

T2P Online Appendix

Online Appendix 1

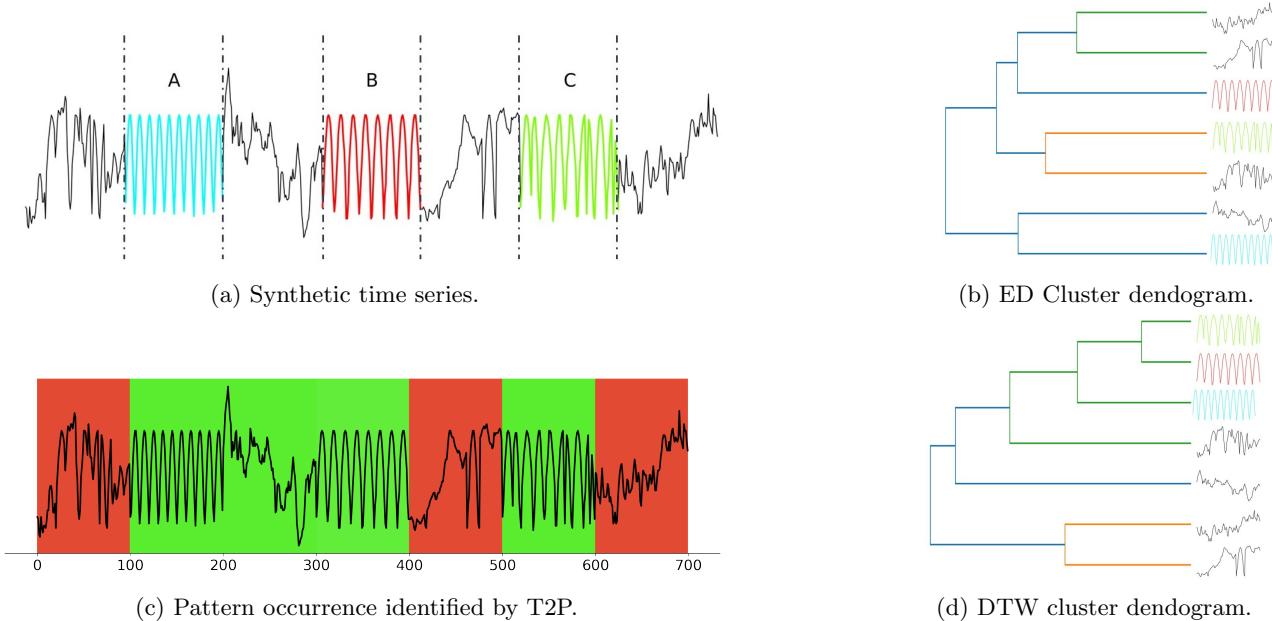


Figure 1: A synthetic time series divided into seven equal-length regions and clustered using complete linkage hierarchical clustering.

Online Appendix 2. BinConcrete PDF

As illustrated in Figure 2, the BinConcrete closely imitates the Bernoulli distribution. BinConcrete models a continuous random variable with a probability density function that is differentiable with respect to its parameters.

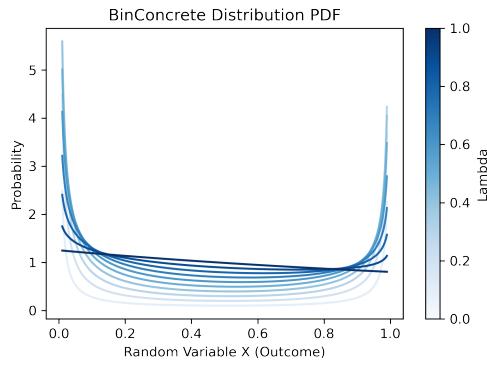


Figure 2: Probability density function of the BinConcrete distribution.

Online Appendix 3. Variational Autoencoder

The variational autoencoder (VAE) is a deep generative model consisting of a true posterior (generator) $p_\theta(x|z)$, a prior $p_\theta(z)$, and an approximate posterior (inference) $q_\phi(z|x)$. Let us consider a dataset $x = \{x^{(i)}\}_{i=1}^N$ consisting of N i.i.d samples of a continuous or discrete variable, x . We would like to compute the true posterior $p_\theta(z|x) = \frac{p_\theta(x,z)}{p_\theta(x)}$, but computing $p_\theta(x) = \int p_\theta(x|z)p(z)dz$ is intractable in many cases. To overcome this obstacle, we can approximate $p_\theta(z|x)$ by selecting a tractable distribution $q(z)$ such as Gaussian. To approximate

$p_\theta(z|x)$, we minimize the KL-divergence between the approximate and true posterior $\text{KL}(q_\phi(z|x) \parallel p_\theta(z|x))$. Thus, the marginal likelihoods can be written as:

$$\log p_\theta(x^{(i)}) = \text{KL}(q_\phi(z|x^{(i)}) \parallel p_\theta(z|x^{(i)})) + \mathcal{L}(\theta, \phi; x^{(i)}) \quad (1)$$

Because the KL divergence is non-negative, the $\mathcal{L}(\theta, \phi; x^{(i)})$ term is called the (variational) lower bound on the marginal likelihood of data point i . Instead of minimizing the KL divergence between the approximate and true posterior, we can maximize the variational lower bound

$$\log p_\theta(x) \geq \mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_\phi(z|x)} [-\log q_\phi(z|x) + \log p_\theta(x, z)] \quad (2)$$

which can also be written as:

$$\mathcal{L}(\theta, \phi; x) = -\text{KL}(q_\phi(z|x) \parallel p_\theta(z)) + \mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)] \quad (3)$$

We need to compute the gradient of the lower bound $\mathcal{L}(\theta, \phi; x)$ w.r.t. the generative parameter θ and the inference parameter ϕ , in order to optimize.

$$\nabla_{\phi, \theta} \mathcal{L}(\theta, \phi; x) = \nabla_{\phi, \theta} \mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)] - \nabla_{\phi, \theta} \text{KL}(q_\phi(z|x) \parallel p_\theta(z)) \quad (4)$$

For the first part of the lower bound, the gradient w.r.t. θ can be easily computed using Monte Carlo sampling

$$\nabla_\theta \mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)] = \mathbb{E}_{q_\phi(z|x)} [\nabla_\theta \log p_\theta(x|z)] \approx \frac{1}{S} \sum_{s=1}^S \nabla_\theta \log p_\theta(x|z^s) \quad (5)$$

However, the gradient w.r.t. ϕ can not be sampled that easily

$$\nabla_\phi \mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)] = \nabla_\phi \int q_\phi(z|x) \log p_\theta(x|z) dz = \int \log p_\theta(x|z) \nabla_\phi q_\phi(z|x) dz. \quad (6)$$

To overcome this problem, we use a reparameterization trick. The function $g_\phi(\cdot)$ is chosen such that it maps a data point $x^{(i)}$ and a random noise vector $\epsilon^{(l)}$ to a sample from the approximate posterior for that datapoint: $z^{(i,l)} = g_\phi(\epsilon^{(l)}, x^{(i)})$ where $z^{(i,l)} \sim q_\phi(z|x^{(i)})$. A vector of latent variables z in a high-dimensional space Z is sampled according to a probability density function (PDF) $P(z)$ defined over Z . Thus, the reparameterized lower bound $\tilde{\mathcal{L}}(p, q; x) \approx \mathcal{L}(p, q; x)$ can be written as

$$\tilde{\mathcal{L}}(p, q; x) = \frac{1}{S} \sum_{s=1}^S \log p_\theta(x|z^s) - \text{KL}(q_\phi(z|x) \parallel p_\theta(z)) \quad (7)$$

The term $\frac{1}{S} \sum_{s=1}^S \log p_\theta(x|z^s)$ represents the negated reconstruction error. The combination of this term and the KL-divergence act as a regularizer in autoencoder parlance.

Online Appendix 3. Biases in Similarity Functions

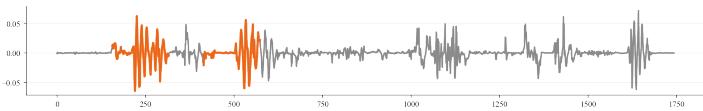
As discussed in the paper, similarity functions can introduce inadvertent biases toward finding only summary patterns that exhibit a predefined characteristic. To investigate this issue, we conducted an experiment in which we computed distances between all non-overlapping subsequences within a time series dataset. This dataset consists of measurements obtained from a hip-worn gyroscope sensor. We employed well-known similarity functions, including z-normalized Euclidean distance, Manhattan distance, Spearman correlation, Cosine similarity, Cross-Correlation, and Dynamic Time Warping, to identify the closest pair of subsequences. Figure 3 illustrates the outcomes generated by each similarity function.

We scrutinize the influence of different similarity functions on the discovery of the most analogous subsequence of a fixed length within the same data set. The results produced by each similarity function are visually presented in Figure 3. As demonstrated in Figures 3b, 3e, 3g, and 3f, Cross correlation, z-normalized Euclidean distance, and Cosine similarity, along with Manhattan distance, tend to select patterns that are flatter and exhibit less warping. Conversely, Spearman correlation and DTW distance pick out subsequences that are more warped as the most similar pairs, as depicted in Figures 3a and 3c. Interestingly, we noticed that the Edit distance did not perform optimally in this scenario - the subsequences identified as most similar in the data set did not appear visually similar, as shown in Figure 3a. This observation underscores the importance of choosing suitable similarity measures in time series analysis, particularly when the objective is to detect patterns that may not conform to a specific predefined shape or structure.

Euclidean Distance is the most common distance metric to measure the similarity between two points in a Euclidean space. This metric is defined as the square root of the sum of the squared differences between corresponding elements of two vectors. This metric is sensitive to the magnitude of the data and is not robust to the presence of outliers [MC88].

$$x_{1_{norm}} = \frac{x_1 - \mu_{x_1}}{\sigma_{x_1}}, \quad x_{2_{norm}} = \frac{x_2 - \mu_{x_2}}{\sigma_{x_2}} \quad (8)$$

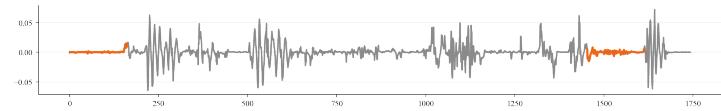
$$\text{z-normalised Euclidean distance}(x_{1_{norm}}, x_{2_{norm}}) = \sqrt{\sum_i (x_{1_{norm}}[i] - x_{2_{norm}}[i])^2} \quad (9)$$



First pair



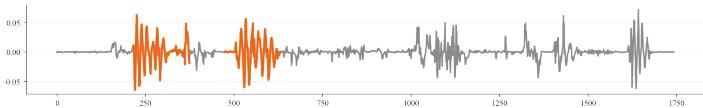
(a) The closest pair using Spearman correlation.



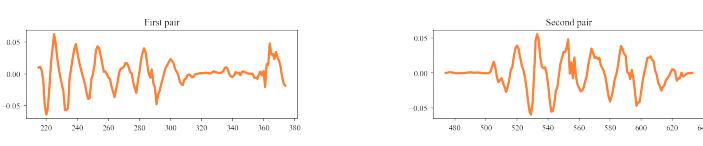
First pair



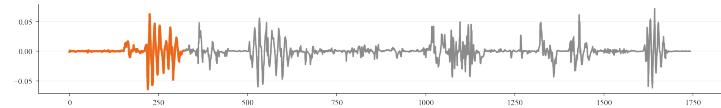
(b) The closest pair using Cross correlation.



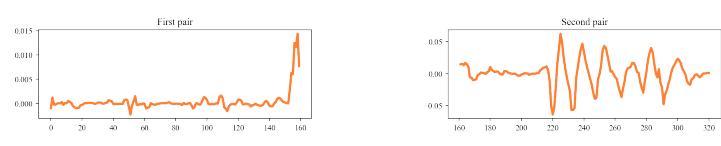
First pair



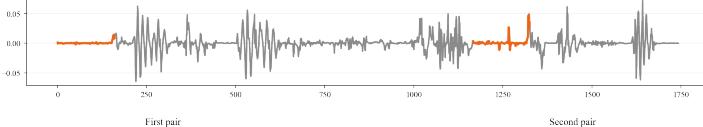
(c) The closest pair using DTW distance.



First pair



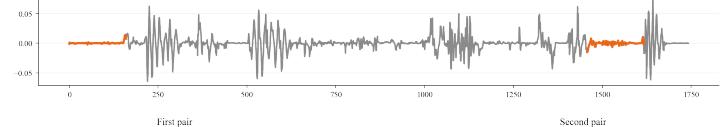
(d) The closest pair using Edit distance.



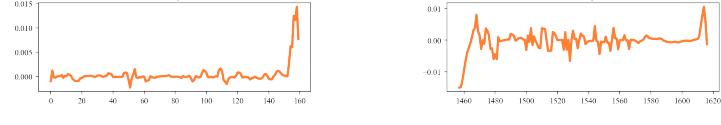
First pair



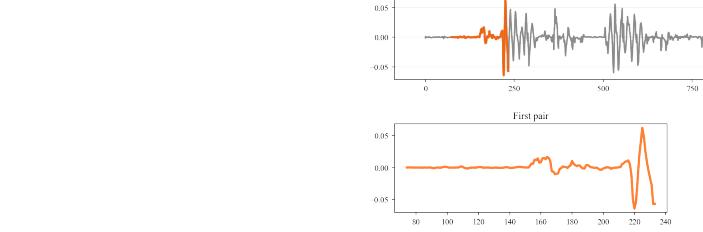
(e) The closest pair using z-normalized Euclidean distance.



First pair



(f) The closest pair using Manhattan distance.



(g) The closest pair using Cosine similarity.

Figure 3: The top panel displays the time series data in gray, with the highlighted orange segments representing the closest pair. The bottom panel presents a side-by-side comparison of the closest pair extracted from the data using the specified similarity function.

Manhattan Distance, also known as L1 or city block distance, is calculated as the sum of the absolute differences between corresponding elements of two vectors. This metric is more robust to outliers compared to Euclidean distance [ZDJ12] but is also sensitive to the magnitude of the data.

$$\text{Manhattan distance}(x_1, x_2) = \sum_i |x_1[i] - x_2[i]| \quad (10)$$

Spearman Correlation measures the strength of the monotonic relationship between two variables based on their ranks. Unlike Euclidean and Manhattan distances, Spearman correlation is invariant to monotonic transformations of the data and is less sensitive to the presence of outliers. However, it only captures the strength of the monotonic relationship and not the actual shape of the relationship between variables.

$$\rho = 1 - \frac{6 \sum_i (r_{1i} - r_{2i})^2}{n(n^2 - 1)} \quad (11)$$

Cosine Similarity is a measure of the cosine of the angle between two non-zero vectors in an inner product space. This metric is invariant to the magnitude of the data and is only concerned with the orientation of the vectors. Cosine similarity is particularly useful when dealing with high-dimensional data, as it tends to be less sensitive to the curse of dimensionality.

$$\text{Cosine similarity}(x_1, x_2) = \frac{x_1 \cdot x_2}{\|x_1\|_2 \cdot \|x_2\|_2} \quad (12)$$

Cross Correlation is a measure of similarity between two signals as a function of the time lag applied to one of them. This metric is sensitive to both the magnitude and phase of the data and is widely used in signal processing and pattern recognition tasks. Cross-correlation

measures the linear dependency between two signals. This is a significant limitation when the underlying relationship is non-linear [DGS16].

$$\text{CrossCorrelation}(x_1, x_2) = \sum_i (x_1[i] * x_2[i + \tau]) \quad (13)$$

Dynamic Time Warping (DTW) is a similarity measure for time series data that allows for non-linear alignment of the sequences. DTW calculates the optimal alignment between two sequences by minimizing the overall distance between them, allowing for stretching or compressing subsequences to accommodate temporal variations. This metric is particularly suitable for time series data with varying speeds or temporal distortions. One of the most strongest disadvantages of DTW is its computational complexity. While the flexibility DTW provides in aligning sequences is generally an advantage, it can sometimes lead to overfitting and non-intuitive warping paths [ZTD15].

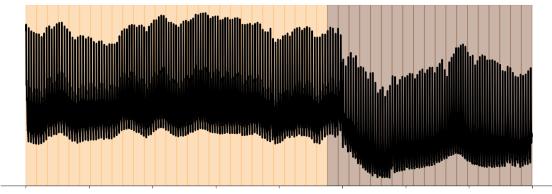
$$DTW(i, j) = |x_1[i] - x_2[j]| + \min(D(i - 1, j), D(i, j - 1), D(i - 1, j - 1)) \quad (14)$$

In conclusion, each similarity function has its strengths and weaknesses, and the choice of an appropriate function depends on the specific requirements of the application and the characteristics of the dataset. Our proposed method, based on information theory, circumvents the need to employ a specific choice of similarity function, thereby reducing the potential biases associated with their use.

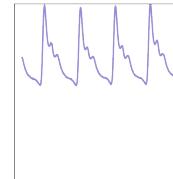
Online Appendix 5. Figures Pertaining to the Experimental Section of the Manuscript

This section encompasses all visualizations related to the experiments presented in the manuscript.

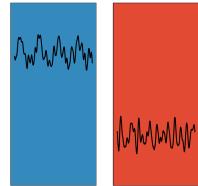
Online Appendix 5.1. Vital Sign



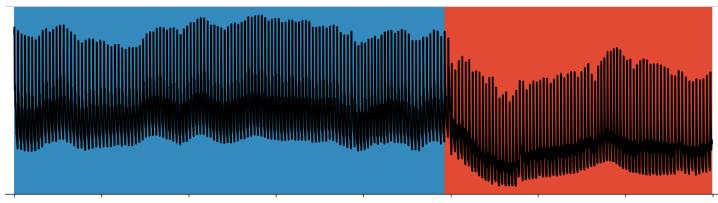
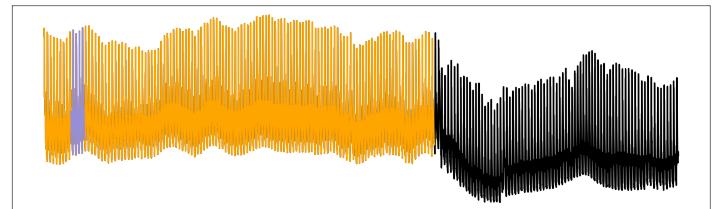
(a) Sample patient ABP.



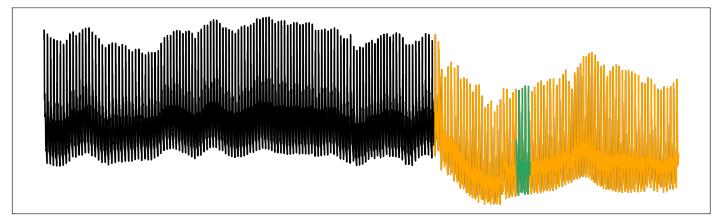
(b) Top two discovered patterns by MP-Snippets.



(c) Two patterns learned by T2P.



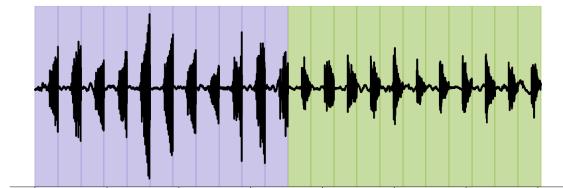
(d) Each subsequence's association with T2P-learned pattern, color mapped.



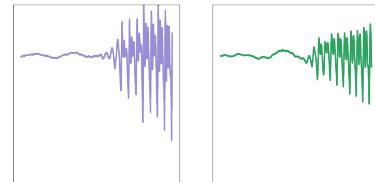
(e) Identified occurrences of MP-Snippets patterns in red, nearest subsequences to each in orange.

Figure 4: Visual inspection of T2P and MP-snippets discoveries and identification of pattern occurrences for patient ABP data.

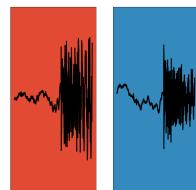
Online Appendix 5.2. Audio MNIST



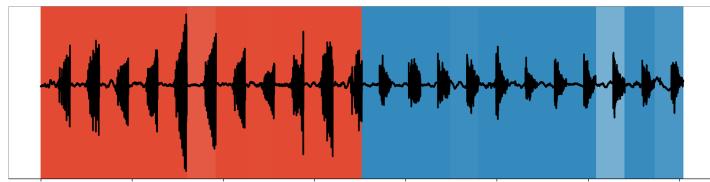
(a) Shows an audio waveform: *one* spoken in purple, *two* in green.



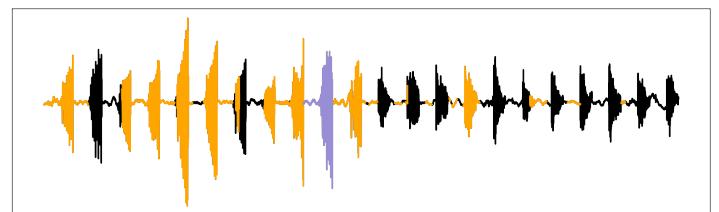
(b) Displays the top two discovered patterns (snippets) by MP-Snippets.



(c) Visualizes the two patterns learned by T2P.



(d) Shows each subsequence's association with T2P-learned patterns, color-coded per pattern.



(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange..

Figure 5: Visual Evaluation of T2P and MP-snippets' performance in identifying *one* and *two* utterances in the given audio waveform.

Online Appendix 5.3. ECG Data

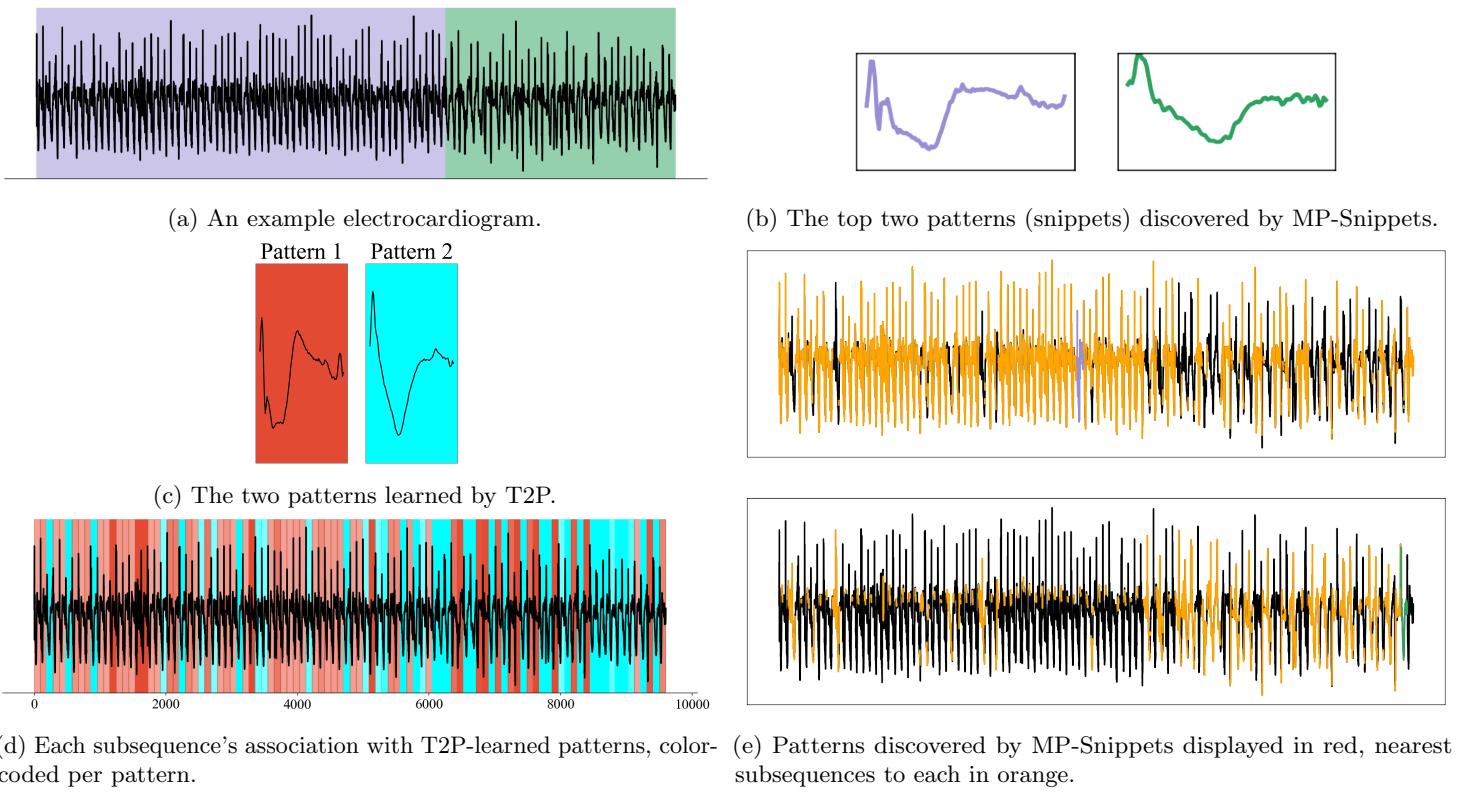


Figure 6: Visual Evaluation of T2P and MP-snippets' performance in identifying *normal heartbeat* and *myocardial infarction* utterances in electrocardiogram data.

Online Appendix 5.4. Plane Data

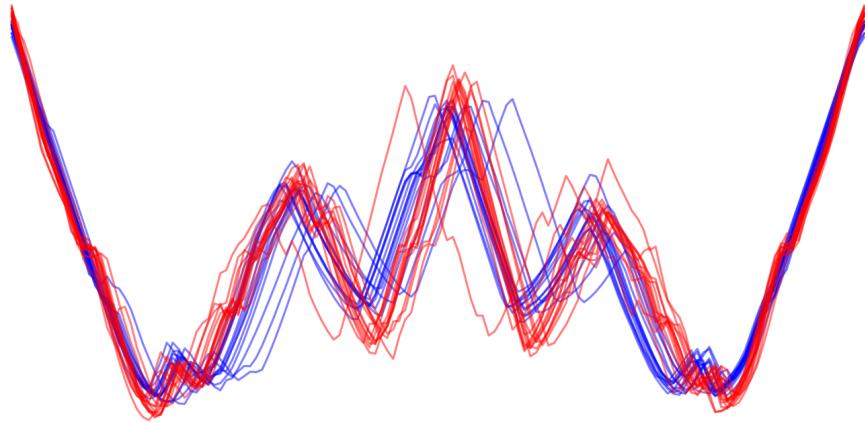
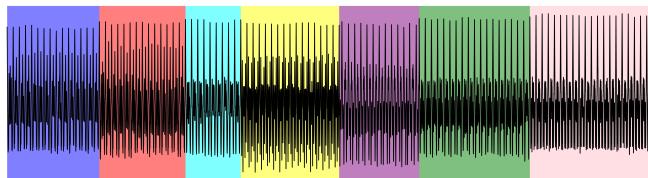
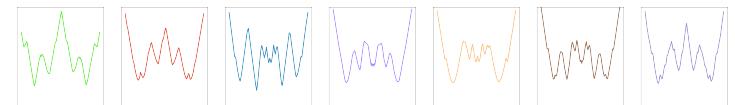


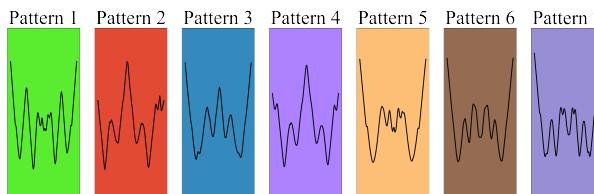
Figure 7: The pseudo time series shape of F-14 with closed wings (blue) and open wings (red)



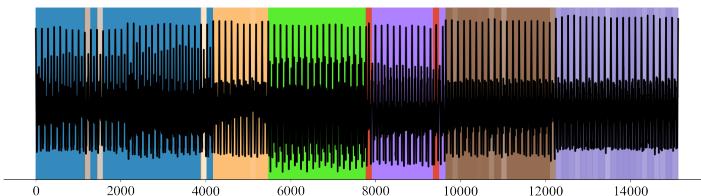
(a) Shows pseudo time series shape of different planes.



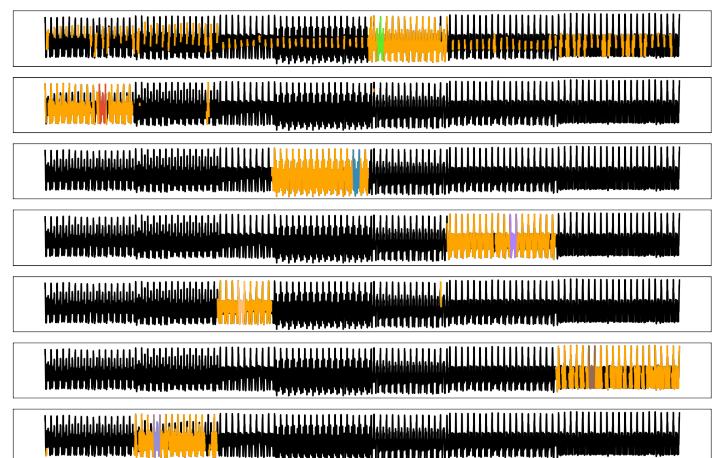
(b) Displays the top seven discovered patterns (snippets) by MP-Snippets.



(c) Visualizes the two patterns of learned by T2P.



(d) Shows each subsequence's association with T2P-learned patterns, color-coded per pattern.

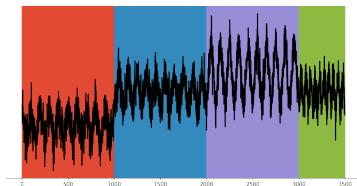


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange..

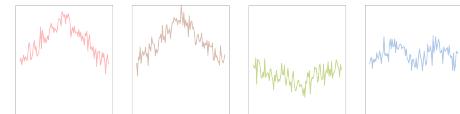
Figure 8: Visual Evaluation of T2P and MP-snippets' performance in identifying *one* and *two* utterances in the given audio waveform.

Online Appendix 6. Dealing with Noisy Conditions in Time Series Pattern Analysis

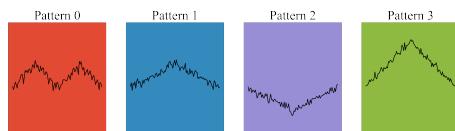
This section encompasses all visualizations corresponding to Appendix B as detailed in the manuscript.



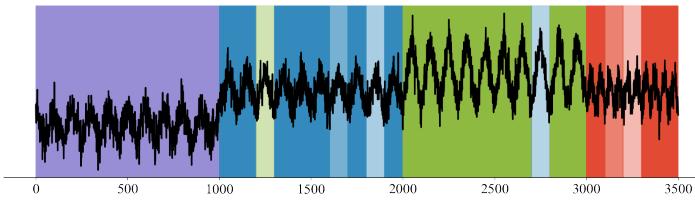
(a) Synthetic data.



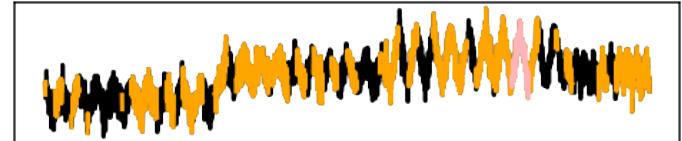
(b) Top four discovered patterns by MP-Snippets.



(c) Top four patterns learned by T2P.

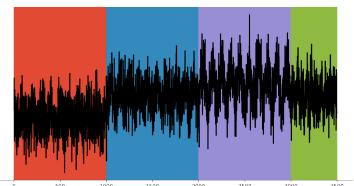


(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

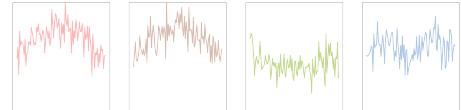


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

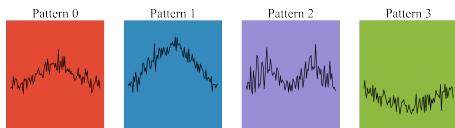
Figure 9: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where 30% noise is injected to the data, utilizing synthetic data.



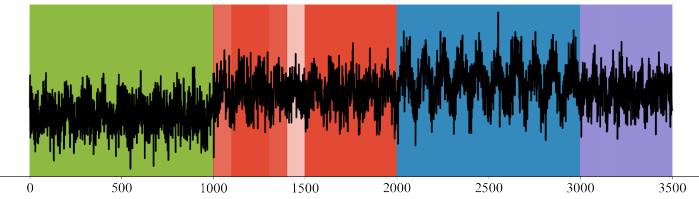
(a) Synthetic data.



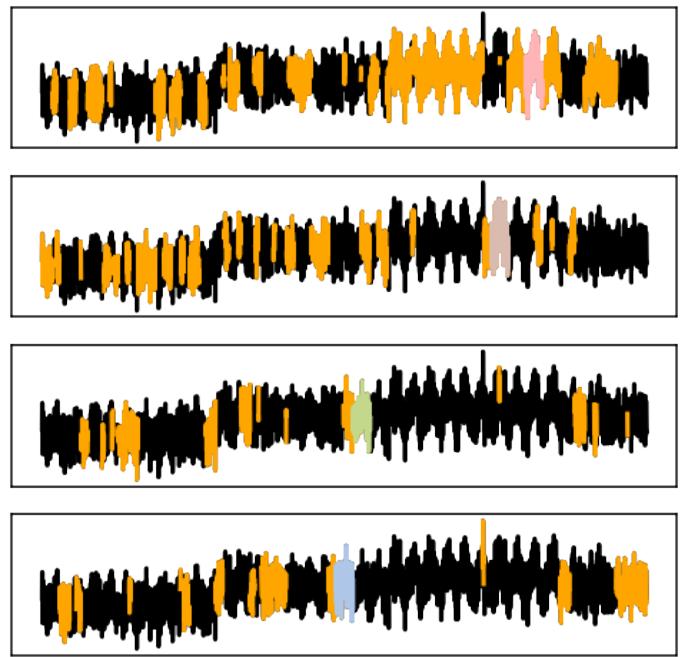
(b) Top four discovered patterns by MP-Snippets.



(c) Top four patterns learned by T2P.

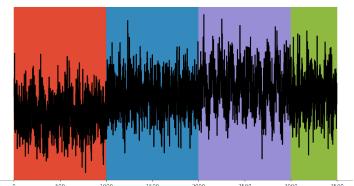


(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

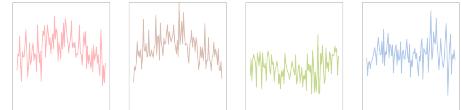


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

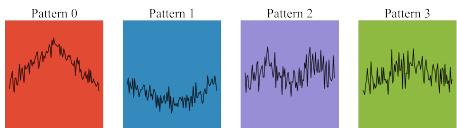
Figure 10: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where 70% noise is injected to the data, utilizing synthetic data.



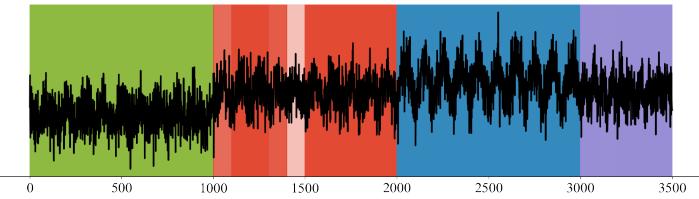
(a) Synthetic data.



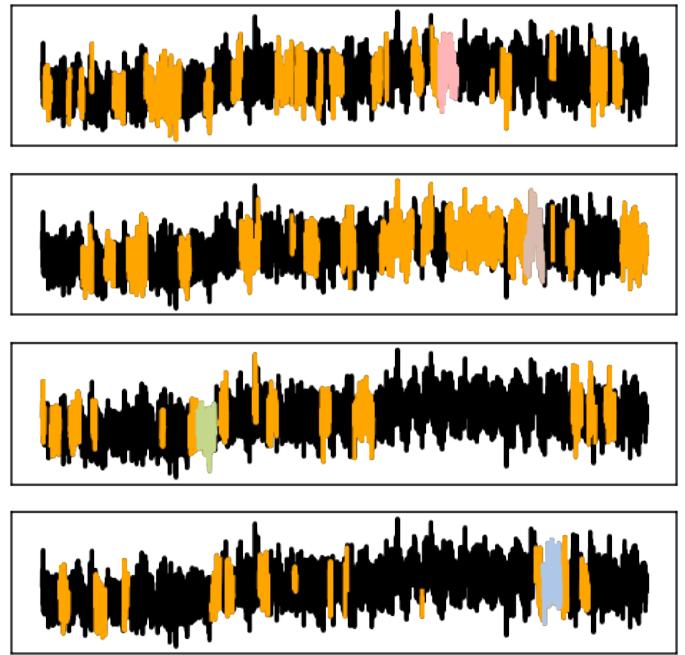
(b) Top four discovered patterns by MP-Snippets.



(c) Top four patterns learned by T2P.

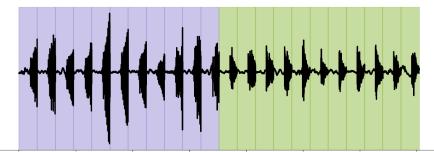


(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

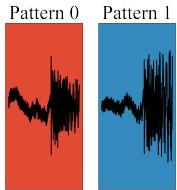


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

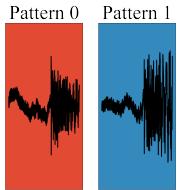
Figure 11: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where 100% noise is injected to the data, utilizing synthetic data.



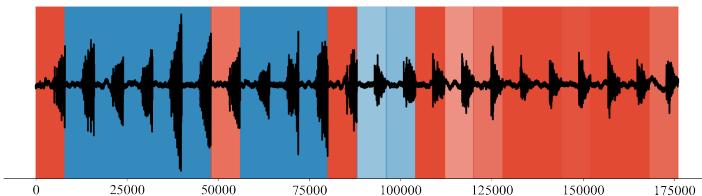
(a) Audio waveforms: *one* spoken in purple, *two* in green.



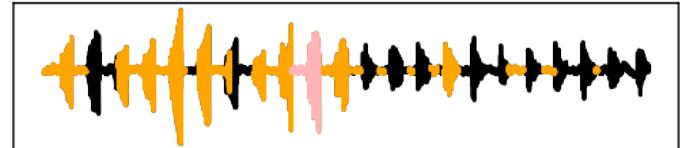
(b) Top two discovered patterns by MP-Snippets.



(c) Top two patterns learned by T2P.



(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.



(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

Figure 12: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where 1% noise is injected to the data, utilizing AudioMNIST data.

Online Appendix 7. Guiding Hyperparameter Selection

This section encompasses all visualizations corresponding to Appendix C as detailed in the manuscript.

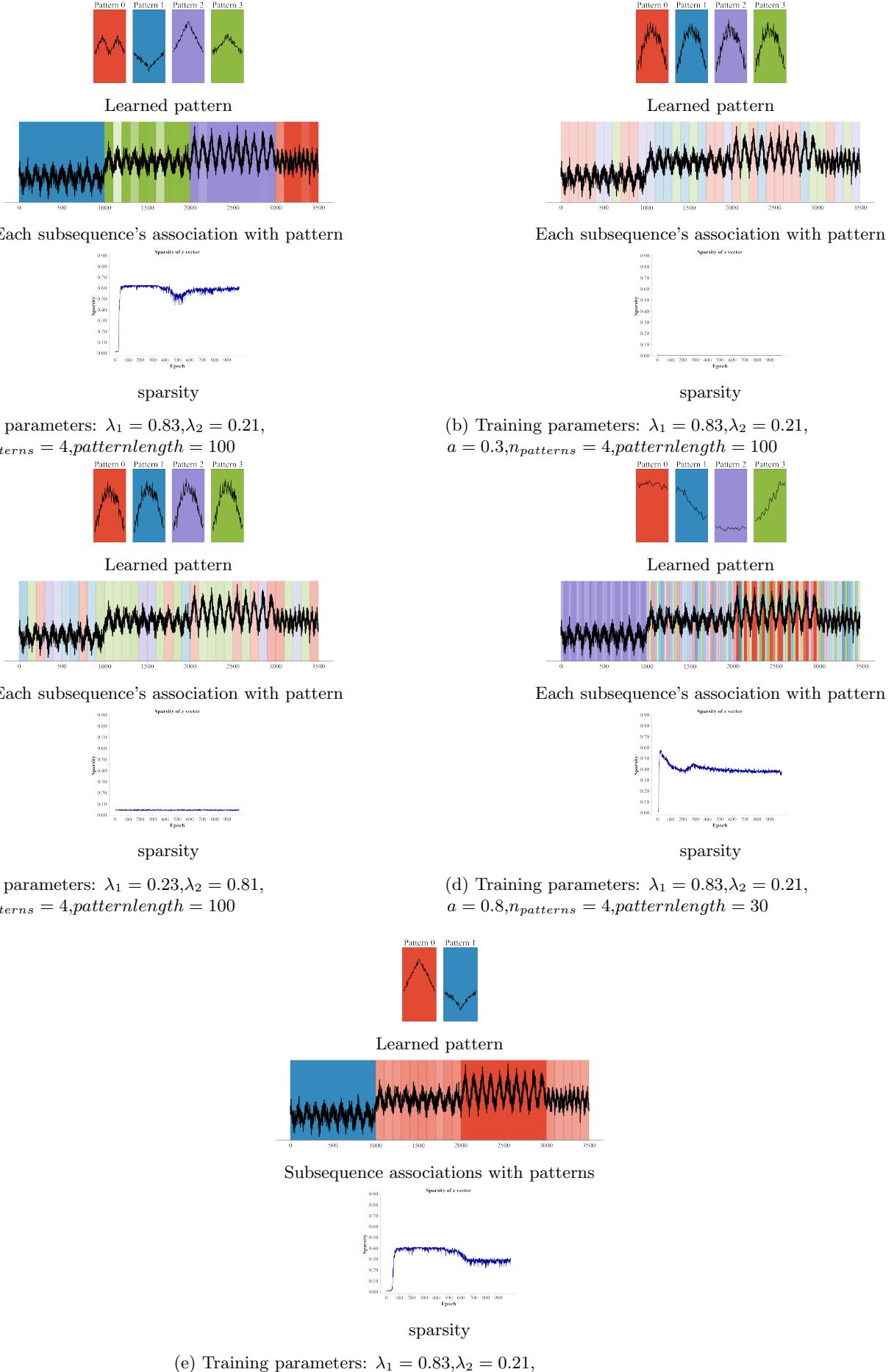


Figure 13: Effect of hyperparameters on latent space sparsity.

Online Appendix 8. Impact of Misspecifying the Number of Patterns

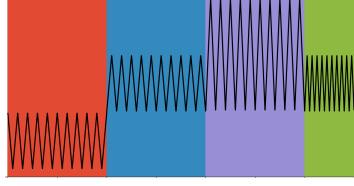
To demonstrate the impact of specifying the model with an incorrect number of patterns, we generate a synthetic time series containing four embedded patterns. We evaluate both T2P and MP-Snippets when they are specified with underestimated ($n_{patterns} = 1$), accurately estimated ($n_{patterns} = 4$), and overestimated ($n_{patterns} = 5$) values of the true pattern cardinality.

Figure 14 illustrates that when the specification underestimates the actual number of patterns, both methods identify the most salient pattern in the data. However, they attribute this pattern to incorrect subsequences due to the constraint that the number of detected patterns is fewer than the true patterns present in the data.

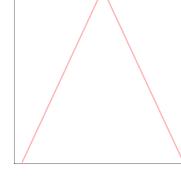
Whether the specification is accurate or an overestimation, T2P consistently outperforms MP-Snippet. It successfully detects all the valid patterns, as corroborated by Figures 15 and 16. This empirical observation underscores the fact that for T2P to be effective, the specified number of patterns must not be an underestimate of the true count.

In the case of an overestimated specification, T2P not only identifies all patterns but also includes a duplicate. On the contrary, MP-Snippets fails to discover new patterns, despite the allowance for extra patterns, as depicted in Figure 16.

One notable observation is that the final three patterns detected by MP-Snippets correspond to the same subsequence in the data. It seems that increasing the number of patterns does not improve MP-Snippets' performance or its ability to discover additional patterns. This is true irrespective of whether the specification is correct, an underestimation, or an overestimation. Therefore, in all scenarios, MP-Snippets is unable to fully discover all the patterns in the synthetic data.



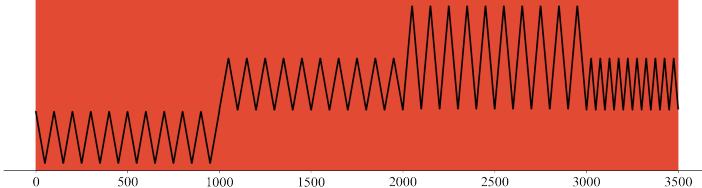
(a) Synthetic data.



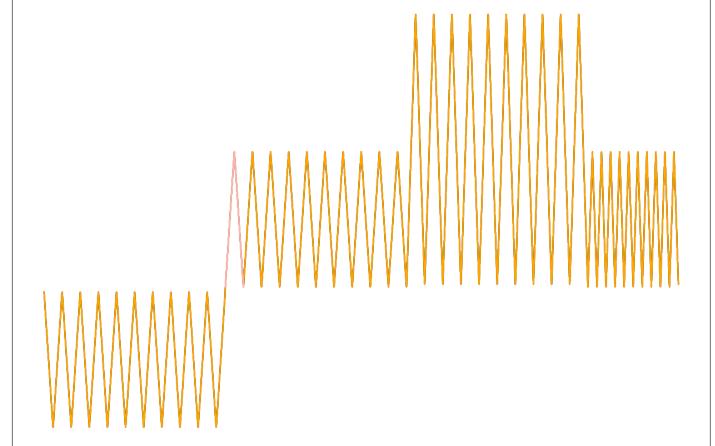
(b) Top one discovered pattern by MP-Snippets.



(c) Top one pattern learned by T2P.

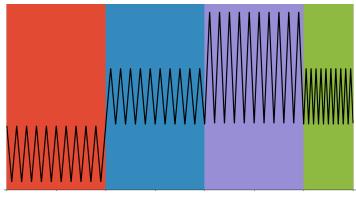


(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

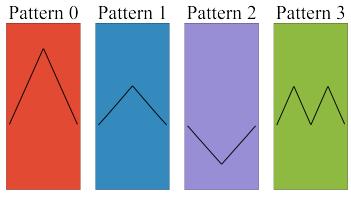


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

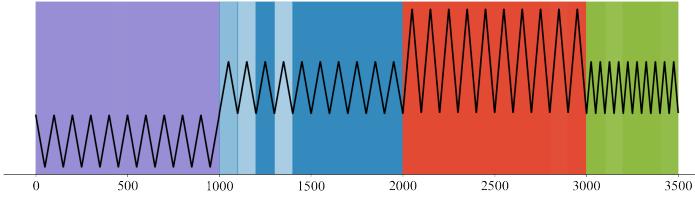
Figure 14: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the number of patterns is underestimated, utilizing synthetic data.



(a) Synthetic data.



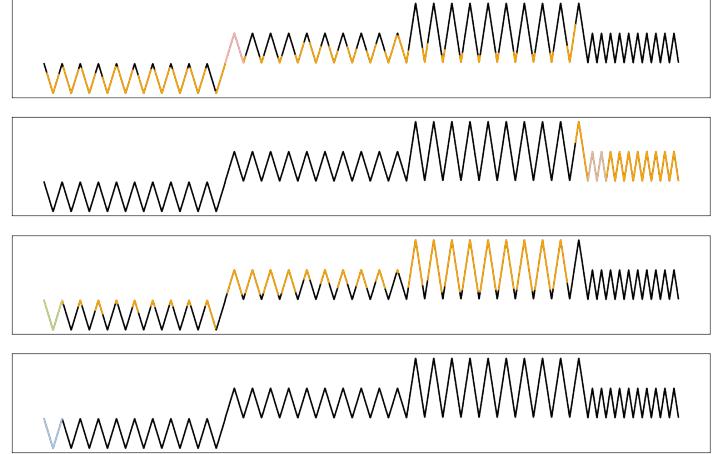
(c) Top-four patterns learned by T2P.



(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

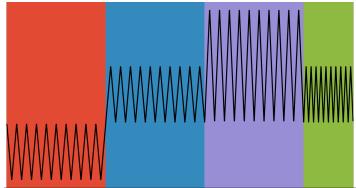


(b) Top-four discovered patterns by MP-Snippets.



(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

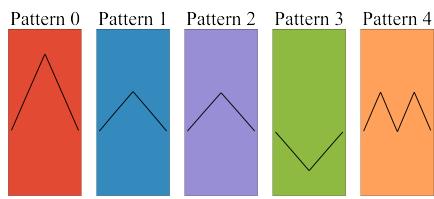
Figure 15: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the number of patterns is correct, utilizing synthetic data.



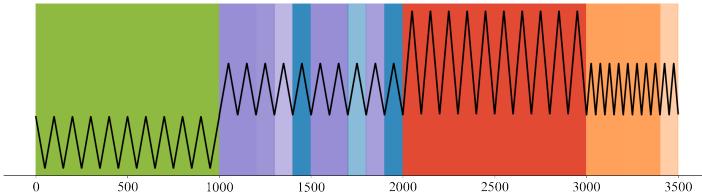
(a) Synthetic data.



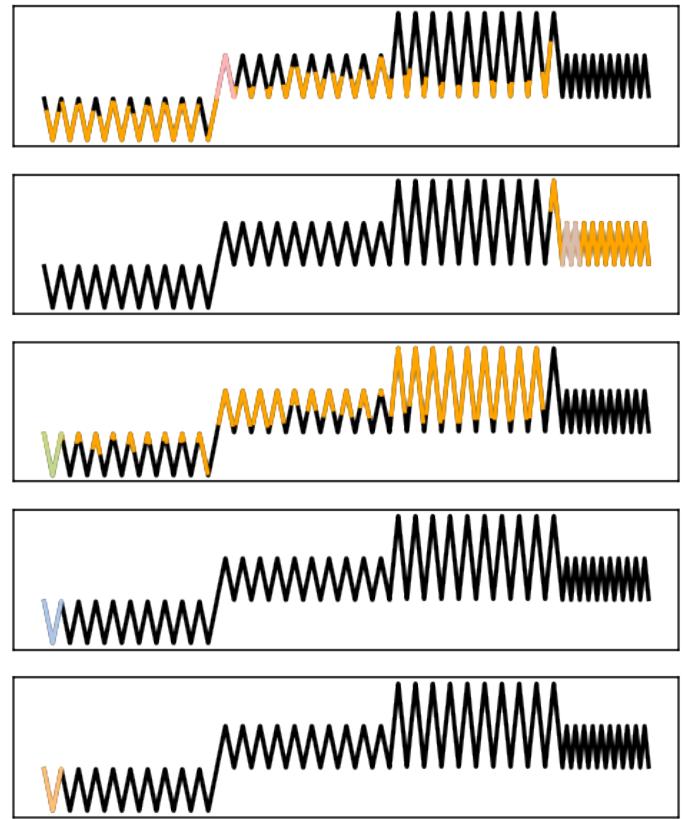
(b) Top-five discovered patterns by MP-Snippets.



(c) Top-five patterns learned by T2P.



(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.



(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

Figure 16: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the number of patterns is overestimated, utilizing synthetic data.

We also investigated the impact of misspecifying the number of patterns on real datasets, using the AudioMNIST dataset as an example. Both T2P and MP-Snippets were tested when the specified number of patterns was an underestimate ($n_{patterns} = 1$), a correct estimate ($n_{patterns} = 2$), and an overestimate ($n_{patterns} = 3$).

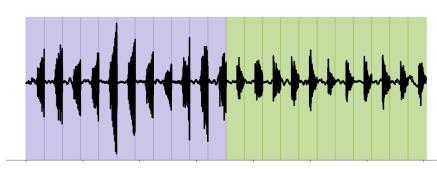
Consistent with our previous findings, both methods demonstrated similar outcomes when the number of patterns was underestimated. Each method identified a primary pattern that best represented the data, yet both made errors when associating this pattern with the correct subsequences in the data. Importantly, T2P, unlike MP-Snippets, learns patterns from the data. Hence, in an underestimation scenario, the pattern identified by T2P can be viewed as a blend of the patterns for *one* and *two*, as demonstrated in Figure 17. Meanwhile, MP-Snippets merely selects one of the data's subsequences.

Similar to the synthetic data, T2P outperformed MP-Snippets when the number of pattern estimations was correct or overestimated. In the case of a correct estimation, T2P successfully identified two correct patterns, while MP-Snippets struggled to detect the second pattern accurately, as discussed in the main paper and depicted in Figure 18.

When overestimating the number of patterns, T2P did not learn exact duplicates due to the data's complexity. Instead the third pattern learned by T2P was a less pronounced pattern appearing later in the dataset, corresponding to the spoken number *two*. Meanwhile, MP-Snippets repeated its earlier mistake of incorrect pattern recognition. But the additional allowance for a pattern, it managed to discover a pattern similar to the thired pattern learned by T2P but continued to associate numerous subsequences with incorrect patterns, as shown in Figure 19.

In general, underestimating the number of patterns can result in missing some existing patterns. Therefore, we advise against underestimating pattern numbers to avoid missing existing patterns. When the correct number of patterns was provided, T2P clearly outperformed MP-Snippets, exhibiting near-perfect performance, whereas MP-Snippets committed several errors. Overestimating the number of patterns led T2P to capture both the true patterns and additional duplicates. Hence, we recommend slightly overestimating the number of patterns as the result of our experiments suggests.

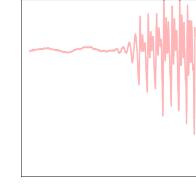
In our future work, we propose augmenting the latent space and decoder capacity of T2P and introducing an strategy to automatically remove duplicate pattern. These enhancements are intended to eliminate the necessity for hyperparameters related to the number of patterns. Possible techniques for eliminating duplicate patterns could include the implementation of a post-processing step like the one described in [Pet+17], the application of group sparsity regularization as detailed in [Bas+16] and [Mac+19], or the incorporation of a KNN process, as suggested by [For+18].



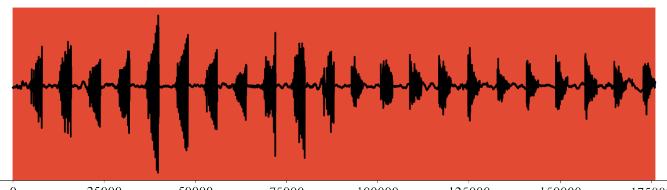
(a) An audio waveform: *one* spoken in purple, *two* in green.



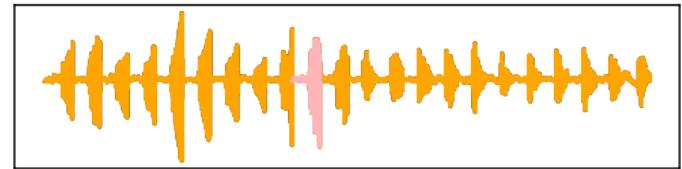
(c) Top-one pattern learned by T2P.



(b) Top one discovered pattern by MP-Snippets.

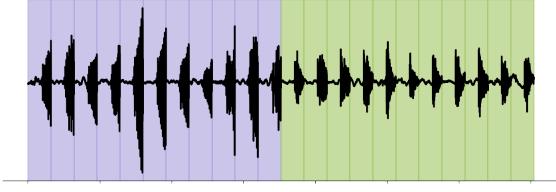


(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

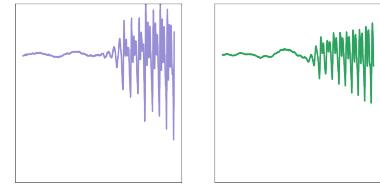


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

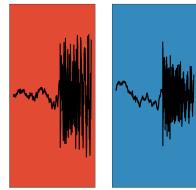
Figure 17: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the number of patterns is underestimated, utilizing AudioMNIST data.



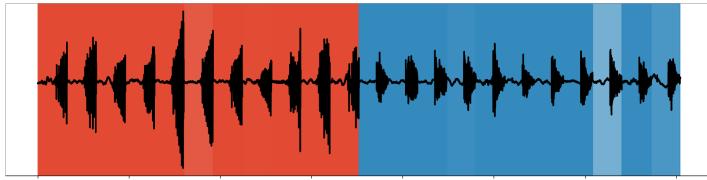
(a) Audio waveforms: *one* spoken in purple, *two* in green.



(b) Top-two discovered patterns by MP-Snippets.



(c) Top-two patterns learned by T2P.

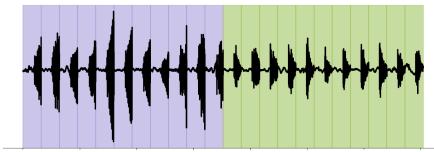


(d) Shows each subsequence's association with T2P-learned patterns, color-coded per pattern.

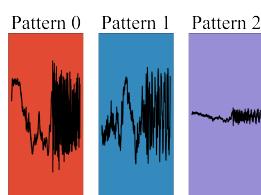


(e) Patterns discovered by MP-Snippets in red, nearest subsequences to each in orange.

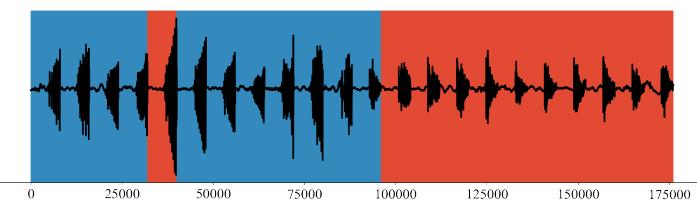
Figure 18: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the number of patterns is correct, utilizing AudioMNIST data.



(a) Audio waveforms: *one* spoken in purple, *two* in green.



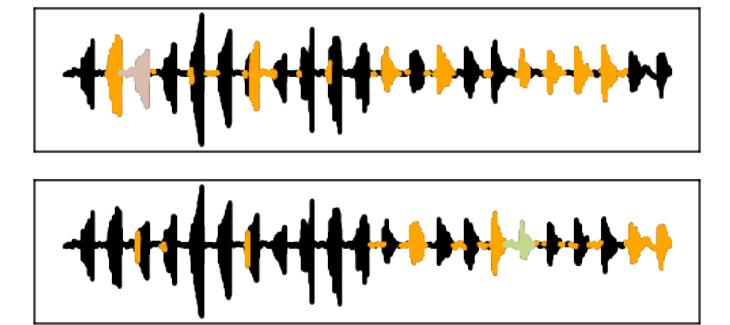
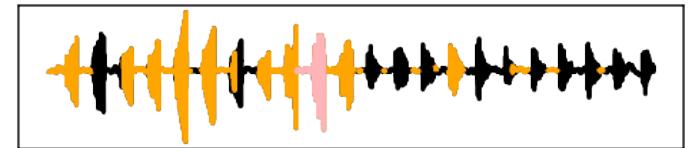
(c) Top-three patterns learned by T2P.



(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.



(b) Top-three discovered patterns by MP-Snippets.



(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

Figure 19: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the number of patterns is overestimated, utilizing AudioMNIST data.

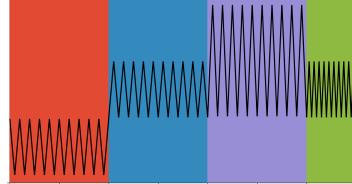
Online Appendix 9. Consequences of Pattern Length Misestimation

To gain a deeper understanding of the impact of mis-specification in pattern length, both T2P and MP-Snippets were put to the test under a variety of conditions on Synthetic and AudioMNIST datasets. For synthetic data, we intentionally specified an underestimated (Length = 50), accurately estimated (Length = 100), and overestimated (Length = 200) value for the pattern length.

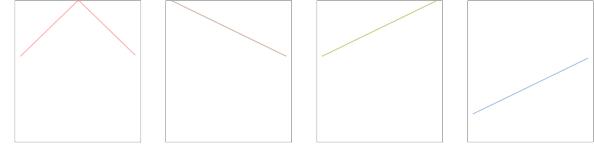
When the length of the pattern is underestimated, implying a greater number of patterns within synthetic data. However, we held constant the number of patterns at four for this experiment. Both T2P and MP-Snippets managed to identify four patterns that exist in the data. However, upon more detailed analysis, the patterns discerned by T2P proved to be a more accurate reflection of the data. This is evident when observing how these patterns are matched with their most analogous subsequences in the data, as depicted in Figure 20.

In situations where the pattern length was correctly estimated, T2P consistently outperformed MP-Snippets. T2P was successful in identifying the correct patterns and accurately associating them with the most closely matching subsequences in the data. On the other hand, MP-Snippets only managed to discover three out of four patterns and frequently mismatched the subsequences with their corresponding patterns, as illustrated in Figure 15.

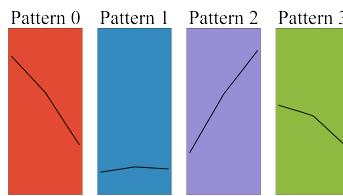
When the pattern length was overestimated, both methods succeeded in identifying the four correct patterns. However, as demonstrated in Figure 21, MP-Snippets did not accurately pair the patterns with their respective subsequences in the data. In contrast, T2P demonstrated perfect precision.



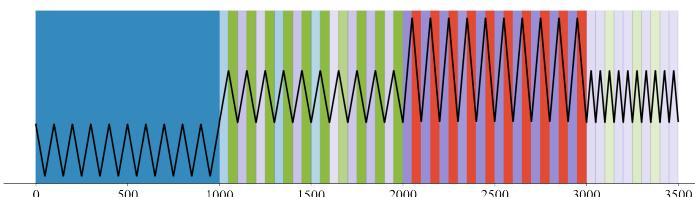
(a) Synthetic data.



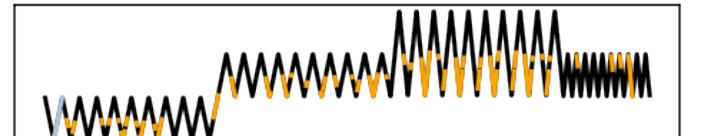
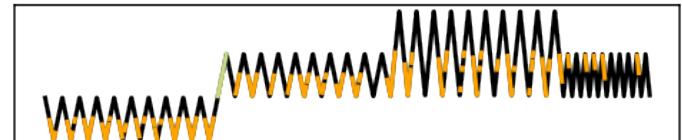
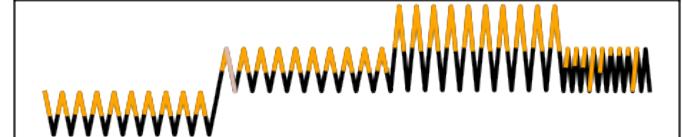
(b) Top four discovered patterns by MP-Snippets.



(c) Top four patterns learned by T2P.

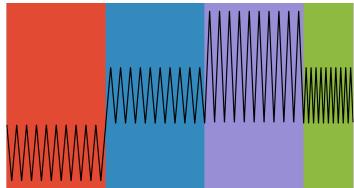


(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

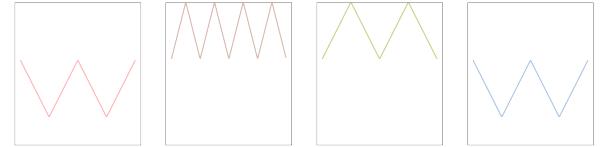


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

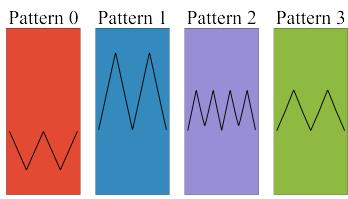
Figure 20: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the length of patterns is underestimated, utilizing synthetic data.



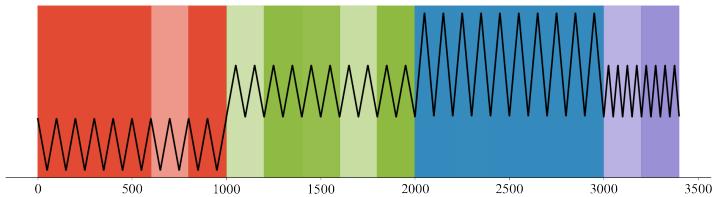
(a) Synthetic data.



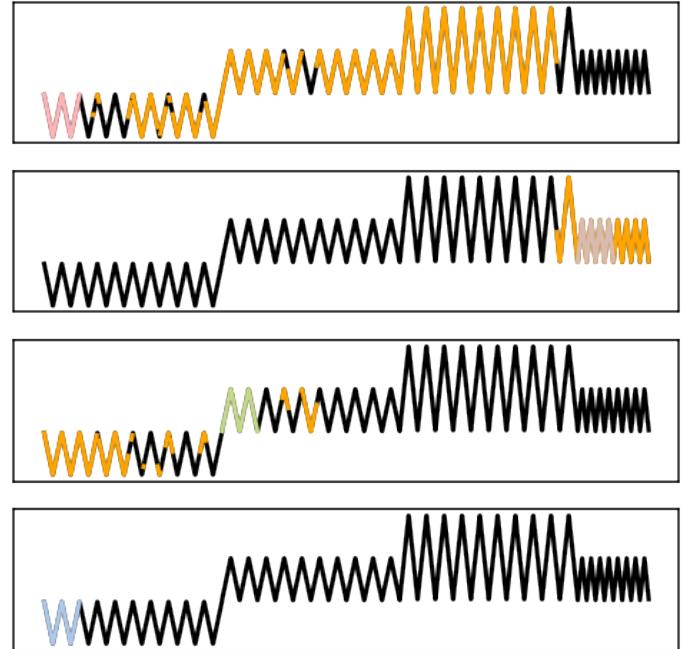
(b) Top four discovered patterns by MP-Snippets.



(c) Top four patterns learned by T2P.



(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.

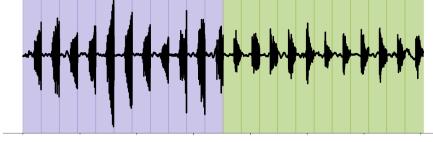


(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

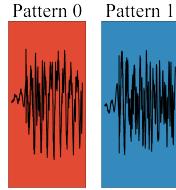
Figure 21: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the length of patterns is overestimated, utilizing synthetic data.

Further investigations were conducted into the consequences of mis-specifying pattern length within audio data. We analyzed scenarios where the pattern length was underestimated (Length = 4000), accurately assessed (Length = 8000), and overestimated (Length = 16000). As evidenced by Figures 22, the underestimation of the pattern length considerably influences the accuracy of the outcomes obtained by both methodologies. However, when the pattern length is correctly identified, T2P surpasses MP-Snippets, delivering flawless results as shown in 18. Overestimating the pattern length seems to assist MP-Snippets in accurately identifying patterns within the AudioMNIST dataset. Yet, under these conditions, T2P requires a more substantial amount of spoken digit samples to perform optimally. By doubling the length, we inadvertently limit the number of samples per spoken digit to five, which is insufficient for effective neural network training. To counter this, we trained T2P on a larger audio dataset where each number was sampled 50 times, compared to the initial 11 times. As exhibited in Figure 23, T2P performed commendably, successfully identifying patterns.¹

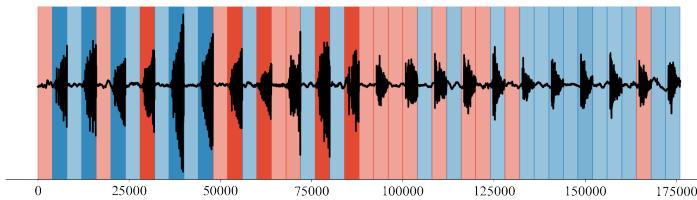
When the pattern length is unknown or when there's uncertainty about the data collection rate, our recommendation is to err on the side of caution by overestimating the pattern length. T2P has demonstrated its competency in detecting partial patterns within larger windows, suggesting its adaptability to handle variations in pattern length.



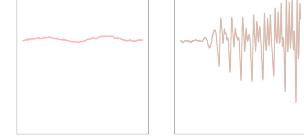
(a) Audio waveforms: *one* spoken in purple, *two* in green.



(c) Top-two patterns learned by T2P.



(d) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.



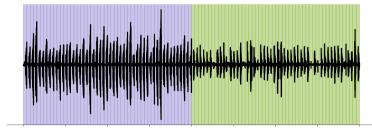
(b) Top-two discovered patterns by MP-Snippets.



(e) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

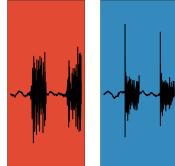
Figure 22: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the length of patterns is underestimated, utilizing AudioMNIST data.

¹Please note that due to the constraints imposed by RAM limitations (32 GiBytes), we were unable to test MP-Snippets on larger datasets.

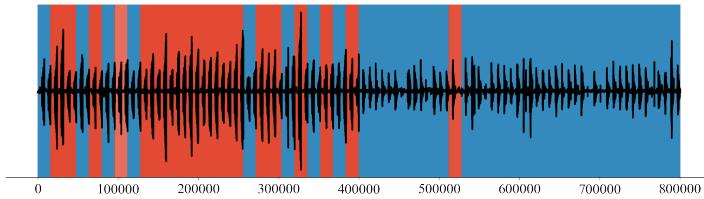


(a) Audio waveforms: *one* spoken in purple, *two* in green.

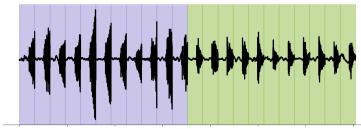
Pattern 0 Pattern 1



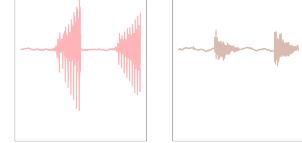
(c) Top-two patterns learned by T2P.



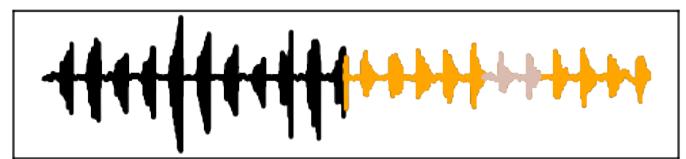
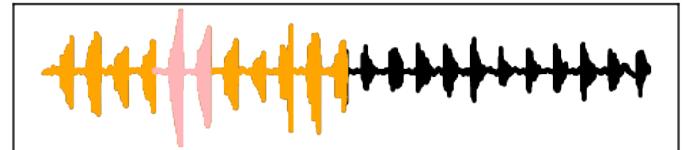
(e) Identified, color-mapped occurrences of T2P's discovered patterns in the original time series data.



(b) Audio waveforms: *one* spoken in purple, *two* in green.



(d) Top-two discovered patterns by MP-Snippets.



(f) Identified, color-mapped occurrences of MP-Snippet's discovered patterns are depicted in the original time series data. Subsequences closely associated with a specific pattern are highlighted in orange.

Figure 23: A visual comparison of pattern discoveries and identifications made by T2P and MP-Snippets under conditions where the length of patterns is overestimated, utilizing AudioMNIST data.

References

- [MC88] Glenn W Milligan and Martha C Cooper. “A study of standardization of variables in cluster analysis”. In: *Journal of Classification* 5 (1988), pp. 181–204.
- [ZDJ12] Yu Zhong, Yunbin Deng, and Anil K Jain. “Keystroke dynamics for user authentication”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE. 2012, pp. 117–123.
- [ZTD15] Zheng Zhang, Ping Tang, and Rubing Duan. “Dynamic time warping under pointwise shape context”. In: *Information sciences* 315 (2015), pp. 88–101.
- [Bas+16] Kevin Bascol et al. “Unsupervised interpretable pattern discovery in time series using autoencoders”. In: *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshop, S+ SSPR 2016, Mérida, Mexico, November 29–December 2, 2016, Proceedings*. Springer. 2016, pp. 427–438.
- [DGS16] Srimonti Dutta, Dipak Ghosh, and Shukla Samanta. “Non linear approach to study the dynamics of neurodegenerative diseases by Multifractal Detrended Cross-correlation Analysis—A quantitative assessment on gait disease”. In: *Physica A: Statistical Mechanics and its Applications* 448 (2016), pp. 181–195.
- [Pet+17] Sven Peter et al. “Sparse convolutional coding for neuronal assembly detection”. In: *Advances in Neural Information Processing Systems* 30 (2017).
- [For+18] Vincent Fortuin et al. “SOM-VAE: Interpretable Discrete Representation Learning on Time Series”. In: *Proceedings of the International Conference on Learning Representations*. 2018.
- [Mac+19] Emily L Mackevicius et al. “Unsupervised discovery of temporal sequences in high-dimensional datasets, with applications to neuroscience”. In: *Elife* 8 (2019), e38471.