Contribution ID: **3** Type: **not specified**

Mean-Field Transformer Dynamics with Gaussian Inputs

Friday 26 September 2025 11:15 (45 minutes)

Transformers, that underlie the recent successes of large language models, represent the data as sequences of vectors called tokens. This representation is leveraged by the attention function, which learns dependencies between tokens and is key to the success of Transformers. However, the dynamics induced by the iterative application of attention across layers remain to be fully understood. To analyze these dynamics, we identify each input sequence with a probability measure, thus handling input sequences of arbitrary length, and model its evolution as a Vlasov equation called Transformer PDE, whose velocity field is non-linear in the probability measure. For compactly supported initial data and several self-attention variants, we show the Transformer PDE is well-posed and is the mean-field limit of an interacting particle system. We also study the case of Gaussian initial data, which has the nice property of staying Gaussian across the dynamics. This allows us to identify typical behaviors theoretically and numerically, and to highlight a clustering phenomenon that parallels previous results in the discrete case.

Presenter: CASTIN, Valérie