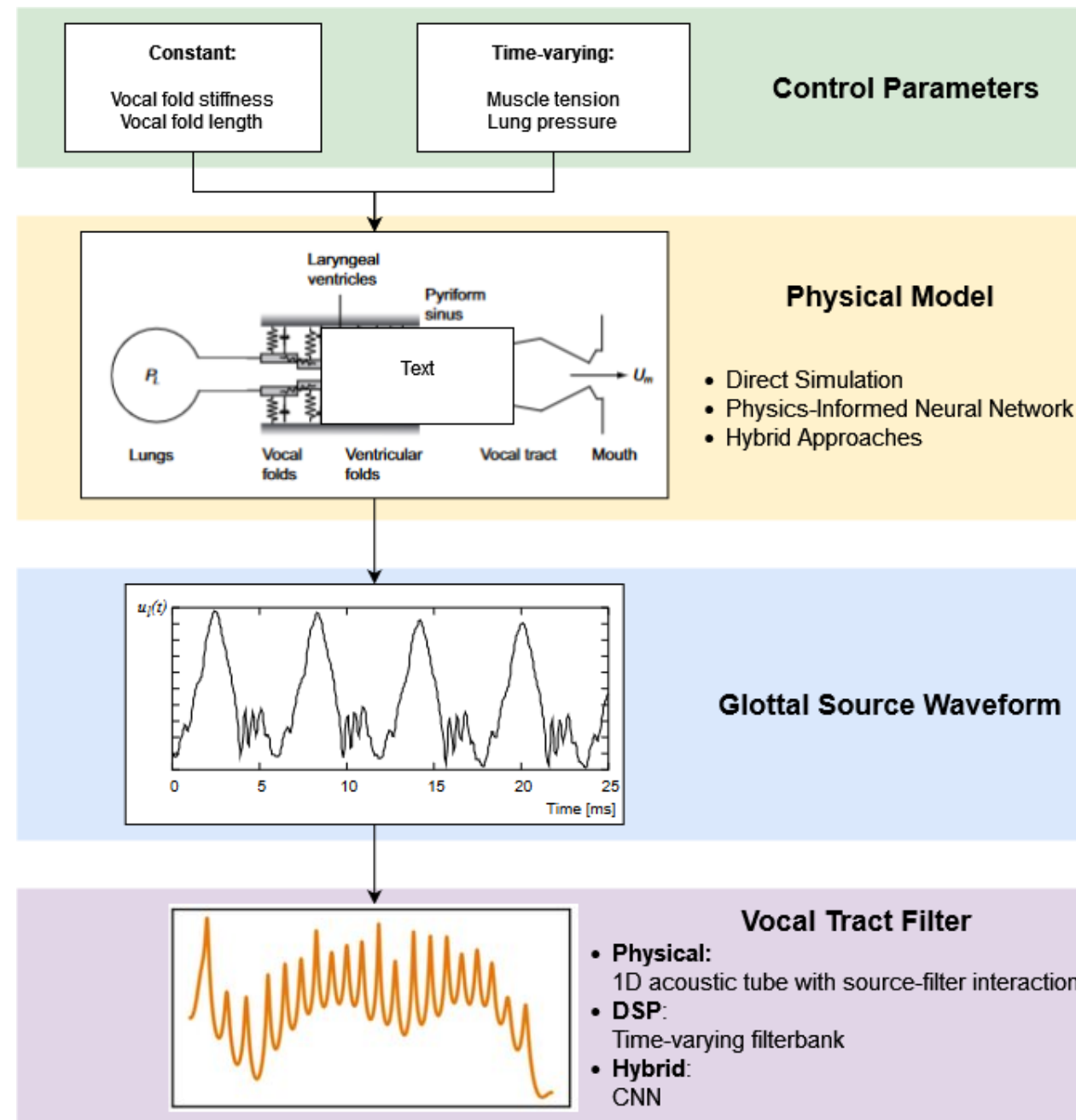


Machine learning of physical models for voice synthesis

David Sūdholt

2022



Finding

More complex and expressive physical models of the vocal folds require non-linear solvers to be ran at every glottal cycle to solve the collision of the vocal folds, putting them generally out of reach for real-time synthesis. We have found that we can train neural proxies for the solvers that make the models real-time feasible.

Question

How can we solve the sound matching problem for these physical glottal models? Given a physical model of the vocal folds and a real voice recording, how can we determine the control parameters such that the model will recreate the sound? Can an inductive bias given by a physical glottal model improve