servicenow

TACTiS

Transformer-**A**ttentional **C**opula for **Ti**me **S**eries

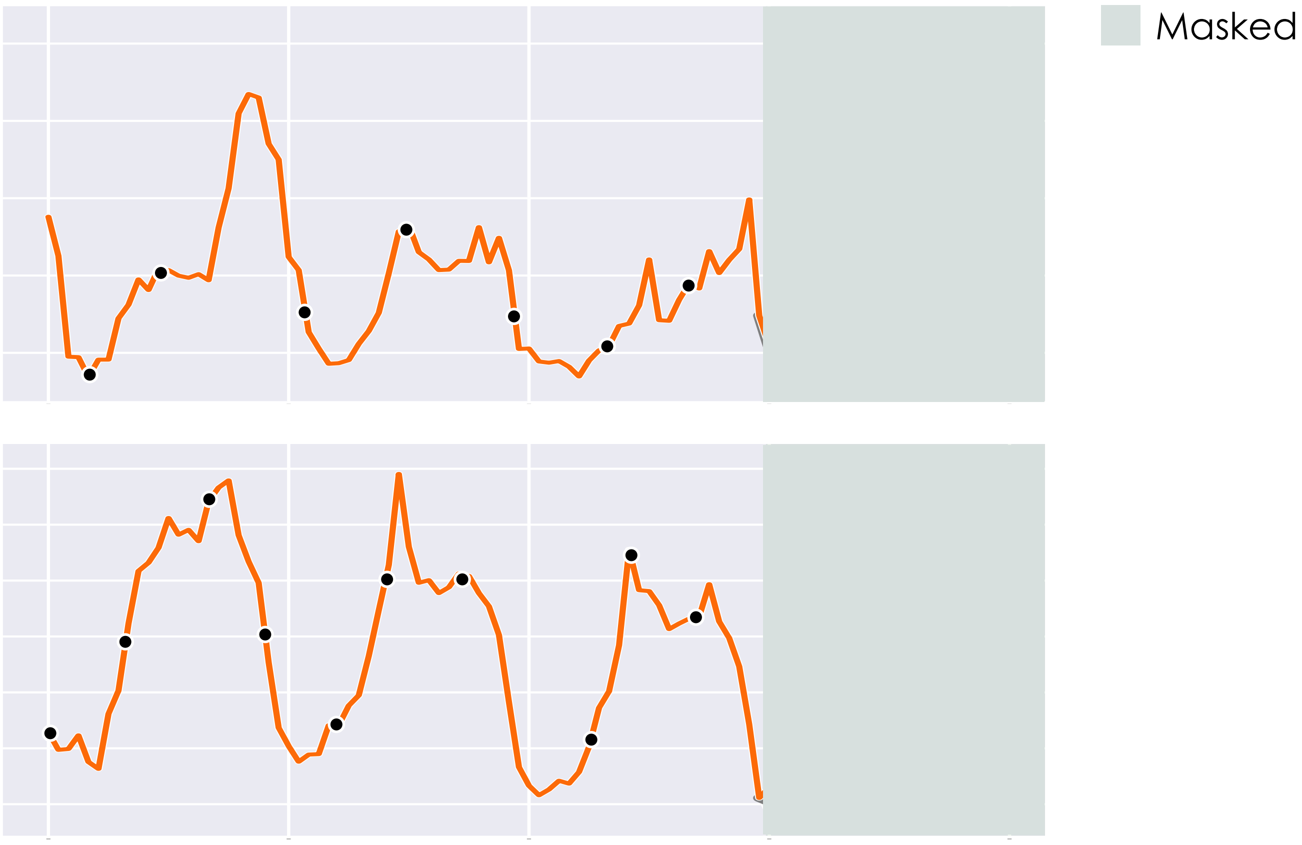
Alexandre Drouin

Étienne Marcotte

Nicolas Chapados

Objective

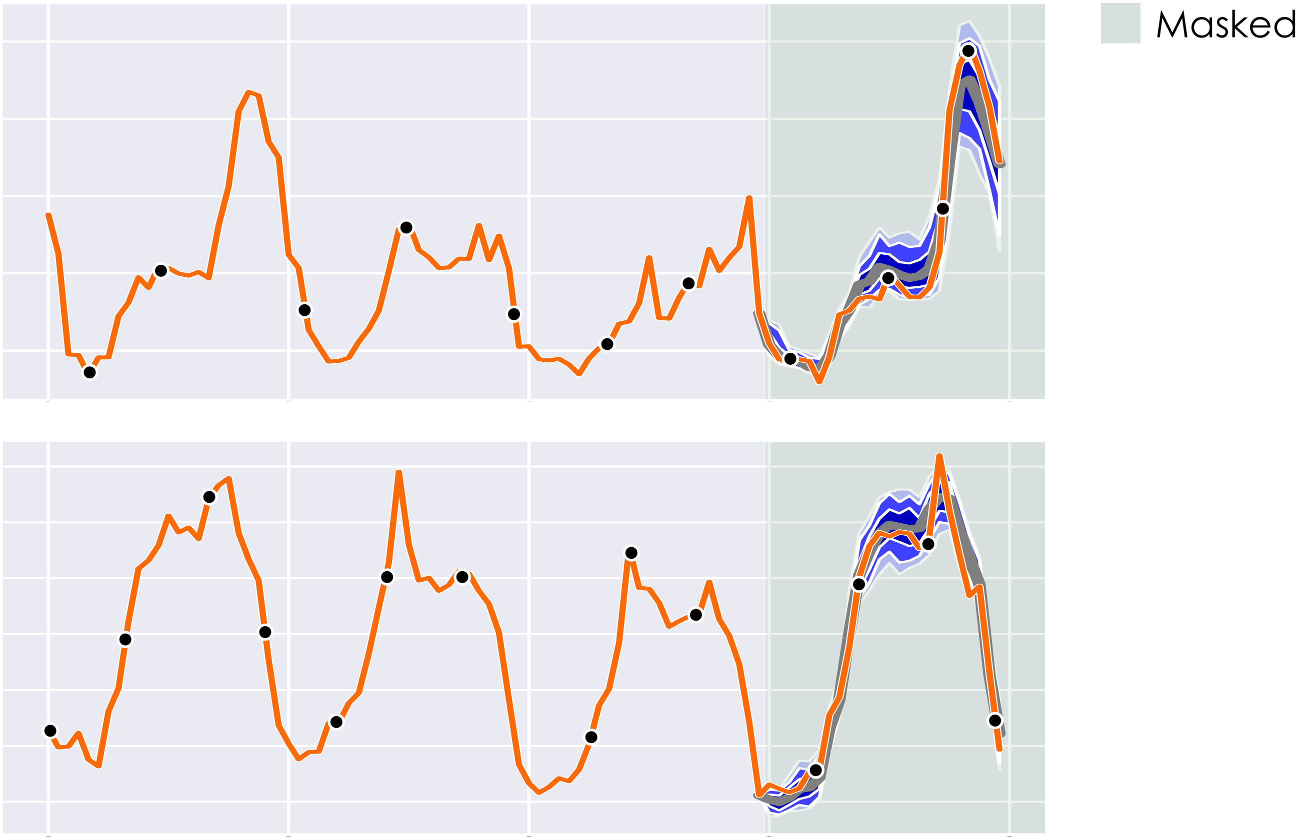
Goal: Infer the joint distribution of masked time points, given the observed time points



Very general: forecasting, interpolation, or arbitrary patterns

Objective

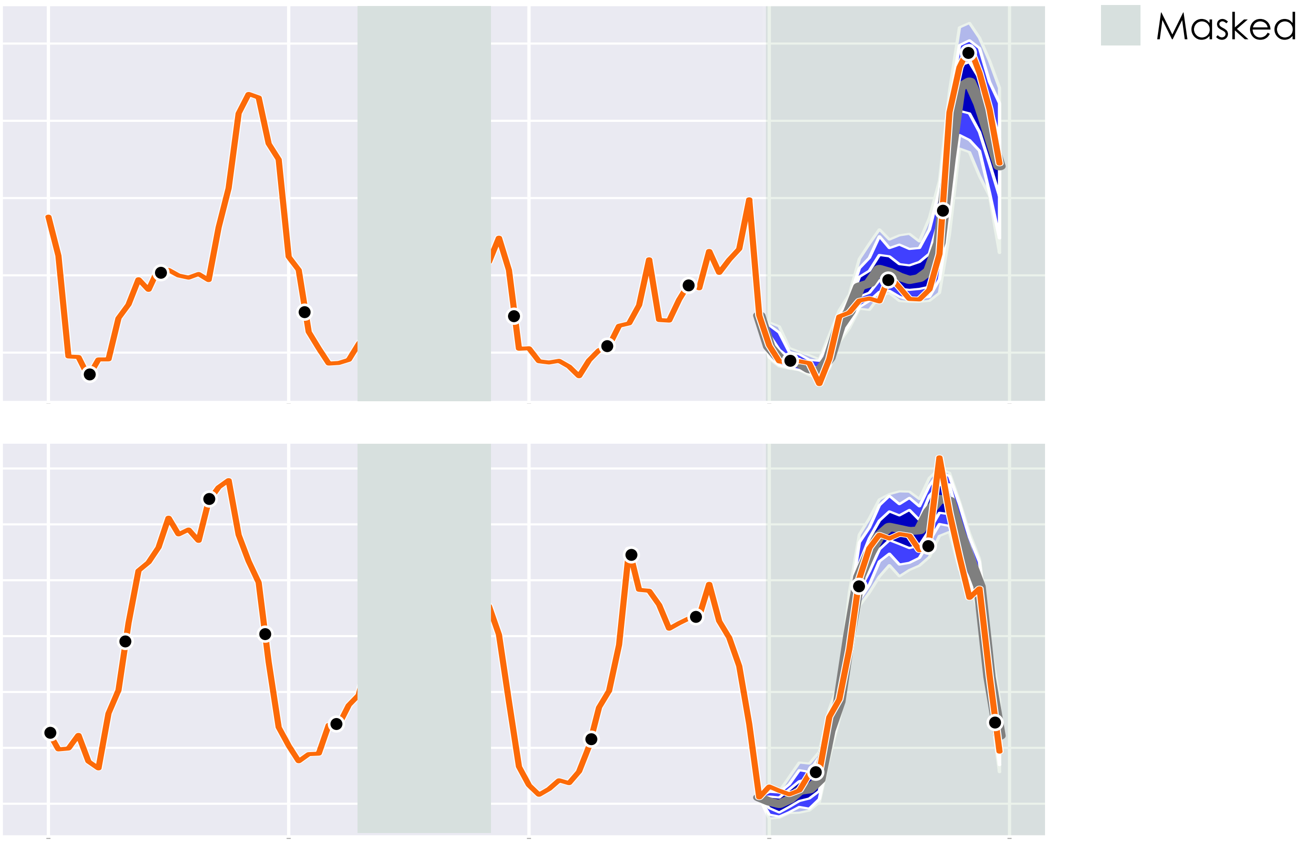
Goal: Infer the joint distribution of masked time points, given the observed time points



Very general: **forecasting**, interpolation, or arbitrary patterns

Objective

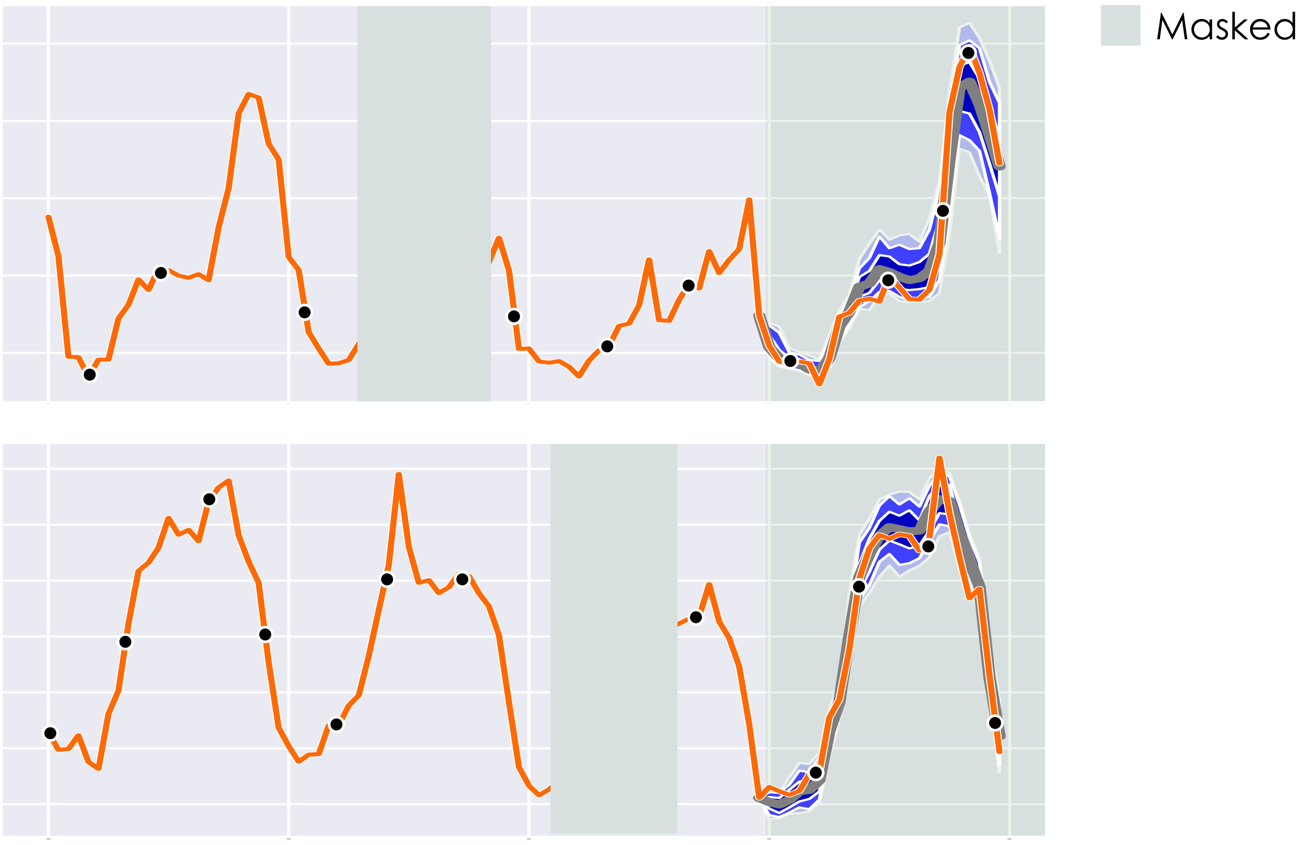
Goal: Infer the joint distribution of masked time points, given the observed time points



Very general: forecasting, **interpolation**, or arbitrary patterns

Objective

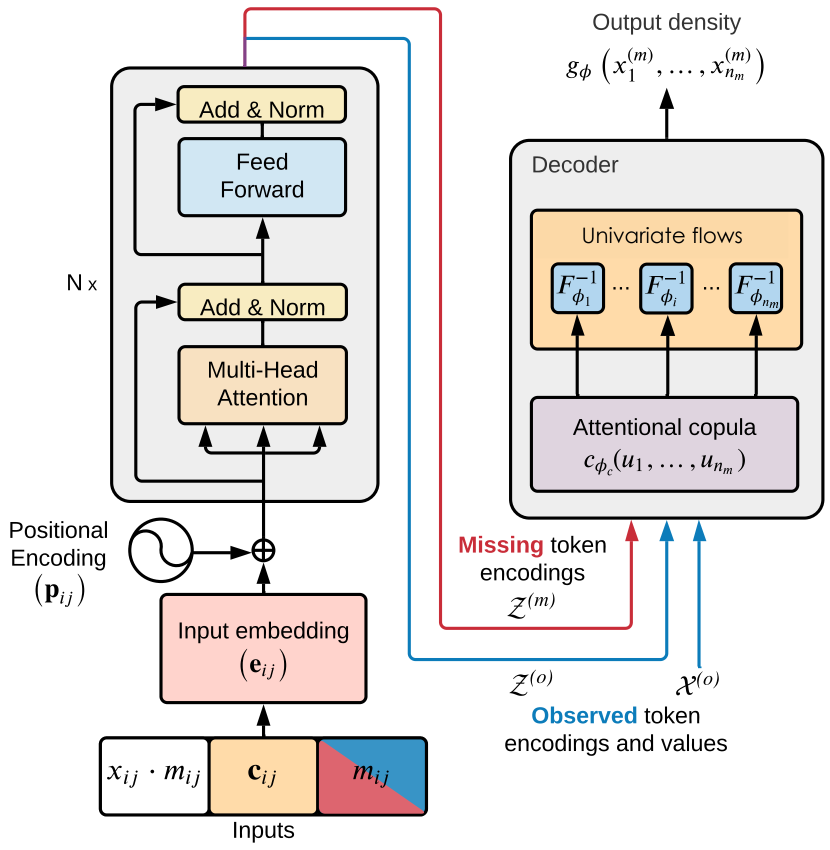
Goal: Infer the joint distribution of masked time points, given the observed time points



Very general: forecasting, interpolation, or **arbitrary patterns**

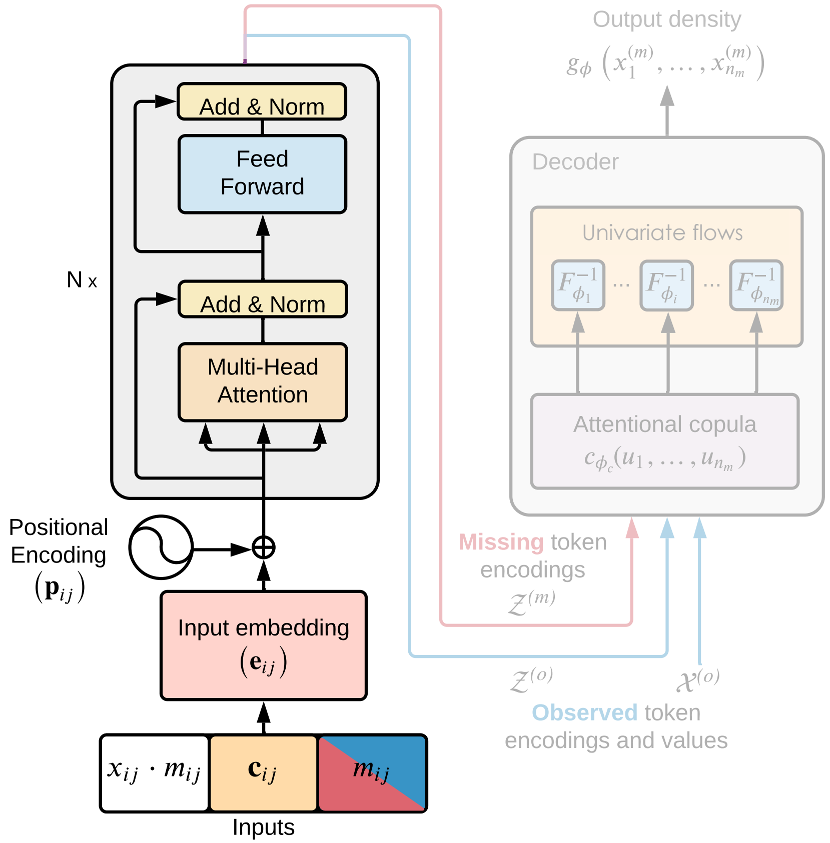
Overview of the model

TACTiS is an encoder-decoder model, similar to standard transformers.

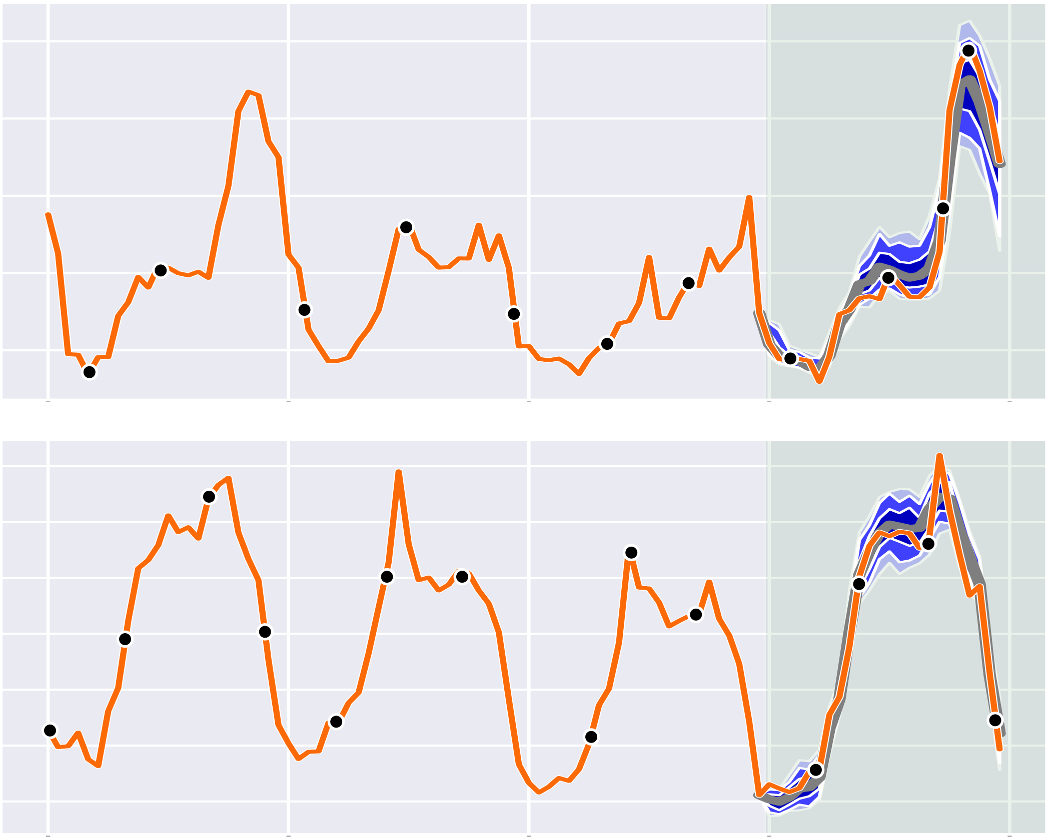


Overview of the model

TACTiS is an encoder-decoder model, similar to standard transformers.

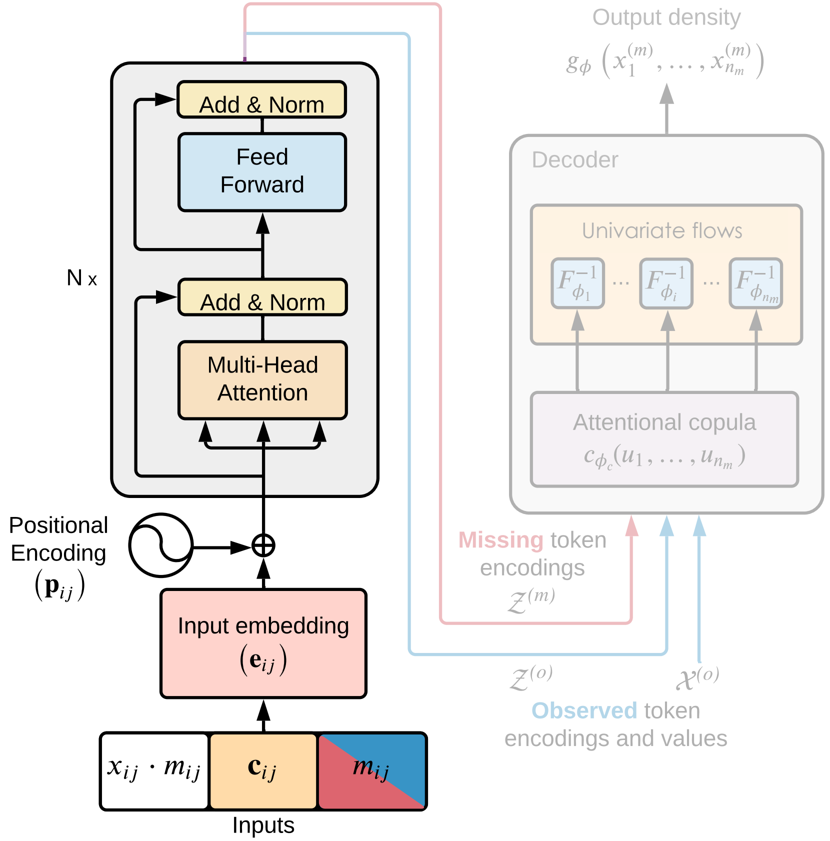


Encoder: each point in each time series is a token

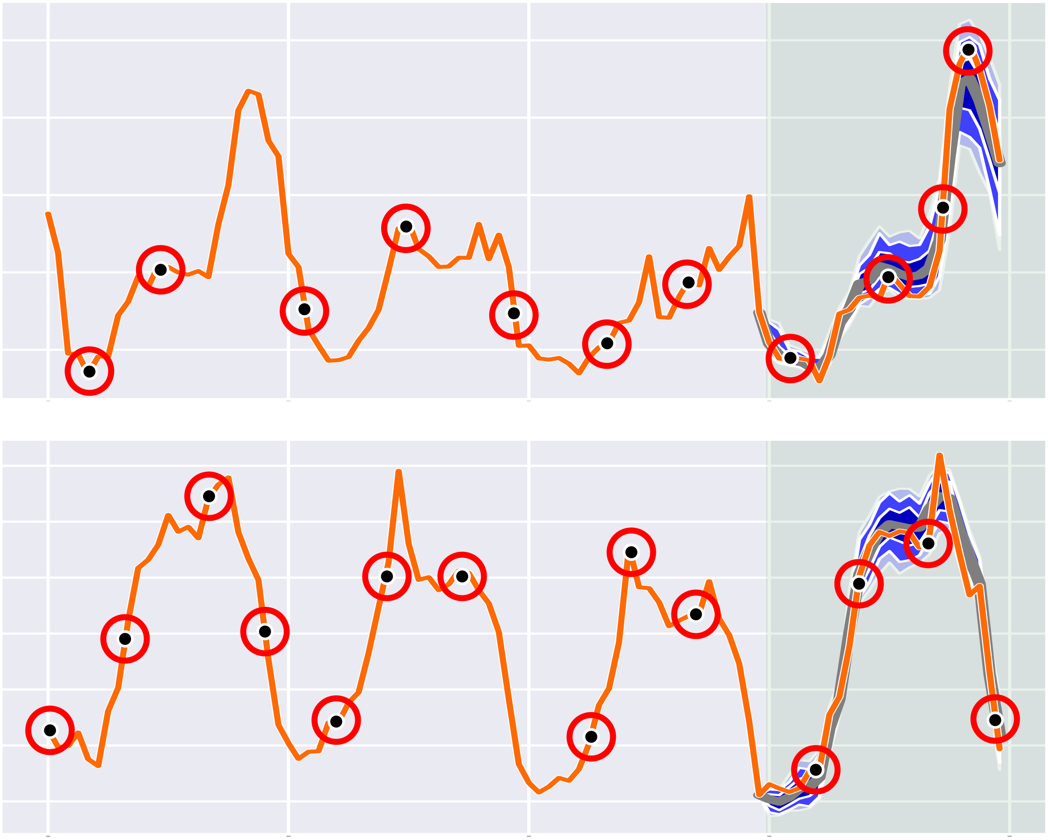


Overview of the model

TACTiS is an encoder-decoder model, similar to standard transformers.

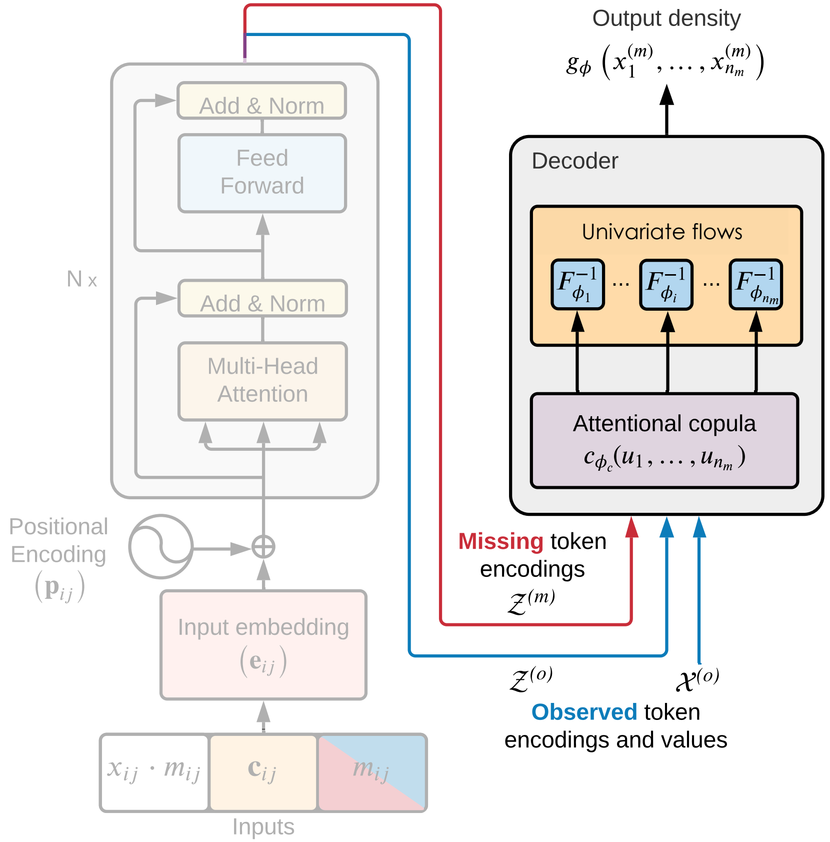


Encoder: each point in each time series is a token



Overview of the model

TACTiS is an encoder-decoder model, similar to standard transformers.



Decoder: a copula-based autoregressive decoder

Theorem (Sklar): any joint distribution can be expressed as a combination of two components:

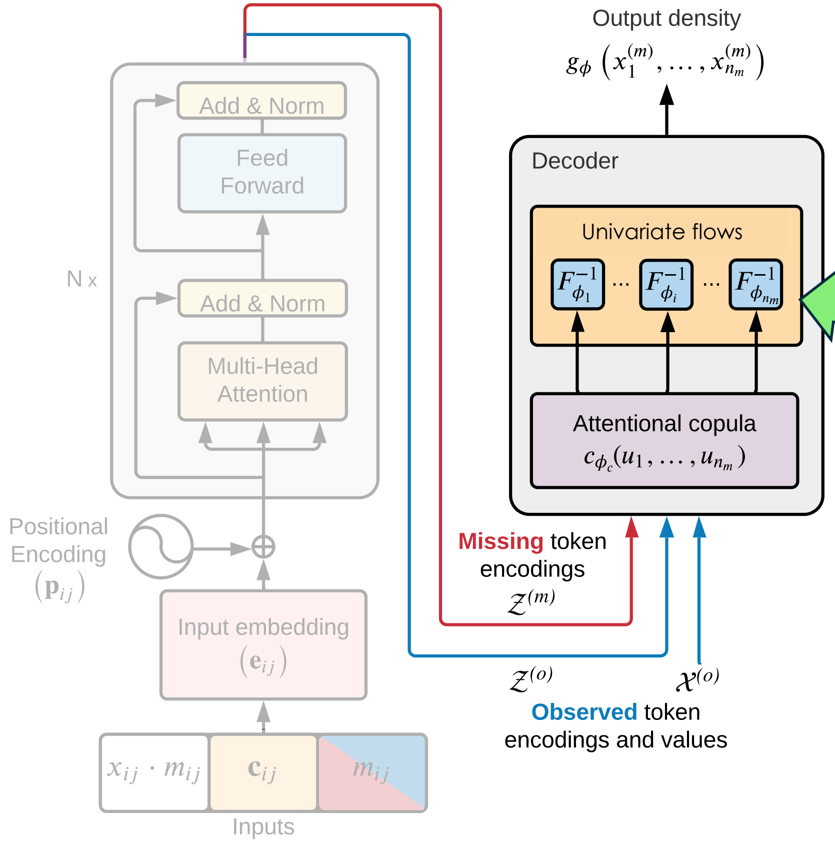
1. **Marginal** distribution of each variable
2. **Copula**: joint distribution on the unit cube

Why?

- Interpretability
- Robustness to distribution shifts

Overview of the model

TACTiS is an encoder-decoder model, similar to standard transformers.



Decoder: a copula-based autoregressive decoder

Theorem (Sklar): any joint distribution can be expressed as a combination of two components:

1. **Marginal** distribution of each variable
2. **Copula**: joint distribution on the unit cube

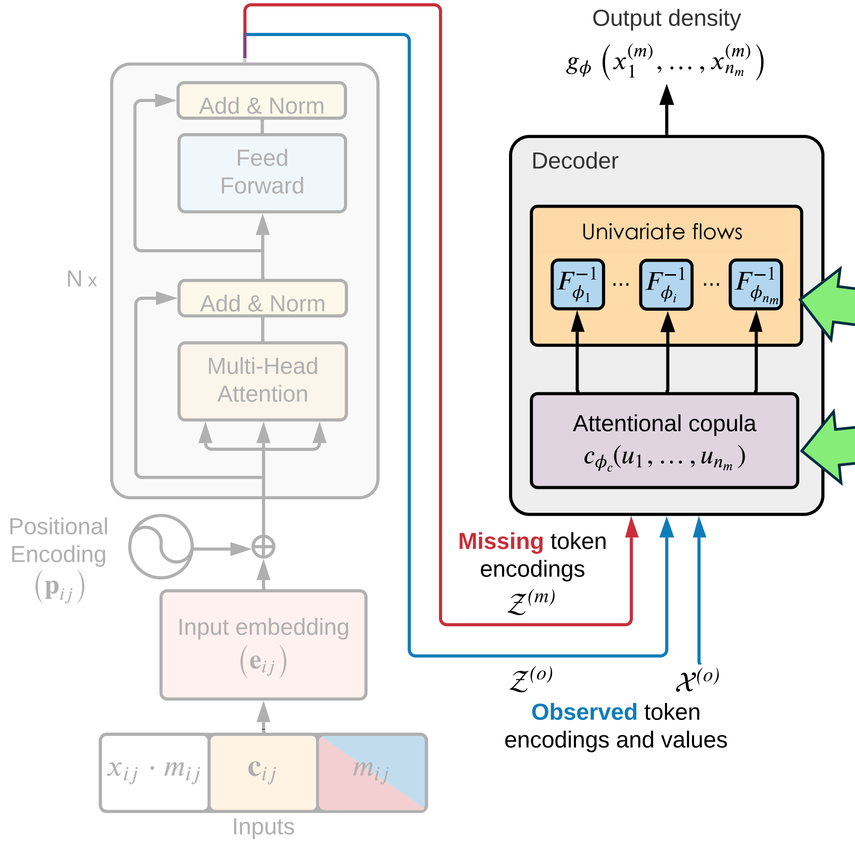
Why?

- Interpretability
- Robustness to distribution shifts

Overview of the model

TACTiS is an encoder-decoder model, similar to standard transformers.

Decoder: a copula-based autoregressive decoder



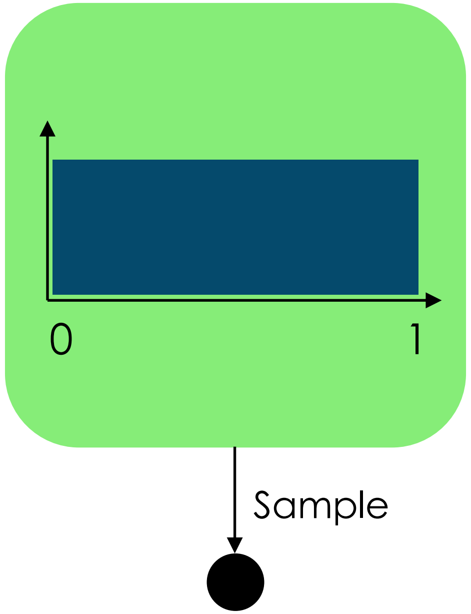
Theorem (Sklar): any joint distribution can be expressed as a combination of two components:

1. **Marginal** distribution of each variable
2. **Copula**: joint distribution on the unit cube

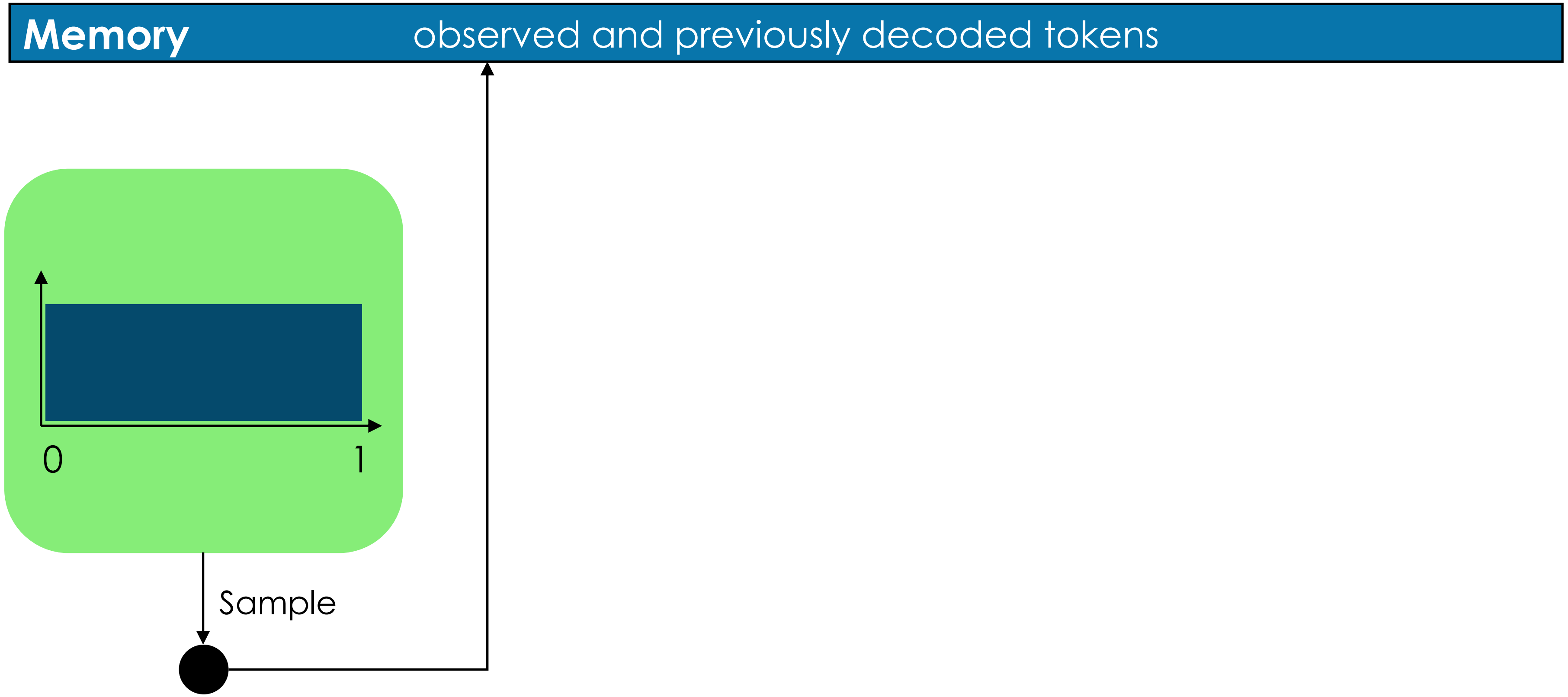
Why?

- Interpretability
- Robustness to distribution shifts

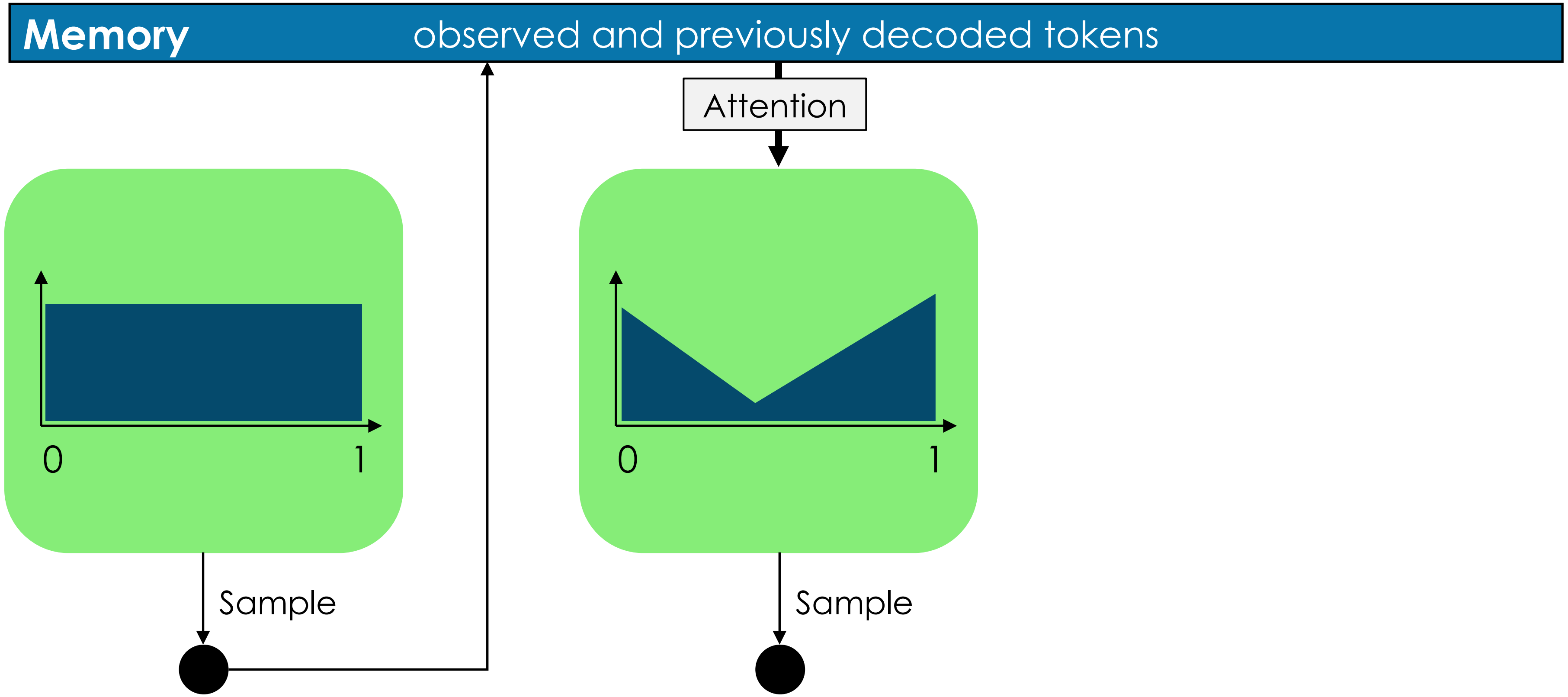
Decoding using attentional copulas



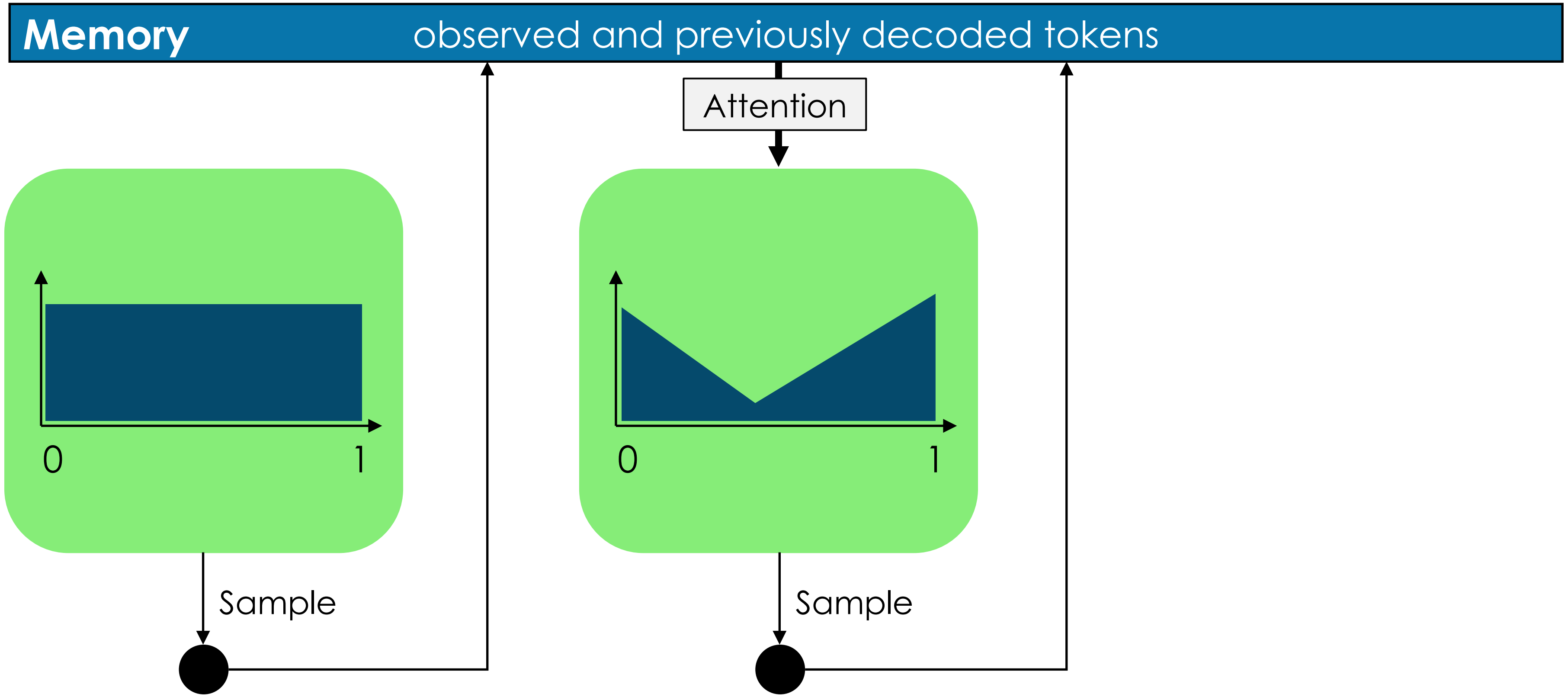
Decoding using attentional copulas



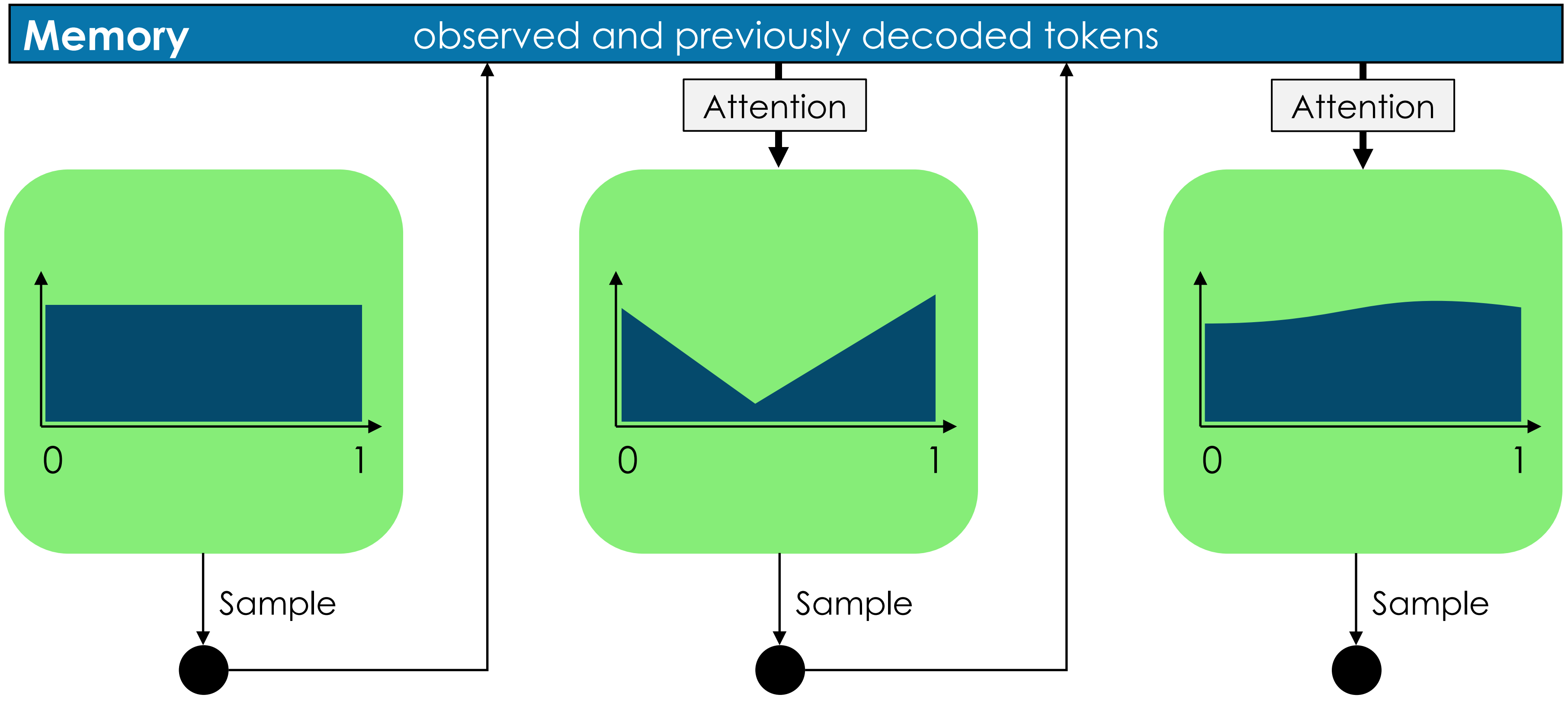
Decoding using attentional copulas



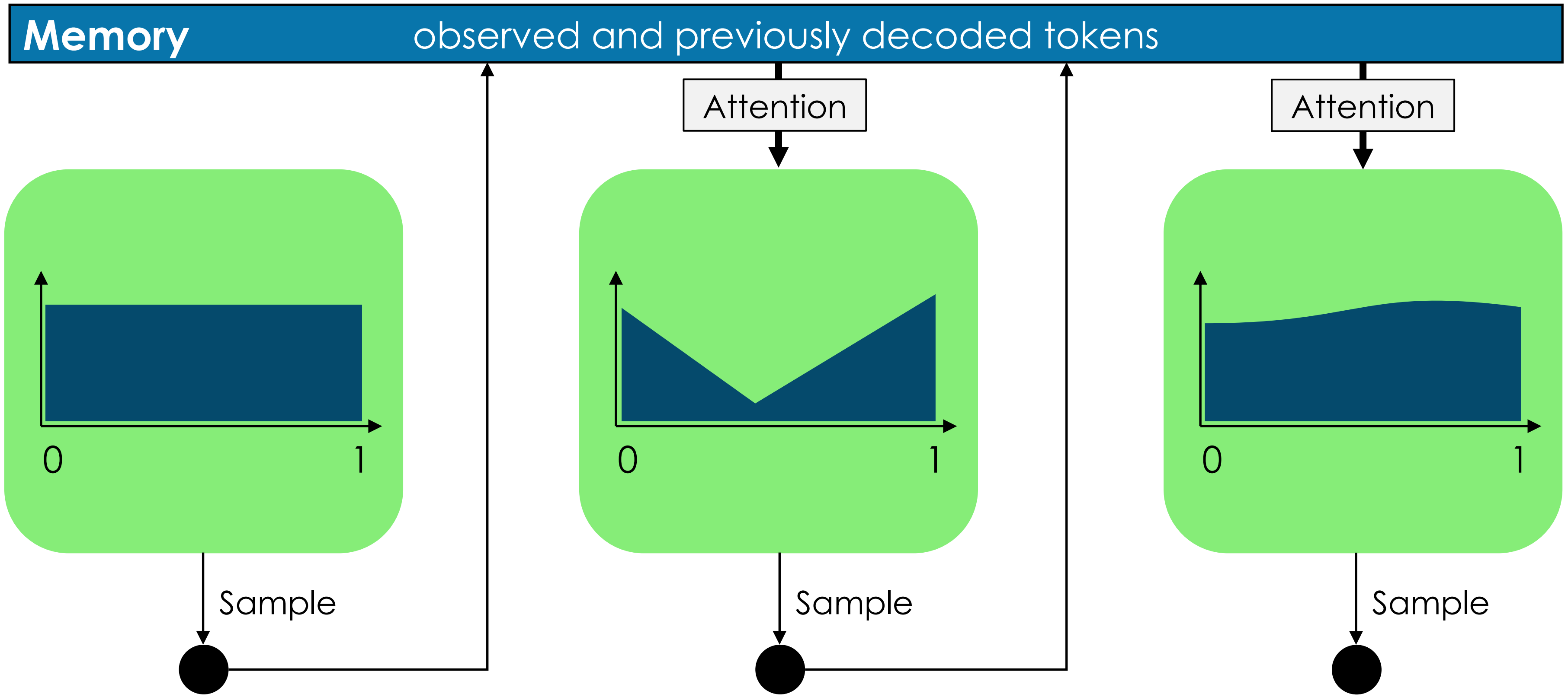
Decoding using attentional copulas



Decoding using attentional copulas



Decoding using attentional copulas



Theorem: decoding in a **random order** guarantees convergence to **valid copulas**

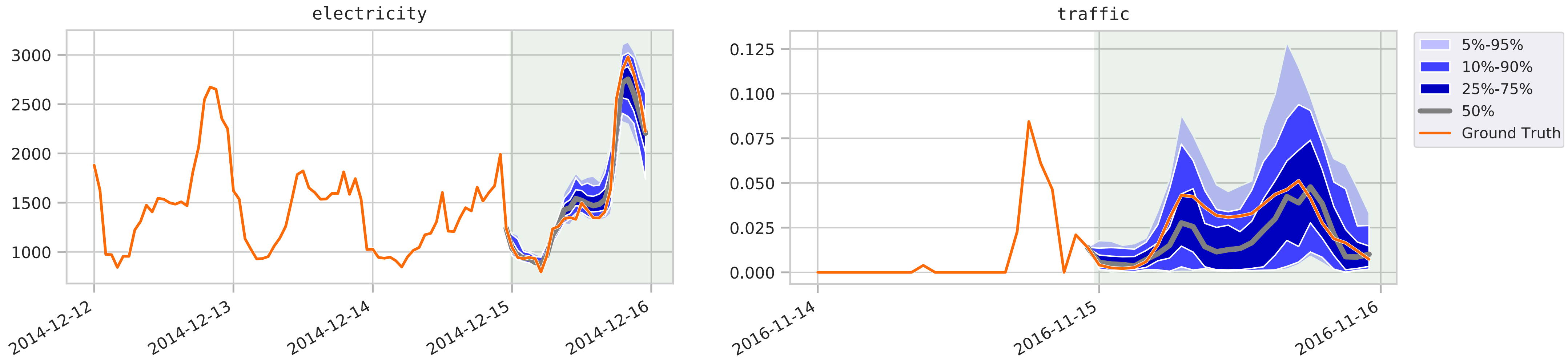
State-of-the-art forecasting performance

CRPS-Sum means (± standard errors). Lower is better. Best results in bold.

Model	electricity	fred-md	kdd-cup	solar-10min	traffic	Avg. Rank
Auto-ARIMA	0.077 ± 0.016	0.043 ± 0.005	0.625 ± 0.066	0.994 ± 0.216	0.222 ± 0.005	4.7 ± 0.3
ETS	0.059 ± 0.011	$\mathbf{0.037 \pm 0.010}$	0.408 ± 0.030	0.678 ± 0.097	0.353 ± 0.011	4.4 ± 0.3
TempFlow	0.075 ± 0.024	0.095 ± 0.004	0.250 ± 0.010	0.507 ± 0.034	0.242 ± 0.020	3.9 ± 0.2
TimeGrad	0.067 ± 0.028	0.094 ± 0.030	0.326 ± 0.024	0.540 ± 0.044	0.126 ± 0.019	3.6 ± 0.3
GPVar	0.035 ± 0.011	0.067 ± 0.008	0.290 ± 0.005	$\mathbf{0.254 \pm 0.028}$	0.145 ± 0.010	2.7 ± 0.2
TACTiS-TT	$\mathbf{0.021 \pm 0.005}$	0.042 ± 0.009	$\mathbf{0.237 \pm 0.013}$	0.311 ± 0.061	$\mathbf{0.071 \pm 0.008}$	$\mathbf{1.6 \pm 0.2}$

TACTiS outperforms **state-of-the-art** models on **real-world datasets** with **hundreds of time series**

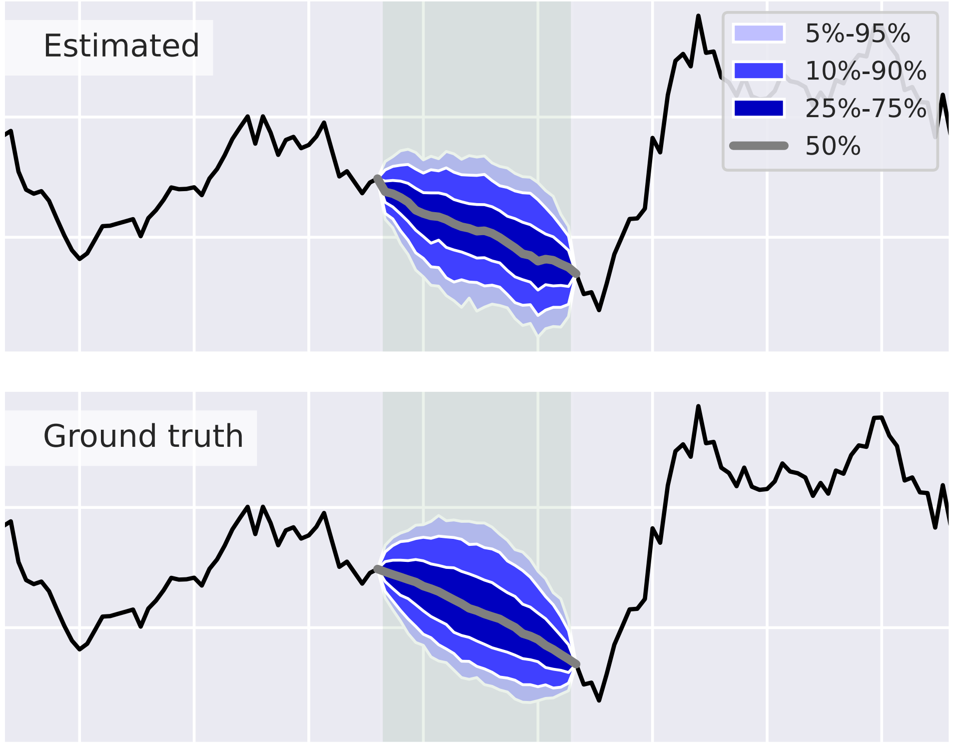
State-of-the-art forecasting performance



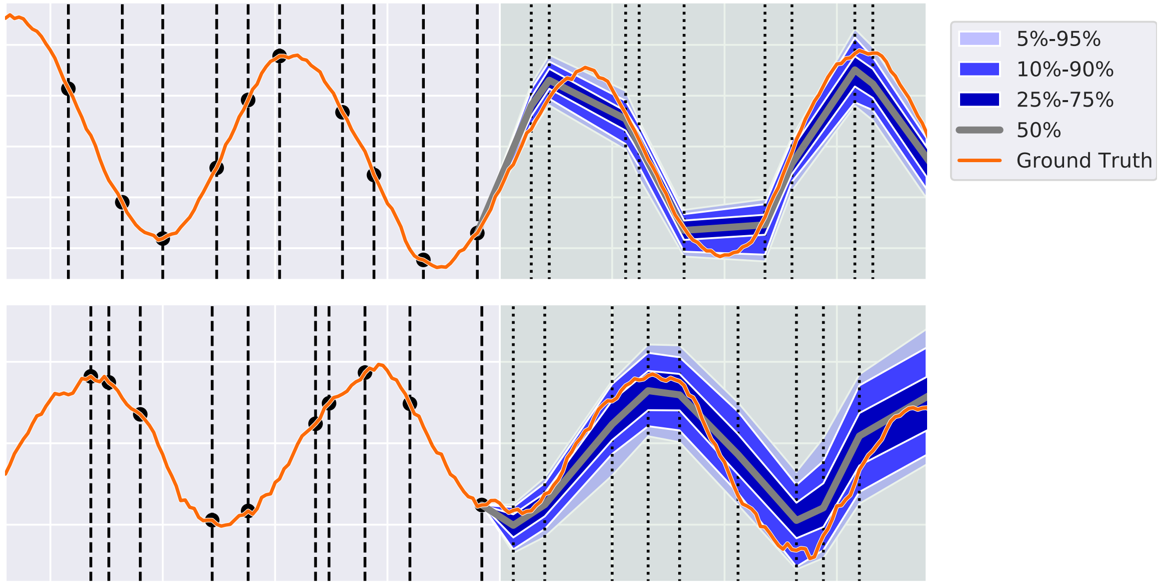
TACTiS outperforms **state-of-the-art** models on **real-world datasets** with **hundreds of time series**

TACTiS is very flexible

Interpolation



Unaligned and non-uniformly sampled data



Thank you!

Please come by our poster!

Code: https://github.com/ServiceNow/TACTiS