

Evaluation von Diskretisierungsansätzen zur Repräsentation von Zeitreihen im Kontext der Mustererkennung

Zwischenvortrag - Bachelorarbeit

Prüfer: Prof. Dr.-Ing. Eric Sax und Prof. Dr. Ralf Reussner

Betreuer: M. Sc. Patrick Petersen und M. Sc. Hanno Stage

von Daniel Schneider

17.04.2023



Institut für Technik der
Informationsverarbeitung (ITIV)



Forschungsbereich Embedded Systems
and Sensors Engineering (ESS)

Motivation

"When it comes to the car of the future, data is literally the new oil"¹



Selbstfahrendes Auto

~4.000 GB pro Tag



Radar

~10-100 KB pro Sekunde



Sonar

~10-100 KB pro Sekunde



Kameras

~20-40 MB pro Sekunde



Lidar

~10-70 MB pro Sekunde



GPS

~50 KB pro Sekunde

¹ <https://newsroom.intel.com/editorials/krzanich-the-future-of-automated-driving/>

Motivation

Messungen über die Zeit liefern Zeitreihendaten, die spezielle Eigenschaften aufweisen

Eigenschaften von Zeitreihen:²



Sehr viele Datenpunkte



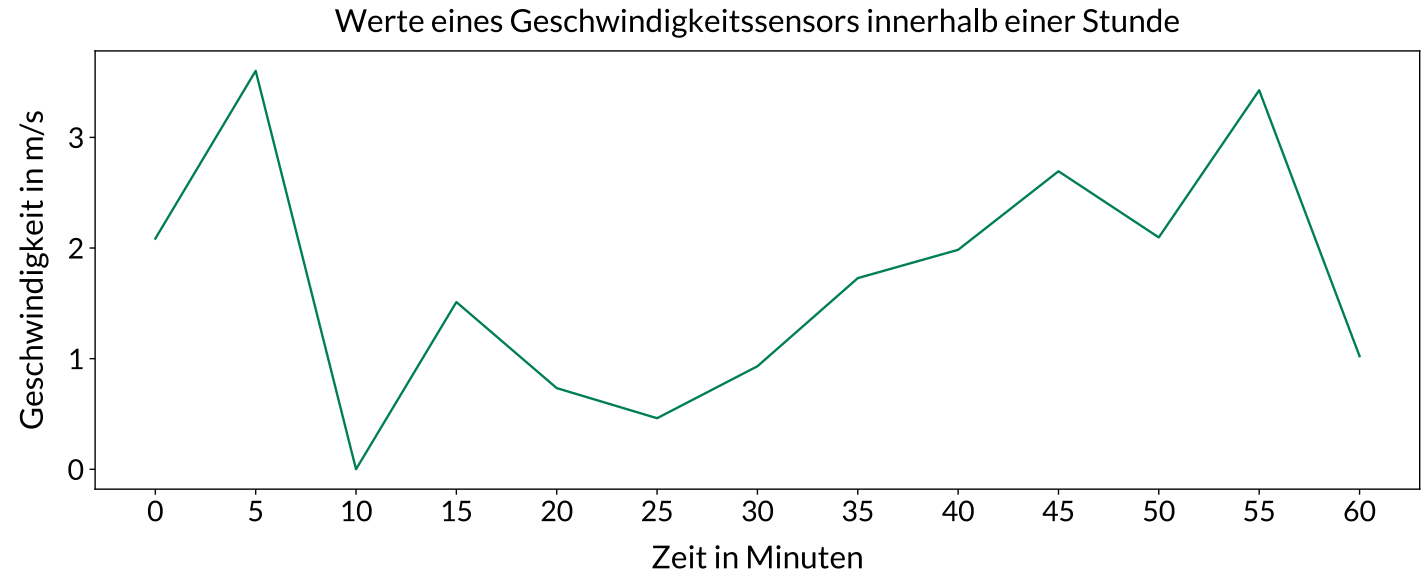
Verrauschte Datenpunkte



Zeitliche Abhängigkeit der Daten



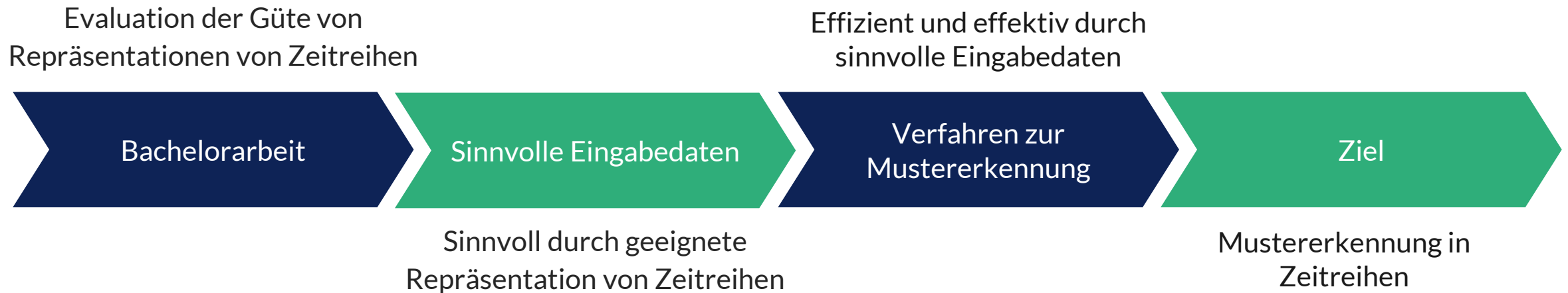
Musterbildung über die Zeit



² Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys (CSUR)*, 45(1), 1-34.

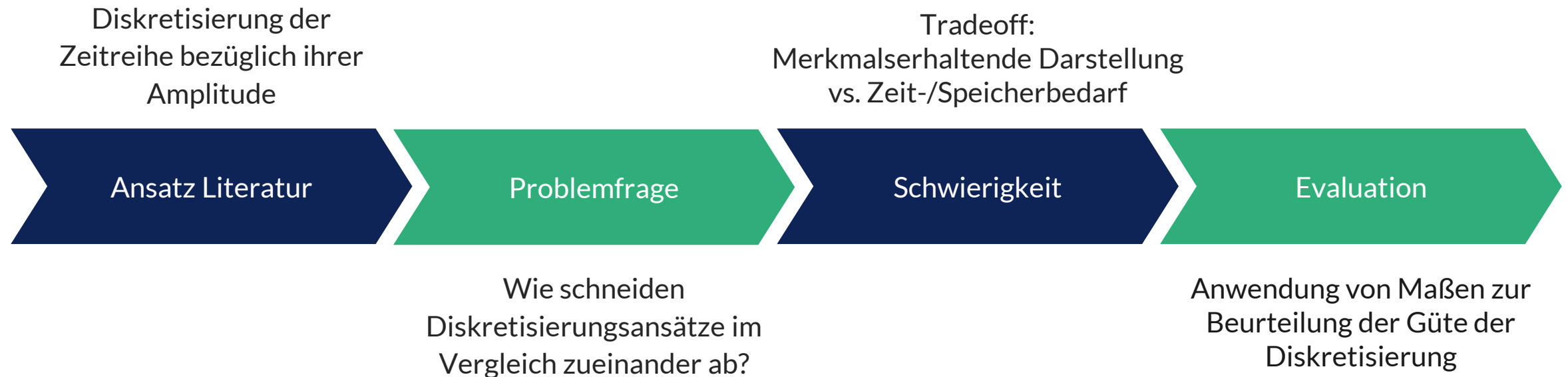
Kontexteinordnung der Bachelorarbeit

Geeignete Repräsentation von Zeitreihen ist notwendig für deren Verarbeitung und Analyse



Herangehensweise der Bachelorarbeit

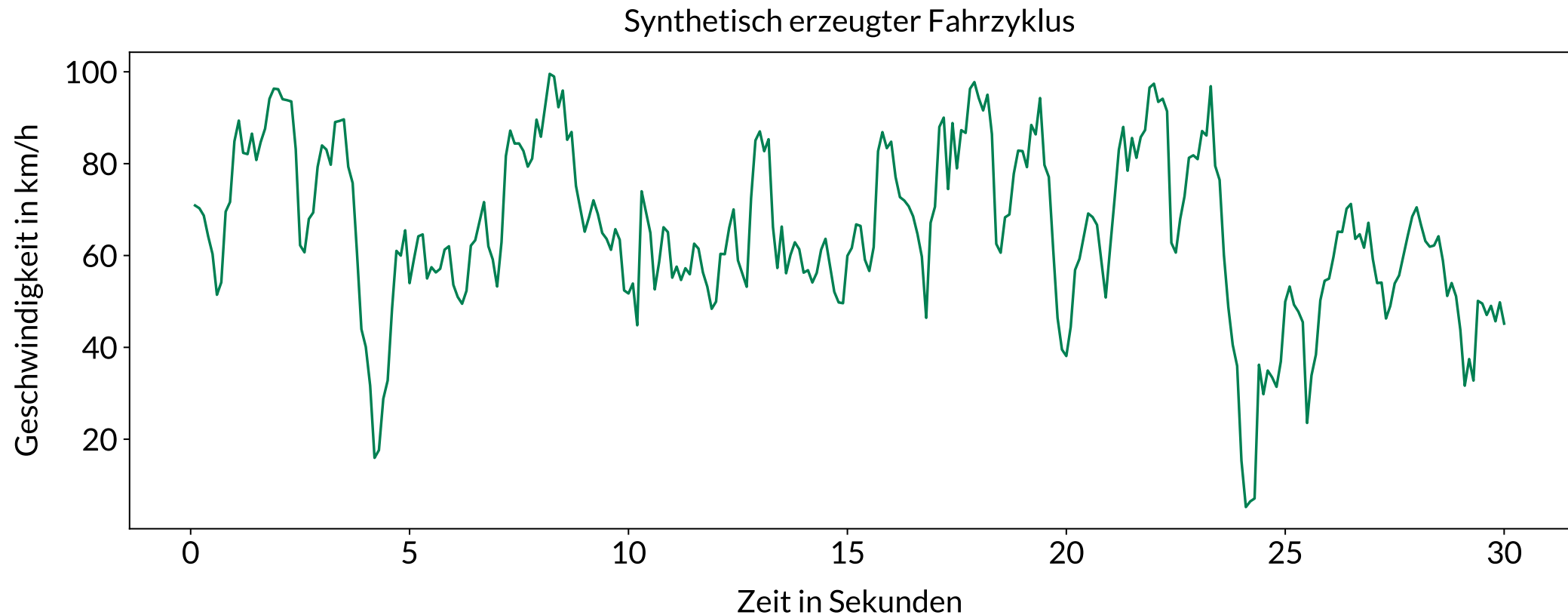
Diskretisierung der Amplitude ist ein Ansatz zur geeigneten Repräsentation von Zeitreihen



Anwendungsbezogener Use Case

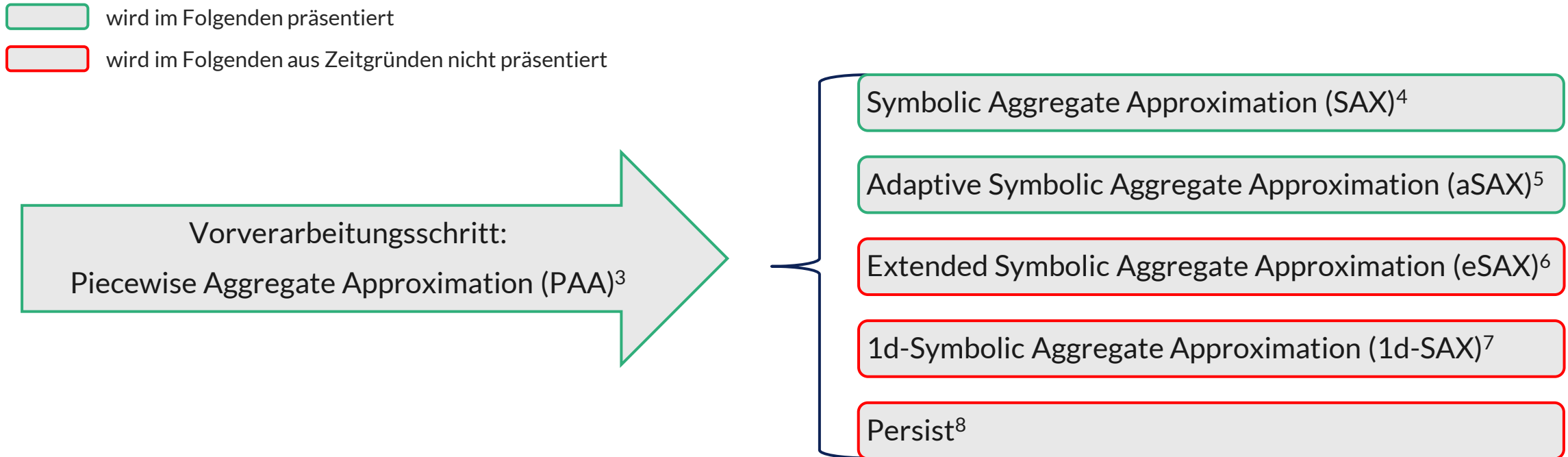
Von der Diskretisierung bis zur Evaluation anhand eines Fahrzyklus

Im Folgenden z-normalisiert, um Vergleichbarkeit zu schaffen



Überblick: Diskretisierungsansätze

Überblick über evaluierte Diskretisierungsansätze im Rahmen der Bachelorarbeit



³ Keogh, Eamonn, et al. "Dimensionality reduction for fast similarity search in large time series databases." *Knowledge and information Systems* 3 (2001): 263-286.

⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

⁵ Pham, Ninh D., Quang Loc Le, and Tran Khanh Dang. "Two novel adaptive symbolic representations for similarity search in time series databases." *2010 12th International Asia-Pacific Web Conference*. IEEE, 2010.

⁶ Lkhagva, Battuguldur, Yu Suzuki, and Kyoji Kawagoe. "New time series data representation ESAX for financial applications." *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. IEEE, 2006.

⁷ Malinowski, Simon, et al. "1d-sax: A novel symbolic representation for time series." *Advances in Intelligent Data Analysis XII: 12th International Symposium, IDA 2013, London, UK, October 17-19, 2013. Proceedings 12*.

⁸ Mörchen, Fabian, and Alfred Ultsch. "Optimizing time series discretization for knowledge discovery." *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. 2005.

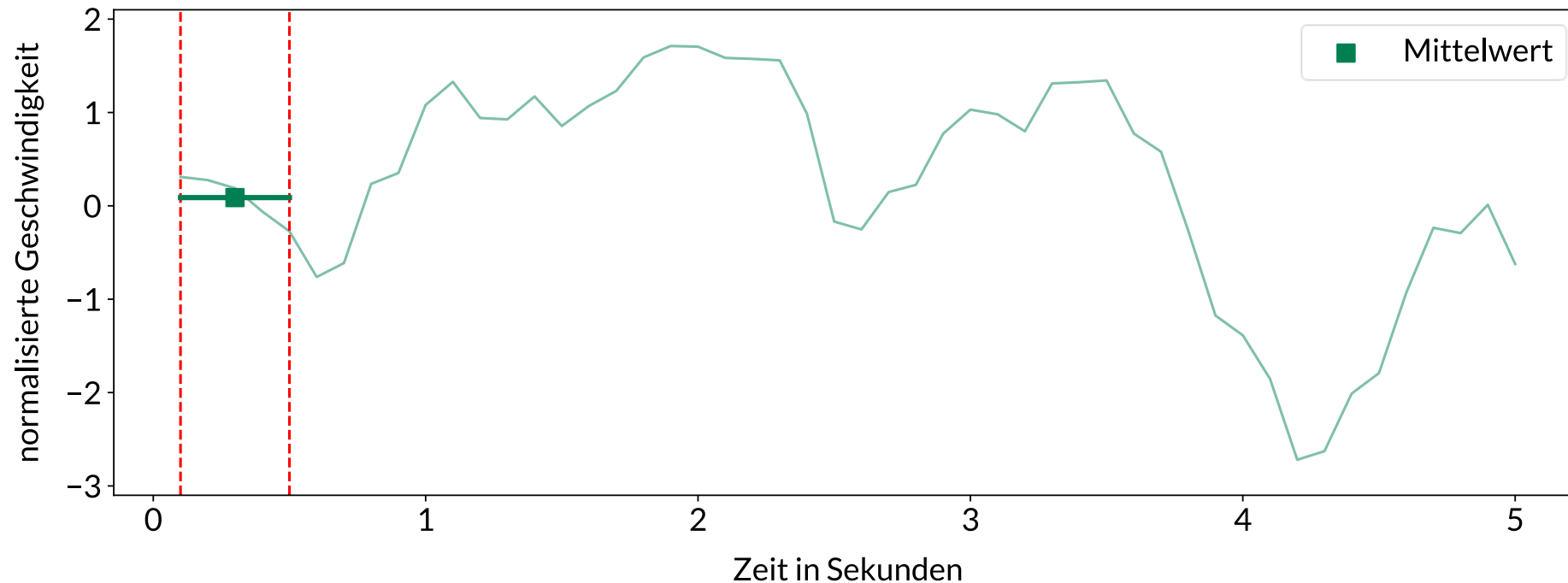
Approximation als Vorverarbeitungsschritt

Sliding-Window-Ansatz: Mittelwert der Punkte eines Fensters dient als Approximation

Ansatz Literatur

Piecewise Aggregate Approximation (PAA)³

Größe des Fensters als Parameter



³ Keogh, Eamonn, et al. "Dimensionality reduction for fast similarity search in large time series databases." *Knowledge and information Systems* 3 (2001): 263-286.

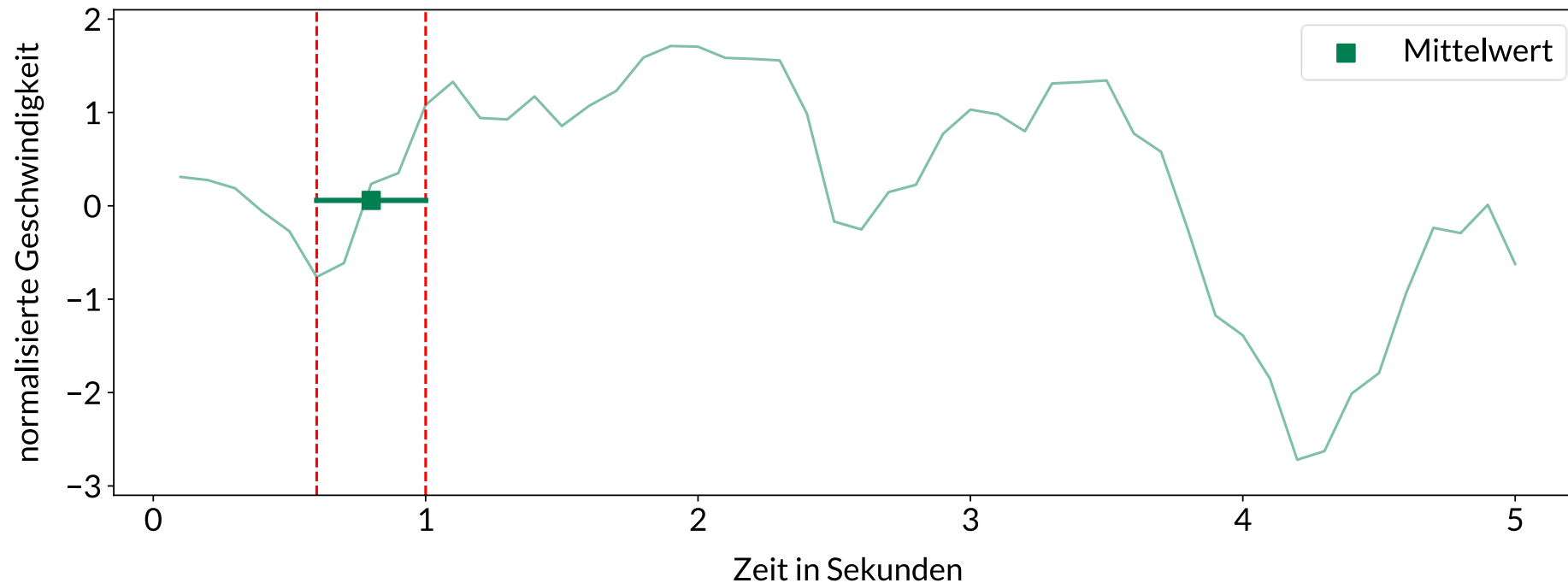
Approximation als Vorverarbeitungsschritt

Sliding-Window-Ansatz: Mittelwert der Punkte eines Fensters dient als Approximation

Ansatz Literatur

Piecewise Aggregate Approximation (PAA)³

Größe des Fensters als Parameter



³ Keogh, Eamonn, et al. "Dimensionality reduction for fast similarity search in large time series databases." *Knowledge and information Systems* 3 (2001): 263-286.

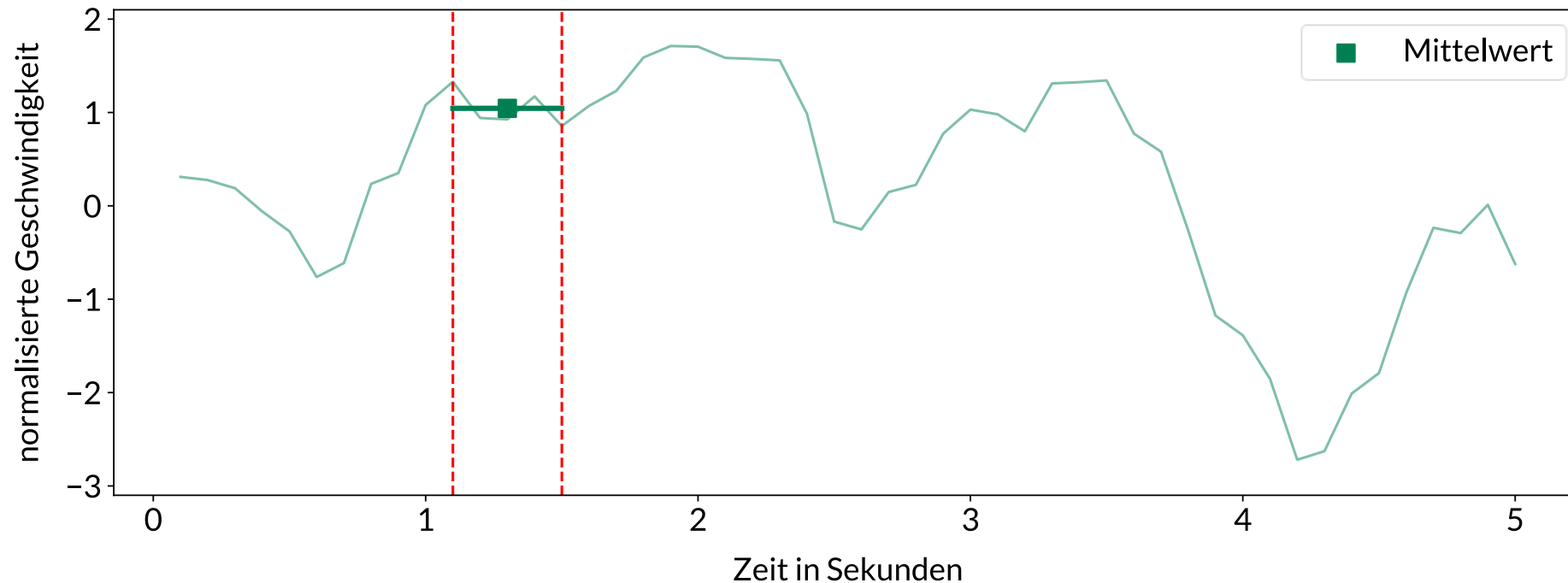
Approximation als Vorverarbeitungsschritt

Sliding-Window-Ansatz: Mittelwert der Punkte eines Fensters dient als Approximation

Ansatz Literatur

Piecewise Aggregate Approximation (PAA)³

Größe des Fensters als Parameter



³ Keogh, Eamonn, et al. "Dimensionality reduction for fast similarity search in large time series databases." *Knowledge and information Systems* 3 (2001): 263-286.

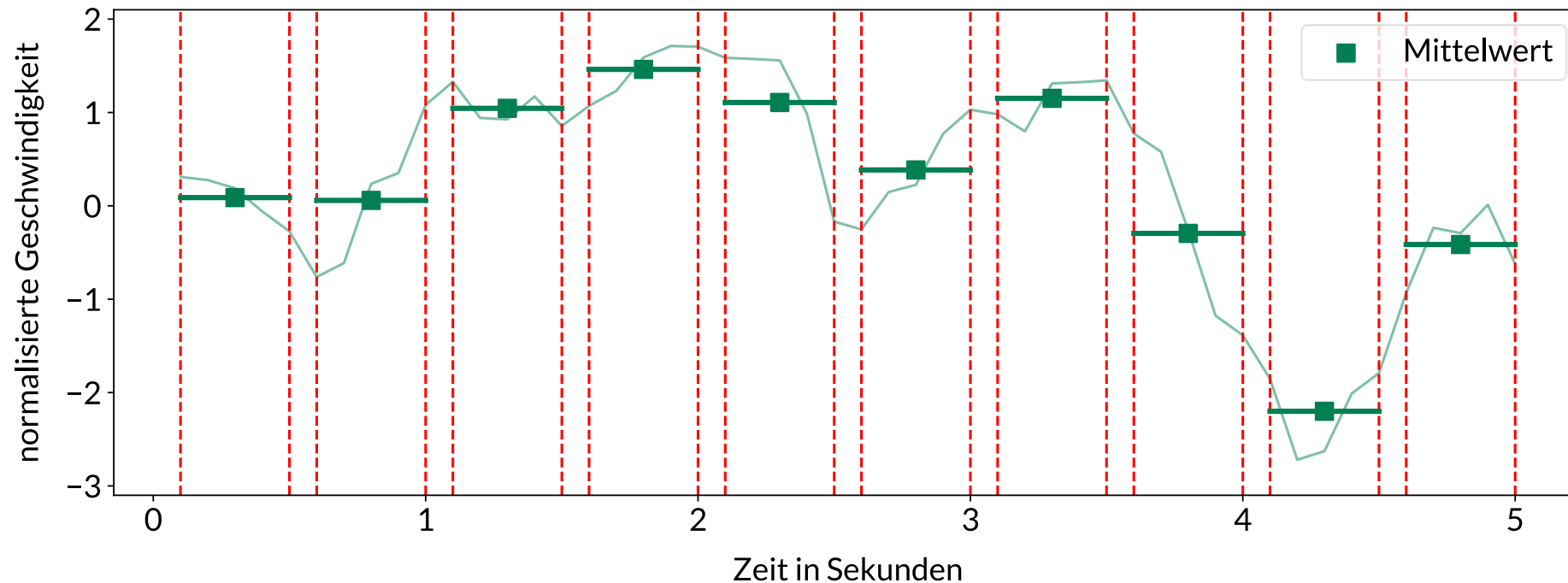
Approximation als Vorverarbeitungsschritt

Sliding-Window-Ansatz: Mittelwert der Punkte eines Fensters dient als Approximation

Ansatz Literatur

Piecewise Aggregate Approximation (PAA)³

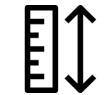
Tradeoff abhängig von Fenstergröße: Merkmalerhaltende Darstellung ↔ Anzahl der benötigten Datenpunkte



³ Keogh, Eamonn, et al. "Dimensionality reduction for fast similarity search in large time series databases." *Knowledge and information Systems* 3 (2001): 263-286.

Tradeoff der Approximation

Abhängig von Fenstergröße: Genauigkeit der Approximation vs. Anzahl Datenpunkte



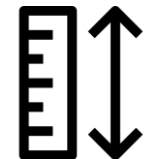
Fenstergröße: 0.5s



Viele Merkmale bleiben erhalten



Repräsentation durch viele Datenpunkten



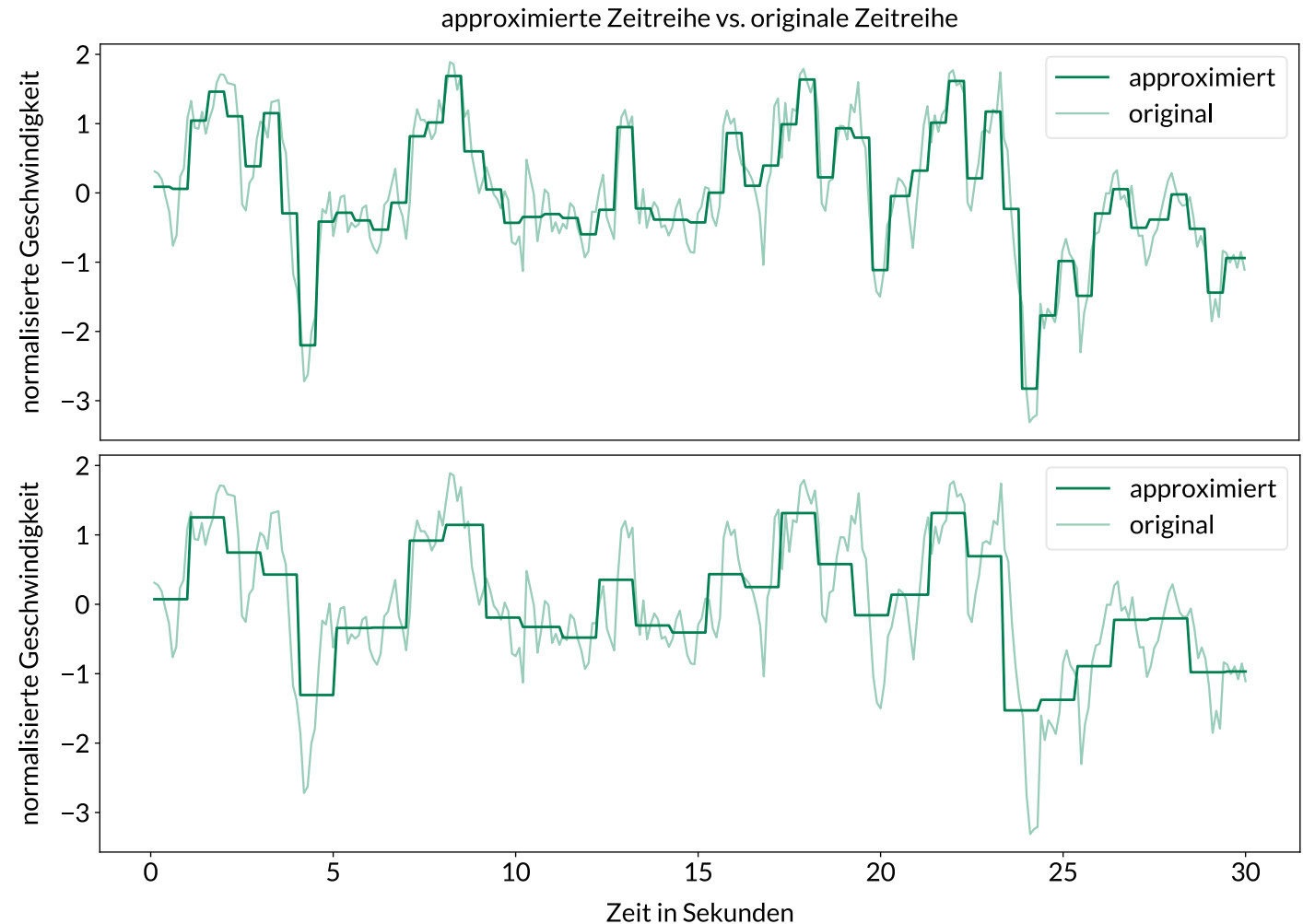
Fenstergröße: 1s



Repräsentation durch wenig Datenpunkten



Wenig Merkmale bleiben erhalten



Diskretisierung basierend auf Approximation

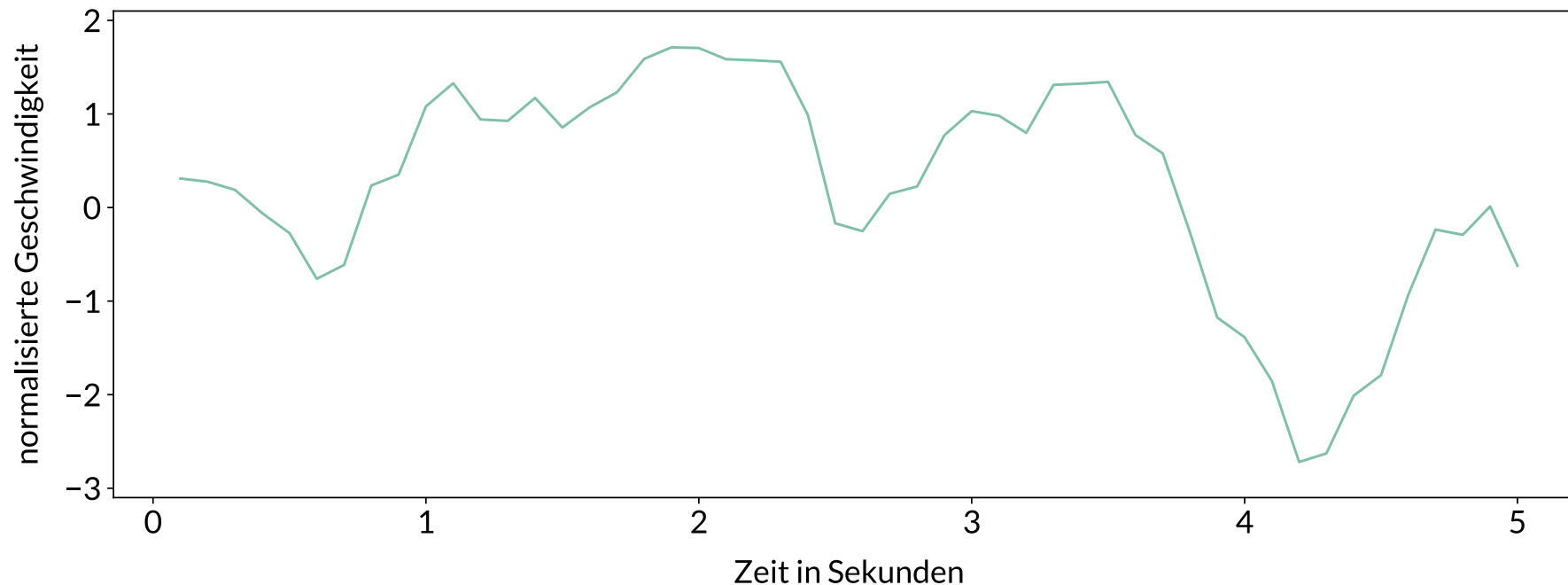
Zuordnung der Mittelwerte zu einem Intervall entlang der Amplitude

Ansatz Literatur

Symbolic Aggregate Approximation (SAX)⁴

Annahme: Normalisierte Zeitreihe folgt Standardnormalverteilung

Intervallgrenzen sind Quantile der Standardnormalverteilung



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

Diskretisierung basierend auf Approximation

Zuordnung der Mittelwerte zu einem Intervall entlang der Amplitude

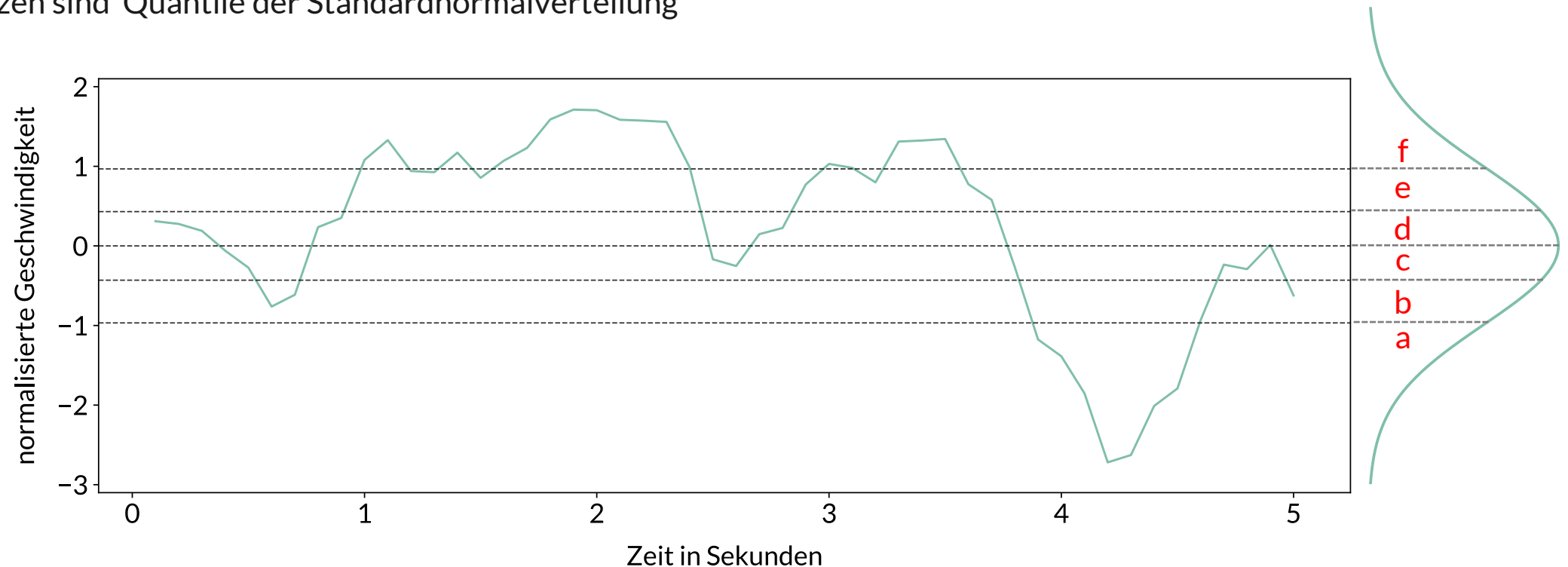
Ansatz Literatur

Symbolic Aggregate Approximation (SAX)⁴

Annahme: Normalisierte Zeitreihe folgt Standardnormalverteilung

Anzahl Intervalle (Alphabetgröße) als Parameter

Intervallgrenzen sind Quantile der Standardnormalverteilung



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

Diskretisierung basierend auf Approximation

Zuordnung der Mittelwerte zu einem Intervall entlang der Amplitude

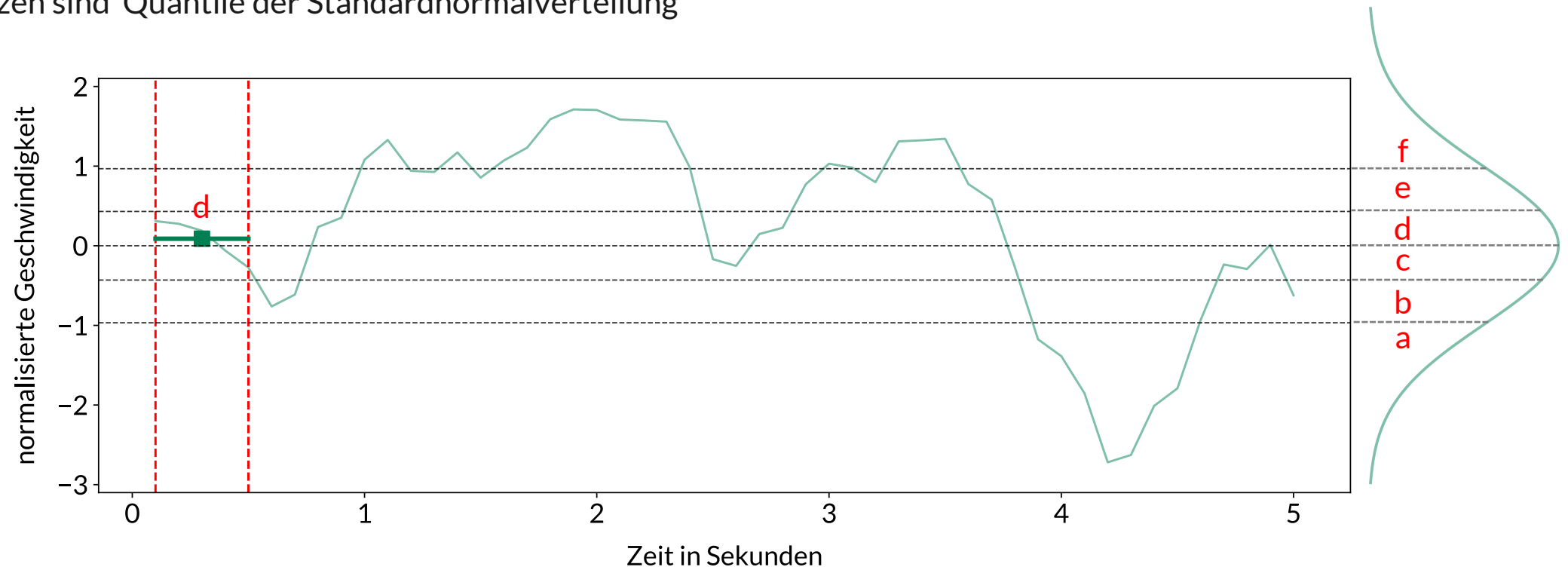
Ansatz Literatur

Symbolic Aggregate Approximation (SAX)⁴

Annahme: Normalisierte Zeitreihe folgt Standardnormalverteilung

Anzahl Intervalle (Alphabetgröße) als Parameter

Intervallgrenzen sind Quantile der Standardnormalverteilung



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

Diskretisierung basierend auf Approximation

Zuordnung der Mittelwerte zu einem Intervall entlang der Amplitude

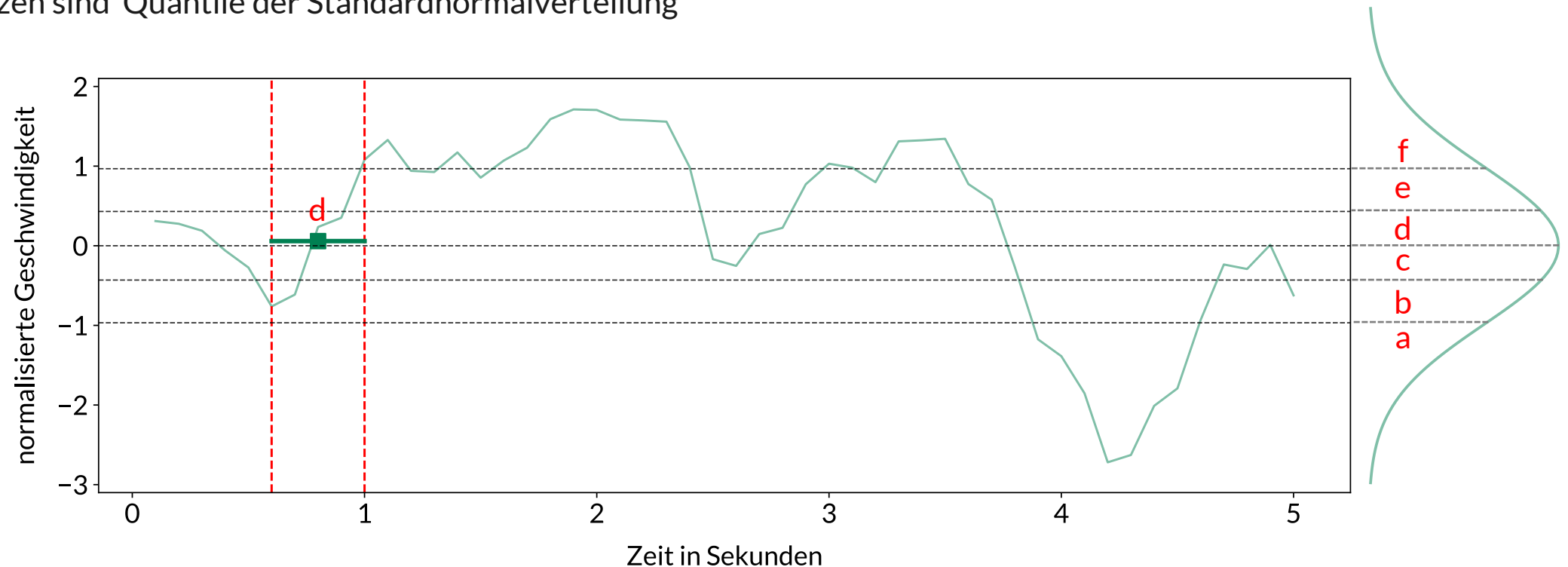
Ansatz Literatur

Symbolic Aggregate Approximation (SAX)⁴

Annahme: Normalisierte Zeitreihe folgt Standardnormalverteilung

Anzahl Intervalle (Alphabetgröße) als Parameter

Intervallgrenzen sind Quantile der Standardnormalverteilung



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

Diskretisierung basierend auf Approximation

Zuordnung der Mittelwerte zu einem Intervall entlang der Amplitude

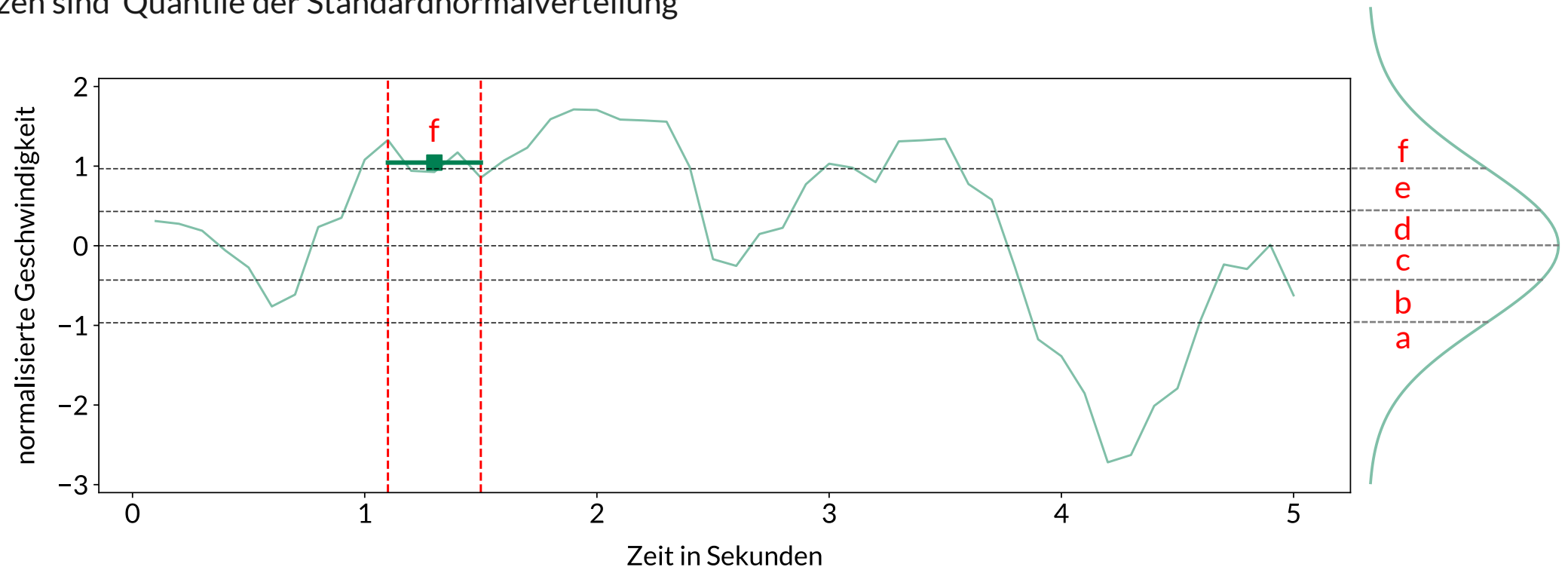
Ansatz Literatur

Symbolic Aggregate Approximation (SAX)⁴

Annahme: Normalisierte Zeitreihe folgt Standardnormalverteilung

Anzahl Intervalle (Alphabetgröße) als Parameter

Intervallgrenzen sind Quantile der Standardnormalverteilung



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

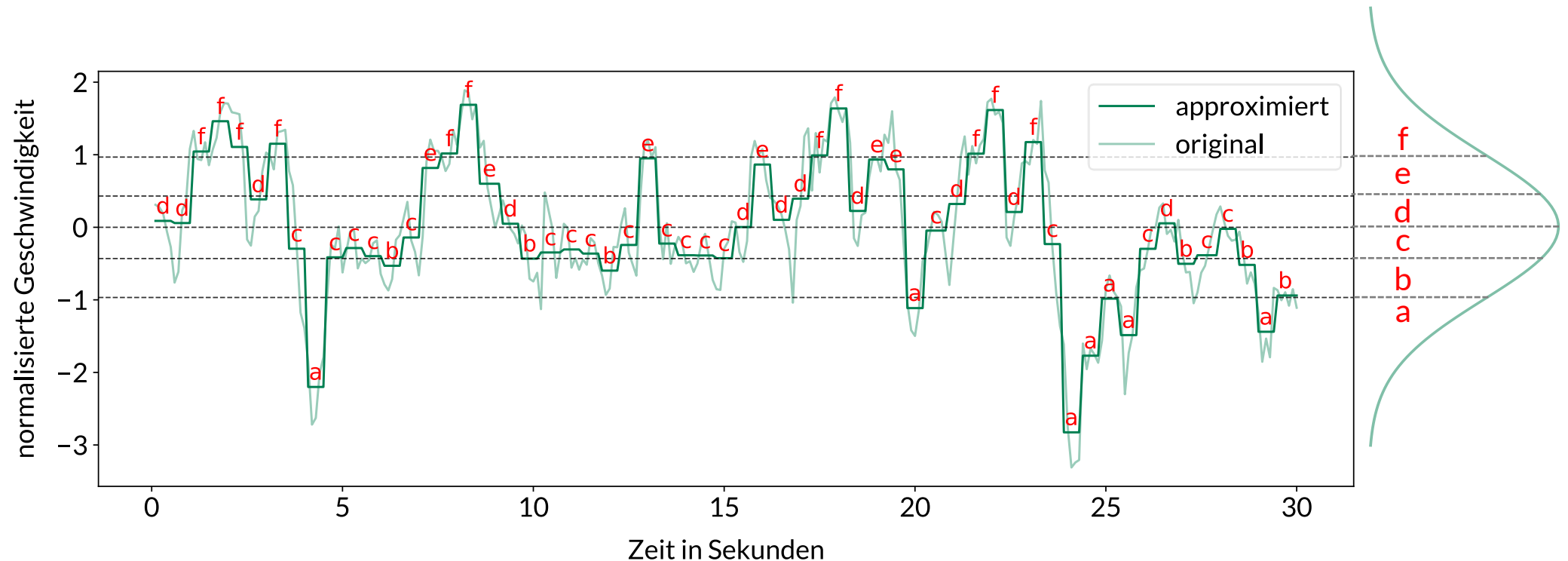
Diskretisierung basierend auf Approximation

Zuordnung der Mittelwerte zu einem Intervall entlang der Amplitude

Ansatz Literatur

Symbolic Aggregate Approximation (SAX)⁴

Diskretisierte Zeitreihe: **ddff fdff accc bcef fedb cccb cecc ccde ddff deea cdff dfca aaac dbcc bab**



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

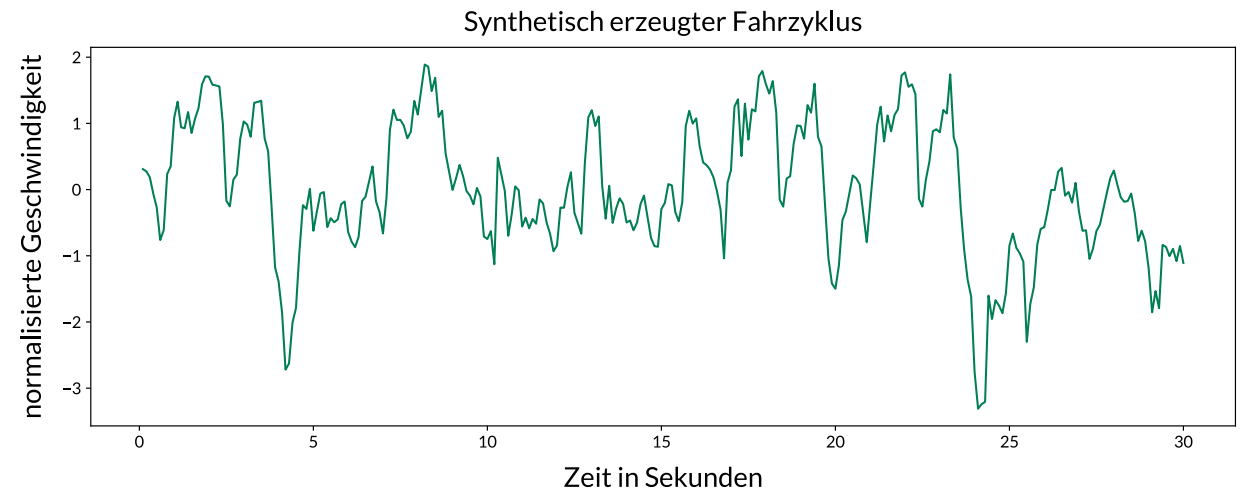
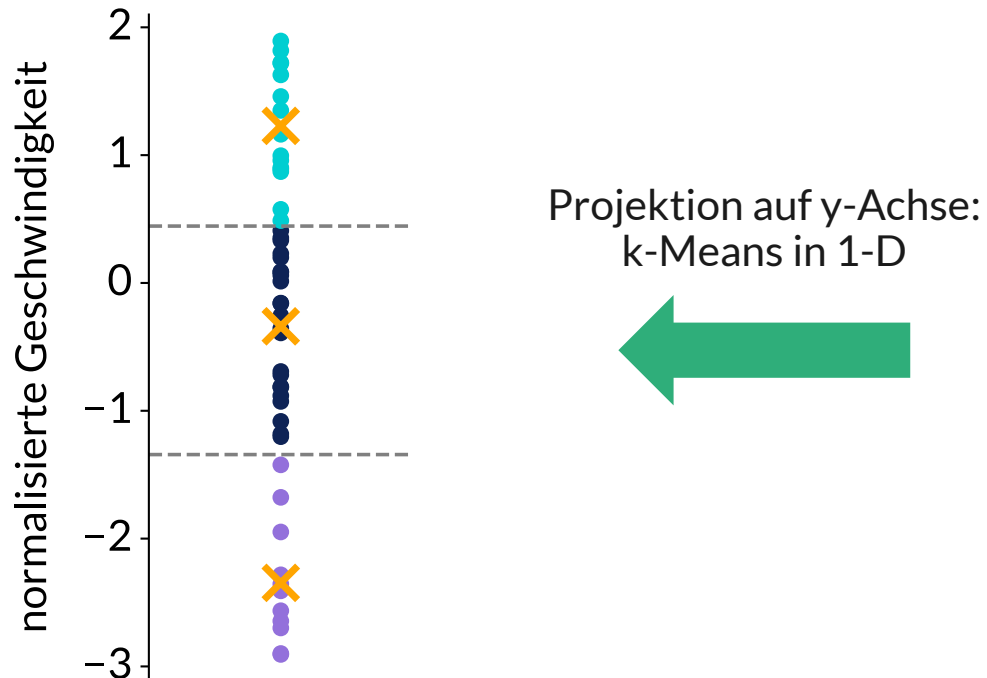
Diskretisierung basierend auf Approximation

Aufhebung der Gauß-Annahme: Intervallgrenzen werden individuell für Zeitreihe bestimmt

Ansatz Literatur

Adaptive Symbolic Aggregate Approximation (aSAX)⁵

Diskretisierung analog zu SAX, aber Intervallgrenzen sind Mittelpunkte zweier Cluster-Schwerpunkte aus k-Means



⁵ Pham, Ninh D., Quang Loc Le, and Tran Khanh Dang. "Two novel adaptive symbolic representations for similarity search in time series databases." *2010 12th International Asia-Pacific Web Conference*. IEEE, 2010.

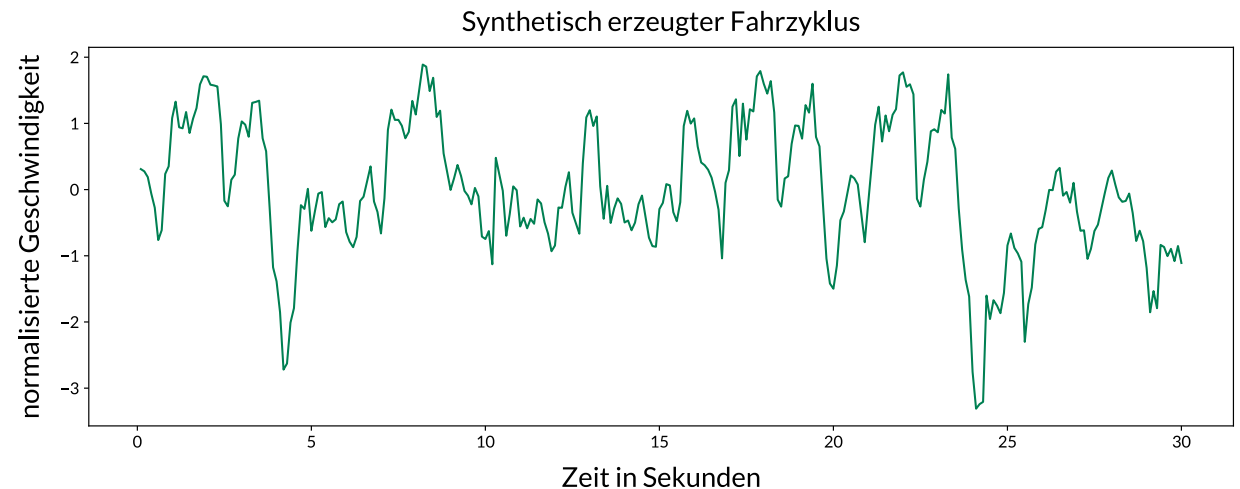
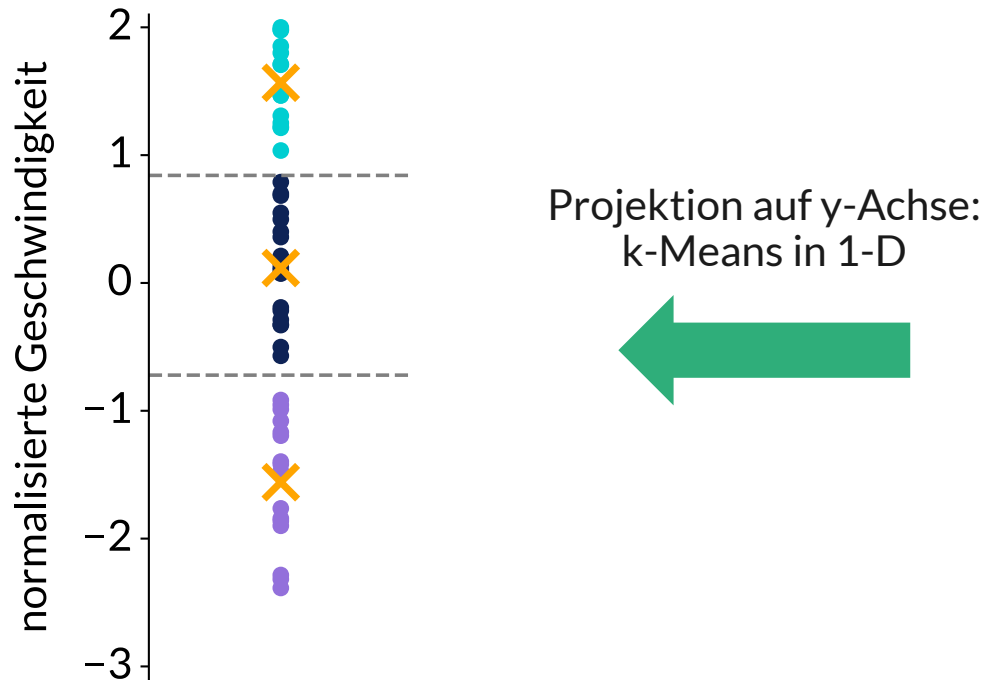
Diskretisierung basierend auf Approximation

Aufhebung der Gauß-Annahme: Intervallgrenzen werden individuell für Zeitreihe bestimmt

Ansatz Literatur

Adaptive Symbolic Aggregate Approximation (aSAX)⁵

Diskretisierung analog zu SAX, aber Intervallgrenzen sind Mittelpunkte zweier Cluster-Schwerpunkte aus k-Means



⁵ Pham, Ninh D., Quang Loc Le, and Tran Khanh Dang. "Two novel adaptive symbolic representations for similarity search in time series databases." *2010 12th International Asia-Pacific Web Conference*. IEEE, 2010.

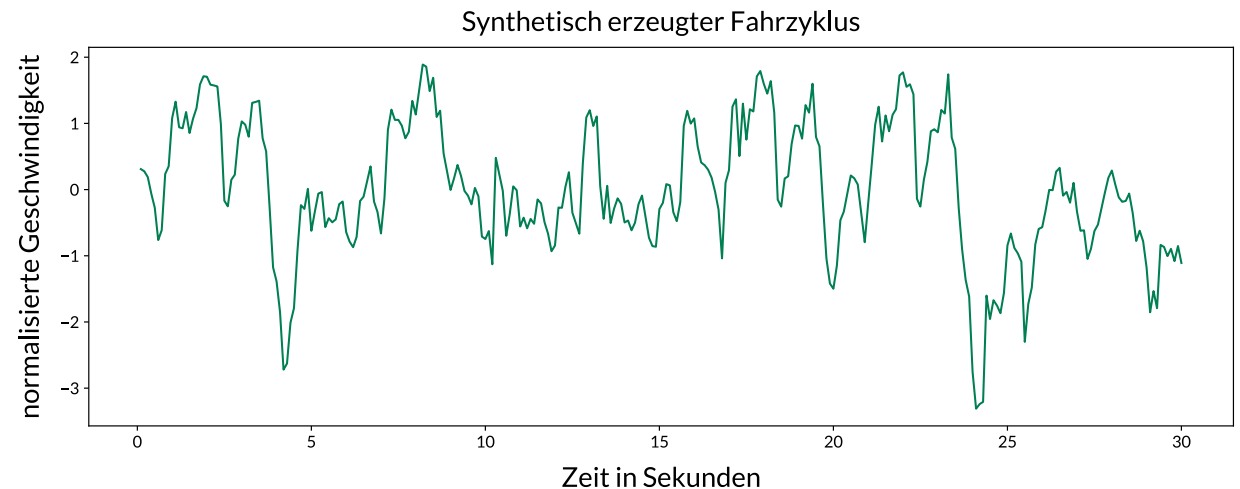
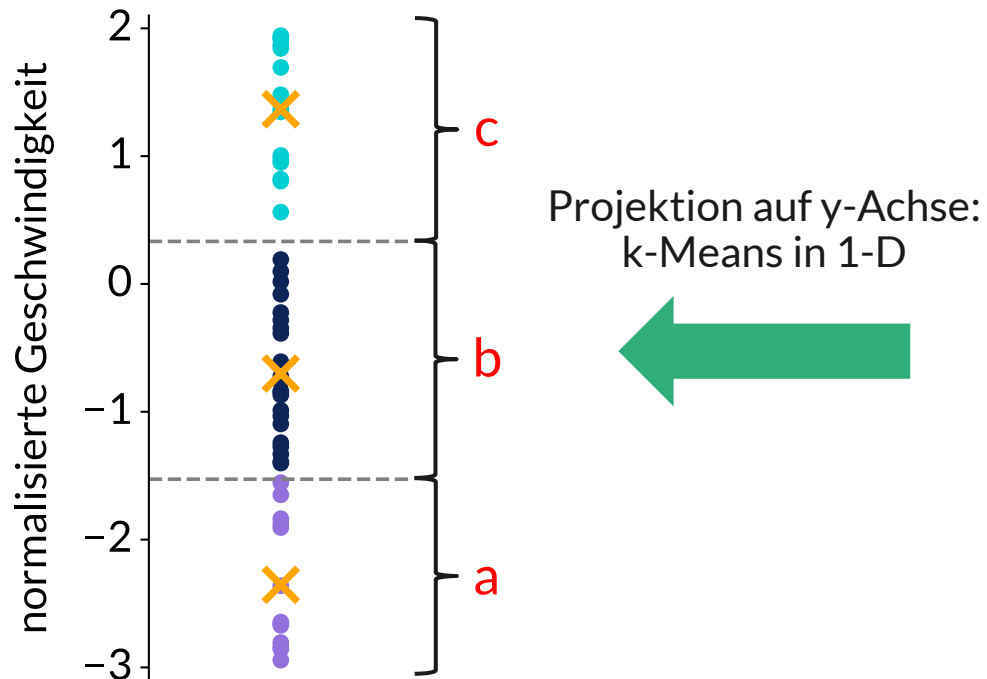
Diskretisierung basierend auf Approximation

Aufhebung der Gauß-Annahme: Intervallgrenzen werden individuell für Zeitreihe bestimmt

Ansatz Literatur

Adaptive Symbolic Aggregate Approximation (aSAX)⁵

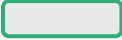

Diskretisierung analog zu SAX, aber Intervallgrenzen sind Mittelpunkte zweier Cluster-Schwerpunkte aus k-Means



⁵ Pham, Ninh D., Quang Loc Le, and Tran Khanh Dang. "Two novel adaptive symbolic representations for similarity search in time series databases." *2010 12th International Asia-Pacific Web Conference*. IEEE, 2010.

Zusammenfassung: Diskretisierungsansätze

Klassifizierende Zusammenfassung der evaluierten Diskretisierungsansätze

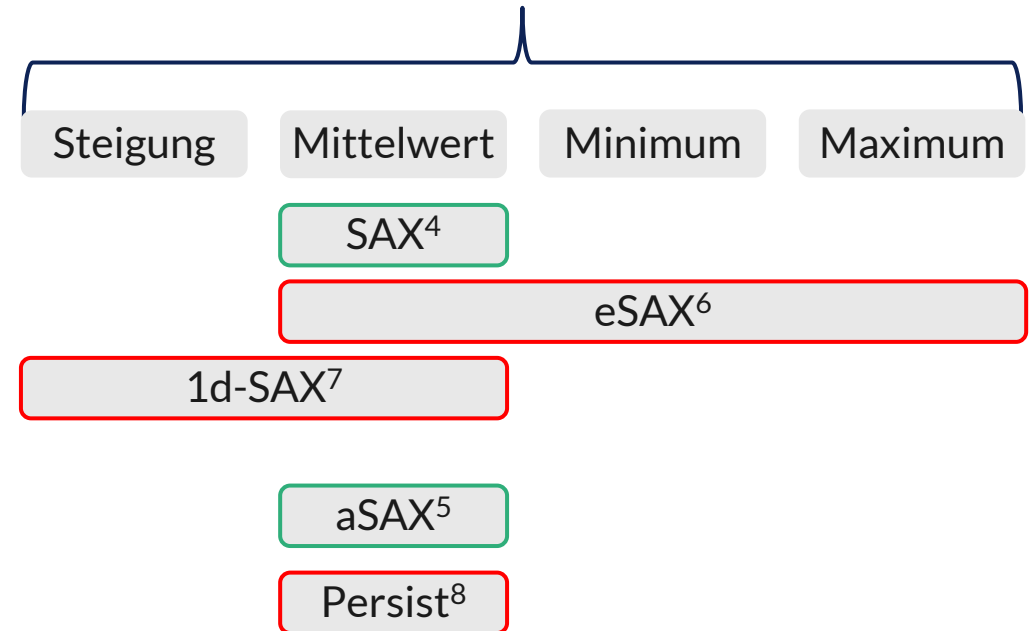
-  wurde zuvor präsentiert
-  wurde zuvor nicht präsentiert

Berechnung der
Diskretisierungsintervalle

feste Intervallgrenzen
basierend auf Gauß-Verteilung

individuelle/adaptive
Intervallgrenzen

Diskretisierung anhand von ... innerhalb Sliding-Window



⁴ Lin, Jessica, et al. "Experiencing SAX: a novel symbolic representation of time series." *Data Mining and knowledge discovery* 15 (2007): 107-144.

⁵ Pham, Ninh D., Quang Loc Le, and Tran Khanh Dang. "Two novel adaptive symbolic representations for similarity search in time series databases." *2010 12th International Asia-Pacific Web Conference*. IEEE, 2010.

⁶ Lkhagva, Battuguldur, Yu Suzuki, and Kyoji Kawagoe. "New time series data representation ESAX for financial applications." *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. IEEE, 2006.

⁷ Malinowski, Simon, et al. "1d-sax: A novel symbolic representation for time series." *Advances in Intelligent Data Analysis XII: 12th International Symposium, IDA 2013, London, UK, October 17-19, 2013. Proceedings 12*.

⁸ Mörchen, Fabian, and Alfred Ultsch. "Optimizing time series discretization for knowledge discovery." *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. 2005.

Überblick: Evaluationsmaße

Überblick über ausgewertete Evaluationsmaße im Rahmen der Bachelorarbeit



wird im Folgenden präsentiert



wird im Folgenden aus Zeitgründen nicht präsentiert

Rekonstruktionsfehler⁹

Güte Motif Discovery¹⁰

Speicherbedarf⁹

Zeitkomplexität¹¹

⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." *2011 4th international congress on image and signal processing*. Vol. 4. IEEE, 2011.

¹⁰ Noering, Fabian Kai Dietrich. *Unsupervised Pattern Discovery in Automotive Time Series: Pattern-based Construction of Representative Driving Cycles*. Vol. 159. Springer Nature, 2022.

¹¹ Garcia, Salvador, et al. "A survey of discretization techniques: Taxonomy and empirical analysis in supervised learning." *IEEE transactions on Knowledge and Data Engineering* 25.4 (2012): 734-750.

Beurteilung der Diskretisierungsansätze

Abbildung der Symbole auf Zahlen ist notwendig für Rekonstruktion der Zeitreihe

Evaluation

Rekonstruktionsfehler⁹

Diskretisierte Zeitreihe: **ddff fdfc accc bcef fedb cccb cecc ccde ddff deea cdff dfca aaac dbcc bab**

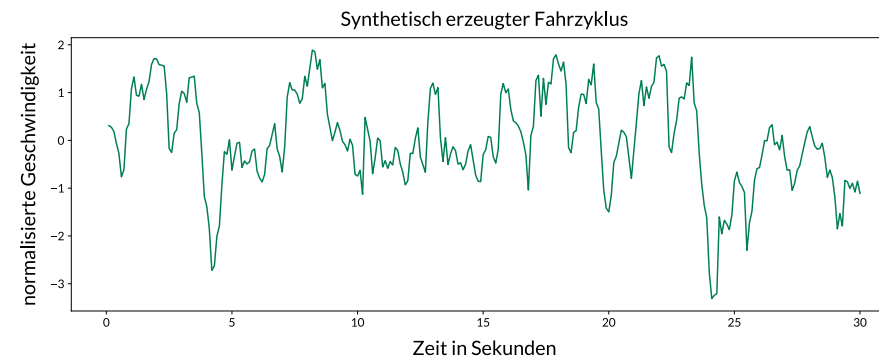


Finde Abbildung der Symbole auf Zahlen

Symbolwert: {a, b, c, d, e, f} -> Zahlen



Rekonstruiere originale Zeitreihe



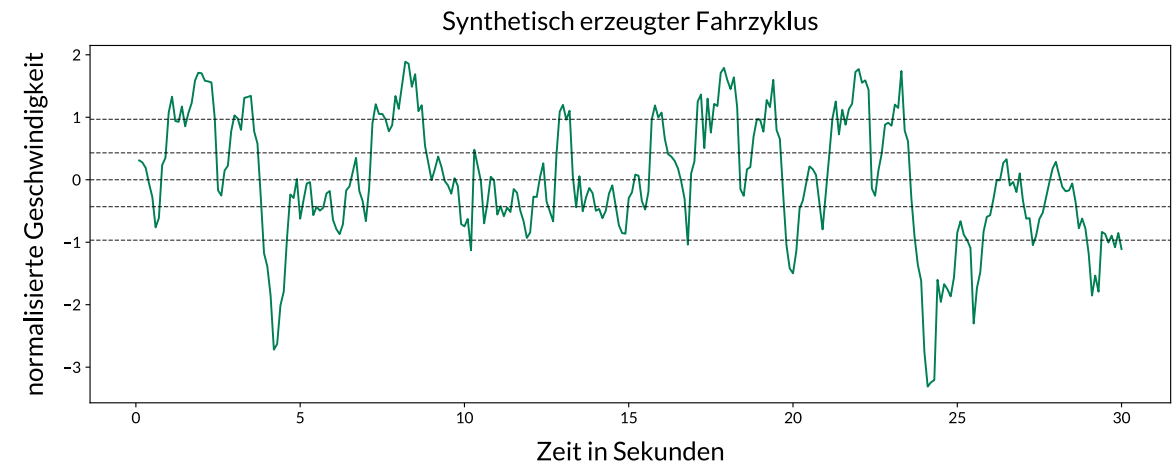
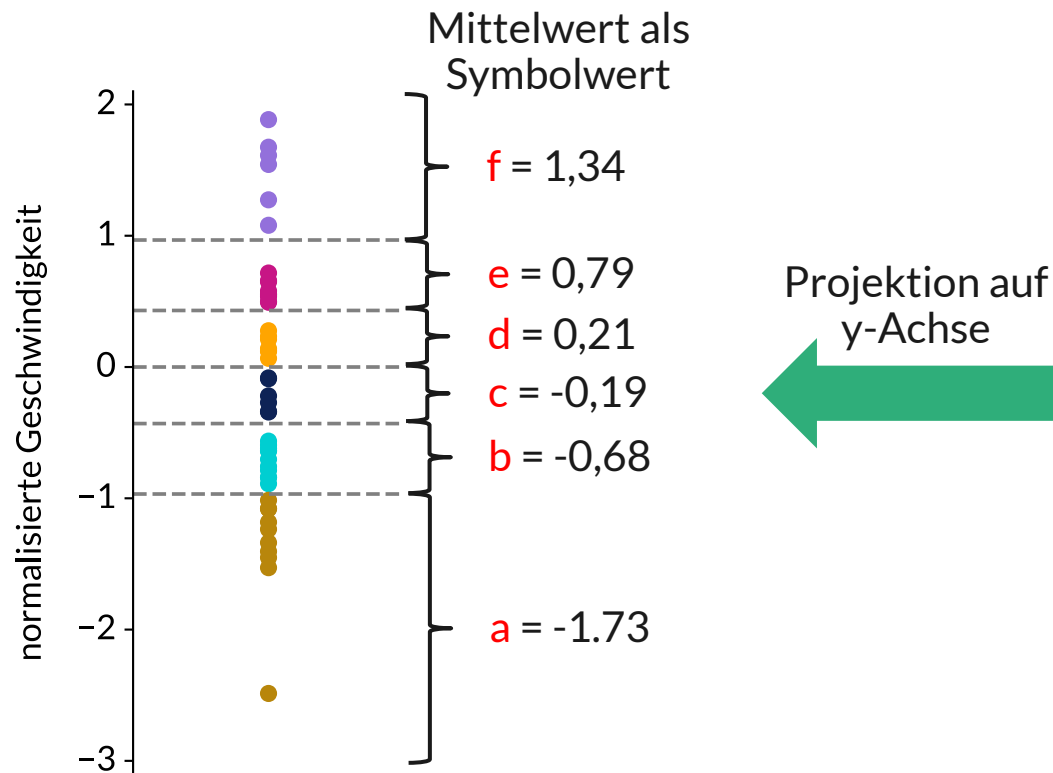
⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." 2011 4th international congress on image and signal processing. Vol. 4. IEEE, 2011.

Beurteilung der Diskretisierungsansätze

Abbildung durch Mittelwert der Punkte der originalen Zeitreihe innerhalb eines Intervalls

Evaluation

Rekonstruktionsfehler⁹



⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." 2011 4th international congress on image and signal processing. Vol. 4. IEEE, 2011.

Beurteilung der Diskretisierungsansätze

Alle Punkte innerhalb eines Fensters bekommen denselben entsprechenden Symbolwert

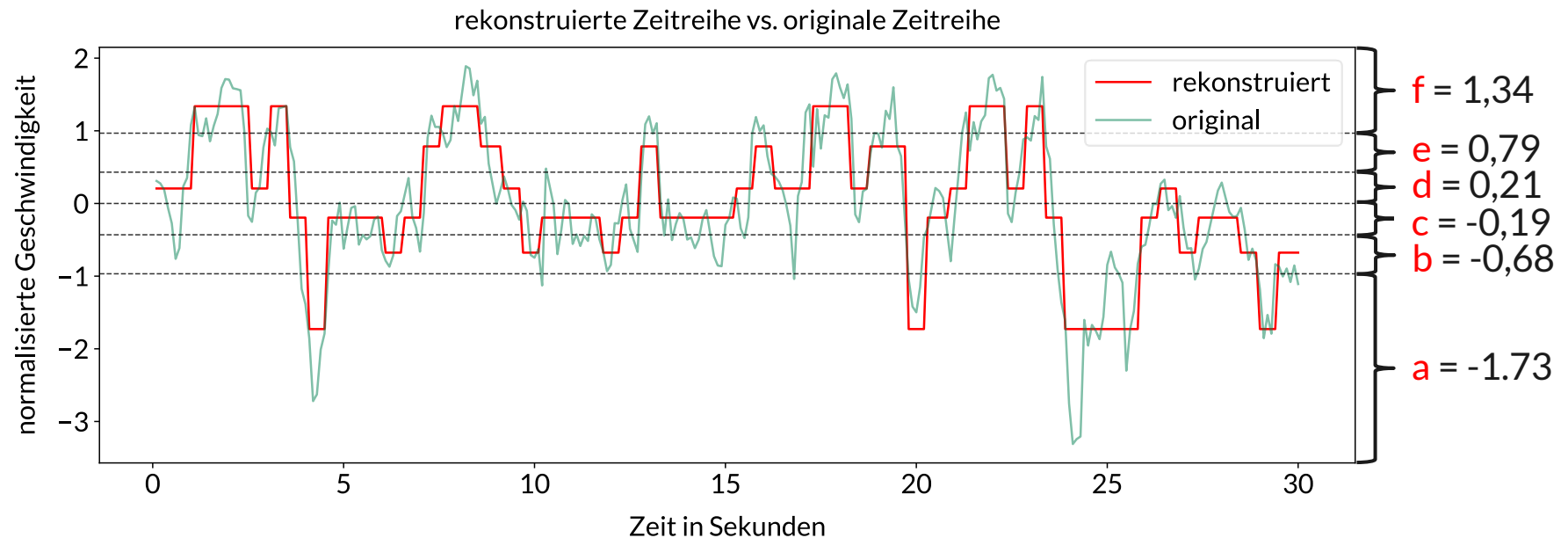
Evaluation

Rekonstruktionsfehler⁹

Diskretisierte Zeitreihe: **ddff fdfc accc bcef fedb cccb cecc ccde ddff deea cdff dfca aaac dbcc bab**



Rekonstruktion der originalen Zeitreihe



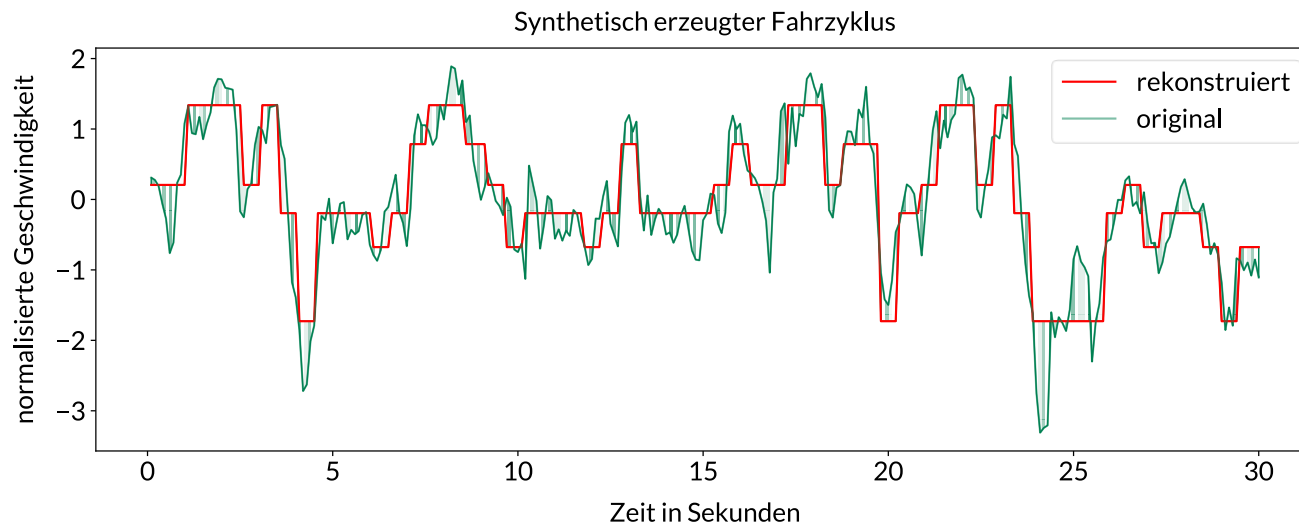
⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." 2011 4th international congress on image and signal processing. Vol. 4. IEEE, 2011.

Beurteilung der Diskretisierungsansätze

Rekonstruktionsfehler wird mittels Punkt-zu-Punkt Distanzen gemessen

Evaluation

Rekonstruktionsfehler⁹



$$\text{Mean Absolute Error (MAE)} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

$$\text{Mean Squared Error (MSE)} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$$

⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." *2011 4th international congress on image and signal processing*. Vol. 4. IEEE, 2011.

Beurteilung der Diskretisierungsansätze

Diskretisierung mit eSAX und 1d-SAX führt zu kleinstem Rekonstruktionsfehler

Evaluation

Rekonstruktionsfehler⁹

eSAX

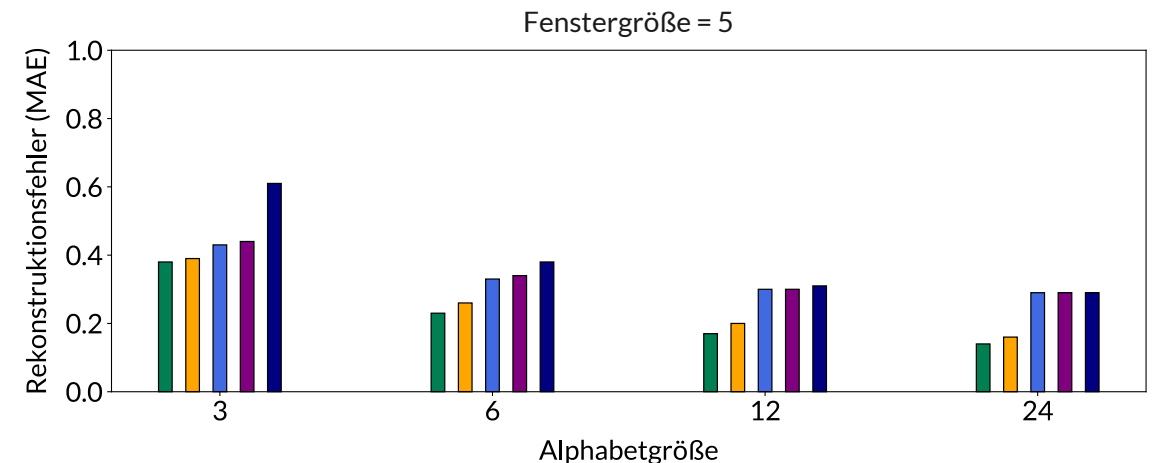
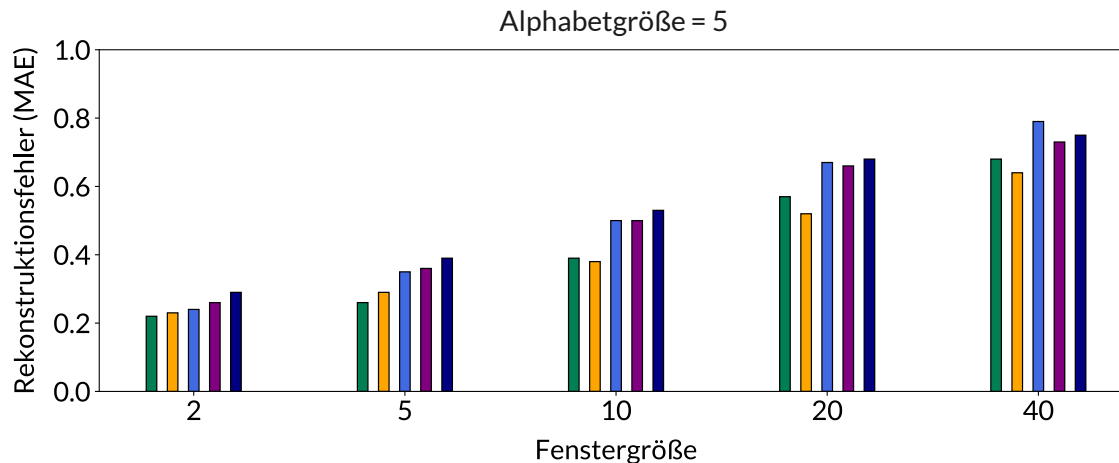
1d-SAX

aSAX

SAX

Persist

4 Datensätze ausgewertet, hier mittlerer MAE für Datensatz mit 700 Zeitreihen mit je 100 bis 3600 Punkten



1d-SAX ist bei größeren Fenstergrößen besser als eSAX
aSAX ist für kleinere Fenstergrößen leicht besser als SAX
Persist schneidet nur auf einem Datensatz so wie aSAX ab

Ergebnisse bzgl. variierender Fenstergröße bleiben erhalten
Konvergenz des Rekonstruktionsfehlers
SAX, aSAX, Persist konvergieren zu selbem Wert

⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." 2011 4th international congress on image and signal processing. Vol. 4. IEEE, 2011.

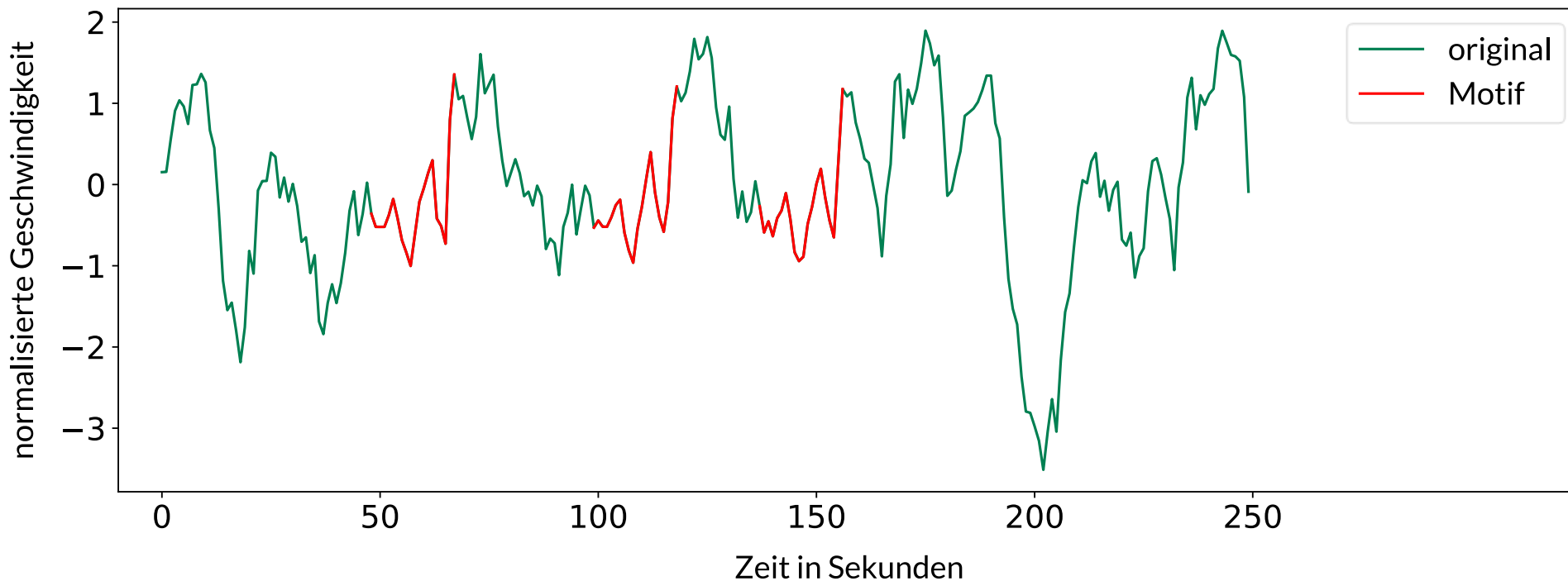
Beurteilung der Diskretisierungsansätze

Ziel Motif Discovery: Finde alle wiederkehrenden Subsequences innerhalb einer Zeitreihe

Evaluation

Motif Discovery¹⁰

Ähnlichkeit von Subsequences ist relativ zu gegebenem Distanzmaß (oft Minkowski-Distanz) und gegebenem Threshold



¹⁰ Noering, Fabian Kai Dietrich. *Unsupervised Pattern Discovery in Automotive Time Series: Pattern-based Construction of Representative Driving Cycles*. Vol. 159. Springer Nature, 2022.

Beurteilung der Diskretisierungsansätze

Messung der Güte von Motif Discovery basierend auf Recall und Precision

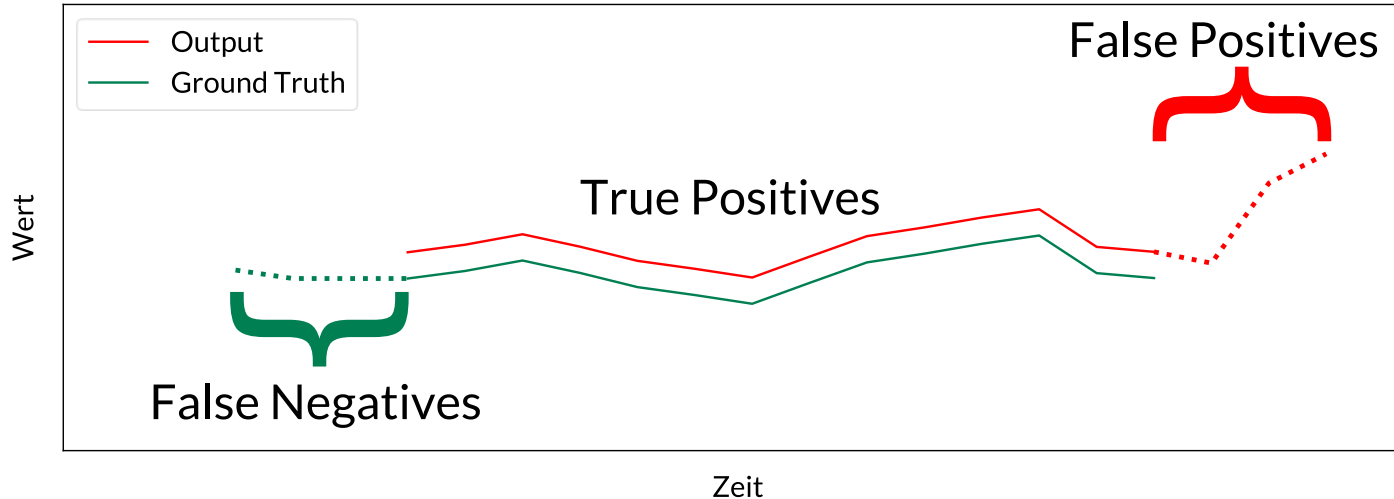
Evaluation

Motif Discovery¹⁰

Output Motif Discovery-Algorithmen: Pro Zeitreihe mehrere Motifs mit jeweils mehreren Subsequences

Messung für eine Zeitreihe: jeweils Mittelwert über Motif Recall und Motif Precision aller Motifs

Messung für einen Datensatz: jeweils Mittelwert über mittleren Motif Recall und mittlere Motif Precision aller Zeitreihen



$$\text{Motif Recall}^{12} = \frac{TP}{TP + FN}$$

$$\text{Motif Precision}^{12} = \frac{TP}{TP + FP}$$

¹⁰ Noering, Fabian Kai Dietrich. *Unsupervised Pattern Discovery in Automotive Time Series: Pattern-based Construction of Representative Driving Cycles*. Vol. 159. Springer Nature, 2022.

¹² Powers, David MW. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." *arXiv preprint arXiv:2010.16061* (2020).

Beurteilung der Diskretisierungsansätze

Diskretisierung mit eSAX ist ungeeignet für Motif Discovery

Evaluation

Motif Discovery¹⁰

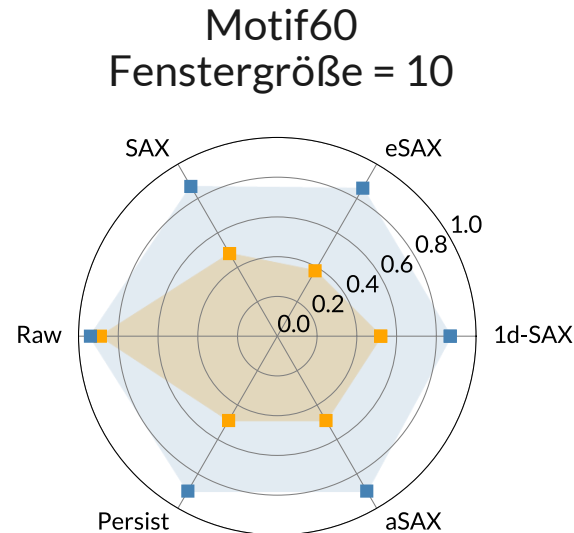
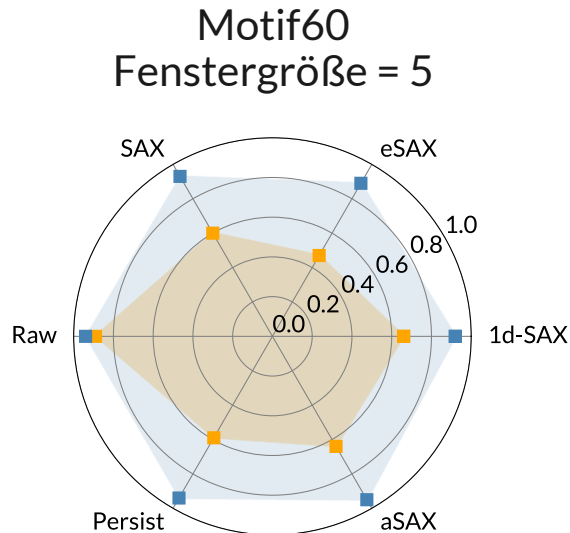
■ Motif Precision

■ Motif Recall

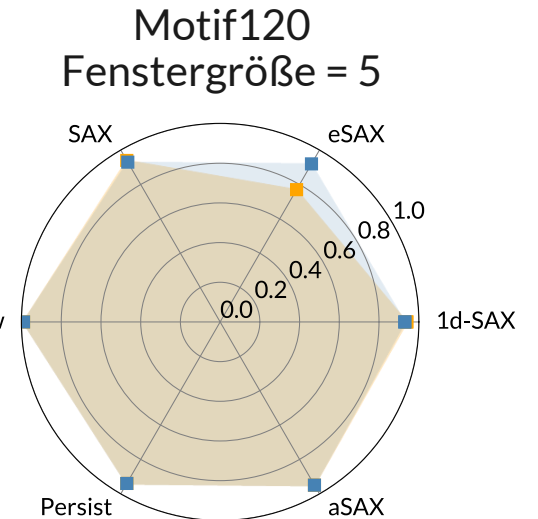
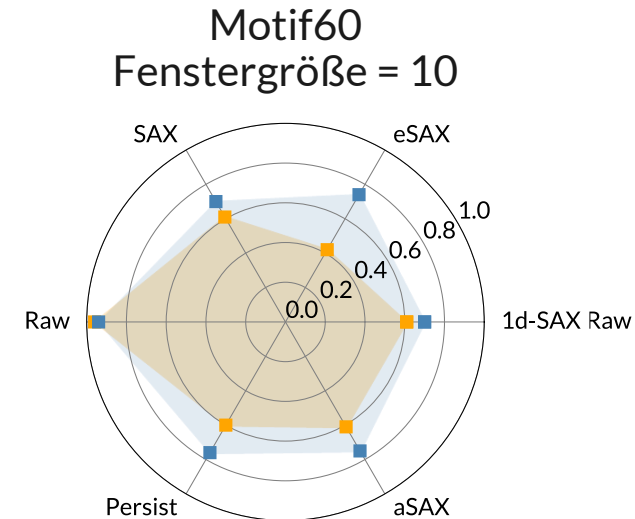
Motif60: Datensatz mit 40 Zeitreihen mit je 6.400 bis 8.200 Punkten und beliebig verteilten Motifs mit je 60 Punkten

Motif120: Datensatz mit 40 Zeitreihen mit je 11.200 bis 14.000 Punkten und beliebig verteilten Motifs mit je 120 Punkten

Brute Force Algorithmus:



Matrix Profile Algorithmus:



¹⁰ Noering, Fabian Kai Dietrich. *Unsupervised Pattern Discovery in Automotive Time Series: Pattern-based Construction of Representative Driving Cycles*. Vol. 159. Springer Nature, 2022.

Beurteilung der Diskretisierungsansätze

Diskretisierungsansätze mit mehreren Symbolen pro Fenster haben höheren Speicherbedarf

Evaluation

Speicherbedarf⁹

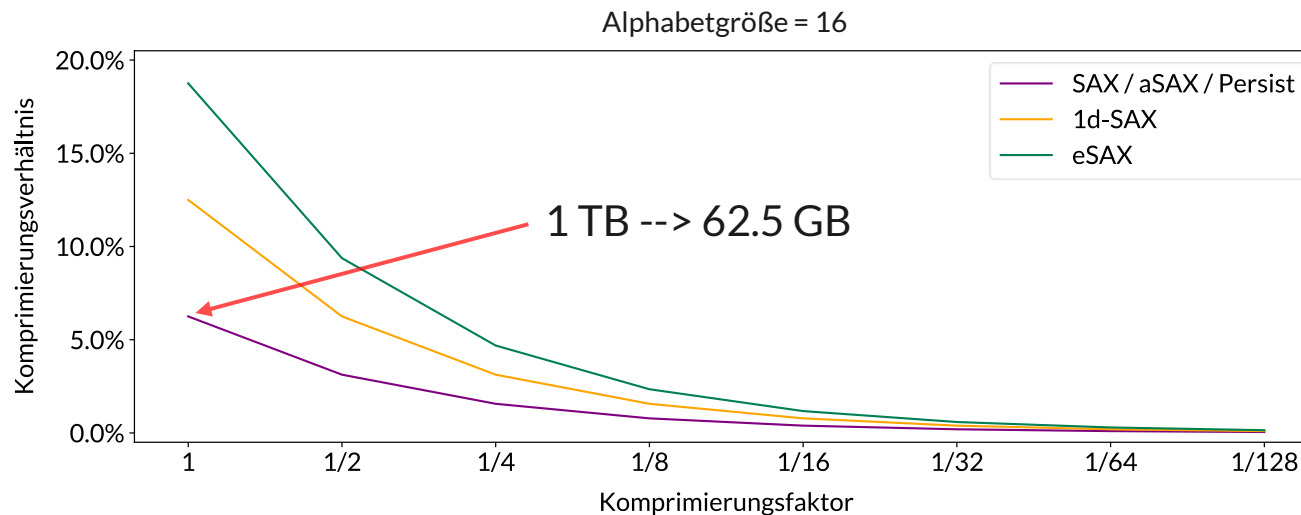
$$\text{Komprimierungsverhältnis} = \frac{\# \text{Bits diskretisierte Zeitreihe}}{\# \text{Bits originale Zeitreihe}} * 100$$

$$\text{Komprimierungsfaktor} = \frac{\# \text{Datenpunkte diskretisierte Zeitreihe}}{\# \text{Datenpunkte originale Zeitreihe}}$$

Annahmen:

64 Bits pro Datenpunkt (doppelte Genauigkeit)

$\lceil \log_2(\text{Alphabetgröße}) \rceil$ Bits pro Symbol



Beachte: Tradeoff der Approximation



Stärkere Komprimierung erhält weniger
Merkmale bei Diskretisierung

⁹ Sant'Anna, Anita, and Nicholas Wickström. "Symbolization of time-series: An evaluation of sax, persist, and aca." 2011 4th international congress on image and signal processing. Vol. 4. IEEE, 2011.

Zusammenfassender Vergleich

	Rekonstruktionsfehler	Motif Discovery	Speicherbedarf
eSAX	gering	nicht geeignet	hoch
1d-SAX	gering	sehr geeignet	mittel
aSAX	mittel	geeignet	gering
SAX	mittel	geeignet	gering
Persist	oft hoch	geeignet	gering

Anhang

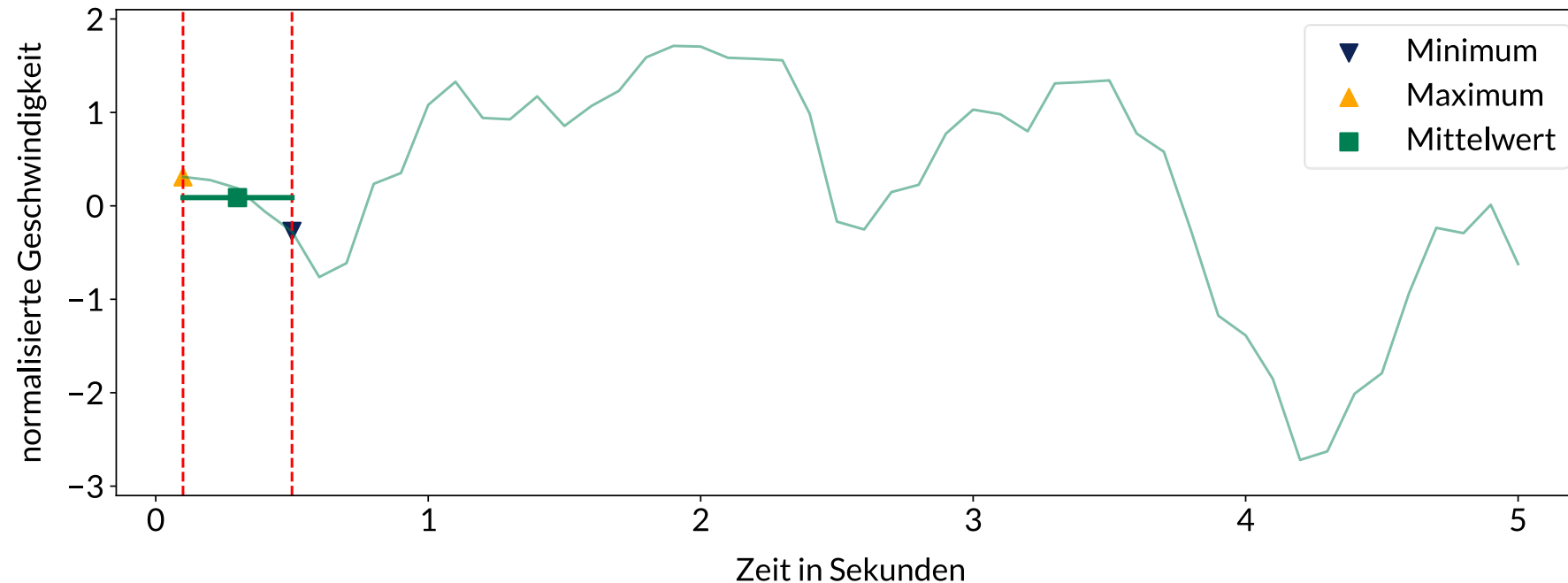
Im Folgenden Inhalte, die aus Zeitgründen nicht präsentiert werden konnten

Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Extrempunkte innerhalb eines Fensters

Ansatz Literatur

Extended Symbolic Aggregate Approximation (eSAX)⁶



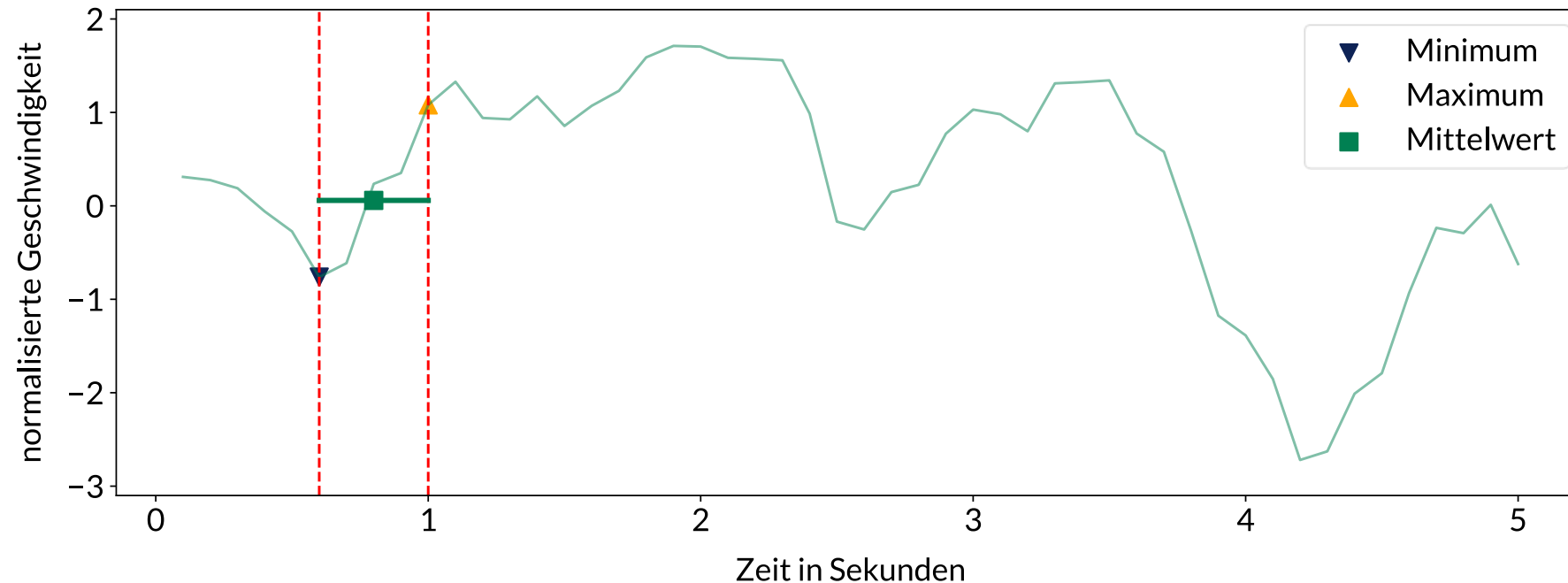
⁶ Lkhagva, Battuguldur, Yu Suzuki, and Kyoji Kawagoe. "New time series data representation ESAX for financial applications." *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. IEEE, 2006.

Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Extrempunkte innerhalb eines Fensters

Ansatz Literatur

Extended Symbolic Aggregate Approximation (eSAX)⁶



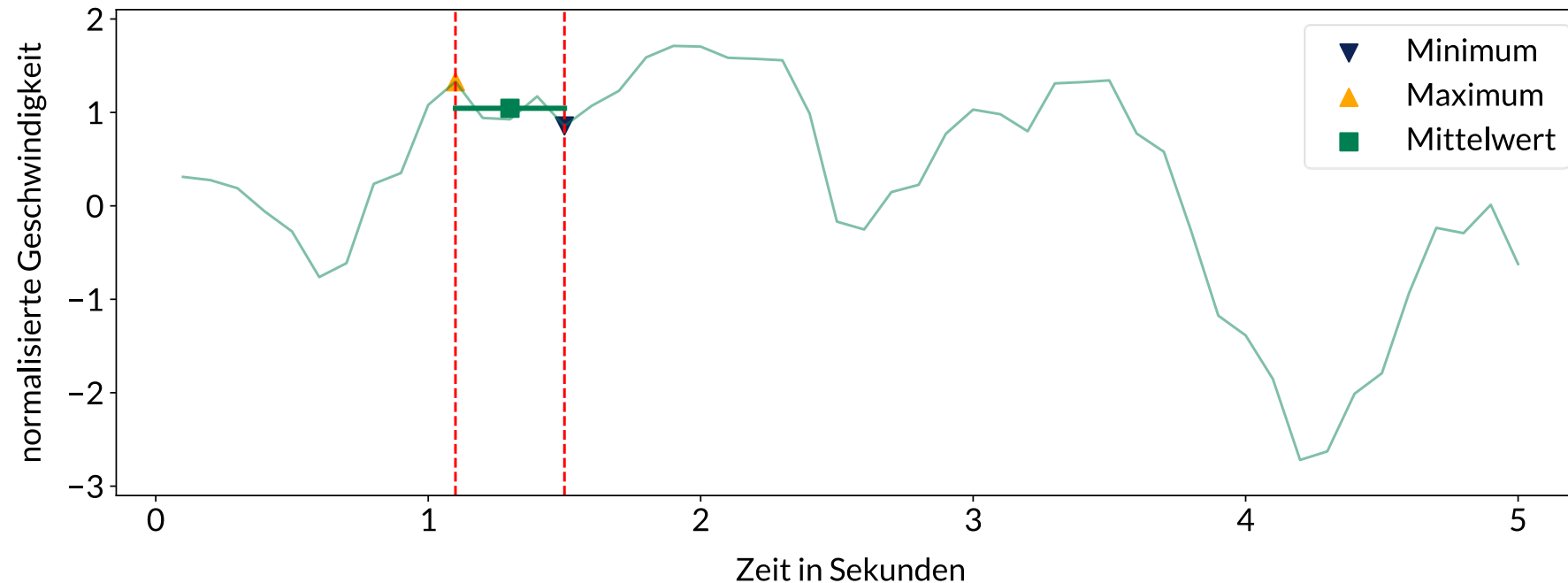
⁶ Lkhagva, Battuguldur, Yu Suzuki, and Kyoji Kawagoe. "New time series data representation ESAX for financial applications." *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. IEEE, 2006.

Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Extrempunkte innerhalb eines Fensters

Ansatz Literatur

Extended Symbolic Aggregate Approximation (eSAX)⁶



⁶ Lkhagva, Battuguldur, Yu Suzuki, and Kyoji Kawagoe. "New time series data representation ESAX for financial applications." *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. IEEE, 2006.

Diskretisierung basierend auf Approximation

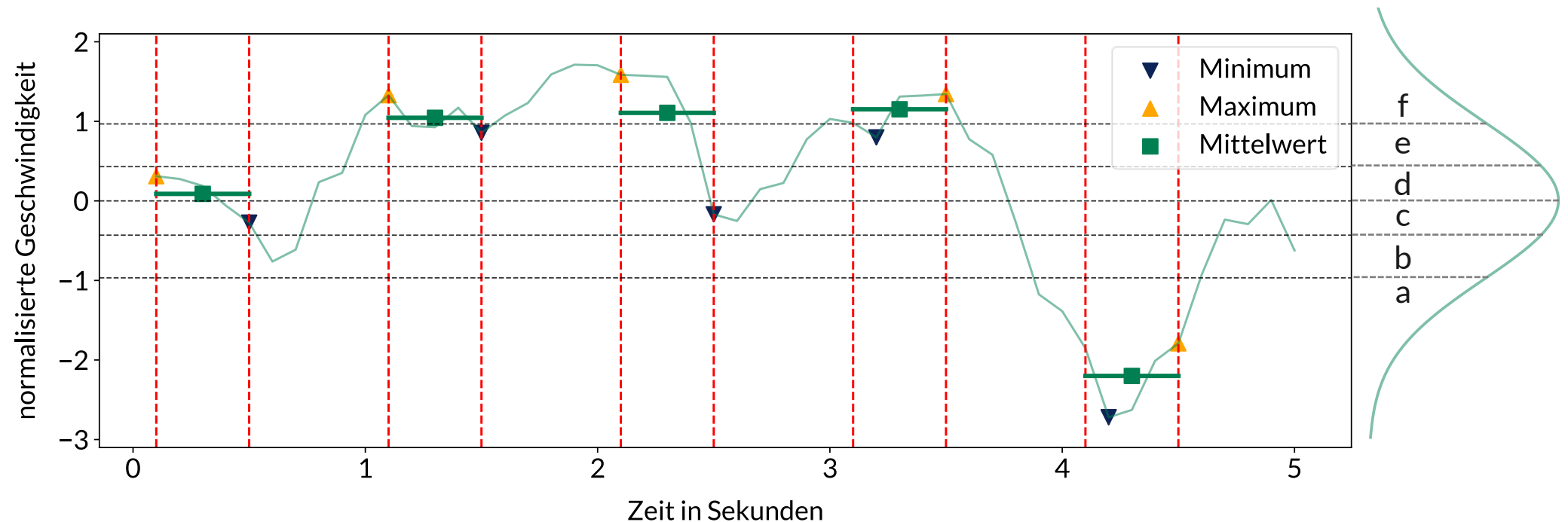
Diskretisierung des Mittelwerts und der Extrempunkte innerhalb eines Fensters

Ansatz Literatur

Extended Symbolic Aggregate Approximation (eSAX)⁶

Annahme der Standardnormalverteilung für normalisierte Zeitreihe bleibt erhalten

Diskretisierte Zeitreihe: ddc ... ffe ... ffc ... eff ... aaa ...



⁶ Lkhagva, Battuguldur, Yu Suzuki, and Kyoji Kawagoe. "New time series data representation ESAX for financial applications." *22nd International Conference on Data Engineering Workshops (ICDEW'06)*. IEEE, 2006.

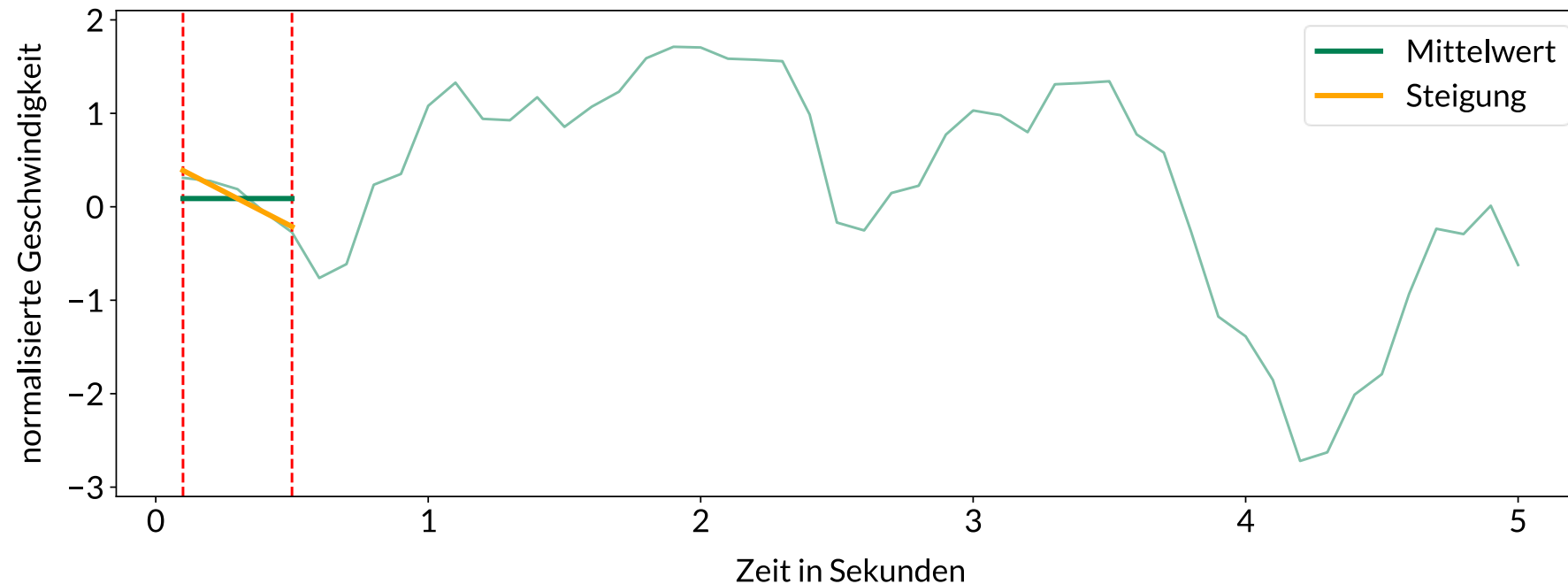
Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Steigung innerhalb eines Fensters

Ansatz Literatur

1d-Symbolic Aggregate Approximation (1d-SAX)⁷

Steigung wird über Lineare Regression mit den Punkten innerhalb eines Fensters berechnet



⁷ Malinowski, Simon, et al. "1d-sax: A novel symbolic representation for time series." *Advances in Intelligent Data Analysis XII: 12th International Symposium, IDA 2013, London, UK, October 17-19, 2013. Proceedings* 12.

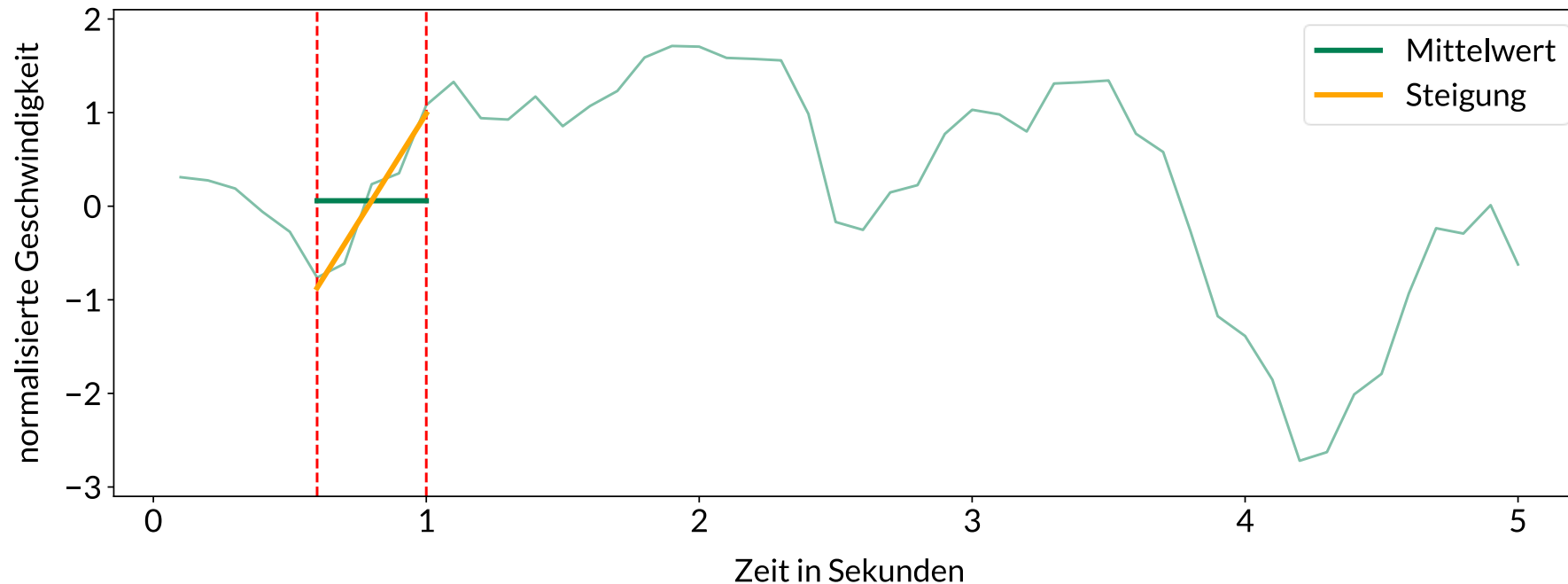
Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Steigung innerhalb eines Fensters

Ansatz Literatur

1d-Symbolic Aggregate Approximation (1d-SAX)⁷

Steigung wird über Lineare Regression mit den Punkten innerhalb eines Fensters berechnet



⁷ Malinowski, Simon, et al. "1d-sax: A novel symbolic representation for time series." *Advances in Intelligent Data Analysis XII: 12th International Symposium, IDA 2013, London, UK, October 17-19, 2013. Proceedings* 12.

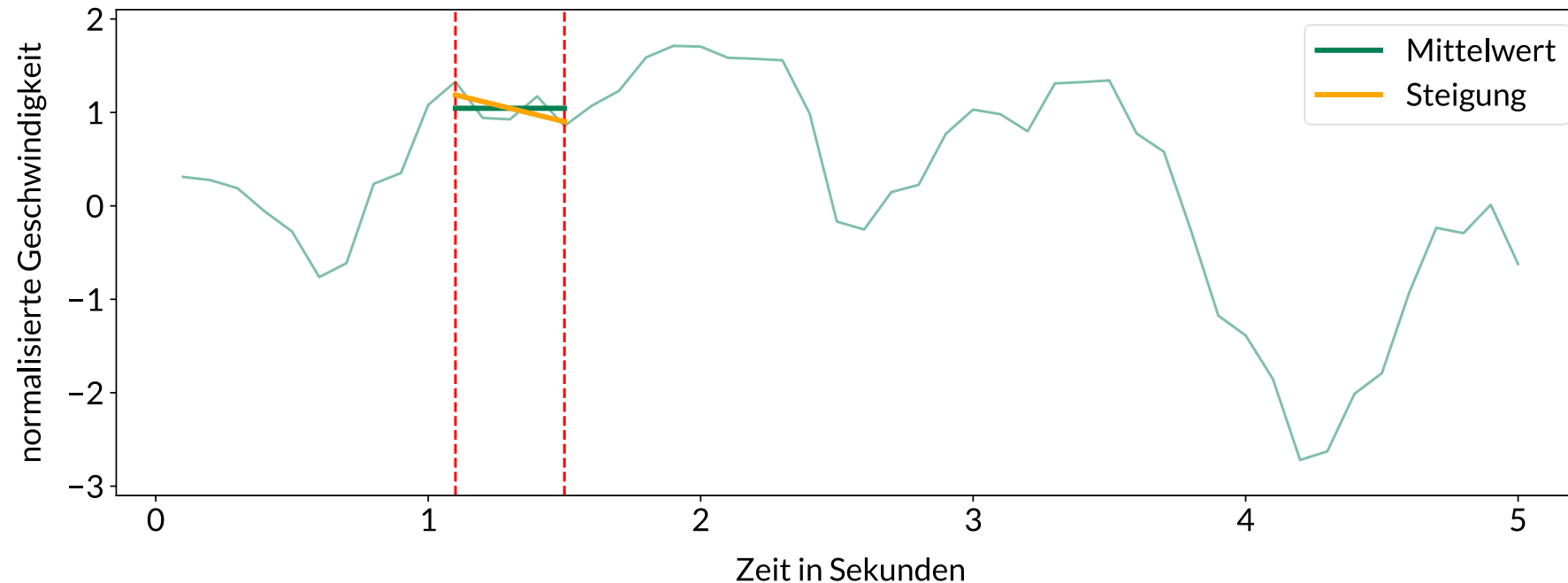
Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Steigung innerhalb eines Fensters

Ansatz Literatur

1d-Symbolic Aggregate Approximation (1d-SAX)⁷

Steigung wird über Lineare Regression mit den Punkten innerhalb eines Fensters berechnet



⁷ Malinowski, Simon, et al. "1d-sax: A novel symbolic representation for time series." *Advances in Intelligent Data Analysis XII: 12th International Symposium, IDA 2013, London, UK, October 17-19, 2013. Proceedings* 12.

Diskretisierung basierend auf Approximation

Diskretisierung des Mittelwerts und der Steigung innerhalb eines Fensters

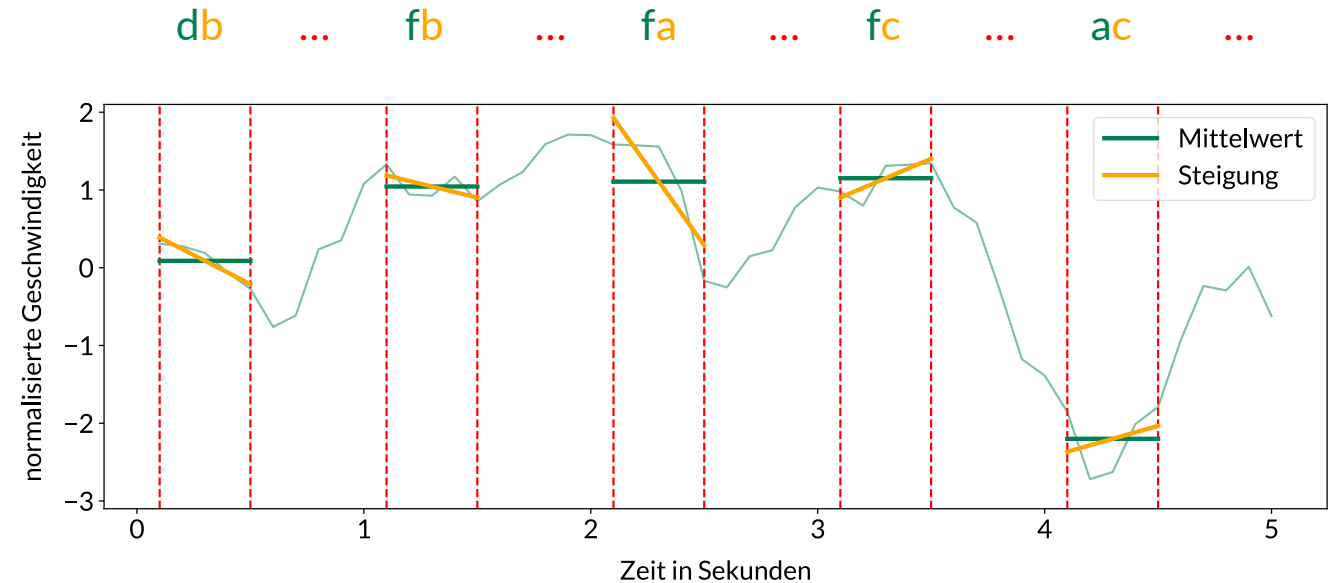
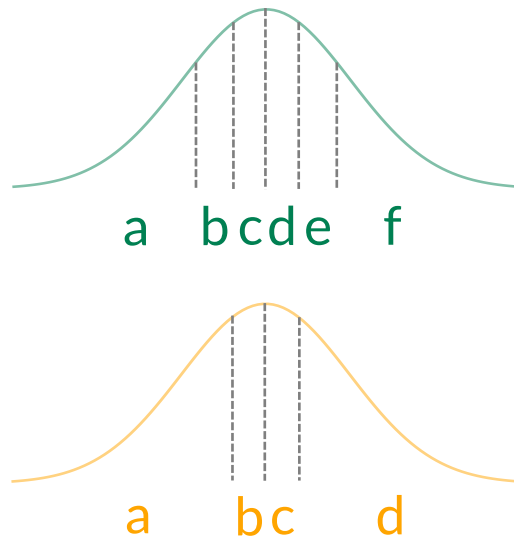
Ansatz Literatur

1d-Symbolic Aggregate Approximation (1d-SAX)⁷

Unterschiedliche Alphabetgrößen/Intervallanzahl

Diskretisierte Zeitreihe:

für Mittelwerte und Steigungen auf Basis
von Normalverteilungsannahmen möglich :



⁷ Malinowski, Simon, et al. "1d-sax: A novel symbolic representation for time series." *Advances in Intelligent Data Analysis XII: 12th International Symposium, IDA 2013, London, UK, October 17-19, 2013. Proceedings 12*.