Scalable Multitask Latent Force Models with Applications to Predicting Lithium-ion Concentration

Daniel Tait
Depts. of Comp. Sci. and Stats.
University of Warwick
The Alan Turing Institute
dtait@turing.ac.uk

Ferran Brosa Planella
WMG
University of Warwick
The Faraday Institution
ferran.brosa-planella@warwick.ac.uk

Theodoros Damoulas
Depts. of Comp. Sci. and Stats.
University of Warwick
The Alan Turing Institute
tdamoulas@turing.ac.uk

W. Dhammika Widanage
WMG
University of Warwick
The Faraday Institution
dhammika.widanalage@warwick.ac.uk

Introduction

A common task in engineering is the reduction of a complete physical description of a system to a more tractable approximation. The question is what is the lost physical information from this reduction?

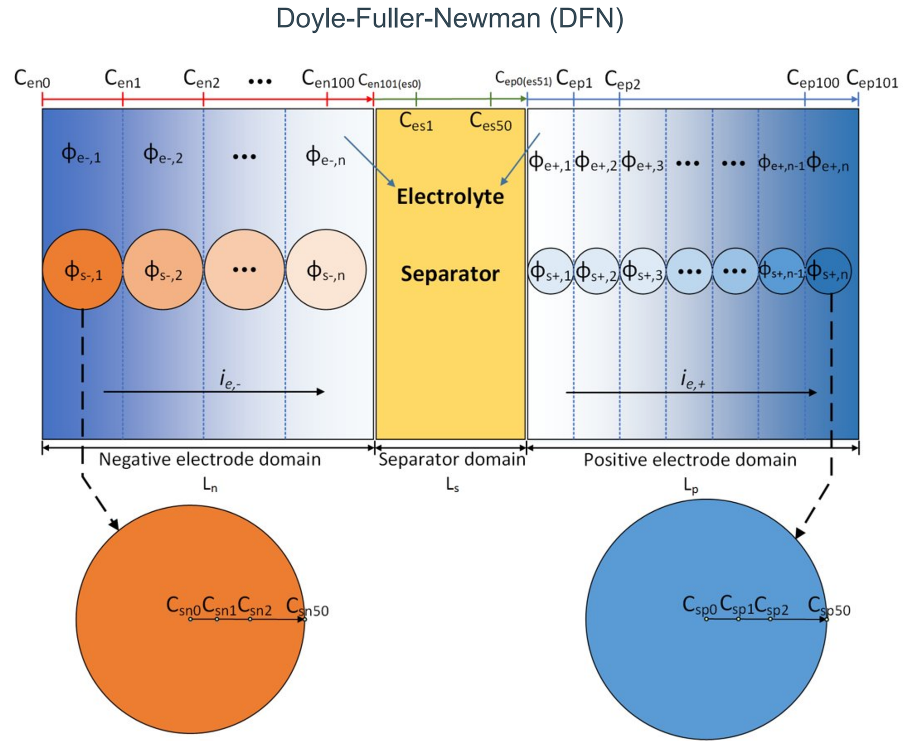
In this work we investigate the use of machine learning and data-driven components to recover this information with the goal of producing hybrid models which are

1. Have the speed of simpler approximations
2. The generalisability of well specified physical models

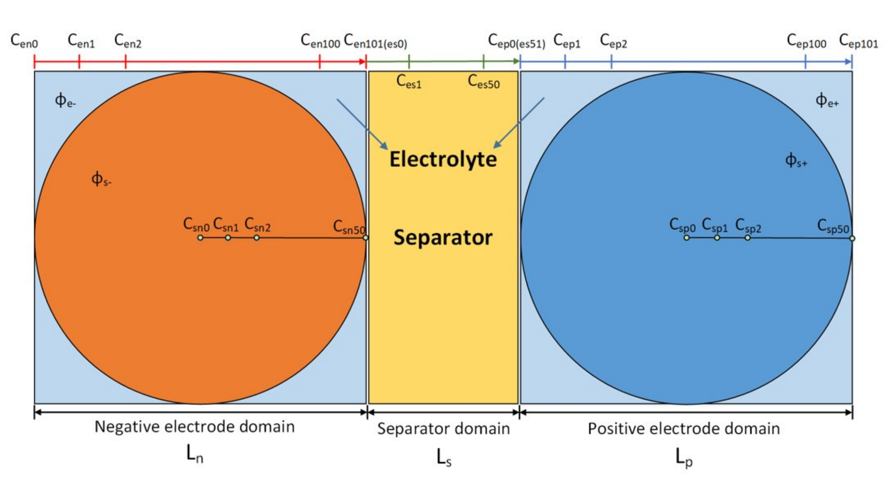
Mathematical reduction

Doyle-Fuller-Newman model which provides a comprehensive description of a cell but requires a complex geometry and systems of differential algebraic equations. Control applications, and easier optimisation, have motivated the creation of *single particle* simplified approximations.

Doyle-Fuller-Newman (DFN)



Approximate electrode as an equivalent particle and simplify electrolyte dynamics



Single Particle Model with electrolyte (SPMe)

The Multitask Model

Information sharing through latent Gaussian Processes

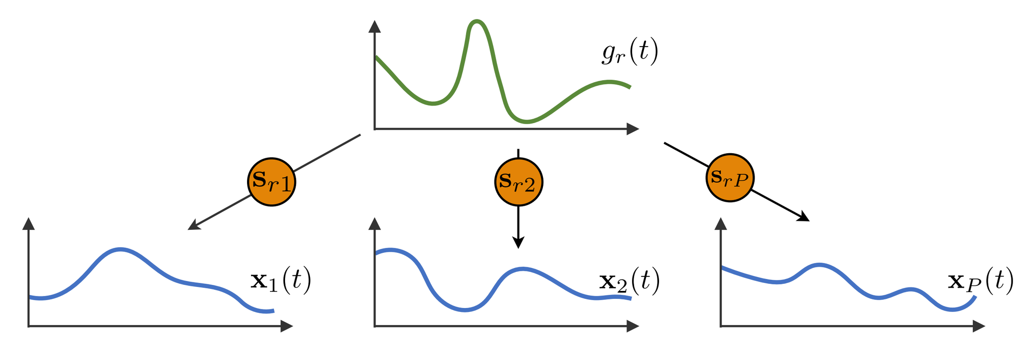
Realisations of the system under different operating conditions "tasks" are connected through a shared set of latent Gaussian forcing functions [Álvarez et al., 2009]

$$\frac{d\mathbf{x}_p(t)}{dt} = \mathbf{A}_p\mathbf{x}_p(t) + \sum_{r=1}^{R}\mathbf{s}_{pr}g_r(t)$$

Simplified mechanistic model

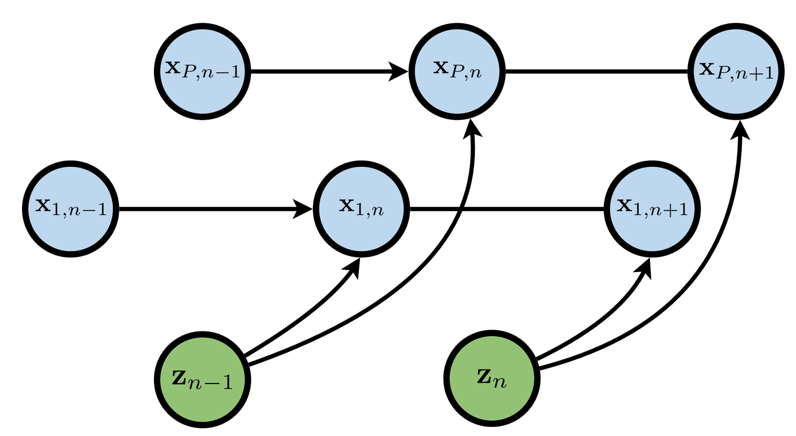
shared forcing functions

The *sensitivity* $\mathbf{s}_{rP}$ distributes common latent forces to each different task

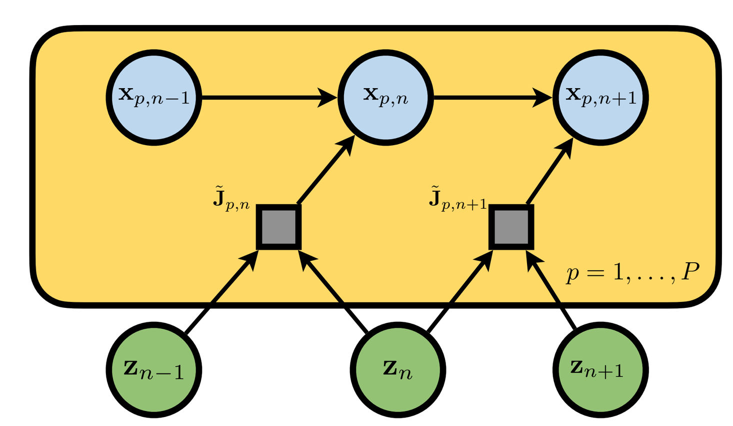


Information is shared through these common forces. If the latent forces also have a state space representation the whole system can be written as a single state space model — but with scaling like $\mathcal{O}(\text{num. timesteps} \times (PD)^3)$

Approximating the posterior



Graph for full model



Graph of decoupled approximation

Discretising we have a state space model

$$\mathbf{x}_{p,n} = e^{A_p\Delta t}\mathbf{x}_{p,n-1} + \mathbf{J}_{p,n} + \mathbf{W}_{p,n}$$
$$\mathbf{z}_n = e^{F\Delta t}\mathbf{z}_{n-1} + \mathbf{W}_{0,n}$$

With noise processes $\mathbf{W}_{p,n}$ independent for different p. Information is shared only through $\mathbf{J}_{p,n}$ which we approximate by a quadrature

$$\mathbf{J}_{p,n} = \int_{t_n}^{t_n+\Delta t} e^{\mathbf{A}_p(t_n+\Delta t-\tau)}\mathbf{S}_p\mathbf{z}(\tau)d\tau \approx \frac{\Delta t}{2}\left(e^{\mathbf{A}_p\Delta t}\mathbf{S}_p\mathbf{z}_{n-1} + \mathbf{S}_p\mathbf{z}_n\right)$$

Conditioned on $\mathbf{J}_{p,n}$ the task models decouple. The resulting approximation to the posterior then becomes a *product of experts*

$$p(\mathbf{X},\mathbf{Z}\mid\mathbf{Y}) \propto p(\mathbf{Y},\mathbf{X},\mathbf{Z}) \approx \prod_{p=1}^{P}\tilde{p}(\mathbf{Y}_p,\mathbf{X}_p,\mathbf{Z}) \propto \prod_{p=1}^{P}\tilde{p}(\mathbf{X}_p,\mathbf{Z}\mid\mathbf{Y})$$

Product of the posteriors for the independent decoupled task models

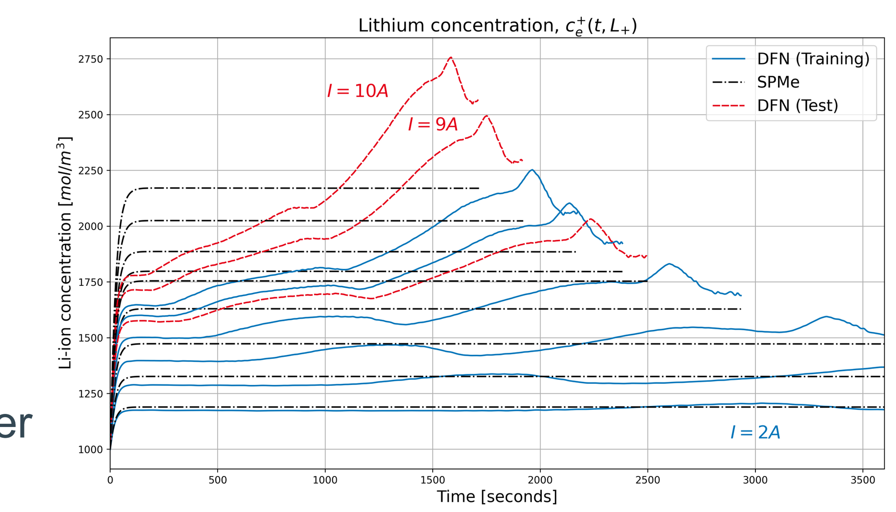
with each individual expert filtered independently leading to a total scaling like $\mathcal{O}(\text{num. timesteps} \times P \times D^3)$

Experiments

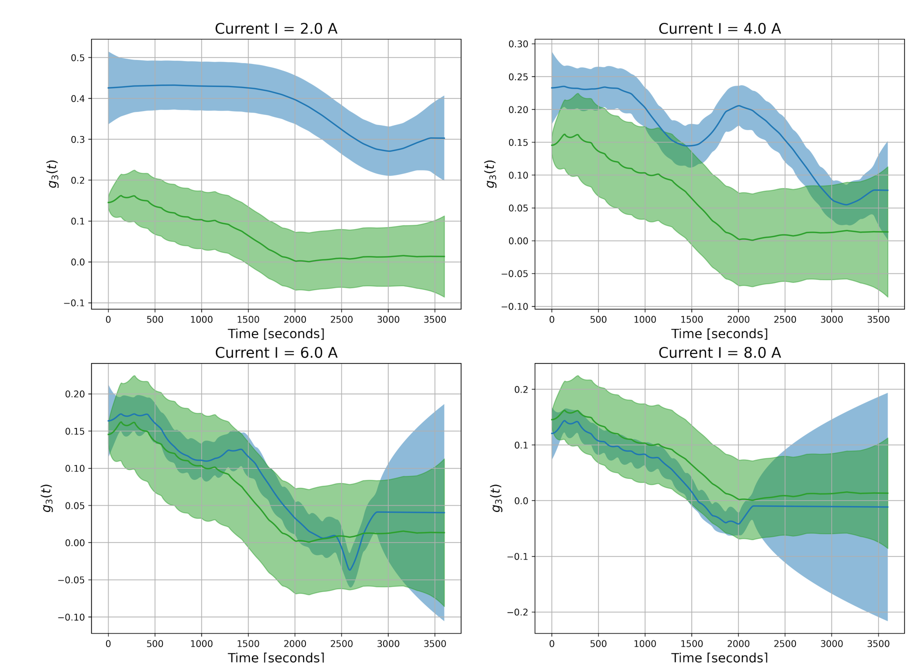
Data generation

Simulate trajectories from the DFN, the true model, at different rates of discharge from 2 — 8 Amps. Also shown are the SPMe trajectories at the same discharge. SPMe works well at slow discharge, but diverges during faster discharging when the behaviour becomes more nonlinear

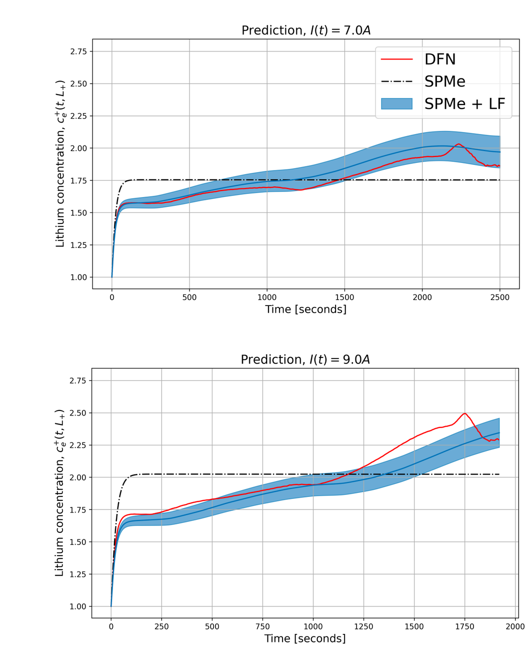
We learn the common latent forces and then test on new currents, both interpolating and the harder problem of extrapolation .



Results



(a) Individual task experts



(b) Predicted concentration

(a) Learned posterior distribution for one of the latent forces represents the predictive distribution for each expert, and the posterior from the joint task expert. (b) Shows the results of the predicted distribution relative to the ground truth DFN, and the simple mechanistic SPMe approximation alone.

Discussion

Our goal has been to create a single model of a dynamical system which can share information at different operating conditions, to maintain tractability we have introduced an approximate decoupling. Future work will aim to quantify this approximation error.

References

Fuller et al (1994). "Simulation and Optimisation of the Dual Lithium Ion Insertion Cell". In: Journal of the Electrochemical Society 141(1), pp. 1–10.

Álvarez et al (2009). "Latent force models." In: PMLR 20, pp. 9–16

Simo Säarkä and Arno Solin (2019) "Applied Stochastic Differential Equations". Cambridge University Press.

Sulzer et al (2020) "Python battery mathematical modelling (PyBaMM)." In: ECSarXiv.