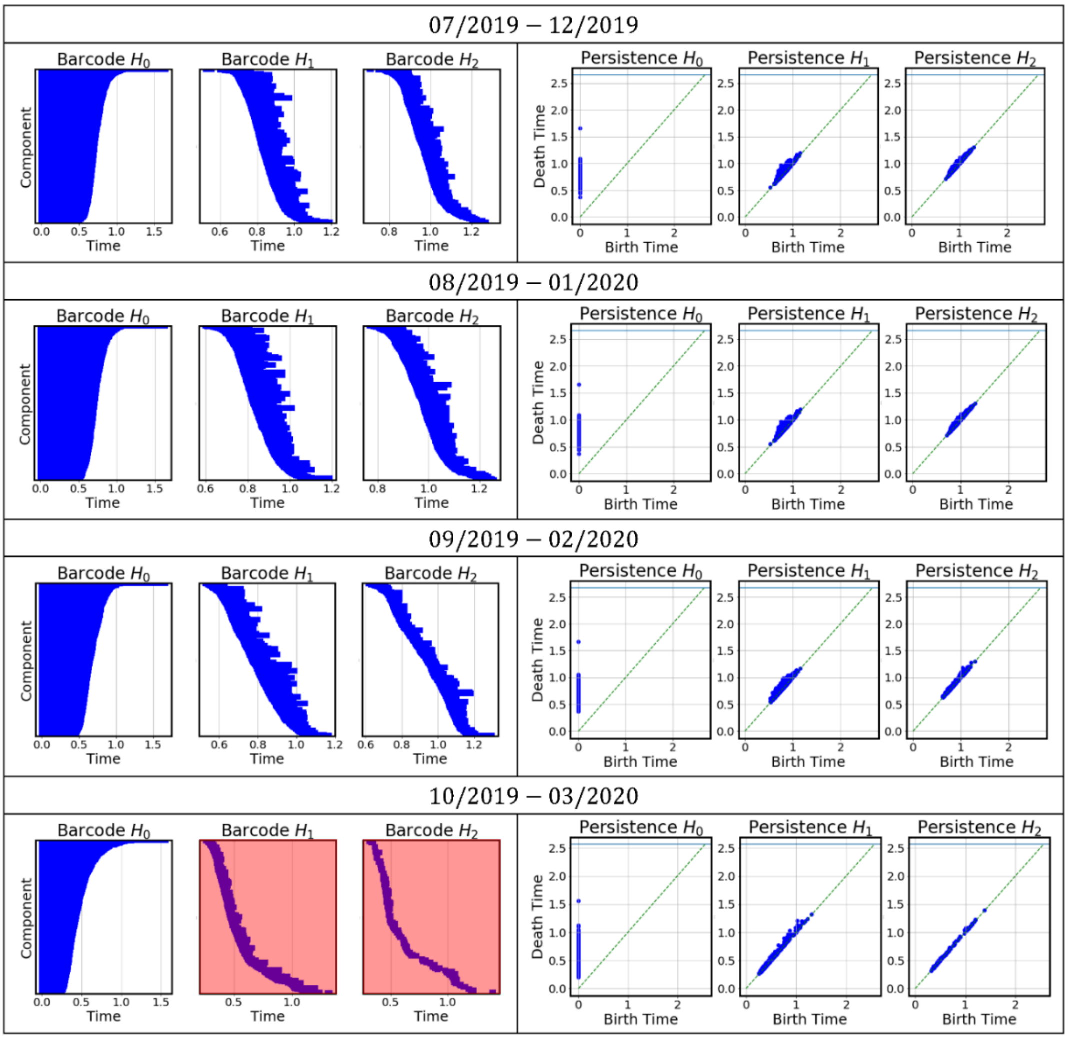
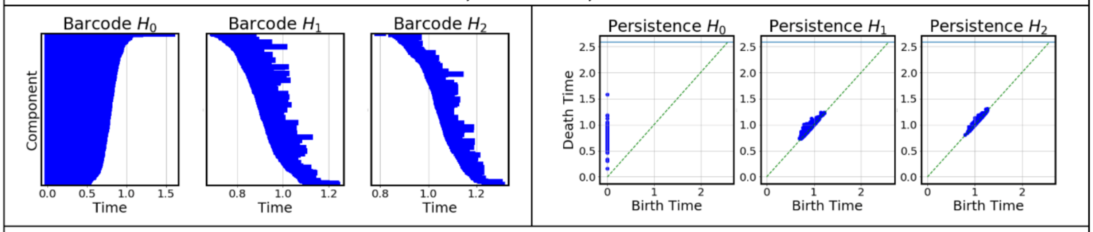
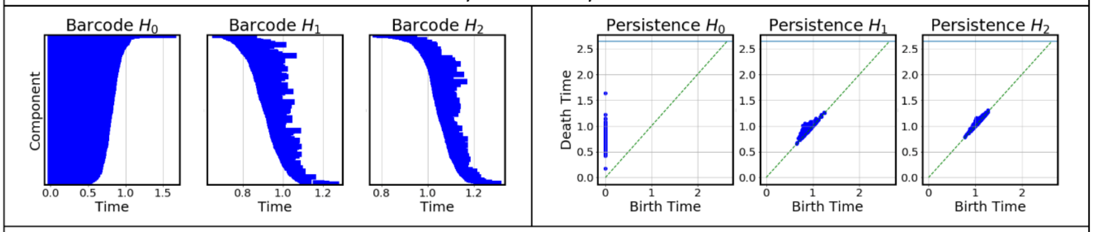


Figure A1. The barcodes and corresponding persistence diagrams for TAIEX, from Jan 2017 to Mar 2020. We use red-shade windows to indicate the barcodes that become less persistent during market crashes.

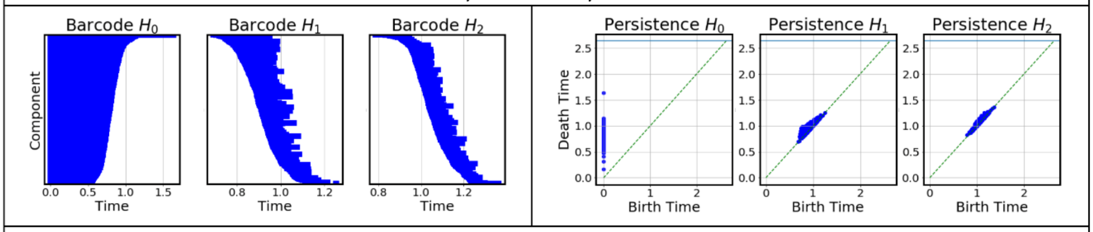
01/2017 – 06/2017



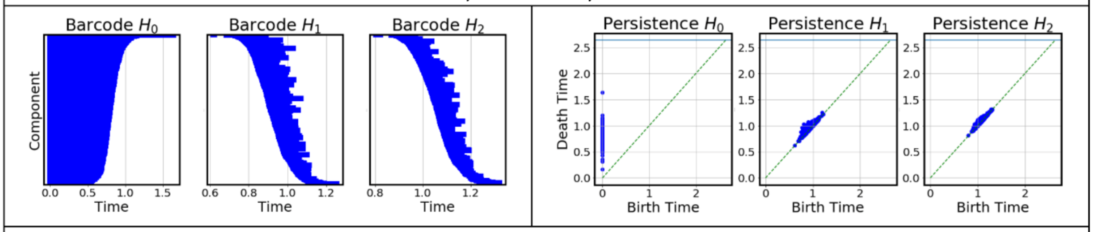
02/2017 – 07/2017



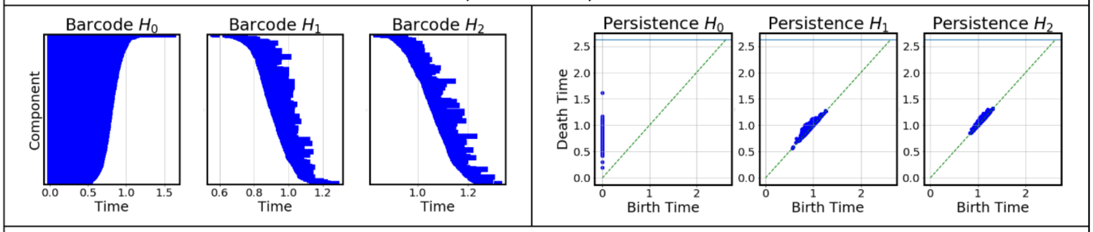
03/2017 – 08/2017



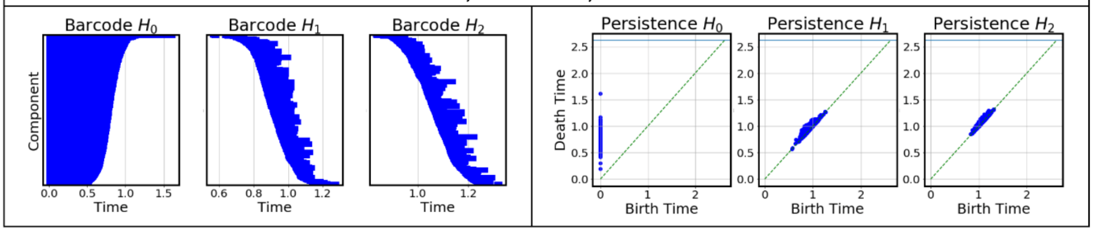
04/2017 – 09/2017

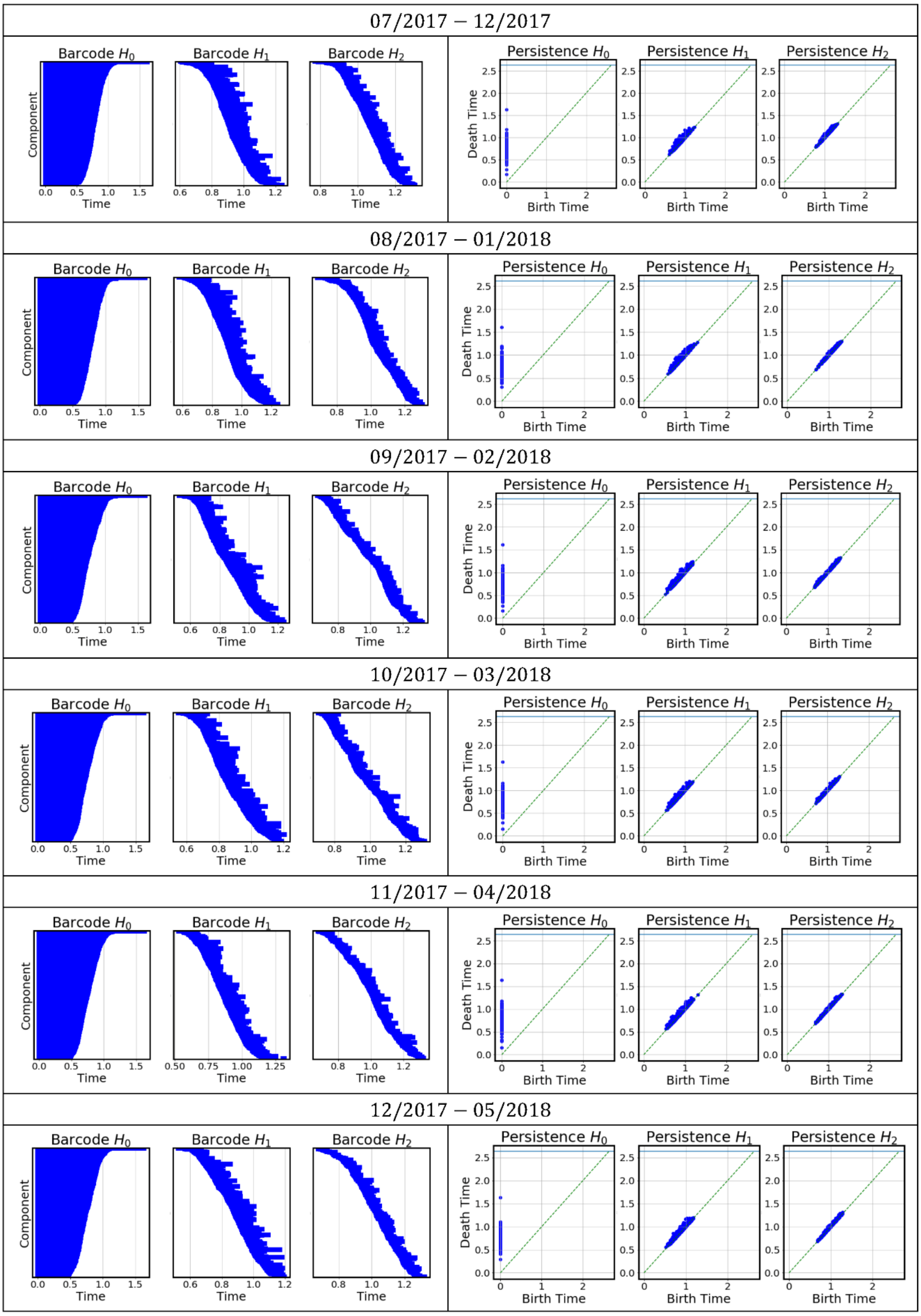


05/2017 – 10/2017

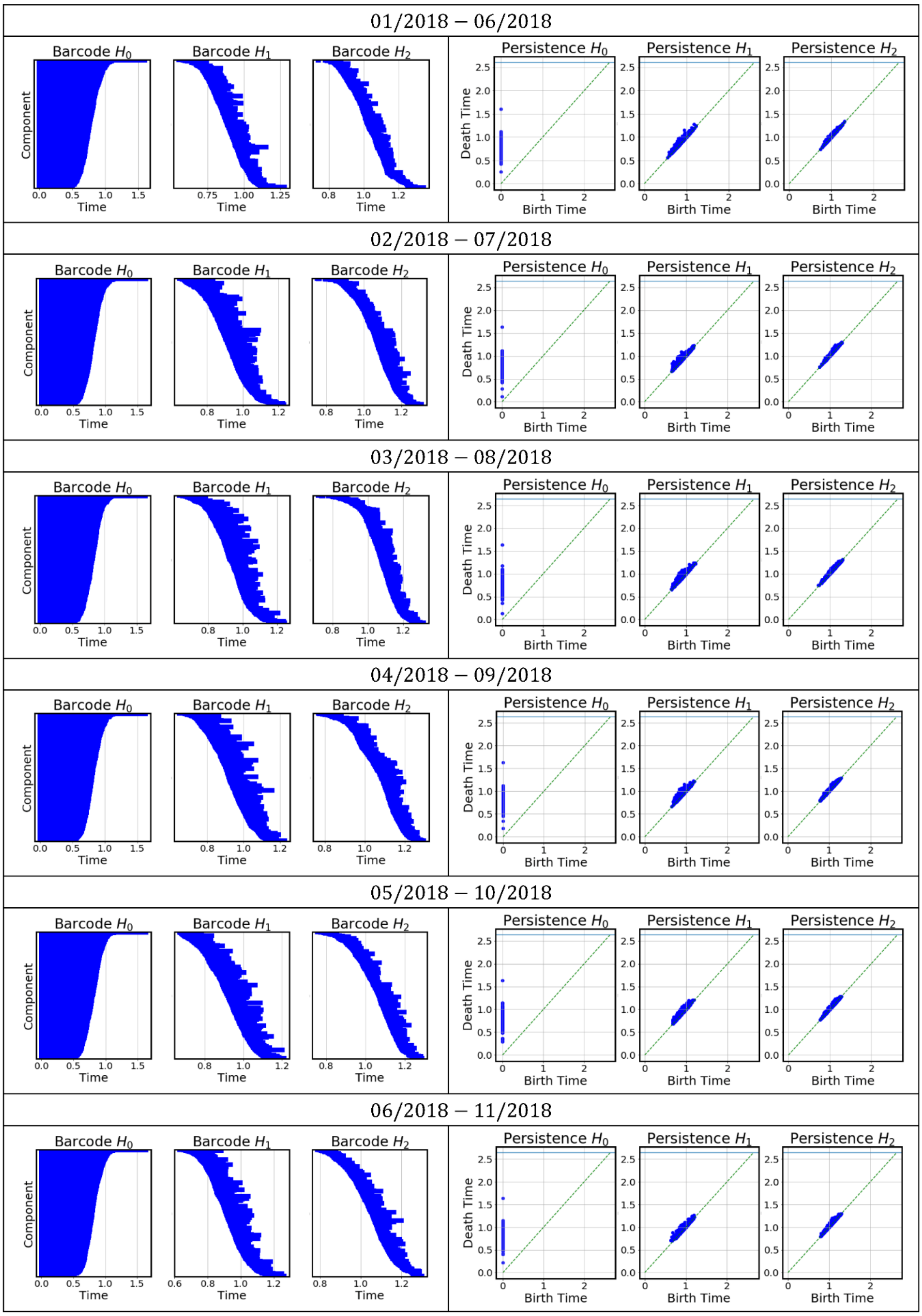


06/2017 – 11/2017





07/2017 – 12/2017
08/2017 – 01/2018
09/2017 – 02/2018
10/2017 – 03/2018
11/2017 – 04/2018
12/2017 – 05/2018
Barcode H_0
Barcode H_1
Barcode H_2
Component
Time
Persistence H_0
Persistence H_1
Persistence H_2
Death Time
Birth Time


01/2018 – 06/2018
02/2018 – 07/2018
03/2018 – 08/2018
04/2018 – 09/2018
05/2018 – 10/2018
06/2018 – 11/2018
Barcode H_0
Barcode H_1
Barcode H_2
Persistence H_0
Persistence H_1
Persistence H_2
Component
Time
Death Time
Birth Time

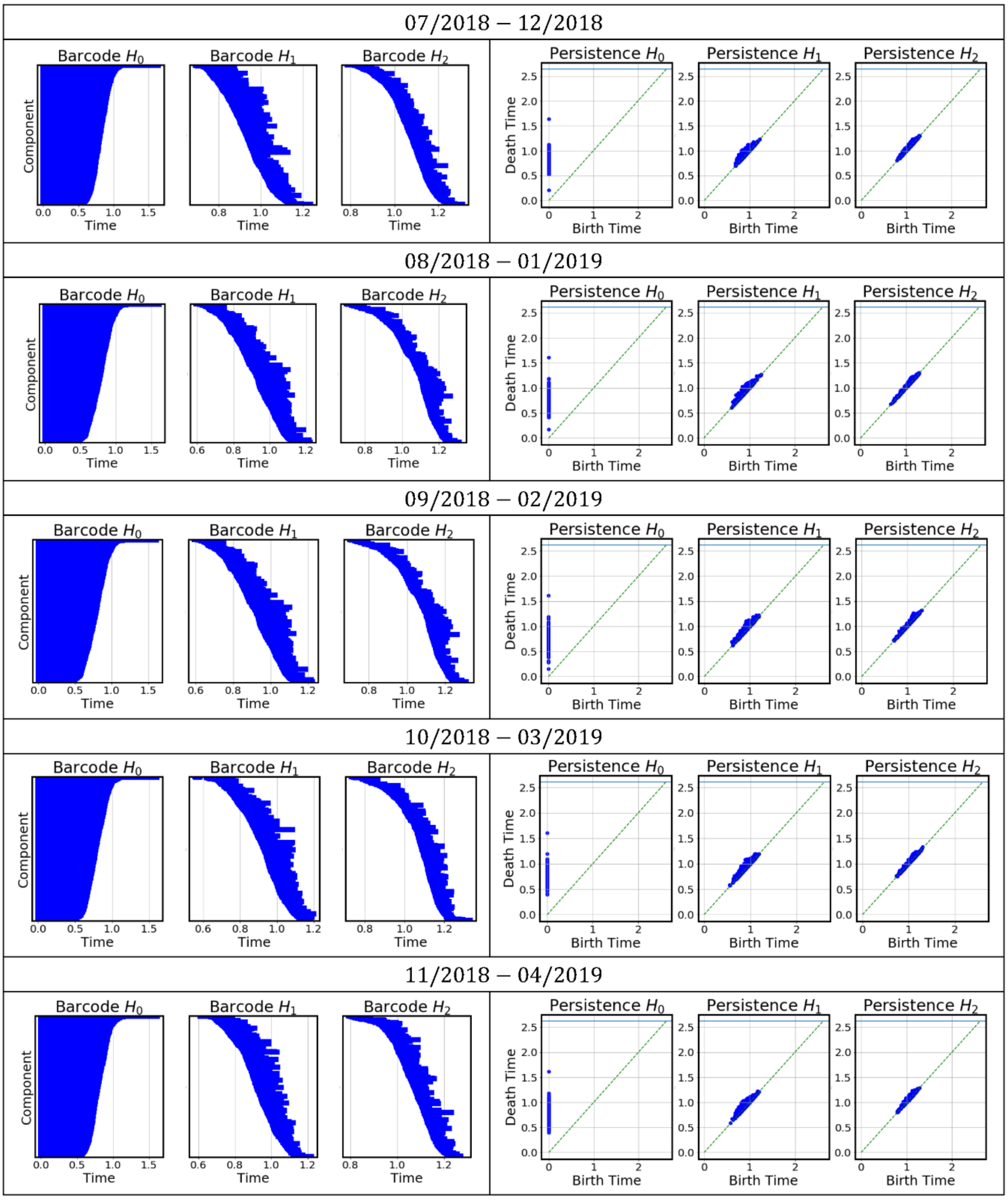


Figure A2 The barcodes and corresponding persistence diagrams for STI, from Jan 2017 to Apr 2019.