Non-uniform decay of predictability in stochastic oscillatory models

N. Barlas\textsuperscript{2}  K. Josić\textsuperscript{2}  S. Lapin\textsuperscript{1}  I. Timofeyev\textsuperscript{2}

\textsuperscript{1}Department of Mathematics
Washington State University

\textsuperscript{2}Department of Mathematics
University of Houston

Jan.23.08 / Applied Math Seminar
Outline

1. Introduction
2. Relative entropy for SDEs
   - Definition of Relative Entropy
   - Properties of Relative entropy
3. The stochastic linear oscillator
   - The decay of relative entropy
   - Return of skill for marginal entropies
4. Nonuniform decay of $R(t)$ in general systems
   - Decay of relative information for nonlinear oscillators
   - Stochastically Perturbed Duffing Equation
The weather prediction is extremely sensitive to the specification of initial conditions

Detailed forecasts are impossible beyond 2 weeks limit

Due to this atmospheric uncertainty statistical prediction involving ensemble of possible projections has become common

The ensemble spread may vary considerably meaning that some predictions are more reliable than others

Need measure to indicate the usefulness of forecast
Definition of relative entropy.

Due to uncertainty the values of initial conditions are given by certain probability distribution $p$. This distribution evolves in time and approaches an equilibrium distribution $q$. Need to measure usefulness of a prediction, i.e. how much information is added by availability of prediction.

The relative entropy gives the information loss sustained by assuming climatology when prediction is available.

$$ R(p(\vec{x}, t), q(\vec{x})) = R(t) = \int p(\vec{x}, t) \log \left( \frac{p(\vec{x}, t)}{q(\vec{x})} \right) d\vec{x}. $$

$q(\vec{x})$ the invariant (climatological) distribution,
$p(\vec{x}, t)$ the probability density corresponding to the ensemble.
Properties of Relative entropy.

Relative entropy satisfies three important properties:

- Invariant under well behaved non-linear transformations of state variables.
- Non-negative.
- Declines monotonically in time for Markov processes.

The fact that relative entropy decreases in time can be interpreted as a decline in the utility of a prediction, or *skill*. 
Relative entropy for Gaussian distribution

\[ R = \frac{1}{2} \left( \log \left( \frac{\det(\sigma_q^2)}{\det(\sigma_p^2)} \right) + \text{Tr}(\sigma_p^2(\sigma_q^2)^{-1}) + (\mu_p)^T(\sigma_q^2)^{-1}(\mu_p) - n \right). \]

For the Gaussian case the relative entropy can be decomposed into

- the signal component \( \Rightarrow \) the difference in the means of the two distributions,
- the dispersion \( \Rightarrow \) the difference in the variances of the two distributions.

The signal and dispersion terms are analogous to the Anomaly Correlation Coefficient and Root Mean Squared Error, respectively.
Fokker-Planck equation

\[
\frac{\partial p(\vec{x}, t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i} [A_i(\vec{x})p(\vec{x}, t)] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij}(\vec{x})p(\vec{x}, t)].
\]

- \(A(\vec{x}, t)\) is the drift vector,
- \(B(\vec{x}, t)\) is the diffusion matrix,
- the distribution \(\rho(\vec{x}, 0)\) provides the initial data for the FP equation.
- assume that the system has a unique equilibrium solution \(q(\vec{x})\).
Recall the definition of Relative entropy:

$$R(t) = \int p(\vec{x}, t) \log \left( \frac{p(\vec{x}, t)}{q(\vec{x})} \right) d\vec{x}. \tag{1}$$

Using the definition we obtain

$$\frac{dR}{dt} = \int d\vec{x} \left[ \frac{\partial p(\vec{x}, t)}{\partial t} \left( \log p(\vec{x}, t) + 1 - \log q(\vec{x}) \right) - \frac{\partial q(\vec{x})}{\partial t} \left( \frac{p(\vec{x}, t)}{q(\vec{x})} \right) \right]. \tag{2}$$
Drift and diffusion contributions to $dR/dt$:

$$
\left( \frac{dR}{dt} \right)_{\text{drift}} = \sum_i \int d\vec{x} \frac{\partial}{\partial x_i} \left[ -A_i p(\vec{x}, t) \log \left( \frac{p(\vec{x}, t)}{q(\vec{x})} \right) \right].
$$

$$
\left( \frac{dR}{dt} \right)_{\text{diff}} = -\frac{1}{2} \sum_{i,j} \int d\vec{x} p(\vec{x}, t) B_{ij} \left[ \frac{\partial}{\partial x_i} \log \frac{p(\vec{x}, t)}{q(\vec{x})} \right] \left[ \frac{\partial}{\partial x_j} \log \frac{p(\vec{x}, t)}{q(\vec{x})} \right].
$$

It can be shown that

$$
\left( \frac{dR}{dt} \right)_{\text{drift}} = 0 \quad \text{and} \quad \left( \frac{dR}{dt} \right)_{\text{diff}} \leq 0.
$$
Several low-dimensional models of El Niño/Southern Oscillation (ENSO) have been proposed in recent years. One common feature among them is the oscillatory behavior of solutions. As a prototype behavior we consider a simple two-dimensional model described in T. Kestin et al. "Time frequency variability of ENSO and stochastic simulations"

\[ dx_1 = \alpha x_1 dt + \beta x_2 dt, \]

\[ dx_2 = \gamma x_1 dt + \delta x_2 dt + \varepsilon dW, \]
The Fokker-Planck equation is

$$\frac{\partial p(\vec{x}, t)}{\partial t} = - \sum_i \frac{\partial}{\partial x_i} [A_i(\vec{x}) p(\vec{x}, t)] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij}(\vec{x}) p(\vec{x}, t)] ,$$

with

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon \end{pmatrix} .$$

This equation can be solved analytically assuming a deterministic initial condition $p(\vec{x}, 0) = \delta_{\vec{x}^0}(\vec{x})$ where $\vec{x}^0 = (x_1^0, x_2^0).$
The solution at time $t$ is a Gaussian with mean

$$
\begin{pmatrix}
\bar{x}_1 \\
\bar{x}_2
\end{pmatrix}
= e^{At}
\begin{bmatrix}
x_1(0) \\
x_2(0)
\end{bmatrix},
$$

and covariance matrix

$$
\begin{bmatrix}
\langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle \\
\langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle
\end{bmatrix}
= \int_0^t dt' e^{A(t-t')}
\begin{pmatrix}
0 & 0 \\
0 & \varepsilon^2
\end{pmatrix}
\begin{pmatrix}
0 & 0 \\
0 & \varepsilon^2
\end{pmatrix}e^{A^T(t-t')}.
$$
Figure: Left: The equilibrium distribution $q(\vec{x})$ for the linear oscillator. Right: Average relative entropy. Parameters: \(\alpha = 0.4573, \beta = 0.2435, \gamma = \delta = -1.0852, \text{ and } \varepsilon = 0.1\).
For the stochastic linear oscillator the diffusion part of the corresponding Fokker-Planck equation reduces to

\[
\left( \frac{dR}{dt} \right)_{\text{diff}} = - \frac{\varepsilon}{2} \int d\vec{x} \ p(\vec{x}, t) \left[ \frac{\partial}{\partial y} \left( \log(p(\vec{x}, t)/q(\vec{x})) \right) \right]^2.
\]
For the stochastic linear oscillator the diffusion part of the corresponding Fokker-Planck equation reduces to

\[
\left( \frac{dR}{dt} \right)_{\text{diff}} = -\frac{\varepsilon}{2} \left[ \int d\vec{x} \ p(\vec{x}, t) \left( \frac{\partial \log p(\vec{x}, t)}{\partial y} \right)^2 \right]_{l_1} - 2 \int d\vec{x} \ p(\vec{x}, t) \frac{\partial \log q(\vec{x})}{\partial y} \right]_{l_2} \right. \\
+ \left. \int d\vec{x} \ p(\vec{x}, t) \left( \frac{\partial \log q(\vec{x})}{\partial y} \right)^2 \right]_{l_3} \equiv l_1 + l_2 + l_3.
\]
Figure: Left: Relative entropy for a particular initial condition \((x_1^0, x_2^0) = (1, 1)\) (solid), and the contribution to the relative entropy due to the signal term (dashed). Right: Behavior of \((\frac{dR}{dt})_{\text{diff}}\) (solid) and \(I_3\) term (dashed) in time for an initial condition \((x_1^0, x_2^0) = (1, 1)\).
Figure: Probability density functions of the transient distribution (dashed line) and the equilibrium distribution (solid line) for times $t=17.5$ (left), $t=25$ (center) and $t=35$ (right).
Return of skill for marginal entropies

The marginal relative entropies do not necessarily decay in time.

Consider *conditional relative entropies*

\[
R_{x_2|x_1}(t) = R(p(x_2|x_1, t), q(x_2|x_1)) \\
= \int p(x_1, t) \int p(x_2|x_1, t) \log \frac{p(x_2|x_1, t)}{q(x_2|x_1)} \, dx_1 \, dx_2.
\]

- \( p(x_2|x_1, t) \) \( \mapsto \) the conditional distribution of \( x_2 \) at time \( t \) given \( x_1 \)
- \( R_{x_1|x_2}(t) \) \( \mapsto \) excess information provided by the marginal distribution \( p(x_2|x_1, t) \) over \( q(x_2|x_1) \).

\[
R(t) = R_{x_2|x_1}(t) + R_{x_1}(t).
\]
**Figure**: Left: means of the variables $|x_1|$ (solid line) and $|x_2|$ (dashed line). Right: marginal relative entropies $R_{x_1}(t)$ (solid line) and $R_{x_2}(t)$ (dashed line) and $R_{x_1|x_2}(t)$ (solid-star line) and $R_{x_2|x_1}(t)$ (dashed-star line) for an initial condition $(x_1^0, x_2^0) = (1, 1)$. 
The nonuniform decay of relative entropy occurs whenever the main mass of the distribution $p(\bar{x}, t)$ approaches, and then diverges from the main mass of the stationary distribution $q(\bar{x})$.

Oscillations in the marginal relative entropies occur when such divergence occurs in the marginal distributions.
Supercritical **Hopf bifurcation**: 

\[
\begin{align*}
    dx_1 &= \mu x_1 \, dt - c\omega x_2 \, dt + \Theta x_1(x_1^2 + c^2 x_2^2) \, dt + \varepsilon dW_1, \\
    dx_2 &= \frac{1}{c} \left(\omega x_1 + c\mu x_2 + c\Theta x_2(x_1^2 + c^2 x_2^2)\right) \, dt + \varepsilon dW_2.
\end{align*}
\]

In the absence of white noise this system has a stable periodic orbit

\[
    x_1(t) = \sqrt{-\frac{\mu}{\Theta}} \cos(\omega t + \phi_0), \quad x_2(t) = \frac{1}{c} \sqrt{-\frac{\mu}{\Theta}} \sin(\omega t + \phi_0).
\]
Introduction
Relative entropy for SDEs
The stochastic linear oscillator
Nonuniform decay of $R(t)$ in general systems
Summary

Decay of relative information for nonlinear oscillators
Stochastically Perturbed Duffing Equation

 transient distribution
Equilibrium distribution

$R(t)$

$R_x(t)$

Barlas, Josić, Lapin, Timofeyev,
Predictability in stochastic oscillatory models
Duffing equations

The Duffing equations.

\[ dx_1 = x_2 \, dt + \varepsilon \, dW_1, \]

\[ dx_2 = (x_1 - x_1^3 - \gamma x_2 + \beta x_1^2 x_2) \, dt + \varepsilon \, dW_2, \]

For \( \varepsilon = 0 \) and parameters \( \gamma = 0.4 \) and \( \beta = 0.497 \) this system has an attracting double homoclinic cycle to the saddle point at the origin.
**Figure:** Left: Homoclinic loop for Duffing equation with $\varepsilon = 0$. Right: Contour plot of probability density function for $\varepsilon = 0.01$. 

Barlas, Josić, Lapin, Timofeyev, Predictability in stochastic oscillatory models
Recall: the decay in relative entropy is due to the \((dR/dt)_{\text{diff}}\) term. For the Duffing model the diffusion term becomes

\[
\left( \frac{dR}{dt} \right)_{\text{diff}} = -\frac{\varepsilon}{2} \int dx_1 dx_2 \ p(x_1, x_2) \sum_{i=1,2} \left[ \frac{\partial}{\partial x_i} \left( \log \frac{p(x_1, x_2)}{q(x_1, x_2)} \right) \right]^2.
\]
Figure: Top: Probability density function for time=5 (left), time=10 (middle) and time=15 (right). Bottom Left: Full relative entropy $R(t)$, Bottom Right: marginal relative entropies $R_{x_1}(t)$ (solid) and $R_{x_2}(t)$ (dashed) in simulations of Duffing equation with initial ensemble centered at $(x_1^0, x_2^0) = (0.25, 0.25)$ and $\varepsilon = 0.01$. 
For the stochastic Duffing equation the behavior of the \( \left( \frac{dR}{dt} \right)_{\text{diff}} \) is more complicated than in the case of linear oscillator. Namely, the value of \( \left( \frac{dR}{dt} \right)_{\text{diff}} \) depends not only on the means, but also on all terms involving variances of the invariant and transient distributions.

**Figure:** Left: mean in \( x_1 \) (dashed line), mean in \( x_2 \) (solid line), Right: variance in \( x_1 \) (dashed line), variance in \( x_2 \) (solid line) in simulations of the stochastic Duffing equation with \( \varepsilon = 0.01 \). (Horizontal lines show equilibrium variances).
The averaged predictability decays exponentially in time for all three considered models.

However, for particular ensemble simulation full relative entropy decays non-monotonically and there is a return of skill for the marginal entropies.
THANK YOU