Chimera: Effectively Modeling Multivariate Time Series with 2-Dimensional State Space Models

Ali Behrouz

Ramin Zabih

Michele Santacatterina

Cornell University **New York University**

State Space Models (SSMs) are linear recurrent models that effectively compress information along the time dimension.

Are effective for modeling sequential data!

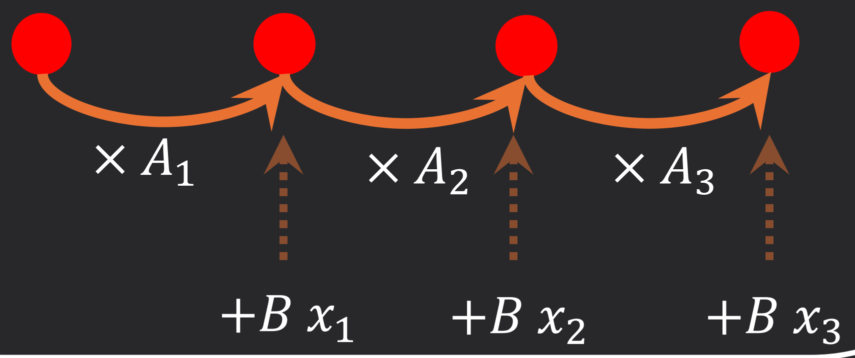
What happens when there 2D dependencies?

$$\frac{\partial}{\partial t} h(t) = A\, h(t) + B\, x(t)$$

$$\Downarrow$$

$$h(t+1) = A_t\, h(t) + B_t\, x(t)$$

$$\Downarrow$$

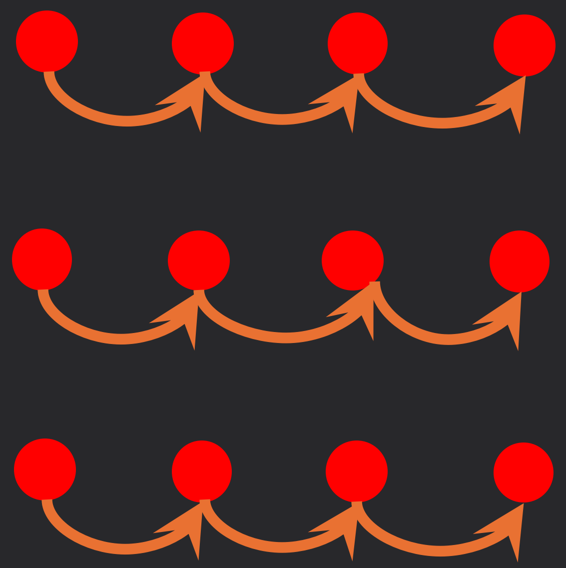


Idea 1:
Each channel is a separate variable!

Missing the dependencies of variates!

$$\frac{\partial}{\partial t_i} h(t_i) = A\, h(t_i) + B\, x(t_i)$$

$$\Downarrow$$



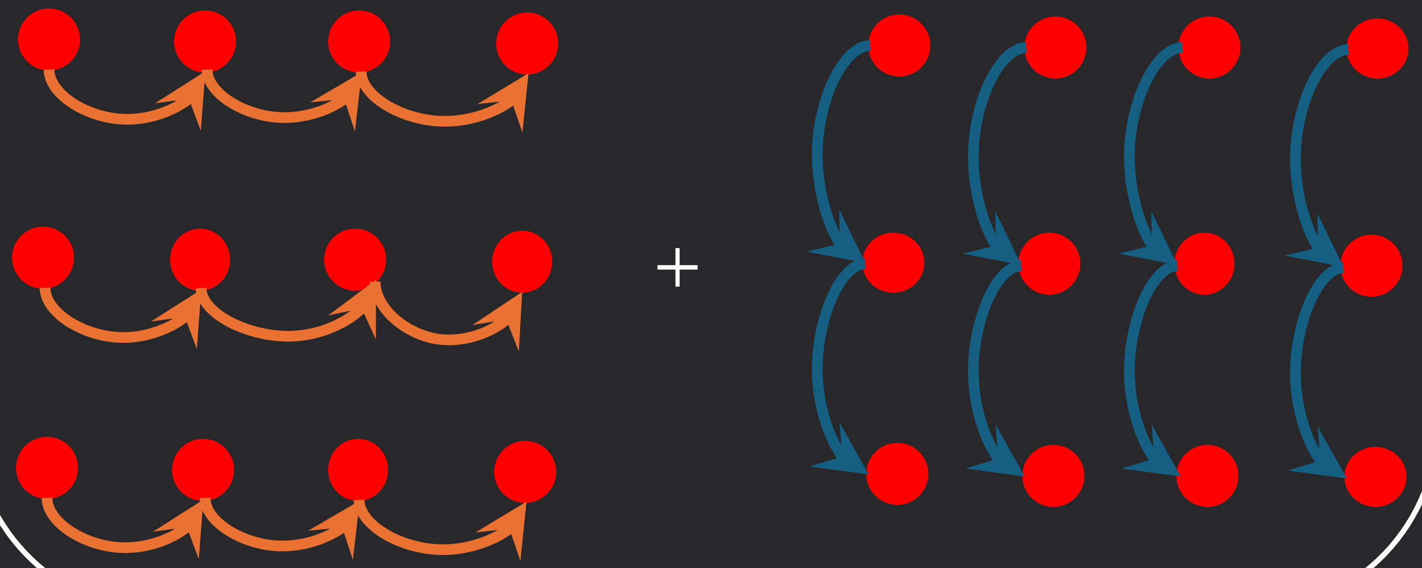
Idea 2:
Capture dependencies across both time and variates but separately!

Missing inter-state dependencies!

$$\frac{\partial}{\partial\, t_i} h^{(1)}(t_i) = A^{(1)} h^{(1)}(t_i) + B^{(1)} x(t_i)$$

$$\frac{\partial}{\partial\, v_i} h^{(2)}(v_i) = A^{(2)} h^{(2)}(v_i) + B^{(2)} x(v_i)$$

$$\Downarrow$$



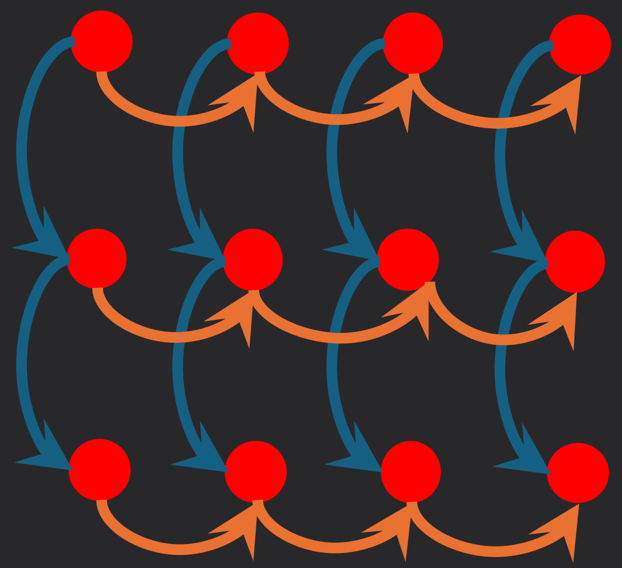
Idea 3:
Capture dependencies across both time and variates (all together)!

Capture all complex dependencies!

$$\frac{\partial}{\partial t} h^{(1)}(t,v) = \left(A^{(1)} h^{(1)}(t,v), A^{(2)} h^{(2)}(t,v)\right) + B^{(1)} x(t,v)$$

$$\frac{\partial}{\partial v} h^{(2)}(t,v) = \left(A^{(3)} h^{(1)}(t,v), A^{(4)} h^{(2)}(t,v)\right) + B^{(2)} x(t,v)$$

$$\Downarrow$$

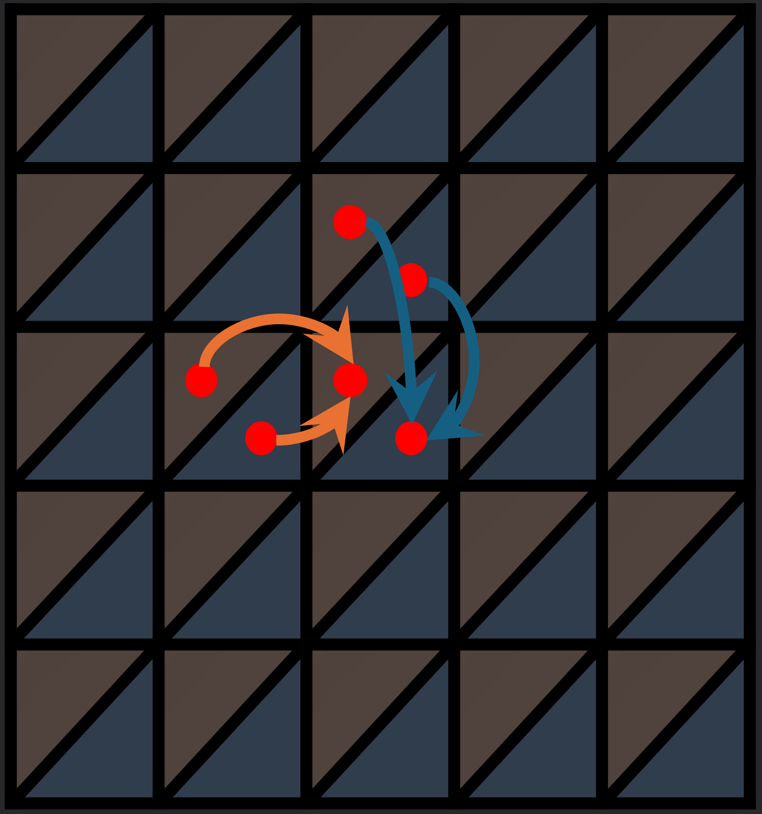


How does 2D recurrence look like?

Each hidden state is responsible to compress information across one dimension

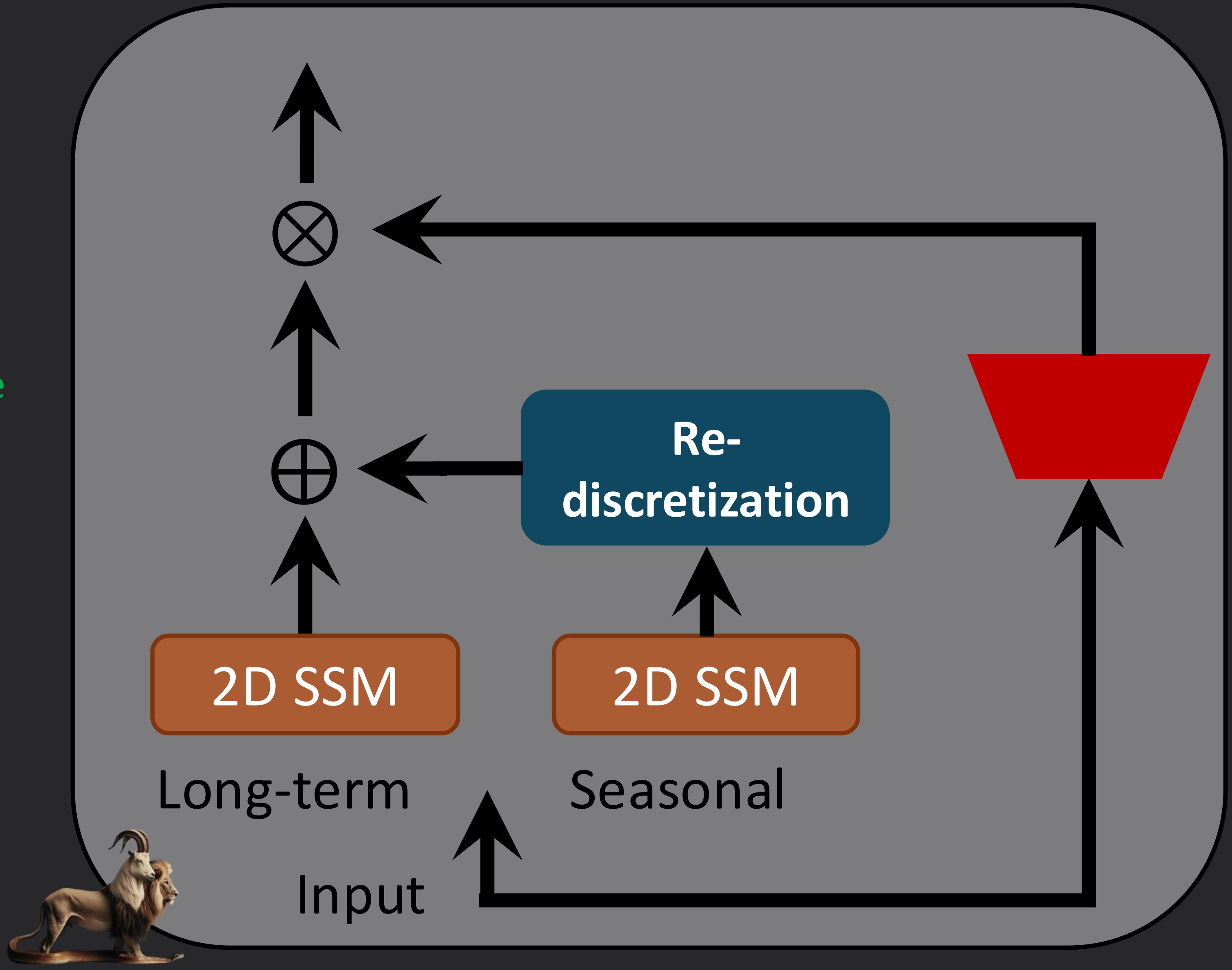
$$h^{(1)}(t+1,v) = A^{(1)}h^{(1)}(t,v) + A^{(2)}h^{(2)}(t,v) + B^{(1)}x(t+1,v)$$

$$h^{(2)}(t,v+1) = A^{(3)}h^{(1)}(t,v) + A^{(4)}h^{(2)}(t,v) + B^{(1)}x(t,v+1)$$



Chimera

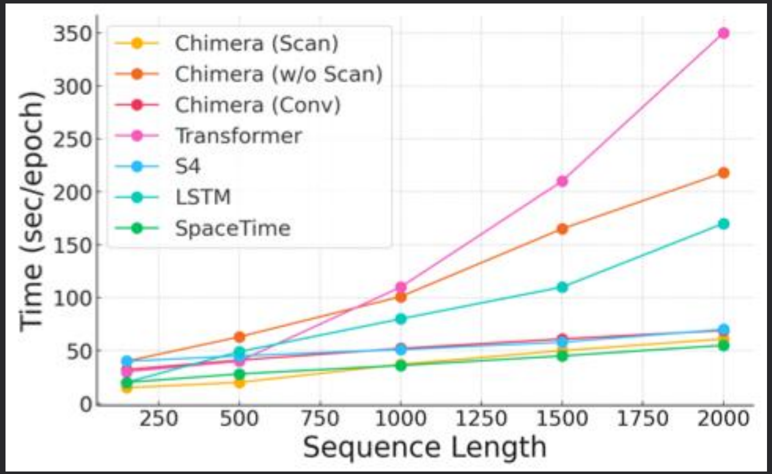
A Three-headed architecture that can capture both long-term and seasonal patterns.



Despite 2D dependencies, this process can be parallelized using the associative sum algorithms with $*$ operators:

$$\begin{pmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \end{pmatrix} * \begin{pmatrix} q_1 & q_2 & q_3 \\ q_4 & q_5 & q_6 \end{pmatrix}$$
$$= \begin{pmatrix} q_1 \odot p_1 & q_2 \odot p_2 & q_1 \otimes p_3 + q_2 \otimes p_6 + q_3 \\ q_4 \odot p_4 & q_5 \odot p_5 & q_4 \otimes p_3 + q_5 \otimes p_6 + q_6 \end{pmatrix}$$

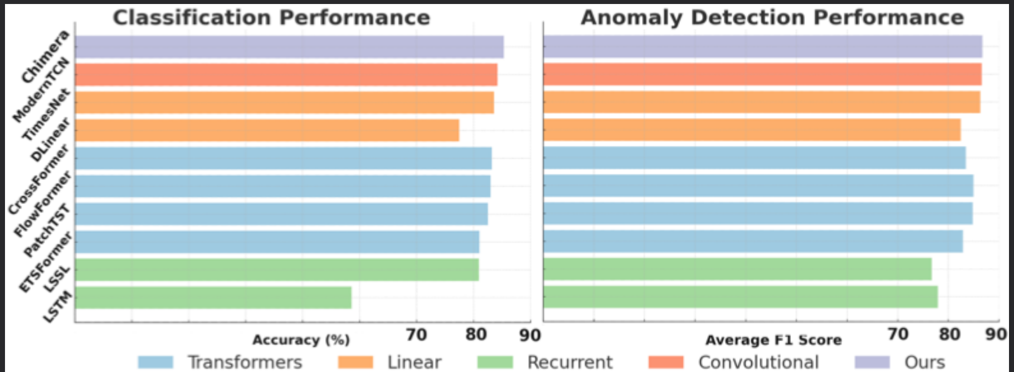
$*$ is associative!



Strong results in long-term and short-term forecasting as well as classification and anomaly detection tasks!

	Chimera (ours)		TSM2 [2024]		Simba [2024]		TCN [2024]		iTransformer [2024]		RLinear [2023]		PatchTST [2023]		Crossformer [2023]		TiDE [2023]		TimesNet [2023]		DLinear [2023]	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	0.355	**0.381**	0.361	-	0.383	0.396	**0.351**	**0.381**	0.407	0.410	0.414	0.407	0.387	0.400	0.513	0.496	0.419	0.419	0.400	0.406	0.403	0.407
ETTm2	**0.252**	0.317	0.267	-	0.271	0.327	0.253	**0.314**	0.288	0.332	0.286	0.327	0.281	0.326	0.757	0.610	0.358	0.404	0.291	0.333	0.350	0.401
ETTh1	0.408	0.425	**0.403**	-	0.441	0.432	0.404	**0.420**	0.454	0.447	0.446	0.434	0.469	0.454	0.529	0.522	0.541	0.507	0.458	0.450	0.456	0.452
ETTh2	**0.321**	**0.377**	0.333	-	0.361	0.391	0.322	0.379	0.383	0.407	0.374	0.398	0.387	0.407	0.942	0.684	0.611	0.550	0.414	0.427	0.559	0.515
ECL	**0.154**	**0.249**	0.169	-	0.185	0.274	0.156	0.253	0.178	0.270	0.219	0.298	0.205	0.290	0.244	0.334	0.251	0.344	0.192	0.295	0.212	0.300
Exchange	0.311	**0.358**	0.443	-	-	-	**0.302**	0.366	0.360	0.403	0.378	0.417	0.367	0.404	0.940	0.707	0.370	0.413	0.416	0.443	0.354	0.414
Traffic	0.403	0.286	0.420	-	0.493	0.291	**0.398**	**0.270**	0.428	0.282	0.626	0.378	0.481	0.304	0.550	0.304	0.760	0.473	0.620	0.336	0.625	0.383
Weather	**0.219**	**0.258**	0.239	-	0.255	0.280	0.224	0.264	0.258	0.278	0.272	0.291	0.259	0.281	0.259	0.315	0.271	0.320	0.259	0.287	0.265	0.317
1st Count	**4**	**5**	1	-	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Models		Chimera (ours)	ModernTCN [2024]	PatchTST [2023]	TimesNet [2023]	N-HiTS [2022]	N-BEATS* [2019]	ETS* [2022]	LightTS [2022]	DLinear [2023]	FED* [2022]	Stationary [2022]	Auto* [2021]	Pyra* [2021]	In* [2021]	Re* [2020]	LSTM [1997]
Weighted Average	SMAPE	**11.618**	11.698	11.807	11.829	11.927	11.851	14.718	13.525	13.639	12.840	12.780	12.909	16.987	14.086	18.200	160.031
	MASE	**1.528**	1.556	1.590	1.585	1.613	1.599	2.408	2.111	2.095	1.701	1.756	1.771	3.265	2.718	4.223	25.788
	OWA	**0.827**	0.838	0.851	0.851	0.861	0.855	1.172	1.051	1.051	0.918	0.930	0.939	1.480	1.230	1.775	12.642



Thank You!