# Stochastically perturbed bred vectors in single-scale systems

Brent Giggins [a] * and Georg A. Gottwald [a]*

[a]*School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia*

*Correspondence to: E-mail: georg.gottwald@sydney.edu.au; brent.giggins@sydney.edu.au

**The breeding method is a computationally cheap procedure to generate initial conditions for ensemble forecasting which project onto relevant synoptic growing modes. Ensembles of bred vectors, however, often lack diversity and align with the leading Lyapunov vector, which severely impacts their statistical reliability. In previous work we developed stochastically perturbed bred vectors (SPBVs) and random draw bred vectors (RDBVs) in the context of multi-scale systems. Here we explore when this method can be extended to systems without scale separation, and examine the performance of the stochastically modified bred vectors in the single scale Lorenz 96 model. In particular, we show that the performance of SPBVs crucially depends on the degree of localisation of the bred vectors. It is found that, contrary to the case of multi-scale systems, localisation is detrimental for applications of SPBVs in systems without scale-separation when initialised from assimilated data. In the case of weakly localised bred vectors, however, ensembles of SPBVs constitute a reliable ensemble with improved ensemble forecasting skills compared to classical bred vectors, while still preserving the low computational cost of the breeding method. RDBVs are shown to have superior forecast skill and form a reliable ensemble in weakly localised situations, but in situations when they are strongly localised they do not constitute a reliable ensemble and are over-dispersive.**

## 1. Introduction

The chaotic nature of the atmosphere and the climate system, and its sensitivity to small uncertainties in the initial conditions may render single forecasts meaningless. Probabilistic forecasts, which instead are derived from an ensemble of forecasts, have become standard in numerical weather forecasting, providing a Monte-Carlo estimate of the probability density function (Epstein 1969; Leith 1974; Leutbecher and Palmer 2008). Such ensemble forecasts issue the most probable forecast alongside measures of its uncertainty. A key question is how to initialise the ensemble members. There exist several methods to generate such ensembles, using singular vectors (Lorenz 1965; Palmer 1993), bred vectors (Toth and Kalnay 1993, 1997), analysis ensembles from ensemble Kalman filters (Evensen 1994; Houtekamer and Mitchell 1998; Wang and Bishop 2003; Buizza *et al.* 2005), and more recently model generated analogs (Atencia and Zawadzki 2017). In this work we consider bred vectors and the so called "breeding method" which constitutes a computationally very attractive method to produce an ensemble of initial conditions introduced by Toth and Kalnay (1993, 1997). In this method initial conditions are generated from finite perturbations, the bred vectors (BVs), which encapsulate information about fast growing modes. Such fast growing initial conditions are then likely to be pre-images of states of high probability. Bred vectors have been successfully implemented for more than a decade since 1992 by the National Centre for Environmental Prediction (NCEP) for their operational 1-15 day ensemble forecasts. Applications range from ENSO prediction (Cai *et al.* 2003; Cheng *et al.* 2010), seasonal-to-interannual forecasting in coupled general circulation models (CGCMs) (Yang *et al.* 2009) to forecasting weather and climate on Mars (Newman *et al.* 2004; Greybush *et al.* 2013).

In the breeding method a control trajectory alongside an ensemble of nearby trajectories is generated. The ensemble members are initialised from perturbed initial conditions with finite perturbation size $\delta$ from the initial condition of the control trajectory. Different as for Lyapunov vectors, all ensemble members are propagated with the full nonlinear model. The perturbed trajectories are periodically rescaled to a specified finite-size distance $\delta$ away from the control trajectory, to avoid saturation of instabilities. Bred vectors are defined as the difference at the time of rescaling of the perturbed trajectories and the control trajectory. The perturbation size is often thought of as a a filter of small scale instabilities, in the sense that BVs are insensitive to very fast growing instabilities

which typically are associated with small scale processes and which nonlinearly saturate at an amplitude smaller than $\delta$. Choosing $\delta$ appropriately allows the forecast to be tuned to specific instabilities of interest. For perturbation sizes of the order of $1-10\%$ of the natural variability in the atmosphere, BVs were found to significantly project onto baroclinic instabilities (Toth and Kalnay 1997; Corazza *et al.* 2003). Moreover, bred vectors allow for the prediction of regime changes for perturbation sizes in a certain range (Peña and Kalnay 2004; Evans *et al.* 2004; Norwood *et al.* 2013)).

Ideally bred vectors constitute a sufficiently diverse ensemble mediated by the provided stochasticity of saturated sub-synoptic processes (Toth and Kalnay 1997). Bred vector ensembles, however, may lack diversity in certain situations and most of the ensemble forecast variability may be captured by a single BV (Wang and Bishop 2003). Indeed, for a range of small perturbation sizes $\delta$ bred vectors align with the leading Lyapunov vector and the ensemble collapses to a single member. When this occurs, this reduction in ensemble diversity is hugely detrimental for the ability of a bred vector ensemble to reliably sample the forecast probability density function. To preserve the attractive features of BVs such as their low computational cost, several methods were proposed to increase the ensemble spread in BVs. In particular by orthogonalizing bred vectors (Annan 2004; Keller *et al.* 2010), by introducing stochasticity either via small random perturbations at each rescaling period (Greybush *et al.* 2013) or via stochastic backscattering (O'Kane and Frederiksen 2008), by rescaling using the geometric rather than the Euclidean norm (Primo *et al.* 2008; Pazó *et al.* 2011, 2013) or by changing the rescaling procedure based on the largest BV (Balci *et al.* 2012).

In recent work, Giggins and Gottwald (2019) proposed a method of stochastically perturbing BVs in the context of ensemble forecasts of the slow dynamics in multi-scale systems to alleviate the problem of small ensemble diversity, introducing stochastically perturbed bred vectors (SPBV). SPBVs were constructed, it was argued, to sample the conditional probability function of the system conditioned on the slow variables by multiplicatively randomising the fast BV components. The localised character of the fast BV components ensures that the perturbed initial conditions relax after a short transient to initial conditions which are close to those of their parent BV, with slow components being close in phase space to those of the original BV and with fast components being dynamically conditioned on the slow components. Hence, these stochastically perturbed BVs generate initial conditions which sample the probability density function conditioned on the slow synoptic state. SPBVs exhibit a markedly increased ensemble dimension, in particular for small but finite perturbation sizes. It was shown that the subsynoptic variability associated with SPBVs generated synoptic variability of the same order as suggested by the analysis fields. The increased diversity of SPBVs lead to an improved forecasting skill when compared to standard BVs. Important for probabilistic forecasts, SPBV ensembles were shown to be reliable in the sense that each ensemble member is equally likely to be closest to the truth. Furthermore, SPBVs were shown to be dynamically consistent and recover characteristic features of the temporal evolution of errors in chaotic dynamical systems. Additionally, random draw bred vectors (RDBVs), which are designed to sample from the marginal equilibrium density of the fast variables (and hence are not conditioned on the slow variables), were introduced. While RDBVs are not dynamically consistent and are typically over-dispersive, they were found to still have improved forecast skill over standard BVs.

In this work we explore if, and under which conditions, the ideas proposed in Giggins and Gottwald (2019) for multi-scale dynamics can be applied to the situation of general dynamical systems without time-scale separation. Single-scaled models such as the quasi-geostrophic equations are often used to study the slow large-scale dynamics of the atmosphere-ocean system. Filtering out fast small-scale processes has the computational advantage of avoiding the numerical difficulties associated with multi-scale systems. We study here in what way stochastically modifying classical BVs may help in using the attractive features of BVs, such as their low computational coast and their dynamic adaptivity in the sense that they resemble realistic error growth. In the realistic situation when the state of the atmosphere is given by the analysis output from a data assimilation procedure, a good forecast ensemble has to satisfy two constraints: It has to evolve into likely future states, and it has to account for the uncertainty of the analysis used to generate the ensemble. In the case of multi-scale dynamics these issues were resolved by generating the necessary small synoptic uncertainty required by the analysis covariance via stochastically perturbing the fast variables, which subsequently quickly relaxed onto the attractor. The situation in single-scale dynamics is more complicated. Whereas localisation of the bred vectors was beneficial in the multi-scale case and allowed for the conditioning of the SPBVs on the slow synoptic dynamical state, localisation of BVs prohibits in the single-scale scenario perturbations outside the localised region. Hence, although the resulting perturbation will be close in phase space to the original BV and appropriately sample the probability density function around it, the resulting initial conditions may not contain sufficient variability in the regions of significant uncertainty of the analysis.

To investigate the performance of stochastically modified bred vectors in a single-scale system we consider the Lorenz 96 model (Lorenz 1996) in two settings, which support strongly localised and weakly localised BVs. Our numerical simulations demonstrate that in both cases SPBVs and RDBVs exhibit significantly increased diversity and ensemble dimension, and that in both cases they provide superior forecast skill and reliability measures when compared to classical BVs. Their forecast skill and reliability, diagnosed by means of error-spread relationships and reliability diagrams, however, crucially depends on the degree of localisation, when the ensemble is centred around an analysed field. In the weakly localised case SPBVs and RDBVs constitute reliable forecast ensembles with forecast skills comparable to a reference ensemble obtained from an ensemble transform Kalman filter (ETKF). In the case of strongly localised bred vectors, however, their reliability is severely impeded. Localisation prevents SPBVs to constitute a reliable ensemble as they are not consistent with the analysis error which may be non-negligible outside the region of significant activity of the SPBV. For the weakly localised case we will show that the strength of the multiplicative noise used for generating SPBV ensembles can be judiciously chosen as a trade-off between providing the most ensemble diversity while preserving dynamic adaptivity in the sense that they resemble realistic error growth comparable to those of Lyapunov vectors. This dynamical consistency is probed by projecting onto the subspace spanned by the dominant covariant Lyapunov vectors. RDBVs will be shown to be dynamically inconsistent but to nevertheless feature improved forecast skill over SPBVs.

The paper is organised as follows. In Section 2 we introduce the Lorenz 96 model (Lorenz 1996). In Section 3 we briefly review the breeding method. Section 4 introduces our stochastically modified bred vector ensemble methods, namely SPBVs and RDBVs,

and shows how they relate to covariant Lyapunov modes as a measure of their dynamic adaptivity. Section 5 introduces the diagnostics used to evaluate the performance and the efficiency of stochastically modified bred vectors. In Section 6 the forecast skill and the reliability of each ensemble type is analysed for ensembles generated from an analysed field, obtained form incorporating imperfect observations in a data assimilation procedure. We conclude in Section 7 with a discussion and an outlook.

## 2. The Lorenz 96 system

The Lorenz 96 (L96) system (Lorenz 1996)

$$\frac{d}{dt}X_k = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F \tag{1}$$

with cyclic boundary conditions $X_{k+K} = X_k$ for $k = 1, \ldots, K$, was introduced as a caricature for the midlatitude atmosphere and has been used as a test bed for numerous studies in atmospheric sciences. The dynamics of the Lorenz 96 system is characterized by energy conserving nonlinear transport, linear damping and forcing. Despite its simplicity the L96 model exhibits many dynamical scenarios also observed in actual geophysical fluid flows such as regimes and transitions between them (Lorenz 2006). The variables $X_k$ can be interpreted as large scale atmospheric fields arranged in the midlatitudes on a latitudinal circle of $30,000$km, such as synoptic weather systems. The classical choice $K = 40$ corresponds to a spacing between adjacent variables of roughly the Rossby radius of deformation of $750$ km. We shall also consider $K = 128$ which implies a spacing between adjacent sites of $234$ km. In both cases we use as forcing amplitude $F = 8$ which implies chaotic dynamics (Lorenz and Emanuel 1998). The setting with $K = 40$ reproduces dynamical patterns with a realistic number of Rossby-like waves and is frequently used in the context of data assimilation. On the other hand, the choice $K = 128$ is used to study intrinsic properties of spatially extended dynamical systems (Pazó *et al.* 2013). For $F = 8$ the climatic variance is estimated as $\sigma^2 = 13.25$ and the decorrelation (*e*-folding) time is $\tau = 0.41$, for both $K = 40$ and $K = 128$. The maximal Lyapunov exponent is measured as $\lambda_{\max} = 1.69$ for $K = 40$ and as $\lambda_{\max} = 1.775$ for $K = 128$. The L96 system is extensive (Karimi and Paul 2010), in the sense that many relevant quantities (such as surface width, attractor dimension, entropy) scale linearly with the size $K$ and the Lyapunov exponents converge to a continuous function in the limit $K \to \infty$. This is illustrated in Figure 1, where we show the Lyapunov spectrum for $K = 40$ and $K = 128$. For $K = 40$ there are 13 distinct positive Lyapunov exponents, while for $K = 128$ there are $42 \approx \frac{128}{40} \times 13$ distinct positive Lyapunov exponents.

To numerically simulate the L96 system we employ a fourth-order Runge-Kutta method with a fixed time step $dt = 0.005$. In our simulations an initial transient time of 5000 time units is discarded to assure that the dynamics has settled on the attractor.
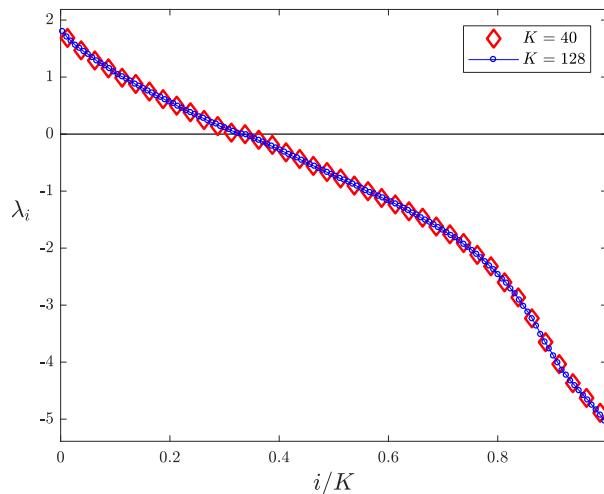


**Figure 1.** Lyapunov exponent spectrum for the L96 model (1) with $F = 8$ for $K = 40$ and $K = 128$.

## 3. Bred vectors and the breeding method

We briefly review the classical breeding method introduced by Toth and Kalnay (1993, 1997). We closely follow the exposition from our previous work (Giggins and Gottwald 2019). BVs are finite-size, periodically rescaled perturbations generated using the full non-linear dynamics of the system. Centred around a control trajectory $\boldsymbol{z}_c(t_i)$ at some time $t_i$, perturbed initial conditions of size $\delta$,

$$\boldsymbol{z}_p(t_i) = \boldsymbol{z}_c(t_i) + \delta \frac{\boldsymbol{p}}{\|\boldsymbol{p}\|},$$

are defined where $\boldsymbol{p}$ is an initial arbitrary random perturbation. The control and the perturbed initial conditions are simultaneously evolved using the full non-linear dynamics for an integration time $T$ until time $t_{i+1} = t_i + T$. At the end of the integration window the difference between the control and the perturbed trajectories

$$\Delta \boldsymbol{z}(t_{i+1}) = \boldsymbol{z}_p(t_{i+1}) - \boldsymbol{z}_c(t_{i+1})$$

is determined, and the bred vector is defined as the difference rescaled to size $\delta$ with

$$\boldsymbol{b}(t_{i+1}) = \delta \frac{\Delta \boldsymbol{z}(t_{i+1})}{\|\Delta \boldsymbol{z}(t_{i+1})\|}.$$

126 The perturbation $\boldsymbol{b}(t_{i+1})$ then determines the initial condition of the perturbed trajectory $\boldsymbol{z}_p(t_{i+1}) = \boldsymbol{z}_c(t_{i+1}) + \boldsymbol{b}(t_{i+1})$ at the start
127 of the next breeding cycle. This process of breeding is repeated for several cycles until the perturbation maintains a sufficiently large
128 growth rate and until the perturbations converge in the sense that at time $t_n$ an ensemble of BVs spans the same subspace as BVs
129 obtained in a breeding cycle which had been initialised further in the past. The characteristic time scales of the instabilities of interest
130 and length of the breeding cycles determine how many breeding cycles are required to achieve convergence (Toth and Kalnay 1993).
131 For the L96 system a breeding cycle length of $T = 0.05$ time units is employed for all simulations, and we employ a spin-up time for
132 the BVs of 500 time units (which amounts to 10000 breeding cycles). An ensemble of $N$ BVs is created by $N$ independent breeding
133 cycles initialised from independent initial perturbations $\boldsymbol{p}$. The resulting converged BV ensemble at time $t_i$ is then employed as initial
134 conditions for ensemble forecasts. The breeding method is conceptually similar to the method for generating Lyapunov vectors.
135 They differ though in that Lyapunov vectors are generated using the linearised dynamics and an infinitesimal perturbation $\delta$, whereas
136 BVs are generated using the full nonlinear model and finite perturbation sizes. In contrast to covariant Lyapunov vectors which are
137 mapped by the linear tangent dynamics onto each other, the dynamics of finite-size BVs is not given by a linear mapping and as such
138 they technically do not form a vector space. Despite the similarities between BVs and Lyapunov vectors, we adopt here the point of
139 view outlined in Giggins and Gottwald (2019) that for probabilistic ensemble forecasts the object of interest are the perturbed states
140 $\boldsymbol{z}_p = \boldsymbol{z}_c + \Delta \boldsymbol{z}$, which constitute the sample points for the Monte-Carlo approximation of the probability density function, rather than
141 the differences $\Delta \boldsymbol{z}$.
142

The bred vectors of the L96 system for system sizes $K = 40$ and $K = 128$ have markedly different spatial structures. This is
illustrated in Figure 2 where snapshots of a typical BV are shown for $K = 40$ and for $K = 128$ for $\delta = 0.1 \approx 0.275\,\sigma_{\mathrm{clim}}$. For the
larger $K = 128$ system BVs are strongly localised with only a well-defined group of sites having significant entries, whereas for
$K = 40$ the localisation is less well defined and the size of the active sites with significant entries almost spans the whole domain.
Note that the number of "active" sites with increased absolute value of the BVs are roughly the same in both cases, reflecting that
BVs capture the same instability, which exhibits the same spatial organization in both cases, but with a better resolution in the larger
domain. As we will see, the degree of localisation plays a crucial role for the performance of stochastically perturbed bred vectors. To
measure the spatial organisation of BVs we consider the $K \times K$ covariance matrix

$$\mathbf{C} = \frac{\boldsymbol{b}(t)\big[\boldsymbol{b}(t)\big]^{\mathsf{T}}}{\|\boldsymbol{b}(t)\|_2 \|\boldsymbol{b}(t)\|_2}, \tag{2}$$

143 and determine its average $\bar{\mathbf{C}} = \langle \mathbf{C} \rangle$, where the average is taken over realisations of independent BVs generated at different points in
144 time. Since all components for the L96 system are statistically equivalent the $k$-th and $(k + l)$-th rows of $\bar{\mathbf{C}}$ are identical up to a shift of
145 $l$ components. Figure 3 shows the row-averaged $\bar{\mathbf{C}}_{k,\cdot}$ of the matrix $\bar{\mathbf{C}}$ for $K = 40$ for some arbitrary component $k$ for the BV depicted
146 in Figure 2 (for $K = 128$ the correlation structure is identical).
147

148 In probabilistic forecasting the aim is to approximate the density $\rho(X, \tau)$ at lead time $\tau$ given an initial density $\rho(X, t = 0)$, describing
149 the current estimate of the system. Adopting our point of view that BVs are designed to represent a good Monte Carlo estimate of
150 $\rho(X, t = 0)$, the property of BVs to capture fast growing dynamically relevant instabilities is translated into the initial conditions
151 associated with BVs which are then likely to be observed at later times $\tau$ representative of the density $\rho(X, \tau)$. The capability of BVs
152 to form an ensemble of independent initial conditions suited for a reliable probabilistic ensemble forecast, depends crucially on the
153 perturbation size $\delta$. For too large perturbation sizes $\delta$, the initial conditions resemble random draws from the attractor (after a typically
154 rapid transition towards it) and the forecast skill deteriorates. Contrary, for too small values of the perturbation size, BVs align with the
155 leading Lyapunov vector (LLV) exhibiting ensemble collapse to a single ensemble member. An ensemble of $N = 20$ BVs with $\delta = 0.1$,
156 which were initialised with different random perturbations, collapses and the ensemble members are indistinguishable by eye from the
157 ones depicted in Figure 2 for both $K = 40$ and $K = 128$. Such a lack of diversity of an ensemble of bred vectors presents a major draw
158 back of bred vectors in ensemble forecasting. In the language of probabilistic forecasts the alignment of BVs with the LLV implies that
159 only a single draw from $\rho(X, t = 0)$ is considered. In the next section we present a method how to overcome this drawback while still
160 preserving the desirable features of BVs such as their low computational cost and their dynamical consistency (Pazó *et al.* 2010).

161 **4. Stochastically perturbed bred vectors**

We review here the method proposed in Giggins and Gottwald (2019) to increase the diversity of BV ensembles for multi-scale systems
and apply it to systems without scale separation. To generate a diverse ensemble of initial conditions conditioned on the current state of
the system, BVs are generated from a parent BV by applying a multiplicative stochastic perturbation to it. The key idea to generating
additional draws from the initial density function $\rho(X, t = 0)$ is to exploit the fact that in spatially extended dynamical systems, BVs
are often localised (as shown in Figure 2 for $K = 128$) or exhibit some non-trivial spatial structure (as shown in Figure 3 for $K = 40$),
corresponding to some degree of spatial organisation of error growth. Stochastically perturbed bred vectors (SPBVs) are designed to
preserve the spatial structure of BVs, which is paramount to conditioning the initial density $\rho(X, t = 0)$ on the current state $X$. SPBVs
are defined as

$$\boldsymbol{b}_{sp} = \delta \frac{(\mathbf{I} + \boldsymbol{\Xi})\boldsymbol{b}}{\|(\mathbf{I} + \boldsymbol{\Xi})\boldsymbol{b}\|}, \tag{3}$$
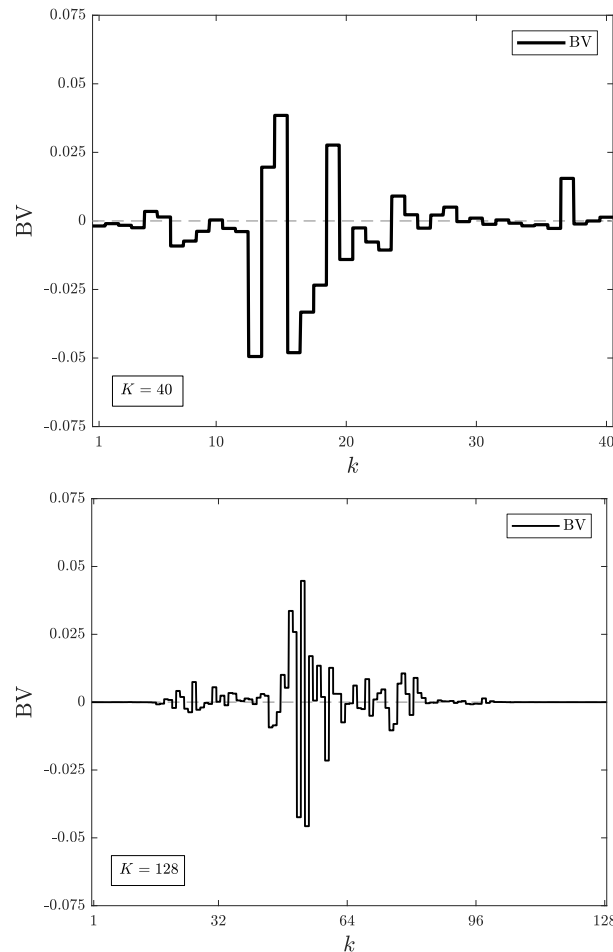
**Figure 2.** Bred vectors of the L96 system with perturbation size $\delta = 0.1$ for different system size $K$. Top: $K = 40$. Bottom: $K = 128$.
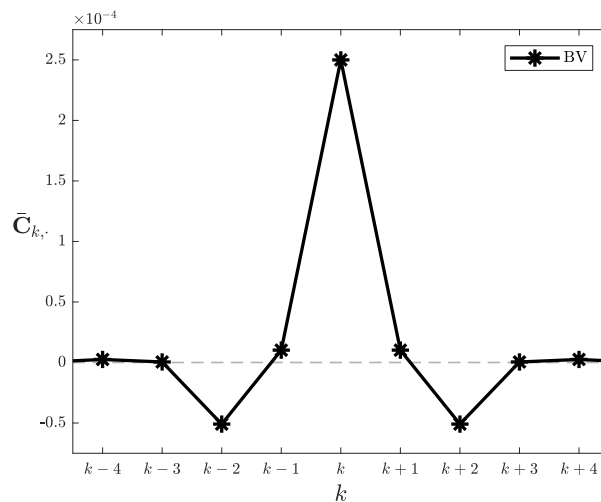


**Figure 3.** Row-averaged $\bar{\mathbf{C}}_{k,\cdot}$ of the covariance matrix (2) for the BV of the L96 system with perturbation size $\delta = 0.1$ displayed in Figure 2 for $K = 40$. For $K = 128$ the plot is indistinguishable by eye.

where $\mathbf{I}$ is the $K \times K$ identity matrix. The diagonal $K \times K$ matrix $\boldsymbol{\Xi}$ with entries $\xi_{jj} \sim \mathcal{N}(0, \sigma^2)$ for $j = 1, \ldots, K$ with variance parameter $\sigma^2$ represents the stochastic perturbation. The stochastic perturbation is performed only once as a post-processing step from a given parent BV when generating initial conditions for a forecast ensemble, and therefore does not significantly add to the computational cost. In Figure 4 we show a realisation of an SPBV, overlaid with their parent BV, for a perturbation size of $\delta = 0.1$ with noise strength $\sigma = 1.25$, for $K = 40$ and for $K = 128$. It is clearly seen that the spatial structure of the perturbation is preserved.

The stochastic perturbations generate initial conditions that are nearby the attractor and after a typically rapid relaxation towards the attractor along the stable manifold, approach the attractor close in phase space to the initial condition associated with the parent BV, which we know is capturing fast error growth. The stochasticity hence allows to sample the phase space on the attractor in the fastest growing region.

173      The noise strength $\sigma$ obviously plays a central role. When $\sigma \to 0$ SPBVs essentially reproduce the parent BVs they were generated
174 from, and the spatial structure is exactly preserved but no diversity is gained. In the other extreme case $\sigma \to \infty$, the behaviour depends
175 on the degree of localisation. For the strongly localised case $K = 128$ with many vanishing BV components (cf. Figure 4 (bottom)),
176 the degree of localisation remains preserved since SPBVs are rescaled to size $\delta$, and the diversity is greatly enhanced. This is the case
177 discussed in the multi-scale setting in Giggins and Gottwald (2019). The weakly localised case when there are no significant regions
178 with vanishing components of the BV (cf. Figure 4 (top)) is more complex. For sufficiently large magnitudes of the noise strength
179 $\sigma$, SPBVs become spatially uncorrelated random perturbations of size $\delta$. This allows for (almost) maximal diversity of the ensemble
180 which, however, comes at the cost of destroying the inherent spatial structure of the dynamically relevant fast growing perturbations.
181 The destruction of the spatial structure implies that we typically do not sample the phase space region locally but instead generate
182 initial conditions as random draws from the attractor, which are not conditioned on the current state. In Figure 6 we illustrate the loss
183 of spatial structure by showing the average of rows $\bar{\mathbf{C}}_{k,\cdot}$ of the covariance matrix (2) for SPBVs for increasing values of the noise
184 strength $\sigma$. It is seen that for $\sigma = 1.0$ and for $\sigma = 1.25$ the nontrivial correlations between adjacent sites are preserved albeit reduced
185 in magnitude, whereas for $\sigma = 5$ the spatial structure is entirely lost and adjacent sites are uncorrelated.

186

187      We also consider the so called random draw bred vectors (RDBVs) introduced in Giggins and Gottwald (2019). An ensemble of
188 RDBVs is generated by randomly selecting classical BVs which were generated from independent initial conditions randomly drawn
189 from the attractor. To avoid storing a huge library of independent BVs, an ensemble of RDBVs is generated on the fly by evolving $N$
190 independent control trajectories started from independent initial conditions, each generating a single BV. Whereas SPBVs are designed
191 to sample the phase space locally, RDBVs are dynamically inconsistent in the sense that they may, after a quick relaxation towards
192 the attractor, evolve into states which are not close in phase space to the current state of the control. Example RDBVs for $K = 40$ and
193 $K = 128$ are shown in Figure 5. We remark that, contrary to SPBVs, RDBVs form an (almost) orthogonal ensemble.
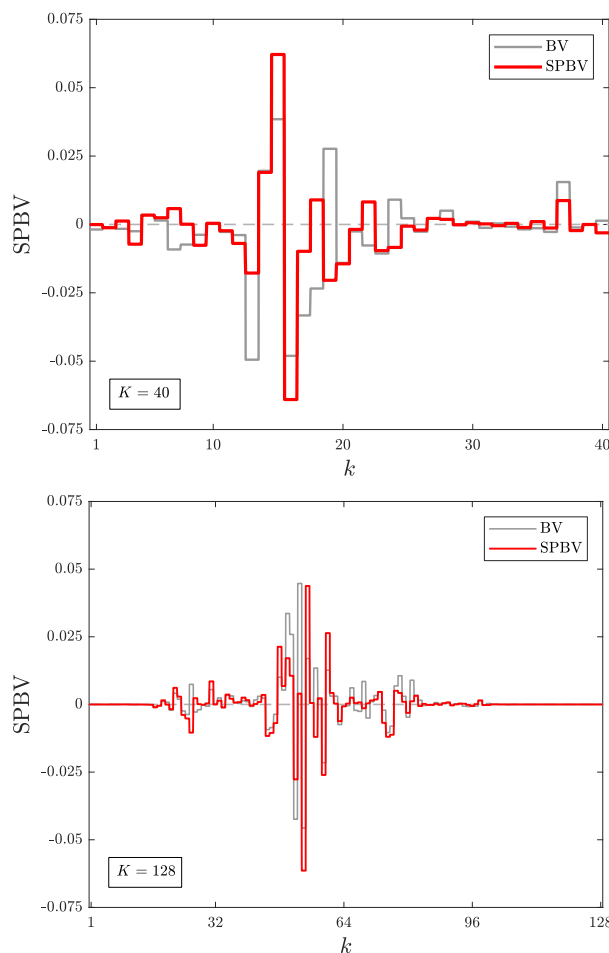194



**Figure 4.** SPBV and its associated parent BV for the L96 model for perturbation size $\delta = 0.1$. Top: $K = 40$. Bottom: $K = 128$.

195 *4.1. Dynamic properties of bred vectors: Backward and covariant Lyapunov vectors*

196 We now probe how bred vectors and their stochastic modifications relate to dynamically relevant modes such as Lyapunov vectors
197 which capture the asymptotic growth of infinitesimal perturbations, and thereby in how far they are dynamically adapted. The dynamic
198 adaptivity of classical BVs was established in Pazó *et al.* (2010). We now show that SPBVs inherit this property from their parent
199 BVs. In particular, we consider the relationship between bred vectors and backward Lyapunov vectors and covariant Lyapunov vectors.
200 Backward Lyapunov vectors are initialised in the asymptotically distant past and are generated by solving the linear tangent model
201 of the dynamical system under a Gram-Schmidt orthogonalisation procedure. The orthogonal backward Lyapunov vectors are not
202 covariant under the linear tangent dynamics and all of them typically evolve under the dynamics into the leading Lyapunov vector
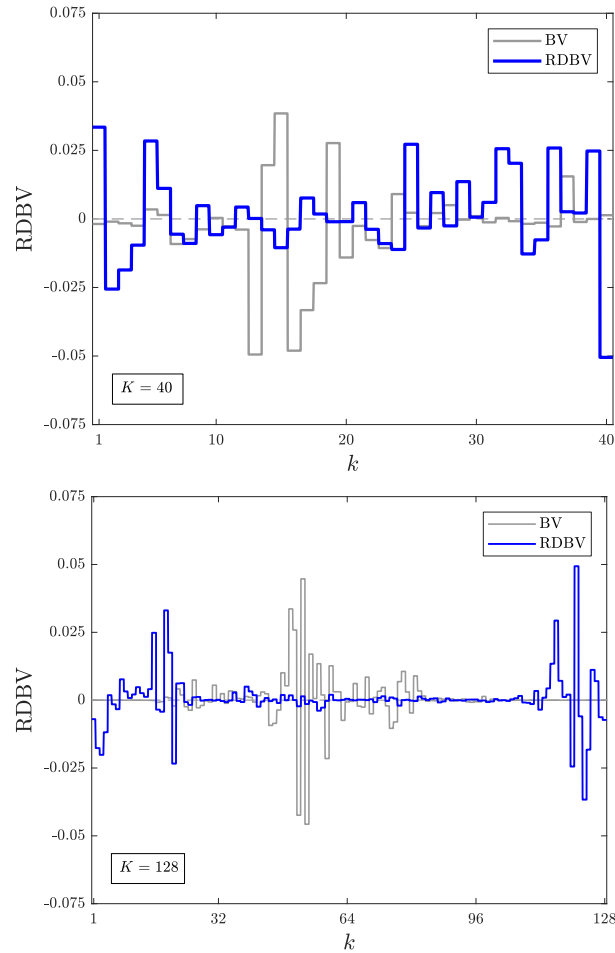
**Figure 5.** RDBV and BV for the L96 model for perturbation size $\delta = 0.1$. The BVs are as in Figure 2. Top: $K = 40$. Bottom: $K = 128$.
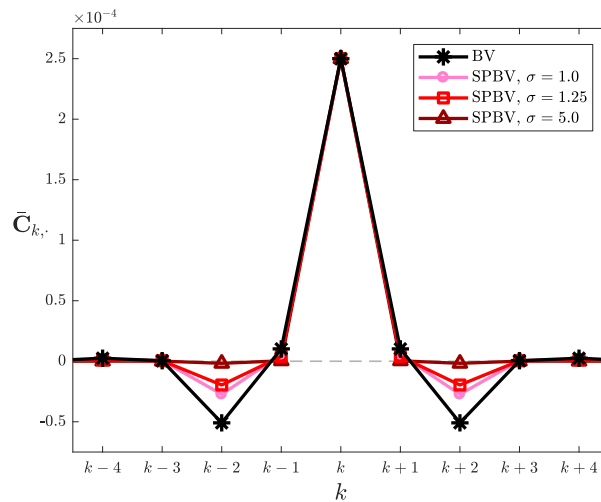


**Figure 6.** Row-averaged $\bar{\mathbf{C}}_{k,\cdot}$ from the covariance matrix (2) for SPBVs of the L96 system with various noise strengths $\sigma = 1.0$, $\sigma = 1.25$ and $\sigma = 5.0$ for $K = 40$. The covariance of BVs is depicted as a reference. For $K = 128$ the plot is indistinguishable by eye.

203 (LLV). Covariant Lyapunov vectors, on the contrary, form a typically non-orthogonal basis of the tangent space and are mapped onto
204 each other by the linearised tangent dynamics. The associated asymptotic growth rates of backward and covariant Lyapunov vectors,
205 the Lyapunov exponents, are shown in Figure 1 for the L96 system. As BVs, the first few leading covariant Lyapunov vectors exhibit a
206 localised spatial structure in the L96 system (not shown), with strong localisation for $K = 128$ and weak localisation for $K = 40$.

We quantify the relationship between the respective BV ensembles and Lyapunov vectors by measuring the average projection of
BV ensembles onto backward and onto covariant Lyapunov vectors, and consider the following measure for the degree of projection

$$\pi_i^n(t) = \left| \frac{\boldsymbol{b}^n(t)}{\|\boldsymbol{b}^n(t)\|} \cdot \frac{\boldsymbol{l}_i(t)}{\|\boldsymbol{l}_i(t)\|} \right|, \tag{4}$$

207 where $\boldsymbol{b}^n(t)$ denotes the $n$th bred vector ensemble member at time $t$ and $\boldsymbol{l}_i(t)$ denotes the Lyapunov vector corresponding to the $i$th
208 largest Lyapunov exponent at time $t$. We report here on the average degree of projection $\bar{\pi}_i$ where we average $\pi_i^n(t)$ over time and over
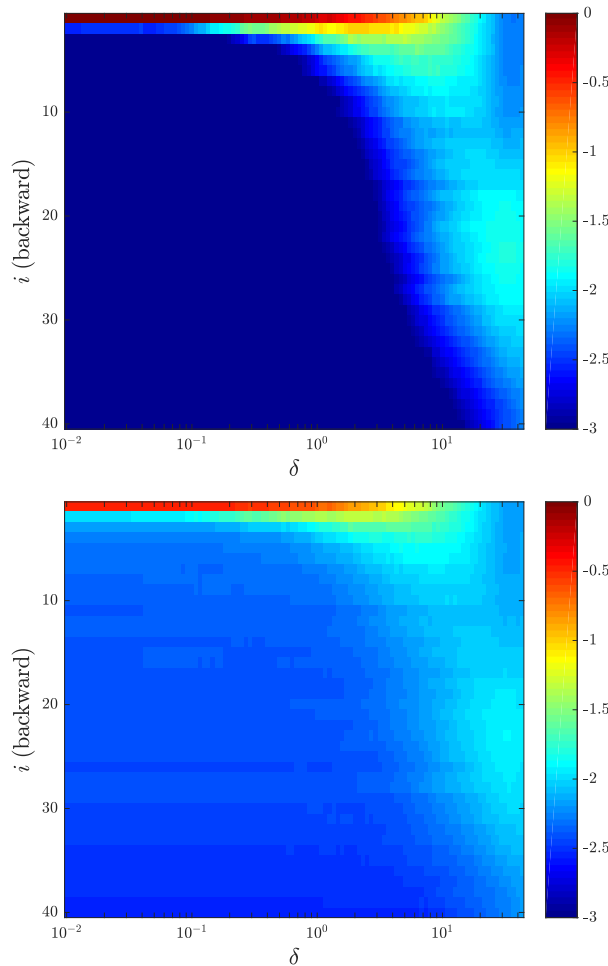
**Figure 7.** Average projection $\bar{\pi}_i$ of backward Lyapunov vectors for $K = 40$. Results are shown on a logarithmic scale. Top: BVs. Bottom: SPBVs with $\sigma = 1.25$.

209  the ensemble members. Note that $\bar{\pi}_i = 1$ corresponds to perfect alignment and $\bar{\pi}_i = 0$ corresponds to (on average) no alignment.

210

211  There exist several efficient numerical algorithms to calculate the covariant Lyapunov vectors (Wolfe and Samelson 2007; Ginelli
212  *et al.* 2007). We use here the algorithm by Ginelli *et al.* (2007) as described in Kuptsov and Parlitz (2012) to numerically calculate
213  covariant Lyapunov vectors. We use a spin-up time of $2,500$ time units to converge to the set of backward Lyapunov vectors evolving
214  forward in time, and a further $2,500$ time units to ensure convergence to the set of expansion coefficients of the covariant Lyapunov
215  vectors that express the covariant Lyapunov vectors in the basis of forward and backward Lyapunov vectors respectively, evolving
216  backward in time. Orthonormalisation of the backward Lyapunov vectors is performed at every time step.

217

218  Figures 7 and 8 display $\bar{\pi}_i$ for all $i = 1, \ldots, K$ backward Lyapunov vectors for $K = 40$ and $K = 128$ respectively, for classical BVs
219  (top) and SPBVs (bottom). It is clearly seen that both, classical BVs and SPBVs, project almost completely onto the first backward
220  Lyapunov vector (the LLV) for small $\delta < 0.1$ and are orthogonal to all other directions for both dimension sizes. When the perturbation
221  size lies between $0.1 \lesssim \delta \lesssim 8$ for $K = 40$ and between $0.1 \lesssim \delta \lesssim 5$ for $K = 128$, BVs also project onto the next few backwards
222  Lyapunov vectors. We shall see in Section 6, that for these perturbation sizes, BV ensembles have collapsed to a single member and
223  have an ensemble dimension (to be defined below in (5)) strictly equal to 1 (cf. Figure 12). For $\delta > 1$ the non-vanishing projections
224  of BVs onto the next Lyapunov vectors stem from increasing fluctuations of the BV ensemble around the LLV. We observe that $\pi_i^n(t)$
225  may strongly fluctuate in time and individual members of a BV/SPBV ensemble may exhibit, locally in time, strong projections on
226  higher Lyapunov vectors. In such cases when BVs/SPBVs do not fully align with the LLV, they lie typically in the subspace spanned
227  by the first few Lyapunov vectors (not shown). For even larger values of the perturbation size, BVs do not significantly project onto the
228  linear Lyapunov vectors as they evolved into truly nonlinear objects.

229  Projections of BVs and SPBVs onto covariant Lyapunov vectors exhibit similar signatures as for backward Lyapunov vectors.
230  Contrary to backward Lyapunov vectors, covariant Lyapunov vectors do not form an orthogonal basis. Furthermore, successive
231  covariant Lyapunov vectors are likely to be localised in similar spatial regions to each other whereas this is not the case for backward
232  Lyapunov vectors due to non-dynamical orthogonality constraint (Herrera *et al.* 2011). Hence fluctuations of BVs cause them to
233  project onto several of the covariant Lyapunov vectors. In particular, we see strong projections of classical BVs onto covariant
234  Lyapunov vectors with index $i \leq 7$ for $K = 40$, and onto those with index $i \leq 12$ for $K = 128$. As for backward Lyapunov exponents,
235  the projection onto dynamically relevant low-index Lyapunov vectors drops off when BVs gain diversity at $\delta \approx 8$ for $K = 40$ and at
236  $\delta \approx 5$ for $K = 128$. SPBVs feature weaker projections onto the higher-index covariant Lyapunov vectors with significant projections
237  in the smaller range $i \leq 3$ and $i \leq 6$ for $K = 40$ and $K = 128$, respectively. This overall stronger projection of SPBVs to the low-index
238  covariant Lyapunov vectors is caused by the ensemble averaging (SPBVs are generated from a single collapsed BV). As for backward
239  Lyapunov vectors, the lower dimensional subspaces onto which BVs and SPBVs project onto can fluctuate over time, in particular for
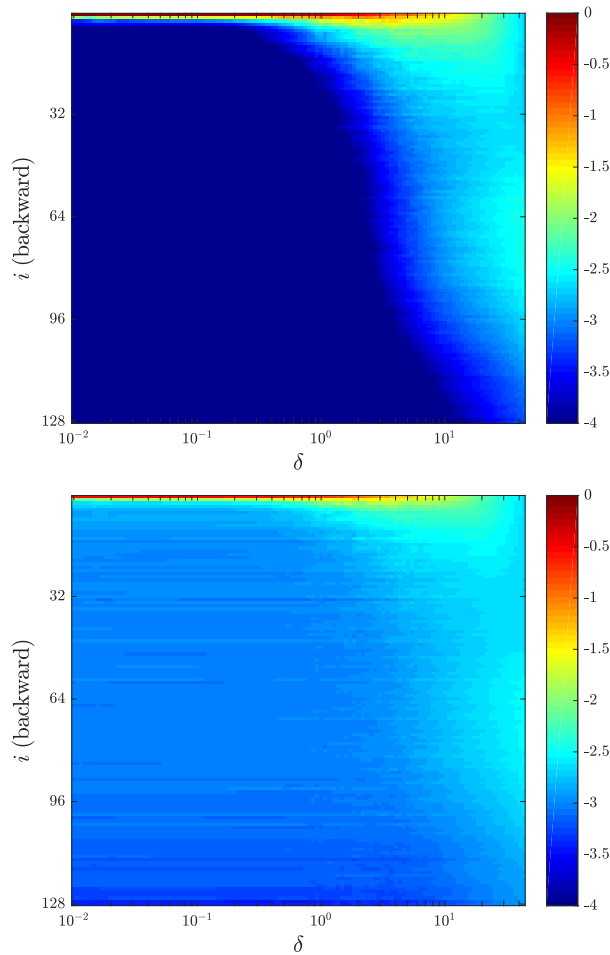
**Figure 8.** Average projection $\bar{\pi}_i$ of backward Lyapunov vectors for $K = 128$. Results are shown on a logarithmic scale. Top: BVs. Bottom: SPBVs with $\sigma = 1.25$.

larger values of the perturbation size $\delta$ (not shown). It is pertinent to notice that the projection of SPBVs onto the Lyapunov vectors weakens as $\sigma$ is increased and saturates for sufficiently large $\sigma$, in particular in the weakly localised case $K = 40$ (not shown). We shall see below that there is a trade-off between providing the most ensemble diversity while preserving dynamic adaptivity for SPBVs.

We conclude that BVs and SPBVs share similar localisation structure to that of the first few covariant Lyapunov vectors, and that for sufficiently small noise strength $\sigma$ SPBVs inherit from BVs the desirable property of dynamical adaptivity. RDBVs, on the other hand, do not exhibit any significant average projections onto any of the backward or covariant Lyapunov vectors as they are unrelated to the local dynamics (not shown). RDBVs are hence dynamically not adapted. We mention that ETKF ensembles are also dynamically adapted. Contrary to BVs and SPBVs which project dominantly onto the first few Lyapunov vectors, ETKF ensembles project homogeneously onto the whole unstable subspace (not shown; see also (Ng *et al.* 2011)).

## 5. Diagnostics

To illustrate how SPBVs can be used as a reliable diverse ensemble with improved forecast skill we now introduce several diagnostics. In particular, we consider the ensemble dimension to measure the diversity of an ensemble, the root-mean-square error to quantify the forecast skill and several reliability measures to probe the probabilistic properties of an ensemble. This set of diagnostics has previously been used to study the performance of SPBVs and RDBVs (Giggins and Gottwald 2019).

### 5.1. Ensemble dimension

We quantify the diversity of an ensemble using the "ensemble dimension" (Bretherton *et al.* 1999; Oczkowski *et al.* 2005), also known as the "bred vector dimension" (Patil *et al.* 2001). The ensemble dimension is a measure for the dimension of the subspace spanned by a set of vectors. For an ensemble of $N$ BVs $\{\boldsymbol{b}^{(n)}(t)\}_{n=1,...,N}$ at a given time $t$, the ensemble dimension is defined as

$$\mathcal{D}_{ens}(t) = \frac{\left(\sum_{n=1}^{N} \sqrt{\mu_n}\right)^2}{\sum_{n=1}^{N} \mu_n}, \tag{5}$$

where the $\mu_n$ are the eigenvalues of the $N \times N$ covariance matrix $\mathbf{C}$ (cf. (2)). The ensemble dimension takes values between $\mathcal{D}_{ens} = 1$ and $\mathcal{D}_{ens} = \min(N, D)$, where $D$ is the total dimension of the dynamical system, depending on whether the ensemble members are all aligned or are orthogonal to each other. We consider in our numerical experiments the temporal average $\bar{\mathcal{D}}_{ens}$ to characterise the diversity of an ensemble.
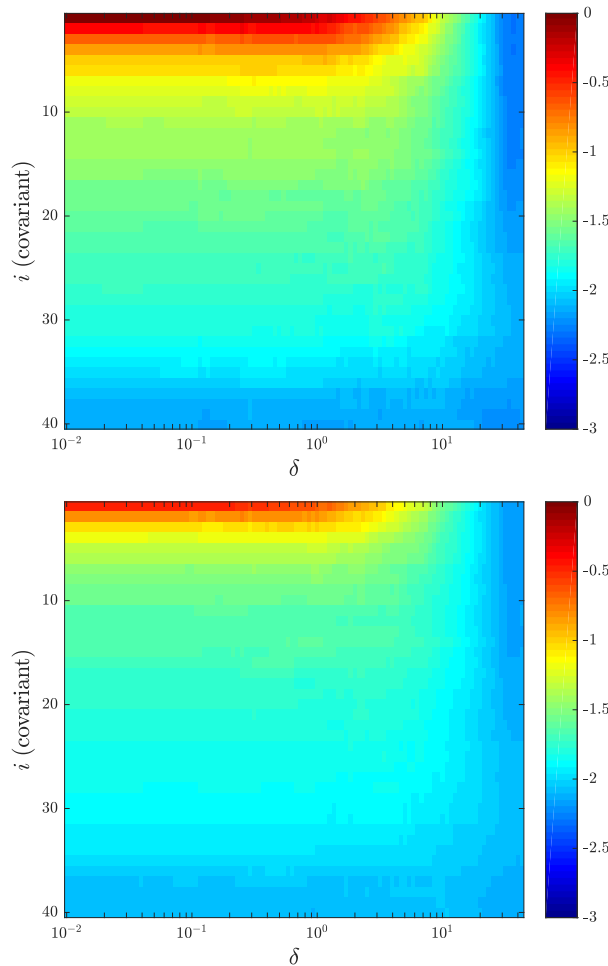
**Figure 9.** Averaged absolute value projection $\bar{\pi}_i$ of covariant Lyapunov vectors for $K = 40$. Results are shown on a logarithmic scale. Top: BVs. Bottom: SPBVs with $\sigma = 1.25$.

### 5.2. Ensemble forecast skill

To measure the forecast skill of an ensemble with $N$ members $X_k^{(n)}$, $n = 1, \ldots, N$, we evaluate the root-mean-square-error (RMS error) between the truth and the ensemble average. Denoting the ensemble average with angular brackets, we introduce the ensemble mean

$$\langle X_k \rangle = \frac{1}{N} \sum_{n=1}^{N} X_k^{(n)} \tag{6}$$

and the site-averaged root-mean-square error between the truth $X_k^{tr}$ and the ensemble average over $M$ realizations as a function of the lead time $\tau$,

$$\mathcal{E}(\tau) = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \frac{1}{K} \sum_{k=1}^{K} \|X_{k,m}^{tr}(\tau) - \langle X_{k,m} \rangle(\tau)\|^2}, \tag{7}$$

where the index $m = 1, \ldots, M$ labels the realisation. Similarly, to quantify the dispersion of the ensemble, we consider the site-averaged root-mean-square spread (RMS spread)

$$\mathcal{S}(\tau) = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \frac{1}{K} \sum_{k=1}^{K} \langle \|X_{k,m}^{(n)}(\tau) - \langle X_{k,m} \rangle(\tau)\|^2 \rangle}. \tag{8}$$

### 5.3. Reliability

The RMS error is not always the appropriate measure to quantify the performance of an ensemble in probabilistic forecasting. For example, if the probability density function has disjoint support, the ensemble average may not have a physical meaning and can result in a poor forecast. For probabilistic forecasts the reliability of an ensemble is more relevant. An ensemble is called *perfectly reliable* if the truth along with each ensemble member are independent draws from the same probability density function $\rho(X)$. In perfect ensembles the ratio between the RMS error and the RMS spread approaches 1 as the ensemble size increases (Wilks 2006; Leutbecher and Palmer 2008). Under-dispersive ensembles, on the other hand, feature a ratio smaller than 1, whereas over-dispersive ensembles
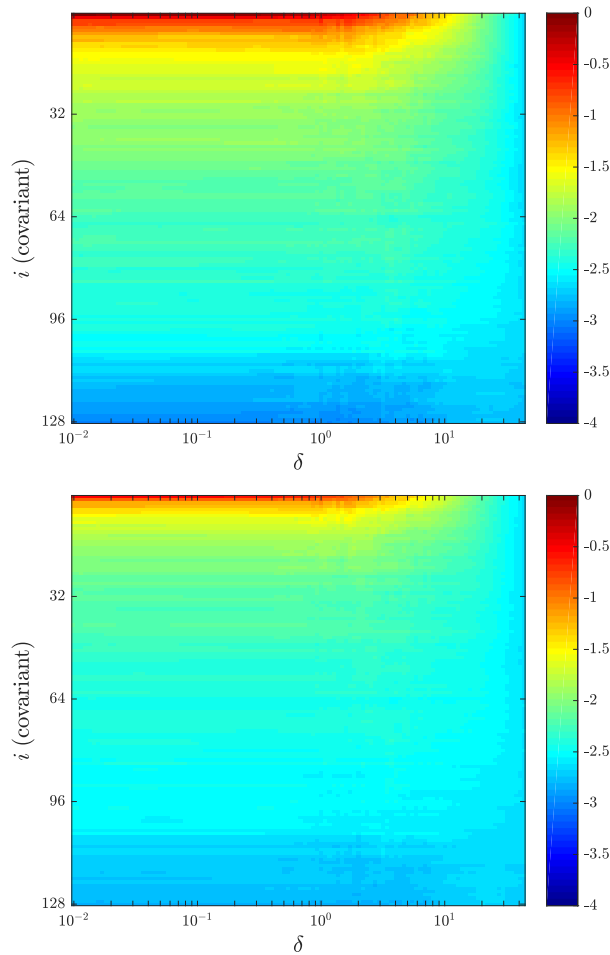
**Figure 10.** Averaged absolute value projection $\bar{\pi}_i$ of covariant Lyapunov vectors for $K = 128$. Results are shown on a logarithmic scale. Top: BVs. Bottom: SPBVs with $\sigma = 1.25$.

feature ratios larger than 1. Furthermore, in reliable ensembles the truth is statistically indistinguishable from any given ensemble member, and each ensemble member has equal probability to be closest to the truth. This property can be probed in Talagrand or Rank histograms (Anderson 1996; Hamill and Colucci 1997; Talagrand 1999). For a given lead time, a Talagrand histogram is created by sorting the $N$ ensemble members in increasing order of their forecast value to form a set of $N + 1$ bins. A histogram of probabilities of the truth falling into a bin $i$ at the given lead time is then produced by counting the frequency that the truth falls into the bin $i$. A reliable ensemble implies a flat histogram as the truth should have equal probability of falling into any given bin. Under-dispersive/over-dispersive ensembles, on the other hand, result in histograms which are convex/concave in shape (Wilks 2006).

## 6. Numerical results

We now present numerical results demonstrating that SPBVs can be used as a reliable diverse ensemble with improved forecast skill in single-scale systems provided that BVs are not strongly localised. We shall present results for the strongly localised case with $K = 128$ and for the weakly localised case with $K = 40$ separately. We examine the ensemble diversity, forecast skill metrics such as the RMS error and RMS spread, as well as the reliability quantified by the error-spread relationship and the Talagrand diagram.

The setup for the numerical simulations is as follows. We employ an Ensemble Transform Kalman Filter to perform the data assimilation and construct the analysis (Tippett *et al.* 2003; Wang *et al.* 2004). The analysis is constructed from a forecast with a perfect model and noisy observations with variance 0.01 (corresponding to observational noise with 2.75% of the climatological standard deviation), following Bowler (2006); Pazó *et al.* (2013). To focus on the performance of the bred vector ensemble rather than on the data assimilation, we use a large ensemble for the ETKF with $K + 1$ members to produce the analysis, preventing filter divergence and avoiding the need for localisation and inflation. The ETKF ensemble is spun-up for 500 time units before commencing the breeding cycles. The average analysis error for the $K = 40$ and $K = 128$ systems is 0.10 and 0.18, respectively, over $2,500$ forecasts. Ideally, the value of $\delta$ that results in a local minima of the RMS forecast error $\mathcal{E}$ matches the size of the analysis error. In practice, however, the perturbation size often needs to be larger to achieve acceptable forecast skill (Toth and Kalnay 1997; Magnusson *et al.* 2008; Giggins and Gottwald 2019).

We employ a breeding cycle length of $T = 0.05$ time units. As is common practice in operational ensemble forecasting, pairs of positive/negative BVs are generated to ensure that the BV forecast ensemble represents the analysis mean at the initial forecast time. We consider BV ensembles consisting of $N$ BV perturbations of size $\delta$. For $K = 40$ we use $N = 10$ and for $K = 128$ we use $N = 20$ ensemble members, which implies 5 and 10 independent breeding cycles for $K = 40$ and $K = 128$, respectively. Each ensemble member is then evolved freely under the L96 dynamics for some lead time $\tau$. Forecasts are run for a total of 5 time units and we report
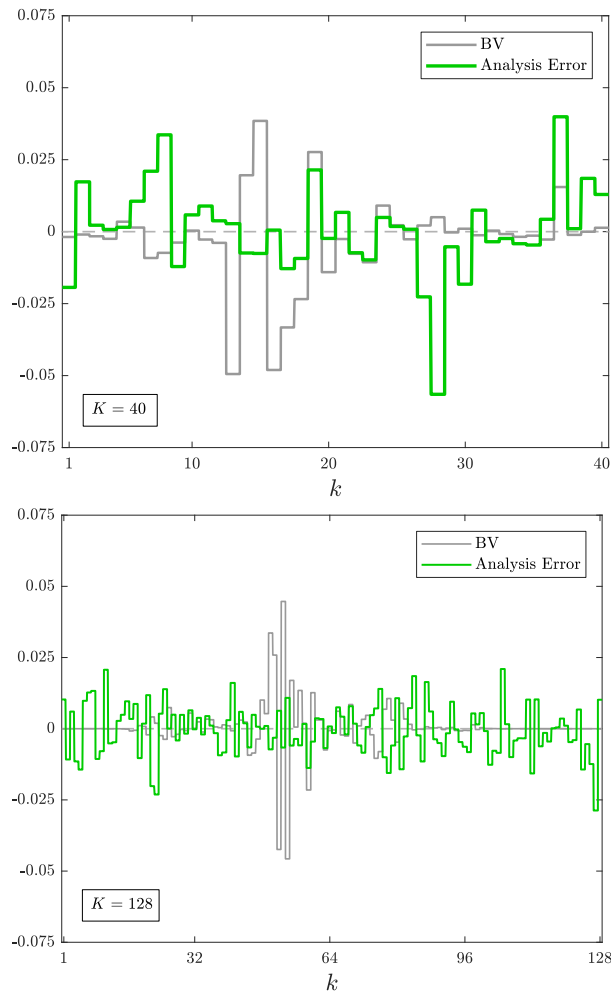
**Figure 11.** The bred vector from Figure 2 and the analysis error at its time of generation. Top: $K = 40$. Bottom: $K = 128$.

the results for lead times $\tau = 2.0$ and $\tau = 4.0$ time units for the forecast metrics presented in Section 5. A new forecast is created each 1.0 time units. SPBVs have been generated using a noise strength of $\sigma = 1.25$. All metrics are averaged over $M = 2,500$ forecasts.

Seeding BV forecast ensembles from the analysis fields poses problems, as a good forecast ensemble not only should evolve into likely future states but also has to account for the uncertainty of the analysis. This is particularly a problem in the case of strongly localised BVs (and SPBVs) for $K = 128$, as the localisation inhibits sampling the uncertainty of the analysis which typically extends outside of the region of localisation and is distributed across the whole domain. Figure 11 shows snapshots of BVs and the analysis error (scaled to have norm equal to 0.1 to facilitate comparison) for the weakly localised case $K = 40$ and the strongly localised case $K = 128$. We remark that by construction, the spatial structure of SPBVs is similar to that of BVs. In the strongly localised case, it is clearly seen that there are large regions of significant uncertainty of the analysis which are not perturbed by the BV. In the weakly localised case $K = 40$ on the other hand, BVs are more evenly distributed over the whole domain, and thus more likely to capture the errors in the analysis.

The lack of activity in sites remote of the region of their spatial localisation is likely to severely inhibit the BV/SPBV ensemble to evolve into states which contain the truth. We shall find below, that the property of localisation is detrimental for the dynamically adapted SPBVs in the L96 system (1) without scale separation, whereas it was essential in the multi-scale case in Giggins and Gottwald (2019). In particular we show that for $K = 128$ strong localisation implies poor reliability of SPBV ensembles. RDBVs, however, despite not being dynamically adapted, exhibit improved reliability and forecast skill compared to classical BVs. In the weakly localised case $K = 40$, SPBVs and RDBVs both constitute a reliable forecast ensemble with superior forecast skill compared to classical BV ensembles.

*6.1. Ensemble Dimension*

Figure 12 shows the average ensemble dimension $\bar{\mathcal{D}}_{ens}$ (5) as a function of $\delta$ for classical BVs, SPBVs with $\sigma = 1.25$, and RDBVs. For classical BVs the average ensemble dimension is $\bar{\mathcal{D}}_{ens} = 1$ for $\delta \lesssim 8$ and for $\delta \lesssim 5$, for $K = 40$ and $K = 128$ respectively, indicating the collapse of BV ensembles. For these perturbation sizes, a BV ensemble typically collapses onto the LLV but can also for $\delta > 1$, when the dynamics of the perturbation begins to feel the nonlinearity of the dynamics, align in a different direction, spanned by the first few leading Lyapunov vectors (cf. Figures 7-8). For even larger perturbation sizes $\delta > 8$ for $K = 40$ and $\delta > 5$ for $K = 128$, the nonlinear dynamics becomes dominant and the ensemble dimension increases rapidly. Perturbation sizes corresponding to $\bar{\mathcal{D}}_{ens} > 1$, however, are unrealistic in the sense that they are much larger than typical analysis errors for the L96 model as reported in Bowler (2006), Ng *et al.* (2011) and Pazó *et al.* (2013). This implies that classical BVs in our setting lack sufficient diversity. We remark that the qualitative behaviour of $\bar{\mathcal{D}}_{ens}$ does not change with the number of independent ensemble members $N$.
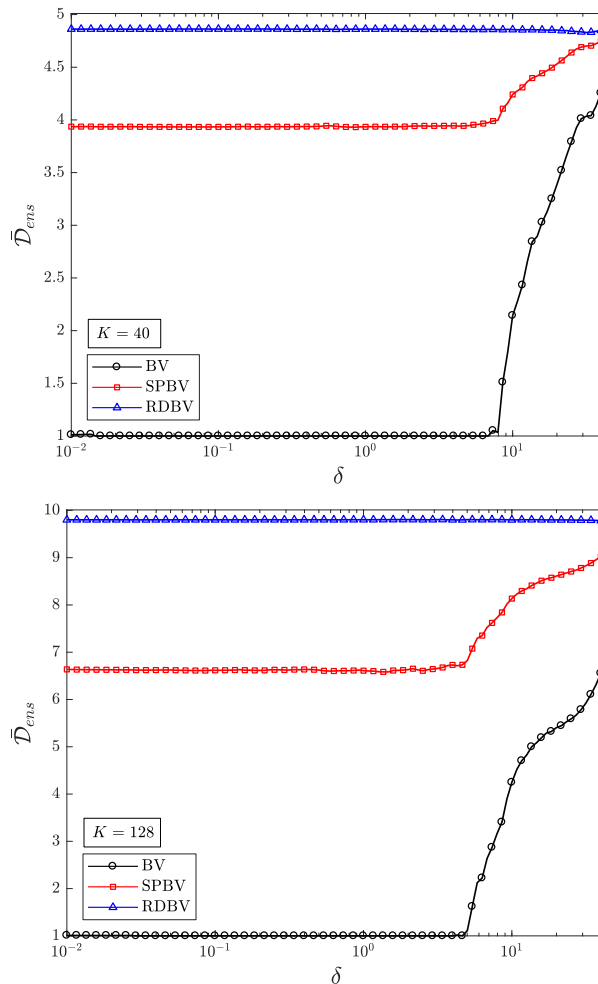
**Figure 12.** Average ensemble dimension $\bar{\mathcal{D}}_{ens}$ as a function of $\delta$ for each ensemble generation method for the L96 system (1). The SPBV ensemble was generated using $\sigma = 1.25$. Top: $K = 40$. Bottom: $K = 128$.

SPBVs and RDBVs exhibit a significant increase in the ensemble dimension. Both methods produce ensembles with a much larger ensemble dimension than the original BVs for all values of $\delta$. SPBVs maintain a consistent ensemble dimension of $\bar{\mathcal{D}}_{ens} = 3.9$ for $K = 40$ and $\bar{\mathcal{D}}_{ens} = 6.6$ for $K = 128$, before increasing in conjunction with the BVs when $\delta$ is large. RDBVs support the highest ensemble dimension as they are independent from each other. They do not attain the maximum ensemble dimension $\bar{\mathcal{D}}_{ens} = N$ since they are not strictly orthogonal. The averaged ensemble dimension of SPBVs is closer to the maximum ensemble dimension of $\mathcal{D}_{ens} = 5$ for $K = 40$ than to $\mathcal{D}_{ens} = 10$ for $K = 128$, reflecting the differing degree of localisation in the two cases; the multiplicative stochastic perturbation can generate a larger ensemble subspace the smaller the degree of localisation.

The ensemble dimension of SPBVs increases for increasing values of the noise strength $\sigma$ as shown in Figure 13 for SPBVs with $\delta = 0.1$. The ensemble dimension approaches a limiting value of $\bar{\mathcal{D}}_{ens} = 4.3$ for $K = 40$ and of $\bar{\mathcal{D}}_{ens} = 7.7$ for $K = 128$. The limiting ensemble dimension is smaller than $N$ for both $K = 40$ and $K = 128$. The difference is larger for the strongly localised case $K = 128$ for the same reason as discussed above. The observed increase of the ensemble dimension with increasing noise strength may suggest that one should use sufficiently large noise strengths $\sigma$ and moreover that the performance is insensitive to changes in $\sigma$ past some threshold value. We will see that this is correct for the forecast skill and the reliability. However, as we have seen in Section 4.1 dynamic adaptivity is lost for too large values of the noise strength $\sigma$.

*6.2. Ensemble Forecast Skill*

Figure 14 shows the RMS error $\mathcal{E}$ for BVs, SPBVs and RDBVs as a function of the perturbation size $\delta$ for lead times $\tau = 2.0$ and $\tau = 4.0$, both for the strongly localised case $K = 128$ and the weakly localised case $K = 40$. We also show as reference the climatic error $\mathcal{E}_{clim} = \sigma_{clim} \approx 3.64$ as well as the RMS error of an ETKF ensemble with a larger ensemble size of $N = K + 1 = 41$ and $N = K + 1 = 129$, respectively, to provide an upper bound for the forecast skill.

Classical BVs exhibit the largest RMS error for both cases and both lead times and for all values of $\delta$. The RMS error exhibits a local minimum for a designated perturbation size $\delta_{min}$. For $K = 40$ we find $\delta_{min} = 0.05$ for BVs for both lead times $\tau = 2.0$ and $\tau = 4.0$. For $K = 128$ we find for BVs $\delta_{min} = 0.08$ and $\delta_{min} = 0.12$ for lead times $\tau = 2.0$ and $\tau = 4.0$, respectively. Hence BV ensembles exhibit their minimal RMS error at perturbation sizes which are not consistent with the average analysis error of 0.10 and 0.18 for $K = 40$ and $K = 128$, respectively. For perturbation sizes around $\delta_{min}$, BVs are collapsed to a single member with ensemble dimension $\bar{\mathcal{D}}_{ens} = 1$ at both lead times. The RMS error assumes unacceptable high values for perturbation sizes which allow for a non-collapsed BV ensemble with $\bar{\mathcal{D}}_{ens} > 1$ for $\delta \gtrsim 8$ ($\delta \gtrsim 5$ for $K = 128$).
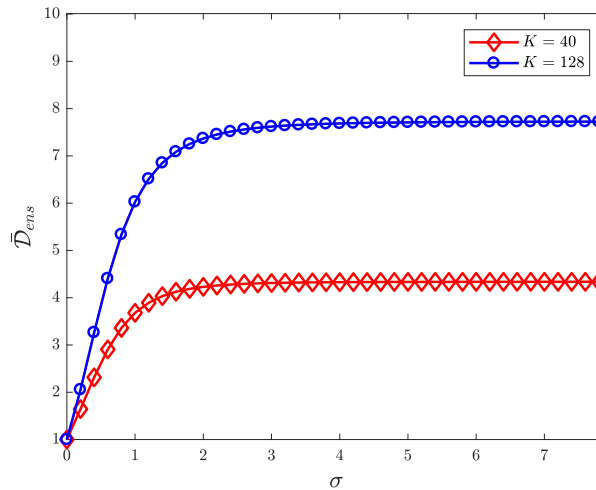
**Figure 13.** Average ensemble dimension $\bar{\mathcal{D}}_{ens}$ of an SPBV ensemble with perturbation size $\delta = 0.1$ as a function of $\sigma$ for the L96 system (1) for $K = 40$ (with $N = 5$ ensemble members) and for $K = 128$ (with $N = 10$ ensemble members).

SPBV and RDBV ensembles exhibit a significant increase in forecast skill and consistently have smaller RMS error for all values of the perturbation size $\delta$ and lead times. For the smaller lead time $\tau = 2$ we observe that all ensembles incur the same RMS error for sufficiently small values of the perturbation size $\delta$. This is because for perturbation sizes $\delta$ which are significantly smaller than the analysis error, all ensembles are highly under-dispersive and the forecast error is dominated by the analysis error. SPBV and RDBV ensembles perform almost identically in the weakly localised case $K = 40$, due to their similar spatial structures, and both ensembles perturb significantly across the whole domain capturing the regions of non-trivial analysis error. Both the SPBV and RDBV ensembles feature an RMS error minimum at approximately $\delta = 0.1$ for all lead times, which matches the size of the average analysis error, almost attaining the forecast skill of an ETKF ensemble with a much larger ensemble size of $N = 41$. We have checked that the minimal RMS error approaches the reference value provided by the ETKF for increasing ensemble sizes of SPBVs and RDBVs. This indicates that the ensembles are well-adapted to capturing the analysis error uncertainties in addition to capturing the dynamic error growth. In the strongly localised case $K = 128$, RDBVs consistently exhibit smaller forecast RMS errors compared to SPBVs. In the strongly localised case, the optimal perturbation size associated with the smallest RMS error depends on the lead time $\tau$ for SPBVs and RDBVs. For lead times $\tau = 2.0$ we find $\delta_{\min} = 0.18$ for SPBVs and $\delta_{\min} = 0.22$ for RDBVs, consistent with the average analysis error of 0.18. The RMS error of RDBVs approaches the reference value provided by the ETKF ensemble for $\tau = 2.0$. For $\tau = 4.0$ we find $\delta_{\min} = 0.39$ for SPBVs and $\delta_{\min} = 0.46$ for RDBVs, which are both inconsistent with the average analysis error. Hence, in the strongly localised case the optimal perturbation size $\delta_{\min}$ of SPBVs and RDBVs does not match the average analysis for all lead times. This suggests that we may not be efficiently capturing the uncertainties of the analysis. The difference between the ETKF reference ensemble and the BV ensembles for $K = 128$ is larger than that for $K = 40$ due to the aforementioned strongly localised nature of the BV ensemble perturbations.

We now discuss how the RMS error for SPBVs changes as the noise strength $\sigma$ is varied. We recall that when $\sigma \to 0$ SPBVs essentially reproduce the original BVs they were generated from, while once $\sigma$ is sufficiently large, the ensemble dimension saturates at some fixed value due to the rescaling back to size $\delta$ (cf. Figure 13). The RMS error of SPBVs deviates rapidly from the value attained by BVs for increasing values of $\sigma$, and then asymptotes to a constant value for large $\sigma$ (not shown). For the weakly localised case $K = 40$ the asymptotic value of the RMS error of SPBV is close to the one of RDBVs - the spatial structure of both ensembles is not related to the current state and their associated initial conditions evolve into random draws from the attractor, so both ensembles have the same statistical properties. For the strongly localised case $K = 128$, on the other hand, the asymptotic RMS error of SPBVs is larger than the one of RDBVs. In the localised case, SPBV ensembles have markedly different statistical properties to RDBVs as they sample locally with all ensemble members exhibiting non-vanishing entries in the same spatial region. We found that increasing $\sigma$ past $\sigma = 1.25$ does not increase the forecast skill in terms of RMS error for the lead times considered here.

The ensemble RMS spread $\mathcal{S}$ (8) for BVs, SPBVs and RDBVs as a function of the perturbation size $\delta$ is shown in Figure 15. For reference we also depict the corresponding RMS spread of an ETKF ensemble. The results are consistent with those of the ensemble dimension and of the RMS error $\mathcal{E}$ shown above. It is clearly seen that classical BVs are deficient in RMS spread. Classical BV ensembles exhibit a non-vanishing spread for small values of $\delta$ despite their ensemble dimension being only $\bar{\mathcal{D}}_{ens} = 1$. This is entirely due to the chosen set-up of using pairs of positive and negative BVs, and is not indicative of any non-trivial diversity of the ensemble. Once $\bar{\mathcal{D}}_{ens} > 1$ (cf. Figure 12) the RMS spread of BVs increases significantly. SPBVs and RDBVs exhibit significantly larger RMS spread compared to BVs. As for the RMS error, the differences betrween the RMS spread of RDBVs and SPBVs is more pronounced in the strongly localised case $K = 128$, reflecting the reduced ensemble space of SPBVs which are generated from a collapsed strongly localised BV by multiplicative perturbations, preserving the localisation. In the weakly localised case $K = 40$ the smaller RMS spread of SPBVs compared to RDBVs implies that SPBVs achieve the same forecast skill with less ensemble spread. In the strongly localised case $K = 128$ the increased ensemble spread of RDBVs positively impacts on their forecast skill and their RMS error.

*6.3. Reliability*

We now use the error-spread ratio and the Talagrand histogram to evaluate if the additional spread acquired by the stochastic modifications of BVs is beneficial in the sense that it leads to a reliable ensemble or whether it causes the ensemble to be simply

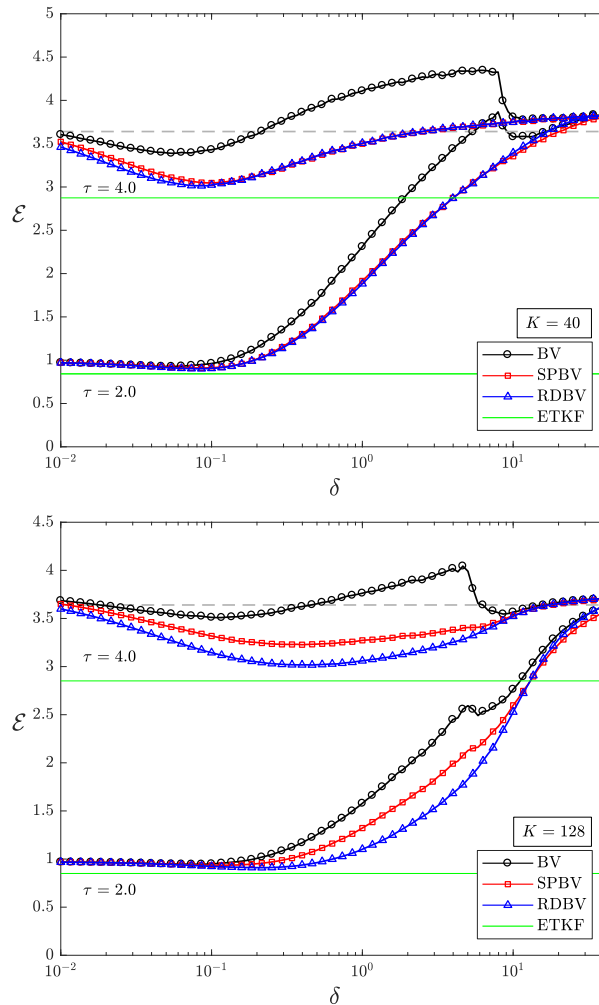**Figure 14.** RMS error $\mathcal{E}$ as a function of $\delta$ for each ensemble generation method for fixed lead times $\tau = 2.0$ and $\tau = 4.0$. The dashed reference lines denote the climatic error $\mathcal{E}_{\text{clim}}$ and the RMS error of an ETKF ensemble with $K + 1$ ensemble members at the respective lead times. Top: $K = 40$. Bottom: $K = 128$.

over-dispersive. The error-spread ratio, parameterised by lead time $\tau$, is shown in Figure 16. The markers indicate the lead times of $\tau = 2.0$, $\tau = 3.0$ and $\tau = 4.0$. The reference ETKF ensembles show a reliable one-to-one ratio. For the bred vectors and their stochastic modifications, each curve was obtained using a different value of the perturbation size $\delta$, which corresponds to the optimal perturbation size producing the smallest RMS error at lead time $\tau = 4.0$ (cf. Figure 14). For the weakly localised case $K = 40$ both SPBVs and RDBVs are close to the ideal error-spread ratio of 1, suggesting a reliable ensemble. Since the ensemble size $N = 10$ is relatively small, the error-spread curves lie just above the one-to-one ratio due to finite-size sampling error effects. On the other hand, for the strongly localised case $K = 128$ SPBV ensembles are over-dispersive for small lead times $\tau \leq 2.0$, becoming under-dispersive for lead times $\tau > 2.0$. RDBV ensembles are seen to be over-dispersive for all lead times $\tau \leq 4.0$.

Talagrand histograms are shown for each of the three forecast ensembles, averaged over all sites, in Figure 17, for lead times $\tau = 2.0$ and $\tau = 4.0$. Each histogram was again obtained using the perturbation size $\delta$ corresponding to the respective minimal RMS error (cf. Figure 14). Consistent with the results on the error-spread ratio above, the Talagrand diagrams show that SPBV and RDBV ensembles are reliable with a flat histogram in the weakly localised case $K = 40$. The reliability of the stochastically modified BV ensembles is linked to the fact that they generate non-trivial variance in regions of non-vanishing analysis error. Consistent with the observed perfect one-to-one error-spread ratio, ETKF ensembles exhibit a flat Talagrand diagram (not shown).

On the other hand, in the strongly localised case $K = 128$, when there is a strong discrepancy between the spatial structure of the analysis error and all of the SPBV ensemble members, SPBV ensembles do not lead to a flat Talagrand histogram, indicating an unreliable under-dispersive ensemble. We observe a high probability for the true state to lie outside the ensemble for both lead times. Remarkably, the interior bins of the histogram are relatively evenly populated and we do not observe the "U" shape typically associated with under-dispersive ensembles. The unusual shape of the Talagrand histogram in the strongly localised case with a flat region embedded between two peaks can be understood as follows. Consider an arbitrary component $k$ of an SPBV away from the localised region of the parent BV, which is not significantly perturbed. If $\bar{\mathcal{D}}_{ens} = 1$, then none of the members of the SPBV ensemble will be able to perturb this site. The initial conditions associated with these SPBVs at site $k$ are therefore approximately equal to the analysis mean at that site. However, typically the true state is much further away in phase space from the analysis mean. After evolving the SPBV ensemble forward in time, for reasonable lead times the ensemble has likely not developed sufficient spread to enclose the truth within its support. Hence, in the corresponding Talagrand diagram the truth falls into one of the exterior bins. This explains the peaks at the edge of the Talagrand histogram of SPBVs observed in Figure 17. On the other hand, the non-trivial components of an SPBV corresponding to the localised region have comparable magnitude to that of the analysis errors. This ensures that there are
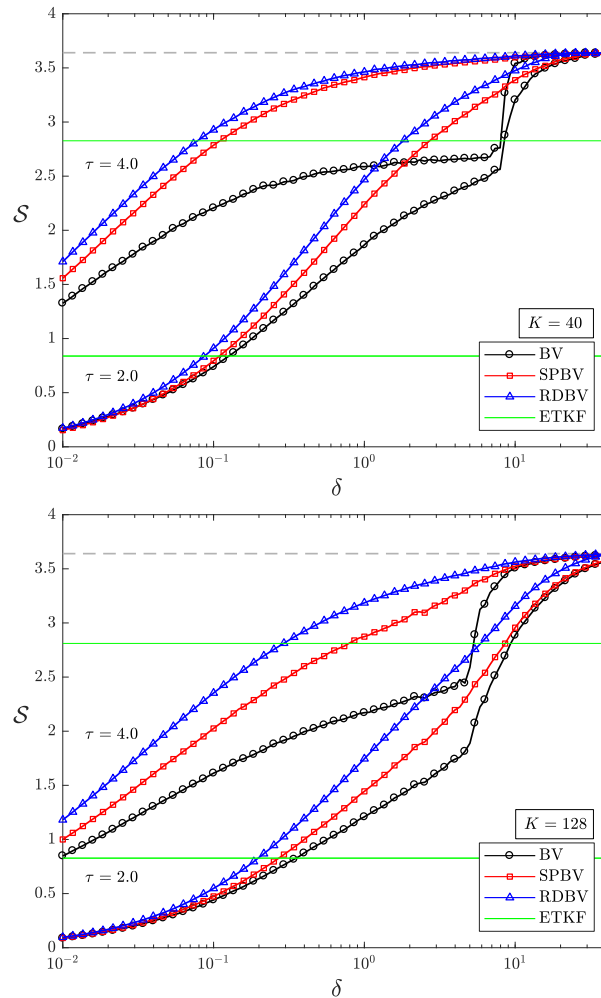
**Figure 15.** RMS spread $\mathcal{S}$ for forecast ensembles as a function of $\delta$ for each ensemble generation method for fixed lead times $\tau = 2.0$ and $\tau = 4.0$. The dashed reference lines denote the climatic error $\mathcal{E}_{\mathrm{clim}}$ and the RMS error of an ETKF ensemble with $K + 1$ ensemble members at the respective lead times. Top: $K = 40$. Bottom: $K = 128$.

426 several components of the L96 model for which the true state is contained within the ensemble, contributing to the evenly distributed
427 tally marks in the middle of the Talagrand histogram.

428     RDBV ensembles display an unusual shape of the rank histogram for $\tau = 2.0$ with two distinct modes in the strongly localised case
429 $K = 128$. This is again linked to the mismatch between the spatial structure of localised BVs and the analysis error. Individual RDBV
430 ensemble members do not efficiently sample the analysis error since each individual RDBV is localised. On the other hand it is likely
431 that each site will be significantly perturbed by at least one of the RDBV members, implying that the true state will rarely be an outlier
432 in the context of a Talagrand histogram. This combination of under- and over-dispersiveness leads to the bimodal structure observed in
433 Figure 17 for RDBVs.

434     We remark that in the strongly localised case $K = 128$, increasing the perturbation size $\delta$ does not mitigate the issue of unreliability.
435 We found that the values of $\delta$ needed to generate a flat Talagrand histogram feature significantly larger forecast errors (not shown).
436 Likewise, improving the accuracy of the observations does not allow for SPBVs and RDBVs to be reliable, but the associated smaller
437 analysis error only causes the poor reliability to occur for an associated smaller optimal perturbation size $\delta$.

439     Naively we may expect that increasing $\sigma$ will lead to more reliable SPBV ensembles. As for the RMS forecast error, we find that
440 reliability measures are insensitive to changes of $\sigma > 1.25$. The reliability of SPBV ensembles saturates for $\sigma \geq 1.25$ and is more
441 under-dispersive for $\sigma < 1.25$ for both dimension sizes $K$ (not shown).

443     We remark that the property of an ensemble to be dynamically adaptive, i.e. their relationship with covariant Lyapunov vectors and
444 that they are conditioned on the current state, does not seem to be necessarily promoting improved forecast skill and reliability. In
445 fact, the dynamically non-adapted RDBVs perform better than the dynamically adapted SPBVs with regards to forecast skills for both
446 $K = 40$ and $K = 128$, and in the weakly localised case $K = 40$ they are also slightly more reliable.

## 7. Discussion and outlook

448 We have explored the framework of stochastically modified bred vectors, developed originally for multi-scale systems in Giggins
449 and Gottwald (2019), for systems without scale separation. We considered two stochastic modifications, SPBVs which preserve any
450 eventual localisation of the their parent BVs and their spatial correlation structure, and RDBVs which do not do so. SPBVs were
451 constructed to sample the probability density function conditioned on the current state whereas RDBVs are not conditioned on the
452 current state but may evolve into future states which do not reliably estimate the probability density function at a given lead time.
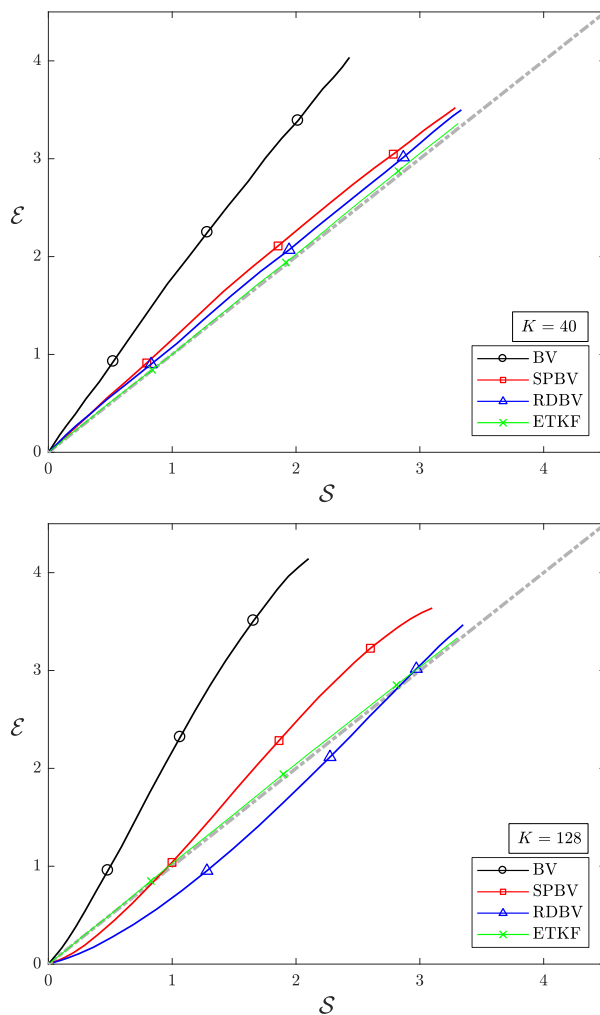
**Figure 16.** RMS error vs RMS spread, parameterised by increasing lead times from $\tau = 0$ to $\tau = 5.0$. Each bred vector ensemble was generated using a perturbation sizes $\delta_{\min}$ corresponding to minimal RMS error for each ensemble type at lead time $\tau = 4.0$ (cf. Figure 14). The markers indicate the specific lead times $\tau = 2.0$, $\tau = 3.0$ and $\tau = 4.0$. The grey dot-dashed line indicates a one-to-one ratio of RMS error and RMS spread, corresponding to a reliable ensemble. Top: $K = 40$. Bottom: $K = 128$.
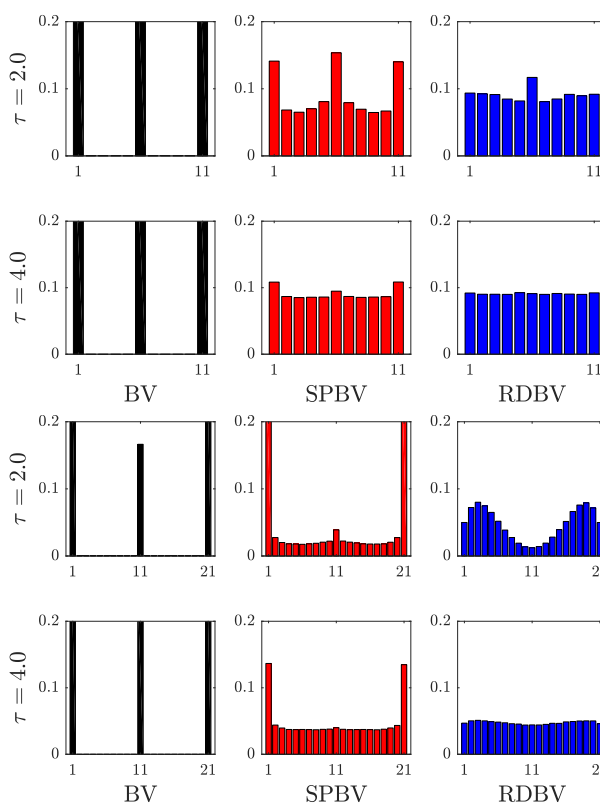


**Figure 17.** Talagrand diagrams for forecast ensembles for lead times $\tau = 2.0$ and $\tau = 4.0$. Each ensemble was generated using a perturbation sizes $\delta_{\min}$ corresponding to minimal RMS error for each ensemble type and lead time (cf. Figure 14). Top: $K = 40$. Bottom: $K = 128$.

The difference in construction renders SPBVs dynamically adapted in the sense that they project onto dynamically relevant covariant Lyapunov vectors whereas RDBVs are not dynamically adapted. Using numerical simulations of the single scale Lorenz 96 model we have shown that SPBVs and RDBVs successfully mitigate the collapse to a single ensemble member of classical BVs with significantly increased ensemble dimension for perturbation sizes $\delta$ in the range of typical analysis errors. Related to this, the forecast skill - as measured by the RMS error – and the ensemble reliability – probed by the error-spread ratio and the Talagrand diagram – are markedly improved by the stochastic modifications.

We identified the property of localisation of fast growing perturbations which is often observed in spatially extended systems to be a crucial aspect for the performance of stochastically modified BVs. Whereas localisation is advantageous to condition on the current state, it is detrimental in allowing the ensemble to perturb spatial regions of non-vanishing analysis error which are outside the localised region. This causes SPBVs to be under-dispersive (independent of the noise strength). RDBVs exhibit the better forecast skill, despite not being dynamically adapted. In the weakly localised case, RDVs and SPBVs perform equally well, and behave (per construction) statistically similarly, and both ensembles significantly improve the forecast skill and reliability of classical BV ensembles. Our results suggest that the applicability of SPBVs to single-scale systems is limited to situations with small degree of localisation.

To counteract the detrimental effect of localisation in SPBVs one could apply additive noise at all sites outside the active localised region, similar to the method proposed in Greybush *et al.* (2013). We tried this in the L96 model but did not find that it overcame the problem. The level of noise required to account for the analysis error was found to be such that the noise to BV-signal ratio was too large and the perturbation would be close to a Gaussian random perturbation. This may be though an artefact of the L96 model and the addition of spatially homogenous noise on SPBVs may still mitigate against the problem of localisation in more complex models.

We increased the diversity by introducing stochasticity directly to the bred vectors. Diversity may also be introduced stochastically by adding noise to the evolution equations generating the bred vectors. This can be done in a dynamically consistent way in the context of multi-scale dynamics (see for example Gottwald *et al.* (2017) and references therein). For multi-scale dynamics exhibiting rapid regime transitions it was moreover shown in Mitchell and Gottwald (2012); Gottwald and Harlim (2013) that stochastically parametrised forecast models for the slow variables significantly improve the analysis of an ETKF as well as the ensemble's reliability. It would be interesting to see if such stochastically perturbed dynamical models can also be used to improve the diversity of bred vector ensembles.

We would like to stress that our work only considers bred vectors here as a method for probabilistic forecasting and is concerned with improving the breeding method. We do not attempt to compare different methods such as ensemble Kalman filter ensembles, singular vectors and other methods, and to determine their individual merits. We used here as a reference ETKF ensembles with much larger ensemble dimension than the bred vector ensembles. The ETKF ensembles perform very well in the L96 model, providing superior forecast skill and reliability while also being dynamically adaptive. However, which ensemble method performs optimally as a forecast ensemble is in fact situation dependent, as pointed out, for example, recently by O'Kane *et al.* (2019). The authors found that in a coupled atmosphere ocean model forecast ensembles initialised using bred vectors with perturbation sizes tuned to capture the tropical Pacific thermocline variability, are best suited for ENSO forecasting, compared to ensembles initialised from ETKFs.

We caution the reader that one cannot simply extrapolate the performance of BVs and their stochastic modifications, SPBV and RDBV, observed in the setting of the L96 toy model to an operational setting. Realistic geophysical fluid models involve the intricate interplay of various processes running on numerous moderately separated time-scales with varying degrees of localisation of error growth, and may exhibit regime transitions between meta-stable states. In realistic operational forecasting scenarios, uncertainty in saturated sub-synoptic processes such as convective events often generate sufficient variability in the synoptic scales, and thereby prevent BVs from collapsing onto a single BV (Toth and Kalnay 1997). On the other hand, regime transitions often occur on fast time scales with rapid error growth, potentially exacerbating the ensemble collapse of classical BVs to a single ensemble member. Our work shows how to mitigate against potential under-dispersive BV ensembles (Palmer 2019) without significant additional computational cost. In particular, our work here on single-scale dynamics and our previous work on multi-scale dynamics (Giggins and Gottwald 2019) have identified the degree of time-scale separation and the degree of localisation as key to the performance of the proposed stochastic modifications; whereas localisation of error growth is crucial for the good performance of SPBVs in multi-scale dynamics, caution needs to be taken in the case of strongly localised error growth in situations without strong scale separation. The respective performance of SPBVs and RDBVs in an operational setting or in other realistic geophysical fluid flow applications will depend on the situation-dependent interplay of (moderate) time-scale separation and localisation.

## Acknowledgements

## References

Anderson JL. 1996. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *Journal of Climate* **9**(7): 1518–1530.

Annan JD. 2004. On the orthogonality of bred vectors. *Monthly Weather Review* **132**(3): 843–849, doi:10.1175/1520-0493(2004)132⟨0843:OTOOBV⟩2.0.CO;2.

Atencia A, Zawadzki I. 2017. Analogs on the Lorenz attractor and ensemble spread. *Monthly Weather Review* **145**(4): 1381–1400, doi:10.1175/MWR-D-16-0123.1, URL https://doi.org/10.1175/MWR-D-16-0123.1.

Balci N, Mazzucato AL, Restrepo JM, Sell GR. 2012. Ensemble dynamics and bred vectors. *Monthly Weather Review* **140**(7): 2308–2334, doi:10.1175/MWR-D-10-05054.1.

Bowler NE. 2006. Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model. *Tellus A* **58**(5): 538–548, doi:10.1111/j.1600-0870.2006.00197.x.

Bretherton CS, Widmann M, Dymnikov VP, Wallace JM, Bladé I. 1999. The effective number of spatial degrees of freedom of a time-varying field. *Journal of climate* **12**(7): 1990–2009.

Buizza R, Houtekamer PL, Pellerin G, Toth Z, Zhu Y, Wei M. 2005. A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Monthly Weather Review* **133**(5): 1076–1097, doi:10.1175/MWR2905.1, URL https://doi.org/10.1175/MWR2905.1.

Cai M, Kalnay E, Toth Z. 2003. Bred vectors of the Zebiak-Cane model and their potential application to ENSO predictions. *Journal of Climate* **16**(1): 40–56.

Cheng Y, Tang Y, Jackson P, Chen D, Deng Z. 2010. Ensemble construction and verification of the probabilistic ENSO prediction in the LDEO5 model. *Journal of Climate* **23**(20): 5476–5497, doi:10.1175/2010JCLI3453.1.

Corazza M, Kalnay E, Patil D, Yang SC, Morss R, Cai M, Szunyogh I, Hunt B, Yorke J. 2003. Use of the breeding technique to estimate the structure of the analysis "errors of the day". *Nonlinear Processes in Geophysics* **10**(3): 233–243.

Epstein ES. 1969. Stochastic dynamic prediction. *Tellus* **21**(6): 739–759, doi:10.1111/j.2153-3490.1969.tb00483.x.

Evans E, Bhatti N, Kinney J, Pann L, na MP, Yang SC, Kalnay E, Hansen J. 2004. Rise undergraduates find that regime changes in Lorenz's model are predictable. *Bulletin of the American Meteorological Society* **85**(4): 520–524, URL http://www.jstor.org/stable/26216961.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans* **99**(C5): 10 143–10 162.

Giggins B, Gottwald GA. 2019. Stochastically perturbed bred vectors in multi-scale systems. *Quarterly Journal of the Royal Meteorological Society* **145**: 642–658.

Ginelli F, Poggi P, Turchi A, Chaté H, Livi R, Politi A. 2007. Characterizing dynamics with covariant Lyapunov vectors. *Phys. Rev. Lett.* **99**: 130 601, doi:10.1103/PhysRevLett.99.130601.

Gottwald GA, Crommelin DT, Franzke CLE. 2017. *Stochastic climate theory*. Cambridge University Press, pp. 209–240, doi:10.1017/9781316339251.009.

Gottwald GA, Harlim J. 2013. The role of additive and multiplicative noise in filtering complex dynamical systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **469**(2155): 20130 096, doi:10.1098/rspa.2013.0096, URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2013.0096.

Greybush SJ, Kalnay E, Hoffman MJ, Wilson RJ. 2013. Identifying Martian atmospheric instabilities and their physical origins using bred vectors. *Quarterly Journal of the Royal Meteorological Society* **139**(672): 639–653, doi:10.1002/qj.1990.

Hamill TM, Colucci SJ. 1997. Verification of Eta/RSM short-range ensemble forecasts. *Monthly Weather Review* **125**(6): 1312–1327.

Herrera S, Pazó D, Fernández J, Rodríguez MA. 2011. The role of large–scale spatial patterns in the chaotic amplification of perturbations in a Lorenz–96 model. *Tellus A* **63**(5): 978–990, doi:10.1111/j.1600-0870.2011.00545.x.

Houtekamer PL, Mitchell HL. 1998. Data assimilation using an ensemble kalman filter technique. *Monthly Weather Review* **126**(3): 796–811, doi:10.1175/1520-0493(1998)126⟨0796:DAUAEK⟩2.0.CO;2, URL https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Karimi A, Paul MR. 2010. Extensive chaos in the lorenz-96 model. *Chaos: An interdisciplinary journal of nonlinear science* **20**(4): 043 105.

Keller JD, Hense A, Kornblueh L, Rhodin A. 2010. On the orthogonalization of bred vectors. *Weather and Forecasting* **25**(4): 1219–1234, doi:10.1175/2010WAF2222334.1.

Kuptsov PV, Parlitz U. 2012. Theory and computation of covariant Lyapunov vectors. *Journal of Nonlinear Science* **22**(5): 727–762, doi:10.1007/s00332-012-9126-5.

Leith CE. 1974. Theoretical skill of Monte Carlo forecasts. *Monthly Weather Review* **102**(6): 409–418, doi:10.1175/1520-0493(1974)102⟨0409:TSOMCF⟩2.0.CO;2.

Leutbecher M, Palmer T. 2008. Ensemble forecasting. *Journal of Computational Physics* **227**(7): 3515 – 3539, doi:https://doi.org/10.1016/j.jcp.2007.02.014.

Lorenz EN. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus* **17**(3): 321–333, doi:10.3402/tellusa.v17i3.9076, URL https://doi.org/10.3402/tellusa.v17i3.9076.

Lorenz EN. 1996. Predictability: A problem partly solved. In: *Proc. Seminar on predictability Vol. 1*, Palmer T (ed). ECMWF: Reading, UK, pp. 1–18.

Lorenz EN. 2006. Regimes in Simple Systems. *Journal of the Atmospheric Sciences* **63**(8): 2056–2073, doi:10.1175/JAS3727.1, URL https://doi.org/10.1175/JAS3727.1.

Lorenz EN, Emanuel KA. 1998. Optimal sites for supplementary weather observations: Simulation with a small model. *Journal of the Atmospheric Sciences* **55**(3): 399–414.

Magnusson L, Leutbecher M, Källén E. 2008. Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system. *Monthly Weather Review* **136**(11): 4092–4104.

Mitchell L, Gottwald GA. 2012. Data Assimilation in Slow-Fast Systems Using Homogenized Climate Models. *Journal of the Atmospheric Sciences* **69**(4): 1359–1377, doi:10.1175/JAS-D-11-0145.1, URL https://doi.org/10.1175/JAS-D-11-0145.1.

Newman CE, Read PL, Lewis SR. 2004. Investigating atmospheric predictability on Mars using breeding vectors in a general-circulation model. *Quarterly Journal of the Royal Meteorological Society* **130**(603): 2971–2989, doi:10.1256/qj.03.209.

Ng GHC, McLaughlin D, Entekhabi D, Ahanin A. 2011. The role of model dynamics in ensemble Kalman filter performance for chaotic systems. *Tellus A* .

Norwood A, Kalnay E, Ide K, Yang SC, Wolfe C. 2013. Lyapunov, singular and bred vectors in a multi-scale system: an empirical exploration of vectors related to instabilities. *Journal of Physics A: Mathematical and Theoretical* **46**(25): 254 021.

Oczkowski M, Szunyogh I, Patil D. 2005. Mechanisms for the development of locally low-dimensional atmospheric dynamics. *Journal of the Atmospheric Sciences* **62**(4): 1135–1156.

O'Kane TJ, Frederiksen JS. 2008. A comparison of statistical dynamical and ensemble prediction methods during blocking. *Journal of the Atmospheric Sciences* **65**(2): 426–447, doi:10.1175/2007JAS2300.1.

O'Kane TJ, Sandery PA, Monselesan DP, Sakov P, Chamberlain MA, Matear RJ, Collier MA, Squire DT, Stevens L. 2019. Coupled data assimilation and ensemble initialization with application to multiyear ENSO prediction. *Journal of Climate* **32**(4): 997–1024, doi:10.1175/JCLI-D-18-0189.1, URL https://doi.org/10.1175/JCLI-D-18-0189.1.

Palmer T. 2019. The ECMWF ensemble prediction system: Looking back (more than) 25?years and projecting forward 25?years. *Quarterly Journal of the Royal Meteorological Society* **145**(S1): 12–24, doi:10.1002/qj.3383, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3383.

Palmer TN. 1993. Extended-range atmospheric prediction and the Lorenz model. *Bulletin of the American Meteorological Society* **74**(1): 49–66, doi:10.1175/1520-0477(1993)074⟨0049:ERAPAT⟩2.0.CO;2, URL https://doi.org/10.1175/1520-0477(1993)074<0049:ERAPAT>2.0.CO;2.

Patil D, Hunt BR, Kalnay E, Yorke JA, Ott E. 2001. Local low dimensionality of atmospheric dynamics. *Physical Review Letters* **86**(26): 5878.

Pazó D, López J, Rodríguez M. 2013. The geometric norm improves ensemble forecasting with the breeding method. *Quarterly Journal of the Royal Meteorological Society* **139**(677): 2021–2032.

Pazó D, Rodríguez M, López J. 2010. Spatio-temporal evolution of perturbations in ensembles initialized by bred, Lyapunov and singular vectors. *Tellus A* **62**(1): 10–23, doi:10.1111/j.1600-0870.2009.00419.x.

Pazó D, Rodríguez MA, López JM. 2011. Maximizing the statistical diversity of an ensemble of bred vectors by using the geometric norm. *Journal of the Atmospheric Sciences* **68**(7): 1507–1512.

Peña M, Kalnay E. 2004. Separating fast and slow modes in coupled chaotic systems. *Nonlinear Processes in Geophysics* **11**(3): 319–327.

Primo C, Rodríguez M, Gutiérrez J. 2008. Logarithmic bred vectors. A new ensemble method with adjustable spread and calibration time. *Journal of Geophysical Research: Atmospheres* **113**(D5).

Talagrand O. 1999. Evaluation of probabilistic prediction systems. In: *Workshop proceedings" Workshop on predictability", 20-22 October 1997, ECMWF, Reading, UK*.

Tippett MK, Anderson JL, Bishop CH, Hamill TM, Whitaker JS. 2003. Ensemble square root filters. *Monthly Weather Review* **131**(7): 1485–1490.

Toth Z, Kalnay E. 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society* **74**(12): 2317–2330.

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

Wang X, Bishop CH. 2003. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *Journal of the Atmospheric Sciences* **60**(9): 1140–1158.

Wang X, Bishop CH, Julier SJ. 2004. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Monthly Weather Review* **132**(7): 1590–1605, doi:10.1175/1520-0493(2004)132⟨1590:WIBAEO⟩2.0.CO;2.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*. Elsevier: Oxford.

Wolfe CL, Samelson RM. 2007. An efficient method for recovering Lyapunov vectors from singular vectors. *Tellus A: Dynamic Meteorology and Oceanography* **59**(3): 355–366, doi:10.1111/j.1600-0870.2007.00234.x.

Yang SC, Keppenne C, Rienecker M, Kalnay E. 2009. Application of coupled bred vectors to seasonal-to-interannual forecasting and ocean data assimilation. *Journal of Climate* **22**(11): 2850–2870.