1 Background

In physical applications, dynamical models and observational data play dual roles in prediction and uncertainty quantification, each representing sources of incomplete and inaccurate information. In data rich problems, first-principle physical laws constrain the degrees of freedom of massive data sets, utilizing our prior insights to complex processes. Respectively, in data sparse problems, dynamical models fill spatial and temporal gaps in observational networks. The dynamical chaos characteristic of these process models is, however, among the primary sources of forecast uncertainty in complex physical systems. Observations are thus required to update predictions where there is sensitivity to initial conditions and uncertainty in model parameters. My research lies in this intersection of data and dynamical instability — under the project, **Reduced subspace in big data treatment: A new paradigm for efficient geophysical Data Assimilation**, I have utilized the interplay of chaotic dynamics with Bayesian inference to develop novel filtering methodology for improved uncertainty quantification in physical applications.

Tracking the state of an evolving random process with incomplete and noisy observations is a central theme of Bayesian inference and sequential data assimilation in the geosciences. The complete solution to this state estimation problem is the posterior error distribution, characterized by the forward evolved prior, updated with respect to its likelyhood given the available observational data at the current time. While the theoretical solution may be obtained by evolving the Chapman-Kolmogorov equation for the prior distribution and updating with the appropriate likelyhood function, this is computationally infeasible for many applications: in numerical weather prediction, the dimension of observational data is typically $\mathcal{O}(10^7)$ and the model state dimension is typically $\mathcal{O}(10^9)$, prohibiting direct computation. Successful approaches estimating the physical state of the atmosphere or ocean have thus focused on the accurate, and computationally efficient, representations of the mean or mode of the posterior, in Kalman filtering methods [12] and variational techniques respectively. Rigorously estimating the uncertainty of these state estimates, and even faithfully representing the covariance, remains intractable due to the extremely large dimension of models and volume of data.

1.1 Assimilation in the Unstable Subspace

While the butterfly effect renders the forecasting problem inherently volatile, chaotic dynamics also put strong constraints on the evolution of errors. It is well understood in the weather prediction community that the growth of forecast uncertainty is confined to a much lower dimensional subspace corresponding to the directions of rapidly growing perturbations [16] — this is characterized by the unstable-neutral manifold of the state being tracked, with dimension equal to the number of non-negative Lyapunov exponents. Anna Trevisan and her collaborators proposed a data assimilation methodology known as Assimilation in the Unstable Subspace (AUS) [3, 17, 18, 13, 14] to take advantage of this feature of physical models: an adequate representation of the forecast uncertainty can be made computationally feasible by reducing this representation to the lower dimensional space in which the uncertainty is concentrated.

My collaborators and I have put forward a mathematical framework for AUS, where we have demonstrated that in linear models, with Gaussian observational error, the forecast error covariance collapses to the span of the unstable and neutral, backward Lypunov vectors [11]. Moreover, we proved that the exact Kalman filter converges to reduced rank filters only representing the unstable-neutral subspace [2]. For applications where model error is negligible, our work validates the central hypothesis of AUS, that the problem of uncertainty quantification can be reduced to a lower dimensional subspace.

1.2 Current Research

1.2.1 Dimensional reduction in data assimilation

My current research extends the framework of AUS to the presence of stochastic model errors, in the form of additive noise. In the presence of model errors, the forecast error covariance will generally be of full rank. However, I have proven that under generic conditions the forecast error in the stable subspace is uniformly bounded in time *independently of filtering*, by the stable dynamics alone. In this work I have likewise provided novel necessary conditions for boundedness of forecast error in the Kalman filter, and explored the fundamental role of transient instability in the stable subspace for filter divergence with sub-optimal filters — these results are currently in review by the SIAM/ ASA Journal on Uncertainty Quantification [6]. I demonstrated that AUS methodology can be extended to stochastic dynamics: if corrections to forecast error account for modes with transient instabilities then the error in the remaining stable subspace can remain unfiltered.

However, a sub-optimal filter which neglects the stable subspace may be overly confident, and the estimated error much smaller than the true error: uncertainty transmits upward through the Lyapunov filtration via the QR dynamics in the backward vectors, and a representation of the error in the unstable-neutral subspace must account for this contribution. It is typically the case for the ensemble Kalman filter that the ensemble dimension is much lower than the full state dimension; assuming a sub-optimal filter as such, I have derived the recursion for the contribution of unfiltered error into the ensemble subspace in linear models. In nonlinear models, the recursion is approximate but gives insight to covariance inflation techniques to account for representation errors of the forecast uncertainty for sup-optimal extended and ensemble Kalman filters. These results are in preparation [8] and will be submitted to the special issue of Nonlinear Processes in Geophysics for the proceedings of the Symposium honoring the legacy of Anna Trevisan. Additional results [7], extending the bounds developed in [6] to the Kalman smoother, are also to be submitted in 2018.

1.2.2 Model reduction for multiscale electric grids

High dimensional data assimilation techniques have a long history in the weather forecasting community, but there is increasing interest their application in other physical sciences. In my graduate internship at the Center for Nonlinear Studies of the Los Alamos National Laboratory, I studied data driven methodology for physics preserving reductions of electric grid models. The multiscale nature of the future power grid, including local distribution of energy through rooftop solar, and dynamic control of microgrids, demands novel approaches to model representation, state estimation and control. In my internship, I developed network reduction algorithms and implemented these on a real test case, multiscale electric grid as a proof of concept. My Matlab repository, Javascript visualizations and the toy data I generated have been published in my Github and our results are currently in review by the SIAM Journal on Multiscale Modeling and Simulation [9]. In our ongoing collaboration, our next objective is to utilize our reduced electric grid model for online state and parameter estimation, applying ensemble data assimilation techniques in this novel application.

1.2.3 Geometric dynamics & stability analysis

While my postdoctoral experience has focused on data assimilation methodology, this was one among several projects I engaged with as a PhD student, including numerical methods for determining the stability of steady states for reaction diffusion equations. Stable waves and solitons describe the long time behavior of physical systems, and are those solutions which are the most realistic models for noisy processes in nature. Determining the spectral stability of steady states has historically been studied with the Evans function, and in my thesis work I proved an analytic reduction to the Evans function formulation in terms of the Hopf and Stiefel bundles [5, 4]. In particular, parallel transport with respect to the system's spectral parameter, varying over a closed contour, can be used to count the multiplicity of the eigenvalues enclosed by the contour. Between June and August 2017, I supervised Armand Vic, a visiting Masters student from the École Normale Supérieure de Rennes under an Erasmus

Plus graduate research training grant. Armand studied numerical extensions of my thesis work, focusing on the numerical calculation of the geometric phase in the Stiefel bundle.

2 Future Research

In my future research program, I intend to follow several different lines, depending on my students' mathematical sophistication, computational aptitude and interest in applications — I will be eager to mentor students in both applied and theoretical research in the newly formed PhD program at UNR.

2.1 A dynamical framework for stochastic filtering

The long term objective in building a framework for AUS is to develop computationally feasible, and rigorous uncertainty quantification for high dimensional models. Both Kalman filters and variational techniques make implicit assumptions about the linearity and Gaussianity of the error statistics which are highly unrealistic for the physical applications they are designed for. However, fully nonlinear data assimilation, e.g. Markov Chain Monte Carlo techniques, suffer from the inability to sample meaningful statistics from the high dimensional spaces in weather forecasting. It is of interest, therefore, to extend my mathematical results to fully Bayesian filters and to rigorously characterize the error statistics for stochastic filters in nonlinear and stochastic dynamical systems. A rigorous description of the nonlinear error statistics in terms of the lower dimensional attractor and its manifolds may improve sampling design and performance of Bayesian filters in high dimensional data assimilation. This project involves my longest collaborators: Alberto Carrassi[1], Marc Bocquet[2], Christopher KRT Jones[3] and Amit Apte[4]. We have coordinated our efforts in re-

[1] Nansen Environmental and Remote Sensing Center

[2] Marc Bocquet, CEREA, joint laboratory École des Ponts ParisTech and EDF R&D

[3] University of North Carolina at Chapel Hill

[4] International Centre for Theoretical Sciences, Bengaluru

search and funding opportunities and we intend to collaborate on parallel proposals for the NSF solicitation PD 16-8069, Computational and Data-Enabled Science and Engineering in Mathematical and Statistical Science, and the Norwegian Research Council - Program FRINATEK/IKTPLUSS to support our ongoing research efforts.

2.2 State estimation and stochastic control in smart grids

Purely statistical machine learning techniques are particularly useful in prediction where data is abundant and models of the process are inferior to observations. However, neither is the case for the electric grid where Phasor Measurement Units (PMUs) offer spatially sparse measurements and electric transmission is guided by physical power flow laws. For these reasons, the rigorous incorporation of nonlinear, time varying process models is strong advantage of data assimilation methodology for the problems of state estimation and fault detection in the electric grid. The ensemble Kalman filter has shown itself to be a robust algorithm for both state and parameter estimation in high dimensional physical systems. Recent research [10, 15] has studied the use of mixed Kalman filter, machine learning methodology for joint state and stochastic parameter estimation where stochastic parametrization represents unresolved physical processes inferred from noisy data. With the incorporation of renewable energy, e.g., wind and solar, optimal power flow and grid control is moving into a stochastic optimization framework [1], where renewable generation is treated as a stochastic parametrization. I am particularly interested in studying the use of joint ensemble Kalman filter/ machine learning algorithms for state and parameter estimation in this context, where dynamics of the electric grid may be explicitly resolved and renewable generation stochastically simulated.

My established results with Michael Chertkov[5] and my ongoing collaboration with the Nansen Environmental and Remote Sensing Center gives a multidisciplinary perspective and complimentary skill base for the problem of state estimation and stochastic optimization for the electric grid. The ensemble Kalman filter was in-

[5]Center for Nonlinear Studies, Los Alamos National Laboratory

vented at the Nansen Center, and the in-house expertise can provide valuable operational perspective to algorithm design; the Center for Nonlinear Studies has likewise pioneered the study of smart grid design and control with machine learning methodology. My program is thus in a unique position to seek funding through the recent NSF solicitation 17-521, Algorithms for Modern Power Systems, and in November I will visit the Center for Nonlinear Studies to meet my collaborators in smart grid research to plan our proposal. With my collaboration network, I will lead a program which utilizes our mutual expertises in designing joint ensemble Kalman filter / machine learning techniques for state estimation and stochastic control in the context of renewable energy.

2.3 Optimal control & autonomous vehicles

While I am new to the application of autonomous vehicle guidance, the techniques of data assimilation, in particular the extended Kalman filter, have a long history in this domain. For this reason, I am particularly excited to establish new collaborations with the Nevada Advanced Autonomous Systems Innovation Center. It will be a top priority in my research program to co-advise students involved in engineering applications with autonomous vehicles and to coordinate course offerings and seminars in data assimilation and optimal control.

2.4 Fit with UNR

The heart of my work is in finding rich mathematical and statistical problems at the intersection of data and models, with applications in the physical sciences — my wide range of scientific interests, collaborations and application domains makes me well suited for the research environment at University of Nevada, Reno. I will be eager to mentor students with diverse interests, and backgrounds, in mathematics and statistics ranging from methodology and algorithm design to the application of these techniques in weather forecasting and engineering domains. My work will compliment the interests of Eric Olson, who is actively involved in the application of data assimilation in chaotic models and Ilya Zaliapin using statistical and dynamical systems techniques in geophysical problems. Methodologically, I share research interests with Yinghan Chen, developing Bayesian inference and sampling methodology, and more broadly, with Deena Schmidt and Paul Hurtado in the analysis of stochastic dynamics in physical systems and networks. I would greatly look forward to developing my research program in such a dynamic department, and to support its continuing growth into a major research center. I thank the committee for its consideration.

References

[1] D. Bienstock, M. Chertkov, and S. Harnett. Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review*, 56(3):461–495, 2014.

[2] M. Bocquet, K.S. Gurumoorthy, A. Apte, A. Carrassi, **C. Grudzien**, and C.K.R.T. Jones. Degenerate Kalman filter error covariances and their convergence onto the unstable subspace. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):304–333, 2017.

[3] A. Carrassi, A. Trevisan, L. Descamps, O. Talagrand, and F. Uboldi. Controlling instabilities along a 3DVar analysis cycle by assimilating in the unstable subspace: a comparison with the EnKF. *Nonlinear Processes in Geophysics*, 15:503–521, 2008.

[4] **C. Grudzien**. The instability of the Hocking–Stewartson pulse and its geometric phase in the Hopf bundle. *Journal of Computational and Applied Mathematics*, 307:162–169, 2016.

[5] **C. Grudzien**, T.J. Bridges, and C.K.R.T. Jones. Geometric phase in the Hopf bundle and the stability of non-linear waves. *Physica D: Nonlinear Phenomena*, 334:4–18, 2016.

[6] **C. Grudzien**, A. Carrassi, and M. Bocquet. Asymptotic forecast uncertainty and the unstable subspace in the presence of additive model error. *arXiv preprint arXiv:1707.08334*, 2017.

[7] **C. Grudzien**, A. Carrassi, and M. Bocquet. 4D posterior bounds for the Kalman smoother with additive model error. *In Preparation*, 2018.

[8] **C. Grudzien**, A. Carrassi, and M. Bocquet. Chatoic dynamics and the role of covariance inflation for reduced rank Kalman filters with model error. *In Preparation*, 2018.

[9] **C. Grudzien**, D. Deka, M. Chertkov, and S.N. Backhaus. Structure-& physics-preserving reductions of power grid models. *arXiv preprint arXiv:1707.03672*, 2017.

[10] D. Dreano, P. Tandeo, M. Pulido, B. Ait-El-Fquih, T. Chonavel, and I. Hoteit. Estimating model-error covariances in nonlinear state-space models using Kalman smoothing and the expectation–maximization algorithm. *Quarterly Journal of the Royal Meteorological Society*, 143(705):1877–1885, 2017.

[11] K.S. Gurumoorthy, **C. Grudzien**, A. Apte, A. Carrassi, and C.K.R.T. Jones. Rank deficiency of Kalman error covariance matrices in linear time-varying system with deterministic evolution. *SIAM Journal on Control and Optimization*, 55(2):741–759, 2017.

[12] PL Houtekamer and F. Zhang. Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 144(12):4489–4532, 2016.

[13] L. Palatella, A. Carrassi, and A. Trevisan. Lyapunov vectors and assimilation in the unstable subspace: theory and applications. *J. Phys. A: Math. Theor.*, 46:254020, 2013.

[14] L. Palatella and A. Trevisan. Interaction of Lyapunov vectors in the formulation of the nonlinear extension of the Kalman filter. *Phys. Rev. E*, 91:042905, 2015.

[15] M. Pulido, P. Tandeo, M. Bocquet, A. Carrassi, and M. Lucini. Stochastic parameterization identification using ensemble Kalman filtering combined with expectation-maximization and Newton-Raphson maximum likelihood methods. *arXiv preprint arXiv:1709.07328*, 2017.

[16] Z. Toth and E. Kalnay. Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, 125(12):3297–3319, 1997.

[17] A. Trevisan, M. D'Isidoro, and O. Talagrand. Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension. *Q. J. R. Meteorol. Soc.*, 136:487–496, 2010.

[18] A. Trevisan and L. Palatella. On the kalman filter error covariance collapse into the unstable subspace. *Nonlinear Processes in Geophysics*, 18(2):243–250, 2011.