EPFL-Inria
INTERNATIONAL LAB

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Inria
INVENTEURS DU MONDE NUMÉRIQUE

**Workshop EPFL-Inria**
**January 20 & 21, 2022 -Lausanne**

# "Implicit Bias of SGD for Diagonal Linear Networks: a Provable Benefit of Stochasticity"

**Loucas Pillaud-Vivien**
**EPFL**

**Abstract**

Understanding the implicit bias of training algorithms is of crucial importance in order to explain the success of overparametrised neural networks. In this paper, we study the dynamics of stochastic gradient descent over diagonal linear networks through its continuous time version, namely stochastic gradient flow. We explicitly characterise the solution chosen by the stochastic flow and prove that it always enjoys better generalisation properties than that of gradient flow. Quite surprisingly, we show that the convergence speed of the training loss controls the magnitude of the biasing effect: the slower the convergence, the better the bias. To fully complete our analysis, we provide convergence guarantees for the dynamics. We also give experimental results which support our theoretical claims. Our findings highlight the fact that structured noise can induce better generalisation and they help explain the greater performances observed in practice of stochastic gradient descent over gradient descent.

**Biography**

Since September 2020 I am a postdoc in the Theory of Machine Learning group of Nicolas Flammarion at EPFL where I study the effect of stochasticity in the optimization algorithms used in machine learning. I did my Ph.D. in the SIERRA Team, at École Normale Supérieure and INRIA under the supervision of Francis Bach and Alessandro Rudi on stochastic approximation for high dimensionnal learning problems. Prior to that, I graduated from École Polytechnique in 2016 and got a master degree from Paris VI and École des Ponts Paris Tech in PDEs (ANEDP). I did my master thesis in Molecular dynamics in the CERMICS, where I was advised by Julien Reygner and Tony Lelievre.