A Statistical Investigation of Long Memory in Language and Music

Alexander Greaves-Tunnell and Zaid Harchaoui

ICML 2019

UNIVERSITY of WASHINGTON

Problem

- How do we define **long-range dependence**?

Problem

- How do we define **long-range dependence**?
- How can it be **estimated** in modern sources of sequence data?

Problem

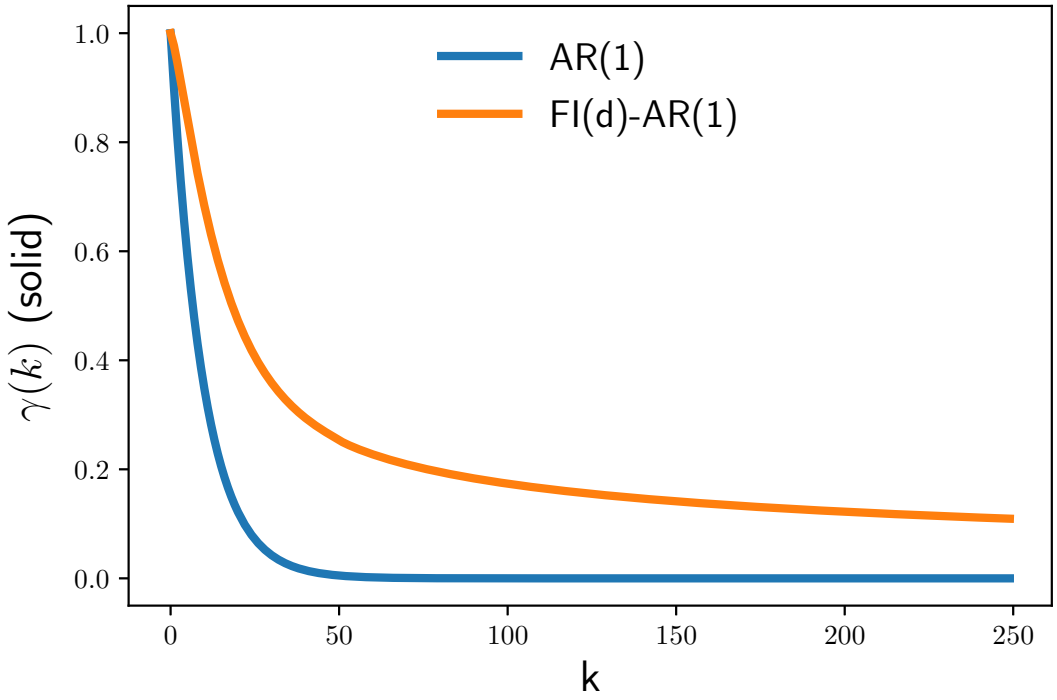
- How do we define **long-range dependence**?
- How can it be **estimated** in modern sources of sequence data?
- How can we **evaluate** if a model has captured this property?

Contributions

- Introduce a framework for evaluation of long-range dependence anchored in the literature of **long memory** stochastic processes.

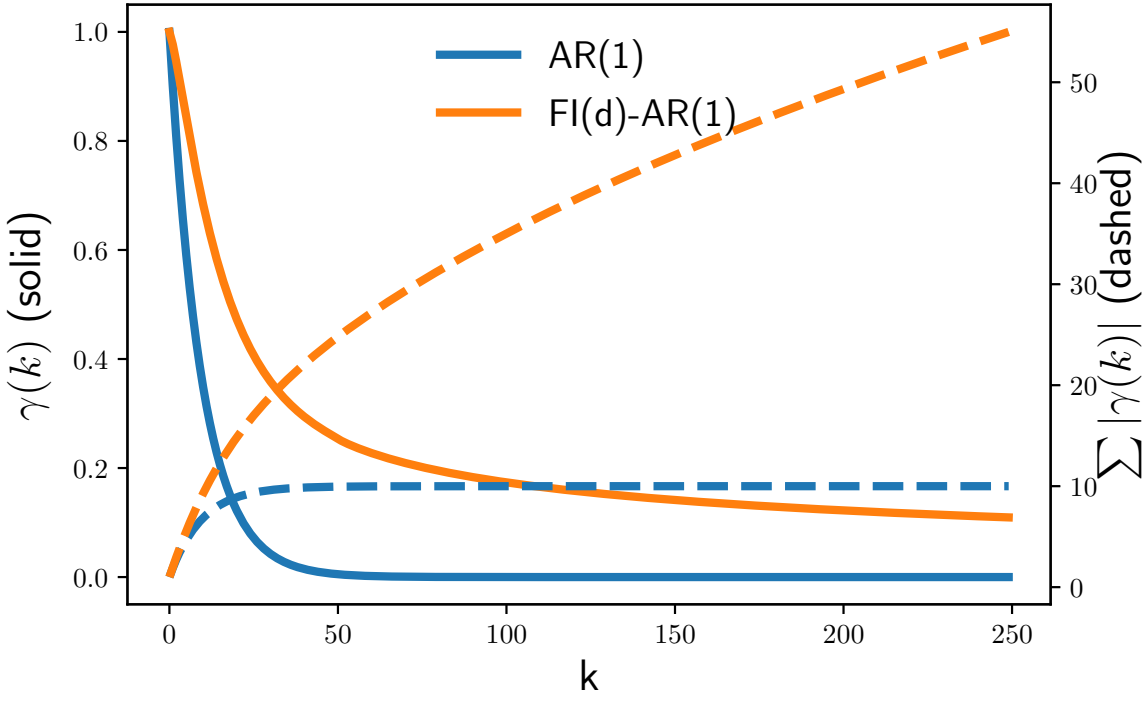
Contributions

- Introduce a framework for evaluation of long-range dependence anchored in the literature of **long memory** stochastic processes.



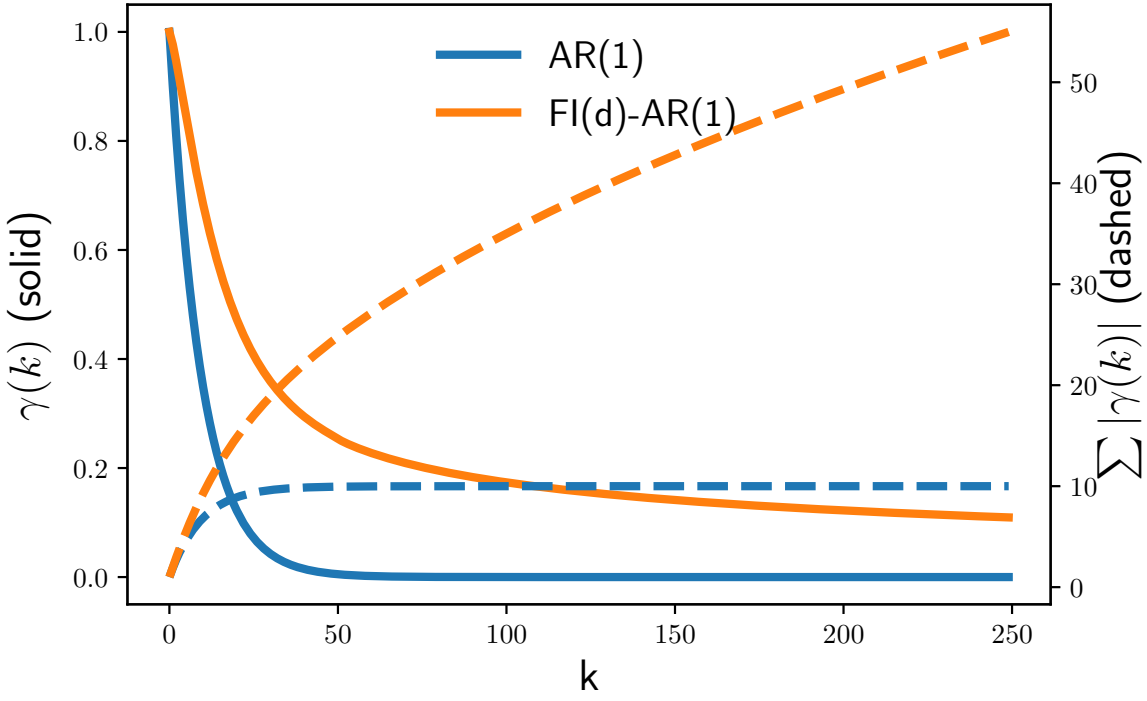
Contributions

- Introduce a framework for evaluation of long-range dependence anchored in the literature of **long memory** stochastic processes.



Contributions

• Introduce a framework for evaluation of long-range dependence anchored in the literature of **long memory** stochastic processes.



• Adapt **semiparametric** statistical methods to define estimation and testing procedure for long memory in high dimensions.

Results: Language and Music

Do language and music data have long memory?

Results: Language and Music

Do language and music data have long memory?

Visual heuristic from differing behavior of partial sums:

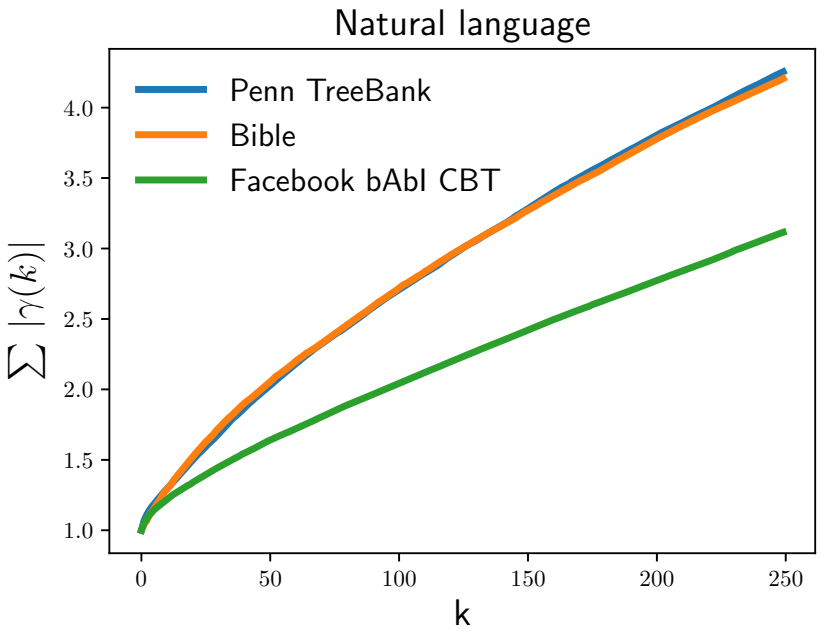
$$\sum_{k=1}^{K} |\gamma(k)| \to \begin{cases} \infty & \textbf{long} \text{ memory} \\ c < \infty & \textbf{short} \text{ memory} \end{cases}, \text{ as } K \to \infty$$

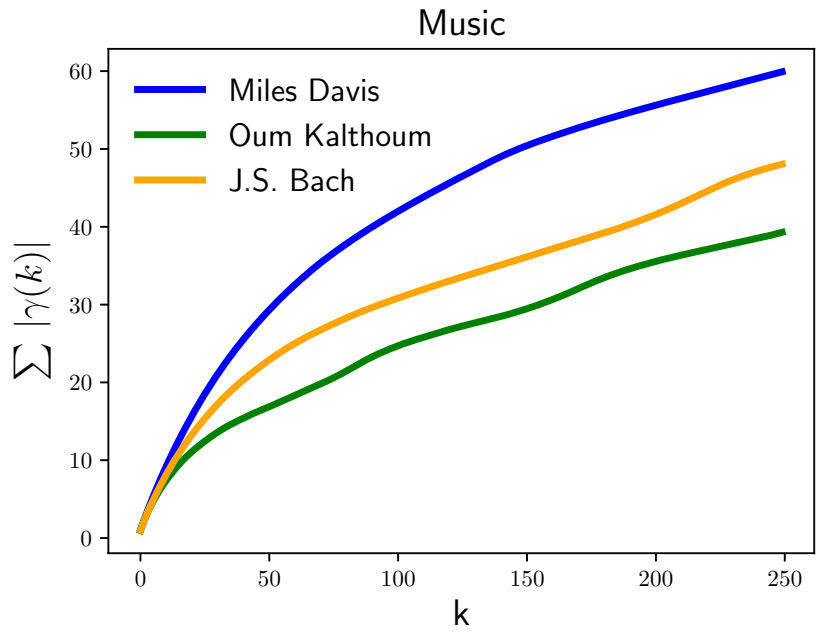
Results: Language and Music

Do language and music data have long memory?

Visual heuristic from differing behavior of partial sums:

$$\sum_{k=1}^{K} |\gamma(k)| \to \begin{cases} \infty & \textbf{long} \text{ memory} \\ c < \infty & \textbf{short} \text{ memory} \end{cases}, \text{ as } K \to \infty$$





Results: RNN Models

Hypothesis test for long memory:

$$\mathcal{H}_0 : \mathbf{d} = 0 \quad \text{vs.} \quad \overbrace{\mathcal{H}_A : \mathbf{d} > 0}^{\textbf{anticipated result}} \quad .$$

Results: RNN Models

Hypothesis test for long memory:

$$\overbrace{\mathcal{H}_0 : \mathbf{d} = 0}^{\textbf{experimental result}} \qquad \text{vs. } \mathcal{H}_A : \mathbf{d} > 0.$$

Model	**memory d**	**p-value**	**Reject $\mathcal{H}_0$?**
LSTM (trained)	-8.59×10^{-4}	0.583	X
LSTM (untrained)	-4.17×10^{-4}	0.572	X
Memory cell	-5.96×10^{-4}	0.552	X
SCRN	2.37×10^{-3}	0.324	X

Results: RNN Models

Hypothesis test for long memory:

$$\overbrace{\mathcal{H}_0 : \mathbf{d} = 0}^{\textbf{experimental result}} \qquad \text{vs. } \mathcal{H}_A : \mathbf{d} > 0.$$

Model	**memory d**	**p-value**	**Reject $\mathcal{H}_0$?**
LSTM (trained)	-8.59×10^{-4}	0.583	X
LSTM (untrained)	-4.17×10^{-4}	0.572	X
Memory cell	-5.96×10^{-4}	0.552	X
SCRN	2.37×10^{-3}	0.324	X

Come see our poster! **Pacific Ballroom # 253**.