A New Perspective on Covariance Propagation for Data Assimilation Applications

by

S. E. Gilpin

B.A., University of California Santa Cruz, 2017

M.S., University of Colorado Boulder, 2022

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Applied Mathematics

2023

Committee Members:

Tomoko Matsuo, Chair

Stephen E. Cohn

Keith Julien

Ian Grooms

Gregory Beylkin

Gilpin, S. E. (Ph.D., Applied Mathematics)

A New Perspective on Covariance Propagation for Data Assimilation Applications

Thesis directed by Associate Professor Tomoko Matsuo

The propagation of the error covariance is an important, but not well-understood, aspect of the statistical estimation of dynamical systems, such as data assimilation. One problem often encountered during covariance propagation in data assimilation is spurious loss of variance, in which the estimated variance under-approximates the exact variance. Current approaches to understanding and combatting this issue are formulated primarily in discrete terms and often fall short of addressing fundamental causes. In an effort to understand the root cause of this spurious loss of variance, this thesis closely examines the underlying continuum covariance dynamics. Motivated by atmospheric data assimilation, the problem for states governed by the continuity and other related hyperbolic partial differential equations is considered. By analyzing the continuum covariance propagation, this thesis is the first to identify a discontinuous change in the continuum covariance dynamics along the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$. Through a series of numerical experiments and error analysis of full-rank covariance propagation, this work then demonstrates that standard methods of variance propagation are inherently inaccurate because of this discontinuous change in the continuum covariance dynamics. The result in discrete space is inaccurate variance propagation, in which full-rank covariance propagation can produce both spurious loss and gain of variance by approximating the incorrect continuum dynamics. Based on the insights gained from the continuum analysis and numerical experiments, an alternative method to mitigate this inaccurate variance propagation is proposed. As part of this attempt, this thesis develops a new correlation function required to implement the proposed alternative method of covariance propagation, and with immediate uses in other areas of data assimilation.

By studying the continuum problem, this research provides a new perspective on covariance propagation that can strengthen our understanding of its practice in data assimilation applications.

Acknowledgements

The author would like to thank the National Science Foundation (NSF) for supporting this work through the NSF Graduate Research Fellowship.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

Data assimilation is an application of Bayesian estimation to a dynamical system. This approach statistically combines a dynamical model and observational data, taking uncertainties into account, to determine the best estimate of initial model states and unknown parameters (e.g., Kalman, 1960; Kalnay, 2003; Evensen, 2009). Data assimilation plays a key role in large-scale geophysical applications and directly impacts society, for instance, through numerical weather prediction (NWP) by improving forecast accuracy (Kalnay, 2003; Bauer et al., 2015).

Uncertainty is an essential component of data assimilation and is quantified through the estimation error covariance. Data assimilation techniques evolve the estimation error covariance along with the model state, either explicitly as in the Kalman filter, implicitly as in variational methods, or using a reduced-rank approximation as in ensemble schemes (Kalman, 1960; Talagrand and Courtier, 1987; Courtier and Talagrand, 1987; Evensen, 1994). To provide context for the problem addressed in this thesis, consider a stochastic model state N-vector $\boldsymbol{q}$ that is propagated forward discretely in data assimilation schemes from time t_{k-1} to t_k as

$$\boldsymbol{q}_k = \boldsymbol{M}_{k,k-1}\boldsymbol{q}_{k-1}, \tag{1.1}$$

where $\boldsymbol{M}_{k,k-1}$ is the deterministic $N \times N$ propagation matrix representing the model dynamics. For simplicity, we consider here the linear case with no forcing, random or otherwise. From the model state we can define the $N \times N$ symmetric positive semi-definite covariance matrix at time t_k,

$$\boldsymbol{P}_k = \mathbb{E}[(\boldsymbol{q}_k - \overline{\boldsymbol{q}}_k)(\boldsymbol{q}_k - \overline{\boldsymbol{q}}_k)^{\mathrm{T}}], \tag{1.2}$$

where $\mathbb{E}[\cdot]$ is the expectation operator, $\overline{\boldsymbol{q}}_k = \mathbb{E}[\boldsymbol{q}_k]$ is the mean state, and superscript T denotes transpose. The basic equation of discrete covariance propagation then follows directly from the discrete state propagation in Eq. (1.1),

$$\boldsymbol{P}_k = \boldsymbol{M}_{k,k-1}(\boldsymbol{M}_{k,k-1}\boldsymbol{P}_{k-1})^{\mathrm{T}}, \tag{1.3}$$

where $\boldsymbol{P}_{k-1}$ and $\boldsymbol{P}_k$ are the covariance matrices at times t_{k-1} and t_k, respectively (Kalman, 1960; Jazwinski, 1970, Ch. 6). We omit a process noise term in Eq. (1.3) and will address this in later discussions.

Covariance propagation, Eq. (1.3), is a challenging mathematical problem that is highly relevant to data assimilation applications. Explicit computation of Eq. (1.3) is computationally expensive, even in modest-sized systems. Data assimilation is typically practiced in large-scale geophysical settings, such as NWP, which currently has dimensionality N on the order of 10^9, for example. In particular, inaccurate covariance propagation is a significant mathematical issue that can have serious consequences in data assimilation applications. One such example is spurious loss of variance, which if not addressed can lead to poor performance of data assimilation techniques and even filter divergence in ensemble-based schemes (Maybeck, 1982; Houtekamer and Mitchell, 1998; Furrer and Bengtsson, 2007). This variance loss is often attributed to reduced-rank approximations of the covariance or to numerical dissipation or diffusion. Thus, current methods to combat variance loss are formulated in discrete terms and can fall short of addressing fundamental causes.

It is typically assumed that if the numerical method accurately propagates the state, it is sufficient for the covariance. Therefore, Eq. (1.3) itself is often overlooked as a source of inaccurate covariance propagation in data assimilation applications. As shown in this thesis for advective dynamics, this assumption is incorrect because of the fundamental nature of the continuum covariance dynamics. In fact, the problem is more than just spurious loss of variance, but rather is inaccurate variance propagation. To demonstrate this requires returning to the continuum formulation of covariance propagation, which is not often done in the data assimilation literature. Through careful analysis of the continuum covariance propagation, my thesis uncovers the fundamental cause of

inaccurate variance propagation in advective systems.

1.1 Background and Motivation

In the past, spurious loss of variance has primarily been discussed in the context of ensemble schemes (e.g., Maybeck, 1982; Tippett et al., 2003; Evensen, 2009, and references therein), where it can be attributed to the use of reduced-rank covariance representations (Furrer and Bengtsson, 2007). As a result, several auxiliary methods have been developed to circumvent spurious loss of variance and prevent filter divergence (e.g., Anderson and Anderson, 1999; Mitchell and Houtekamer, 2000; Tippett et al., 2003; Houtekamer and Mitchell, 2005; Cohn, 2010; Berner et al., 2017; El Gharamti, 2018). It has been observed, however, that even full-rank covariance propagation can result in loss of variance, particularly for advective dynamics (Ménard and Chang, 2000; Ménard et al., 2000; Lyster et al., 2004; Ménard et al., 2021). Figure 1.1 is one such example, where Ménard et al. (2000) observes variance loss caused by full-rank covariance propagation (i.e., Eq. 1.3) in the case of the two-dimensional advection equation. Panel (a) of Fig. 1.1 plots the variances extracted from full-rank covariance propagation (Eq. 1.3) in contrast to panel (b) which computes the variance directly.

The difference between panels (a) and (b) of Fig. 1.1 suggests the loss of variance may actually be caused by the dynamics and numerical propagation, rather than from reduced-rank representations of the covariance. The loss of variance in Fig. 1.1 occurs globally but is most pronounced in the southern hemisphere. In this region, strong horizontal wind shear elongates correlation lengths along the direction of the flow while shrinking correlation lengths orthogonal to the flow. As these correlation lengths orthogonal to the flow shrink, gradients near the covariance diagonal become sharp and variance loss is observed (Ménard et al., 2000, pp. 2659–2660). Other works, such as Lyster et al. (2004) and Ménard et al. (2021), observe a similar phenomenon: as correlation lengths shrink, variance loss tends to becomes worse. As a consequence, the spurious loss of variance observed during full-rank covariance propagation is attributed to numerical dissipation or diffusion (Ménard et al. 2000, p. 2667; Pannekoucke et al. 2021, p. 8), or the inability to resolve

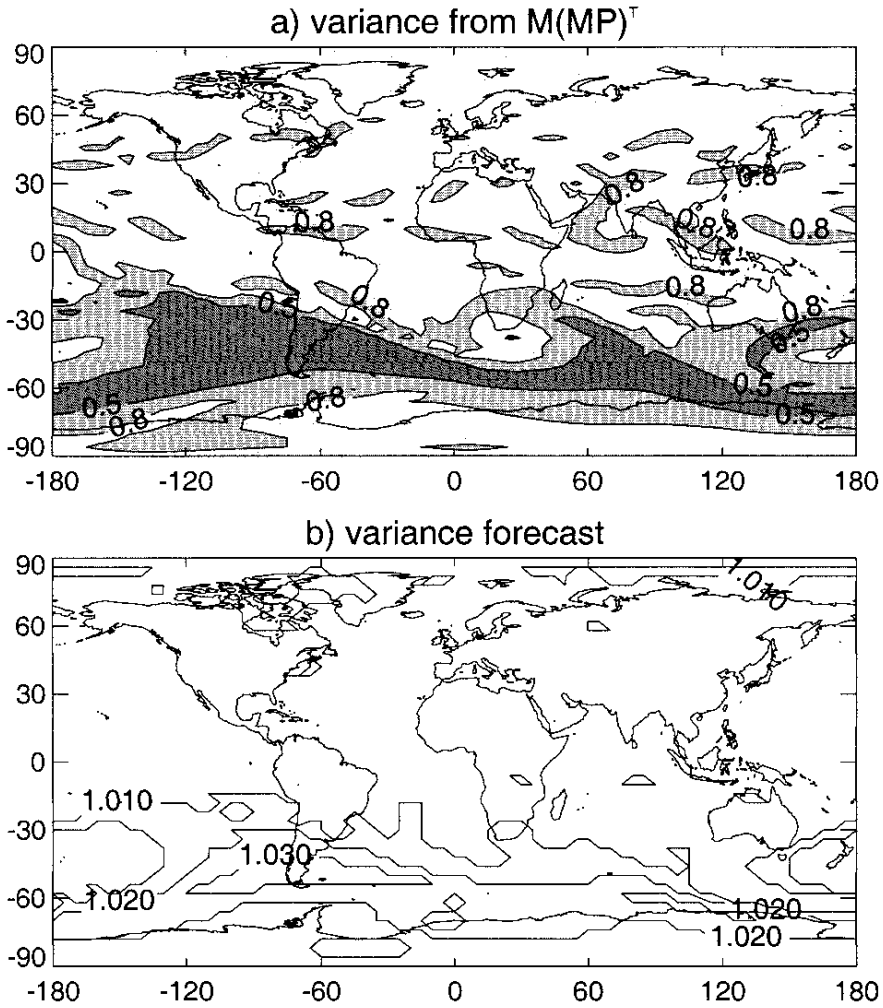

Figure 1.1: Figure 1 from Ménard et al. (2000), illustrating the variances σ^2 associated with covariances P on an isentropic surface of Earth's atmosphere governed by the advection equation (Eq. 1.6 for $b_1 = b_2 = 0$). Panel (a) plots the variances extracted from the full-rank covariance propagation, Eq. (1.3), propagated forward in time 4 days with unit initial variance. Panel (b) plots the variance obtained by numerically solving the variance equation, Eq. (1.8) for $b = 0$, directly up to the same time as in panel (a) with unit initial variance.

correlation lengths (Lyster et al., 2004, p. 2327).

Up to this point, the underlying cause of spurious loss of variance during the propagation of covariances has been unclear. Most often it is attributed to reduced-rank representation of the covariances, or in the case of advective systems, numerical dissipation during full-rank propagation. My thesis shows that the fundamental source of variance loss in advective systems is not simply due to low-rank approximations or numerical dissipation, but is caused by the continuum covariance dynamics itself. My thesis demonstrates that standard methods of variance propagation, namely that associated with Eq. (1.3), are inherently inaccurate because the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$ is a characteristic surface for advective dynamics. This characteristic surface results in a discontinuous change in the continuum covariance dynamics, which then causes problems when evolving variances according to Eq. (1.3). Therefore, the spurious loss of variance observed during full-rank covariance propagation is actually a manifestation of inaccurate variance propagation.

We can gain an intuitive sense for why covariance propagation can cause inaccurate variance propagation by interpreting the problem through the lens of differential equations. Standard methods of covariance propagation, i.e., Eq. (1.3), treat the covariance as two problems in N space dimensions, where N is the dimension of the state: the propagation matrix $\mathbf{M}_{k,k-1}$ is derived from the state dynamics in N space dimensions, and then applied twice to the covariance, first along its columns and then along its rows. The covariance, however, is a quantity in $2N$ space dimensions with properties that cannot necessarily be captured in this way, such as the characteristic surface $\boldsymbol{x}_1 = \boldsymbol{x}_2$ in advective systems. In addition, the discontinuity along the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$ is not treated carefully for standard implementations of Eq. (1.3), whereby acting along the columns and rows, Eq. (1.3) approximates across this discontinuity. My thesis shows that the inherent structure of Eq. (1.3), coupled with the discontinuous change in the continuum dynamics along $\boldsymbol{x}_1 = \boldsymbol{x}_2$, is the underlying cause of inaccurate variance propagation. This is not a problem of numerics, per se, but rather is a problem that occurs when passing from the continuous to discrete problem.

The objective of my thesis is to carefully present the problem of inaccurate variance propagation associated with covariance propagation, identify its source in advective systems, and open the

door to potential solutions. By approaching covariance propagation from the continuum, my thesis is able to rigorously demonstrate that the fundamental problem is caused by the continuum dynamics itself, affecting full-rank covariance propagation and therefore its low-rank approximations. In its totality, this work provides a new perspective on covariance propagation for data assimilation applications.

1.2 Research Objectives and Scope of Work

The results of this thesis directly impact geophysical data assimilation applications that rely on advective dynamics, such as general circulation models and chemical transport models. Thus, this work is scientifically relevant and impactful within the data assimilation community while remaining mathematically rigorous.

Motivated by atmospheric and other large-scale, geophysical data assimilation applications, consider states $q = q(\boldsymbol{x}, t)$ governed by the following hyperbolic partial differential equation (PDE) on the surface of the sphere of radius $r > 0$, $\boldsymbol{x} \in \mathbb{S}^2_r$,

$$\begin{aligned} q_t + \boldsymbol{v} \cdot \boldsymbol{\nabla} q + bq &= 0, \\ q(\boldsymbol{x}, t_0) &= q_0(\boldsymbol{x}), \end{aligned} \tag{1.4}$$

where the initial state q_0 is stochastic, while scalar $b = b(\boldsymbol{x}, t)$ and velocity field $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x}, t)$ are deterministic. For $b = \boldsymbol{\nabla} \cdot \boldsymbol{v}$, Eq. (1.4) represents mass conservation (e.g. Holton and Hakim, 2013, Ch. 2.5), and $b = 0$ yields the advection equation. The covariance $P = P(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ is defined as

$$P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = \mathbb{E}\big\{[q(\boldsymbol{x}_1, t) - \overline{q}(\boldsymbol{x}_1, t)][q(\boldsymbol{x}_2, t) - \overline{q}(\boldsymbol{x}_2, t)]\big\}, \tag{1.5}$$

where $\mathbb{E}\{\cdot\}$ is the expectation operator and $\overline{q} = \mathbb{E}\{q\}$ denotes the mean state. From the state equation, Eq. (1.4), and the covariance definition in Eq. (1.5), the covariance satisfies the following hyperbolic PDE for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{S}^2_r$,

$$\begin{aligned} P_t + \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_1 P + \boldsymbol{v}_2 \cdot \boldsymbol{\nabla}_2 P + (b_1 + b_2)P &= 0, \\ P(\boldsymbol{x}_1, \boldsymbol{x}_2, t_0) &= P_0(\boldsymbol{x}_1, \boldsymbol{x}_2). \end{aligned} \tag{1.6}$$

Subscripts $1, 2$ denote quantities evaluated with respect to $\boldsymbol{x}_1, \boldsymbol{x}_2$, respectively, with $\boldsymbol{\nabla}_i$ denoting the gradient with respect to $\boldsymbol{x}_i$ for $i = 1, 2$.

Both the model state and covariance equations given in Eq. (1.4) and (1.6) are hyperbolic PDEs, the former in two space dimensions and the latter in four. For the covariance equation, the characteristic (or trajectory) equations that describe the coordinate vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ of parcels located initially at coordinates $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$, respectively, both satisfy the same ordinary differential equation,

$$\begin{aligned} \frac{d\boldsymbol{x}}{dt} &= \boldsymbol{v}(\boldsymbol{x}(t), t), \\ \boldsymbol{x}(t_0) &= \boldsymbol{s}, \end{aligned} \tag{1.7}$$

(Shearer and Levy, 2015, following their Ch. 3 notation). The coordinate vectors $\boldsymbol{x}_i$ for $i = 1, 2$ can be written as $\boldsymbol{x}_i = \boldsymbol{x}(t; \boldsymbol{s}_i)$, which is the solution to Eq. (1.7) that represents the arrival point $\boldsymbol{x}_i$ at time t of the characteristic (trajectory) departing from the point $\boldsymbol{s}_i$ at t_0. In the case that initial parameters for $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are equal, $\boldsymbol{s}_1 = \boldsymbol{s}_2$, then it follows from Eq. (1.7) that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are also equal, $\boldsymbol{x}_1 = \boldsymbol{x}(t; \boldsymbol{s}_1) = \boldsymbol{x}(t; \boldsymbol{s}_2) = \boldsymbol{x}_2$. Therefore, initial covariances that start on the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$ (i.e., $\boldsymbol{s}_1 = \boldsymbol{s}_2$) remain on $\boldsymbol{x}_1 = \boldsymbol{x}_2$ for all time, implying that the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane is everywhere characteristic (Cohn, 1993, p. 3130).

As shown in Ch. 2, there is a discontinuous change in solutions to Eq. (1.6) along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane in the limit as correlation lengths tend to zero, for example in the vicinity of sharp gradients in the velocity field which can arise naturally, as seen for instance in Lyster et al. (2004). For initial states q_0 that are spatially correlated, meaning initial covariances P_0 have nonzero initial correlation lengths, the dynamics along the covariance diagonal satisfy the variance equation, $\sigma^2 = \sigma^2(\boldsymbol{x}, t)$ for $\boldsymbol{x} \in \mathbb{S}^2_r$,

$$\begin{aligned} \sigma^2_t + \boldsymbol{v} \cdot \boldsymbol{\nabla}\sigma^2 + 2b\sigma^2 &= 0, \\ \sigma^2(\boldsymbol{x}, t_0) &= \sigma^2_0(\boldsymbol{x}). \end{aligned} \tag{1.8}$$

However, for initial states q_0 that are spatially uncorrelated, we will see in Ch. 2, that the dynamics

along the diagonal instead satisfy the continuous spectrum equation, for $P^c = P^c(\boldsymbol{x}, t)$ and $\boldsymbol{x} \in \mathbb{S}^2_r$,

$$P^c_t + \boldsymbol{v} \cdot \boldsymbol{\nabla} P^c + \big(2b - \boldsymbol{\nabla} \cdot \boldsymbol{v}\big) P^c = 0, \tag{1.9}$$
$$P^c(\boldsymbol{x}, t_0) = P^c_0(\boldsymbol{x}),$$

which is distinct from the variance equation for divergent velocity fields (i.e., $\boldsymbol{\nabla} \cdot \boldsymbol{v} \neq 0$), as is the case in most geophysical applications.

This discontinuous change in the continuum dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane as correlation lengths tend to zero, as demonstrated in this thesis, is the direct cause of the inaccurate variance propagation associated with Eq. (1.3). Chapter 3 illustrates numerically that as correlation lengths become small, the discrete dynamics along the diagonal of the covariance matrix change and do not approximate either Eq. (1.8) or Eq. (1.9). The inaccurate variance propagation observed during discrete covariance propagation is not an issue of the numerical method. As shown in Ch. 4, even for a numerical scheme that is chosen to preserve conservation properties of the state, variance, and covariance, inaccurate variance propagation is still significant. These examples are highly relevant in geophysical applications, particularly for chemical transport models, where correlation lengths orthogonal to the flow shrink in regions of wind shear (e.g., Fig. 1.1). Increasing the spatial resolution will not necessarily mitigate the resulting inaccurate variance propagation, as correlation lengths in these regions will continue to shrink regardless of the grid resolution.

The cause of inaccurate variance propagation can be seen explicitly by analyzing how Eq. (1.3) approximates the continuum dynamics along the covariance diagonal. For standard spatial discretizations, Eq. (1.3) will approximate diagonal elements with off-diagonal elements. Recall that by positive semi-definiteness of the covariance, off-diagonal elements are necessarily smaller than the diagonal elements: $\mathbf{B}^2_{ij} \leq \mathbf{B}_{ii}\mathbf{B}_{jj}$ for any symmetric positive semi-definite matrix $\mathbf{B}$ and any arbitrary row and column indices i, j. Thus, as correlation lengths shrink or, more generally, as sharp gradients form near the diagonal, one would expect that Eq. (1.3) would poorly approximate the covariance diagonal. Since there is a discontinuous change in dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane for advective systems, approximating diagonal elements with off-diagonal elements is even more

problematic. Through careful error analysis of Eq. (1.3), we will see in Ch. 5 that the continuum dynamics approximated along the diagonal defined by Eq. (1.3) change as a function the grid length and correlation length. When the ratio of the grid length to the correlation length approaches unity, error terms arise that completely alter the dynamics being approximated along the covariance diagonal.

Motivated by the error analysis of Ch. 5, Ch. 6 explores a method of building a correction to full-rank propagation from the continuum. Since the error terms involve powers of the ratio of the grid length to the correlation length, correcting Eq. (1.3) is not straightforward. In light of this, my thesis considers an alternative method of propagating covariances that directly addresses the observed inaccurate variance propagation. Rather than evolving the covariance at full rank or approximating its evolution using reduced-rank methods, the variance and correlation length are evolved directly and covariances are reconstructed using a parametric correlation function. First termed "local covariance evolution" in Cohn (1993, Sec. 4), this method mitigates inaccurate variance propagation due to Eq. (1.3) by evolving the variance directly, is not plagued with rank-deficiency issues and spurious correlations by construction, and is computationally efficient. Chapters 7 and 9 describe this method in detail and present preliminary results. For this method, a new correlation function is derived in Ch. 8, that in addition to being used for local covariance evolution, can be applied to other areas of data assimilation, such as for covariance modeling and localization.

The outline of this thesis is as follows. Chapter 2 establishes the foundation for this work by deriving the discontinuous change in the continuum covariance dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane in the limit as correlation lengths tend to zero. Chapter 3 explores the impact of this discontinuous change in continuum dynamics during discrete covariance propagation in a simple one-dimensional example. This is followed by Ch. 4, which demonstrates that even when an appropriate numerical scheme is chosen to preserve conservation properties, inaccurate variance propagation is still quite severe. The results of Chs. 3 and 4 then motivate Ch. 5, where the continuum dynamics being approximated along the covariance diagonal by Eq. (1.3) is derived and

its dependence on correlation length is revealed. Chapter 5 concludes with a discussion of recent work by Pannekoucke et al. (2021) in the context of the error analysis in this chapter. Chapter 6 explores the potential of using this error analysis to correct full-rank propagation and addresses a correction proposed by Ménard et al. (2021). The alternative method of covariance propagation is introduced in Ch. 7, followed by its implementation in one space dimension in Ch. 9 and preliminary results. Chapter 8 describes the new parametric correlation function whose derivation is motivated by Ch. 7. Concluding remarks and discussion of future work is given in Ch. 10. Appendices A – D provide additional information that supplement the work presented in the main text.

Chapter 2

Continuum Covariance Propagation

Chapter 2 serves as the theoretical foundation upon which this thesis is built, published in Sec. 2 of Gilpin et al. (2022). From the continuum covariance equation in Eq. (1.6), we will identify the discontinuous change in dynamics along the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$ as correlation lengths tend to zero and will derive the continuous spectrum equation, Eq. (1.9). To derive this discontinuous change in dynamics requires recasting the initial value problem using integral operators and classical results from functional analysis and spectral theory. Novel to this work is the derivation of the continuum analogue of Eq. (1.3) in Sec. 2.1 and the use of the continuum polar decomposition of the fundamental solution operator, Eq. (2.24). This continuum approach is necessary to provide important insights into the behavior of the fundamental solution and covariance operators and thus derive the discontinuous change in dynamics. Appendix A includes proofs and an additional derivation of the continuous spectrum solution to supplement the discussion presented in this chapter.

For this analysis, we will consider the state and covariance equations as PDEs with solutions in the Hilbert space $L^2(\mathbb{S}_r^2)$ and $L^2(\mathbb{S}_r^2 \times \mathbb{S}_r^2)$, respectively, define the associated linear operators, and use tools from functional analysis to study these equations and operators.

2.1 Preliminaries

Let $\Omega = \mathbb{S}_r^2$ and take $\boldsymbol{x} \in \Omega$ and $t \geq t_0$. We will consider the generalized advection equation for the model state $q = q(\boldsymbol{x}, t)$ given by Eq. (1.4) in Ch. 1, reprised here,

$$\begin{aligned} q_t + \boldsymbol{v} \cdot \boldsymbol{\nabla} q + bq &= 0, \\ q(\boldsymbol{x}, t_0) &= q_0(\boldsymbol{x}). \end{aligned} \tag{1.4 revisited}$$

Recall $b = b(\boldsymbol{x}, t)$ is a scalar, noting that setting $b = \boldsymbol{\nabla} \cdot \boldsymbol{v}$ yields the continuity equation. From Eq. (1.4) we have

$$\frac{d}{dt} \int_\Omega q^2 d\boldsymbol{x} + \int_\Omega (2b - \boldsymbol{\nabla} \cdot \boldsymbol{v}) q^2 d\boldsymbol{x} = 0, \tag{2.1}$$

which is derived via integration by parts, where we assume that $2b - \boldsymbol{\nabla} \cdot \boldsymbol{v} \in L^\infty(\Omega)$ and $q_0 \in L^2(\Omega)$ so that Eq. (1.4) has a unique solution $q \in L^2(\Omega)$ for all time (e.g. using energy arguments applied to Eq. 2.1 similar to those presented in Shearer and Levy, 2015, Sec. 5.3). We write this solution as

$$q(\boldsymbol{x}, t) = (\boldsymbol{\mathcal{M}}_t q_0)(\boldsymbol{x}), \tag{2.2}$$

where $\boldsymbol{\mathcal{M}}_t \colon L^2(\Omega) \mapsto L^2(\Omega)$ is the solution operator of Eq. (1.4),

$$(\boldsymbol{\mathcal{M}}_t f)(\boldsymbol{x}) = \int_\Omega M(\boldsymbol{x}, t; \boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{2.3}$$

The subscript t on all operators denoted using the calligraphy font style, as in $\boldsymbol{\mathcal{M}}_t$ for example, indicates the operator evaluated at time t, not the time derivative. The kernel of the operator $\boldsymbol{\mathcal{M}}_t$, $M = M(\boldsymbol{x}, t; \boldsymbol{\xi})$, is the fundamental solution of Eq. (1.4),

$$\begin{aligned} M_t + \boldsymbol{v} \cdot \boldsymbol{\nabla} M + bM &= 0, \\ M(\boldsymbol{x}, t_0; \boldsymbol{\xi}) &= \delta(\boldsymbol{x}, \boldsymbol{\xi}), \end{aligned} \tag{2.4}$$

where the initial condition is the Dirac delta. Here, we simply view the Dirac delta as the kernel of the identity operator $\boldsymbol{\mathcal{I}} \colon L^2(\Omega) \mapsto L^2(\Omega)$,

$$(\boldsymbol{\mathcal{I}} f)(\boldsymbol{x}) = f(\boldsymbol{x}) = \int_\Omega \delta(\boldsymbol{x}, \boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{2.5}$$

for all $f \in L^2(\Omega)$.

Equation (2.2) is analogous to the discrete state propagation computed in data assimilation schemes. We can propagate our discrete state $\boldsymbol{q}$ in Eq. (1.1) from time t_0 to t_k,

$$\boldsymbol{q}_k = \underbrace{\boldsymbol{M}_{k,k-1}\boldsymbol{M}_{k-1,k-2}...\boldsymbol{M}_{2,1}\boldsymbol{M}_{1,0}}_{\boldsymbol{M}_{k,0}}\boldsymbol{q}_0, \tag{2.6}$$

which is the discrete version of Eq. (2.2) evaluated at time $t = t_k$; the operator $\boldsymbol{\mathcal{M}}_{t_k}$ is the continuum version of the propagation matrix $\boldsymbol{M}_{k,0}$.

With the model state now defined, we can derive the corresponding covariance evolution equation for $P = P(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ with $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Omega$ and $t \geq t_0$, given by Eq. (1.6) in Ch. 1, reprised here,

$$\begin{aligned} P_t + \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_1 P + \boldsymbol{v}_2 \cdot \boldsymbol{\nabla}_2 P + (b_1 + b_2)P = 0, \\ P(\boldsymbol{x}_1, \boldsymbol{x}_2, t_0) = P_0(\boldsymbol{x}_1, \boldsymbol{x}_2), \end{aligned} \tag{1.6 revisited}$$

where again, $\boldsymbol{\nabla}_i$ refers to the gradient with respect to $\boldsymbol{x}_i$, and $\boldsymbol{v}_i = \boldsymbol{v}(\boldsymbol{x}_i, t),\ b_i = b(\boldsymbol{x}_i, t)$ for $i = 1, 2$.

The solution of the covariance evolution equation Eq. (1.6) can be expressed using the fundamental solution operator $\boldsymbol{\mathcal{M}}_t$ and its adjoint $\boldsymbol{\mathcal{M}}_t^*$. The adjoint fundamental solution operator is defined using the inner product over the Hilbert space $L^2(\Omega)$,

$$(\boldsymbol{\mathcal{M}}_t^* f, g)_2 = (f, \boldsymbol{\mathcal{M}}_t g)_2 \quad \forall f, g \in L^2(\Omega). \tag{2.7}$$

The adjoint operator, $\boldsymbol{\mathcal{M}}_t^* \colon L^2(\Omega) \mapsto L^2(\Omega)$, can be expressed as an integral operator whose kernel M^* is the solution to the adjoint final value problem associated with Eq. (1.4),

$$(\boldsymbol{\mathcal{M}}_t^* f)(\boldsymbol{\xi}) = \int_\Omega M^*(\boldsymbol{\xi}; \boldsymbol{x}, t) f(\boldsymbol{x}) d\boldsymbol{x}. \tag{2.8}$$

With respect to the method of characteristics, the fundamental solution operator $\boldsymbol{\mathcal{M}}_t$ propagates the solution forward along the characteristics determined by departure points, and the adjoint fundamental solution operator $\boldsymbol{\mathcal{M}}_t^*$ propagates the solution backwards along the characteristics determined by arrival points. This yields the symmetry property (Courant and Hilbert, 1962, p. 729) that at any fixed time t the kernels satisfy

$$M(\boldsymbol{x}, t; \boldsymbol{\xi}) = M^*(\boldsymbol{\xi}; \boldsymbol{x}, t). \tag{2.9}$$

Using Eq. (2.9), we can express the covariance in terms of the kernels of the fundamental solution and adjoint fundamental solution operators,

$$P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = \int_\Omega \int_\Omega M(\boldsymbol{x}_1, t; \boldsymbol{\xi}_1) P_0(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) M^*(\boldsymbol{\xi}_2; \boldsymbol{x}_2, t) d\boldsymbol{\xi}_2 d\boldsymbol{\xi}_1, \tag{2.10}$$

or simply

$$\boldsymbol{\mathcal{P}}_t = \boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{P}}_0 \boldsymbol{\mathcal{M}}_t^*, \tag{2.11}$$

where $\boldsymbol{\mathcal{P}}_0 \colon L^2(\Omega) \mapsto L^2(\Omega)$ is the operator whose kernel is P_0,

$$(\boldsymbol{\mathcal{P}}_0 f)(\boldsymbol{x}_1) = \int_\Omega P_0(\boldsymbol{x}_1, \boldsymbol{x}_2) f(\boldsymbol{x}_2) d\boldsymbol{x}_2, \tag{2.12}$$

and $\boldsymbol{\mathcal{P}}_t \colon L^2(\Omega) \mapsto L^2(\Omega)$ is the resulting operator at time t,

$$(\boldsymbol{\mathcal{P}}_t f)(\boldsymbol{x}_1) = \int_\Omega P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) f(\boldsymbol{x}_2) d\boldsymbol{x}_2. \tag{2.13}$$

Thus, the covariance evolution equation Eq. (1.6) is interpreted as the evolution equation for the kernel of the covariance operator $\boldsymbol{\mathcal{P}}_t$. As with the state propagation, Eq. (2.11) evaluated at time t_k is the continuum version of the discrete covariance propagation,

$$\boldsymbol{P}_k = \boldsymbol{M}_{k,0} \boldsymbol{P}_0 \boldsymbol{M}_{k,0}^{\mathrm{T}}, \tag{2.14}$$

following from Eqs. (1.3) and (2.6).

2.2 The Polar Decomposition

We next define the (left) polar decomposition of the fundamental solution operator $\boldsymbol{\mathcal{M}}_t$, which will bring to light important properties of the covariance evolution that will be discussed in Sec. 2.3. The polar decomposition is a canonical form for all bounded linear operators on Hilbert spaces (Reed and Simon, 1972, p. 196 – 198). It is the unique decomposition $\boldsymbol{\mathcal{M}}_t = \boldsymbol{\mathcal{D}}_t \boldsymbol{\mathcal{U}}_t$ where $\boldsymbol{\mathcal{D}}_t = (\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*)^{1/2}$ and $\boldsymbol{\mathcal{U}}_t$ is a partial isometry.

To derive the polar decomposition for $\boldsymbol{\mathcal{M}}_t$, we first decompose the fundamental solution M into

$$M(\boldsymbol{x}, t; \boldsymbol{\xi}) = d(\boldsymbol{x}, t) u(\boldsymbol{x}, t; \boldsymbol{\xi}) \tag{2.15}$$

where $d = d(\boldsymbol{x}, t)$ and $u = u(\boldsymbol{x}, t; \boldsymbol{\xi})$ satisfy the following PDEs,

$$d_t + \boldsymbol{v} \cdot \boldsymbol{\nabla}\, d + \left(b - \frac{1}{2} \boldsymbol{\nabla} \cdot \boldsymbol{v} \right) d = 0,$$
$$d(\boldsymbol{x}, t_0) = 1, \tag{2.16}$$

$$u_t + \boldsymbol{v} \cdot \boldsymbol{\nabla}\, u + \frac{1}{2} (\boldsymbol{\nabla} \cdot \boldsymbol{v}) u = 0.$$
$$u(\boldsymbol{x}, t_0; \boldsymbol{\xi}) = \delta(\boldsymbol{x}, \boldsymbol{\xi}), \tag{2.17}$$

The solution u of Eq. (2.17) is quadratically conservative,

$$\frac{d}{dt} \int_\Omega u^2(\boldsymbol{x}, t; \boldsymbol{\xi}) d\boldsymbol{x} = 0, \tag{2.18}$$

and therefore defines a bounded linear operator $\boldsymbol{\mathcal{U}}_t \colon L^2(\Omega) \mapsto L^2(\Omega)$ whose kernel is the solution to Eq. (2.17),

$$(\boldsymbol{\mathcal{U}}_t f)(\boldsymbol{x}) = \int_\Omega u(\boldsymbol{x}, t; \boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{2.19}$$

The operator $\boldsymbol{\mathcal{U}}_t$ is an invertible isometry and therefore unitary.

To obtain the operator $\boldsymbol{\mathcal{D}}_t$, consider the operator $\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*$. From Eq. (2.11), the operator $\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*$ is just the covariance operator $\boldsymbol{\mathcal{P}}_t$ when the initial covariance operator $\boldsymbol{\mathcal{P}}_0$ is the identity operator Eq. (2.5); equivalently, the kernel of the operator $\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*$ is the solution to the covariance equation Eq. (1.6) when the initial covariance $P_0(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the Dirac delta, $\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$. We show in Appen. A.1 that the solution of Eq. (1.6) with the initial condition $P_0(\boldsymbol{x}_1, \boldsymbol{x}_2) = \delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is $P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = d^2(\boldsymbol{x}_1, t) \delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$, where d is the solution to Eq. (2.16). In other words, according to Eq. (2.11), $\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*$ is in fact a multiplication operator. A multiplication operator $\boldsymbol{\mathcal{K}} \colon L^2(\Omega) \mapsto L^2(\Omega)$ is defined as one for which

$$(\boldsymbol{\mathcal{K}} f)(\boldsymbol{x}) = k(\boldsymbol{x}) f(\boldsymbol{x}) = \int_\Omega k(\boldsymbol{\xi}) \delta(\boldsymbol{x}, \boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{2.20}$$

where $k(\boldsymbol{x}) \in L^\infty(\Omega)$ is the multiplication function. The operator $\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^* \colon L^2(\Omega) \mapsto L^2(\Omega)$,

$$(\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^* f)(\boldsymbol{x}) = d^2(\boldsymbol{x}, t) f(\boldsymbol{x}), \tag{2.21}$$

is a multiplication operator, where the multiplication function satisfies the following differential equation,

$$d_t^2 + \boldsymbol{v} \cdot \boldsymbol{\nabla}\, d^2 + (2b - \boldsymbol{\nabla} \cdot \boldsymbol{v})\, d^2 = 0,$$
$$d^2(\boldsymbol{x}, t_0) = 1. \tag{2.22}$$

As the operator $\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*$ is non-negative, its square root exists, and we define the operator $\boldsymbol{\mathcal{D}}_t \colon L^2(\Omega) \mapsto L^2(\Omega)$ as this square root,

$$(\boldsymbol{\mathcal{D}}_t f)(\boldsymbol{x}) = ((\boldsymbol{\mathcal{M}}_t \boldsymbol{\mathcal{M}}_t^*)^{1/2} f)(\boldsymbol{x}) = d(\boldsymbol{x}, t) f(\boldsymbol{x}), \tag{2.23}$$

with the multiplication function for $\boldsymbol{\mathcal{D}}_t$ being the solution to Eq. (2.16). Note that because the operator $\boldsymbol{\mathcal{D}}_t$ is a multiplication operator with a real-valued multiplication function, it is self-adjoint.

The decomposition Eq. (2.15) of the kernel of the fundamental solution operator gives us the (left) polar decomposition of $\boldsymbol{\mathcal{M}}_t$,

$$\boldsymbol{\mathcal{M}}_t = \boldsymbol{\mathcal{D}}_t \boldsymbol{\mathcal{U}}_t, \tag{2.24}$$

with $\boldsymbol{\mathcal{D}}_t$ and $\boldsymbol{\mathcal{U}}_t$ defined above. In the next section, we use this polar decomposition to study the continuum covariance propagation.

Remark 1. *The polar decomposition is a canonical form for all bounded linear operators on Hilbert spaces (Reed and Simon, 1972, pp. 196 – 198) and is crucial for deriving the continuous spectrum equation Eq. (1.9) in Sec. 2.3. Related to the polar decomposition is the continuum singular value decomposition (SVD), which is a canonical form for compact linear operators on Hilbert spaces (Reed and Simon, 1972, Thm. VI.17). The fundamental solution operator $\boldsymbol{\mathcal{M}}_t$ defined for the advective dynamics presented in this work, Eq. (1.4), is not compact, therefore the continuum SVD does not apply. If a diffusion term were added to Eq. (1.4), say $\kappa \boldsymbol{\nabla}^2 q$ for $\kappa > 0$ a diffusion coefficient and $\boldsymbol{\nabla}^2$ the Laplacian operator, the corresponding fundamental solution operator would become compact and a continuum SVD would exist. The discrete SVD is used widely in matrix analysis, and particularly in NWP data assimilation literature (Lorenz, 1965; Buizza and Palmer, 1995; Ehrendorfer and Tribbia, 1997; Barkmeijer et al., 1998).*

2.3 The Discontinuous Change in the Continuum Covariance Dynamics

The solution of the continuum covariance evolution equation Eq. (1.6) depends on the initial covariance $P_0(\boldsymbol{x}_1, \boldsymbol{x}_2)$, and we show in this section how the behavior of the solution changes as the initial correlation length tends to zero. To do so, we will interpret the covariance function $P(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ as the kernel of the operator $\boldsymbol{\mathcal{P}}_t$ defined in Eq. (2.13), whose evolution can be written in terms of the fundamental solution operator, its adjoint, and the initial covariance, as given in Eq. (2.11).

Consider two cases for the initial covariance $P_0(\boldsymbol{x}_1, \boldsymbol{x}_2)$. First, assume the initial covariance is continuous on $\Omega \times \Omega$, and denote it as $P_0^d(\boldsymbol{x}_1, \boldsymbol{x}_2)$. Since P_0^d is continuous, the solution to the covariance equation Eq. (1.6) is a strong solution and has a bounded L^2-norm. Therefore the corresponding covariance operator, which we will denote as $\boldsymbol{\mathcal{P}}_t^d$, is a self-adjoint Hilbert-Schmidt operator (Reed and Simon, 1972, pp. 210 – 211). Hilbert-Schmidt operators are a subclass of the compact operators, and it follows from spectral theory that self-adjoint Hilbert-Schmidt operators only contain eigenvalues in their spectrum (Hunter and Nachtergaele, 2001, pp. 230 – 232) with the possible exception of zero in the continuous spectrum. The set of eigenvalues is often referred to as the discrete spectrum, hence we can refer to the covariance operator $\boldsymbol{\mathcal{P}}_t^d$ as the discrete spectrum covariance operator.

Now, consider the case where the initial state q_0 is spatially uncorrelated, whose covariance we represent by $P_0^c(\boldsymbol{x}_1)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$, and assume that the function P_0^c is continuous on Ω. The Dirac delta in the initial covariance reduces the initial covariance operator Eq. (2.12) to a multiplication operator Eq. (2.20) with multiplication function P_0^c; denote this operator as $\boldsymbol{\mathcal{P}}_0^c$. We can see how this impacts the corresponding covariance operator, which we denote as $\boldsymbol{\mathcal{P}}_t^c$, by applying the polar decomposition Eq. (2.24) to Eq. (2.11) with $\boldsymbol{\mathcal{P}}_0 = \boldsymbol{\mathcal{P}}_0^c$,

$$\boldsymbol{\mathcal{P}}_t^c = \boldsymbol{\mathcal{D}}_t \boldsymbol{\mathcal{U}}_t \boldsymbol{\mathcal{P}}_0^c \boldsymbol{\mathcal{U}}_t^* \boldsymbol{\mathcal{D}}_t = \boldsymbol{\mathcal{D}}_t \tilde{\boldsymbol{\mathcal{P}}}_t \boldsymbol{\mathcal{D}}_t, \tag{2.25}$$

where

$$\tilde{\boldsymbol{\mathcal{P}}}_t = \boldsymbol{\mathcal{U}}_t \boldsymbol{\mathcal{P}}_0^c \boldsymbol{\mathcal{U}}_t^*. \tag{2.26}$$

From the definition of $\boldsymbol{\mathcal{U}}_t$ in Eq. (2.19), it follows that the operator $\tilde{\boldsymbol{\mathcal{P}}}_t \colon L^2(\Omega) \mapsto L^2(\Omega)$ has kernel $\tilde{P} = \tilde{P}(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ given by

$$\begin{aligned} \tilde{P}_t + \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_1 \tilde{P} + \boldsymbol{v}_2 \cdot \boldsymbol{\nabla}_2 \tilde{P} + \frac{1}{2}(\boldsymbol{\nabla}_1 \cdot \boldsymbol{v}_1 + \boldsymbol{\nabla}_2 \cdot \boldsymbol{v}_2)\tilde{P} = 0, \\ \tilde{P}(\boldsymbol{x}_1, \boldsymbol{x}_2, t_0) = P_0^c(\boldsymbol{x}_1)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2), \end{aligned} \tag{2.27}$$

whose solution is shown in Appen. A.1 to be

$$\tilde{P}(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = \tilde{P}^c(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2) \tag{2.28}$$

where

$$\begin{aligned} \tilde{P}_t^c + \boldsymbol{v} \cdot \boldsymbol{\nabla} \tilde{P}^c = 0, \\ \tilde{P}^c(\boldsymbol{x}, t_0) = P_0^c(\boldsymbol{x}). \end{aligned} \tag{2.29}$$

Thus, $\tilde{\boldsymbol{\mathcal{P}}}_t$ is a multiplication operator, and it follows from Eqs. (2.16) and Eq. (2.29) that $\boldsymbol{\mathcal{P}}_t^c$ in Eq. (2.25) is also a multiplication operator,

$$(\boldsymbol{\mathcal{P}}_t^c f)(\boldsymbol{x}_1) = P^c(\boldsymbol{x}_1, t) f(\boldsymbol{x}_1) = \int_\Omega P^c(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2) f(\boldsymbol{x}_2) d\boldsymbol{x}_2, \tag{2.30}$$

where $P^c = d^2\tilde{P}$ is the solution to the continuous spectrum equation

$$\begin{aligned} P_t^c + \boldsymbol{v} \cdot \boldsymbol{\nabla} P^c + (2b - \boldsymbol{\nabla} \cdot \boldsymbol{v})P^c = 0, \\ P^c(\boldsymbol{x}, t_0) = P_0^c(\boldsymbol{x}), \end{aligned} \tag{1.9 revisited}$$

(see Appen. A.2 for alternative derivation). Since multiplication operators contain only a continuous spectrum (Hunter and Nachtergaele, 2001, pp. 219, 240), we will refer to $\boldsymbol{\mathcal{P}}_t^c$ as the continuous spectrum covariance operator and to P^c as the continuous spectrum solution.

Thus, we have shown that the solution of the covariance evolution equation Eq. (1.6) for initial condition $P_0^c(\boldsymbol{x}_1)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is $P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = P^c(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$; white noise evolved under the state dynamics Eq. (1.4) remains white. Further, we can see explicitly that the dynamics of the covariance along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane are governed by the continuous spectrum equation Eq. (1.9)

for spatially uncorrelated initial states, rather than by the variance equation

$$\sigma_t^2 + \boldsymbol{v} \cdot \boldsymbol{\nabla} \sigma^2 + 2b\sigma^2 = 0,$$

$$\sigma^2(\boldsymbol{x}, t_0) = \sigma_0^2(\boldsymbol{x}) = P_0^d(\boldsymbol{x}, \boldsymbol{x}), \qquad \text{(1.8 revisited)}$$

which follows directly from Eq. (1.6) when the initial covariance is continuous on $\Omega \times \Omega$. Only in cases where the velocity field is divergence-free are the continuous spectrum equation and variance equation identical. The dynamics of the covariance along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane are governed by the variance equation Eq. (1.8) for all continuous initial covariances, independently of nonzero initial correlation lengths, but at zero correlation length the dynamics change abruptly to those of the continuous spectrum equation Eq. (1.9).

2.4 Conclusions and Discussion

Covariances associated with spatially uncorrelated initial states, which correspond to covariances with zero initial correlation length scales, are limiting cases in our analysis, as well as in practice. Atmospheric wind fields, for example, have sharp vertical correlation structures relative to long horizontal correlations that need to be represented in covariances for NWP models (Phillips, 1986), and horizontal wind shear leads to tracer correlations that shrink in the direction perpendicular to the flow such that they are no longer spatially resolved (Lyster et al., 2004). Through careful analysis and use of the polar decomposition, we are able to derive the discontinuous change in dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane as the correlation length approaches zero for covariances governed by Eq. (1.6), which is not readily apparent when first considering the model state and covariance equations nor when only considering these equations in discrete space.

The analysis performed in this chapter serves as the foundation for understanding covariance propagation in both the continuum and discrete space. We will see in the next chapter thats knowing the full, continuum behavior of the covariance will be essential to properly interpret the numerical experiments.

Chapter 3

Numerical Demonstration Part I: Inaccurate Variance Propagation

With the discontinuous change in the continuum dynamics established in Ch. 2, one may then ask how this behavior impacts covariance propagation in discrete space, and in particular how it relates to the spurious loss of variance observed in Lyster et al. (2004) and Ménard et al. (2000, 2021), for example. In this chapter, published in Sec. 3 of Gilpin et al. (2022), we begin to answer this question through a series of numerical experiments. We will consider a simple, one-dimensional example so that the numerical results can be compared to a known exact solution. We will find that inaccurate variance propagation is the actual problem, resulting in both loss and gain of variance relative to the exact solution. This is also the case when initial covariances are diagonal, as the covariance matrix diagonals extracted from Eq. (1.3) hardly resemble the continuous spectrum solution. Thus, this chapter demonstrates that standard methods of variance propagation, i.e. those associated with Eq. (1.3), are inherently inaccurate because of the discontinuous change in the continuum covariance dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane.

3.1 Experimental Setup

For these experiments, we will consider the one-dimensional version of the continuity equation, Eq. (1.4) for $b = v_x$, over the unit circle $\mathbb{S}^1_1$. We take the velocity field to be independent of time and spatially-varying,

$$v(x) = \sin(x) + 2. \tag{3.1}$$

The exact solution to the one-dimensional continuity equation with this velocity field can be obtained explicitly using the method of characteristics and is used as reference for our experiments; see Appen. B.1 for further discussion.

The spatial domain $\mathbb{S}^1_1$ is discretized on a uniform grid, $x_i = i\Delta x,\ i = 0, 1, ..., N-1$ where $N = 200$ and $\Delta x = \frac{2\pi}{N}$. The time discretization is given by $t_k = k\Delta t$ where the time step Δt is determined from the Courant-Friedrichs-Lewy number λ,

$$\lambda = \max_{x\in[0,2\pi]} |v(x)|\frac{\Delta t}{\Delta x} = 3\frac{\Delta t}{\Delta x} \leq 1. \tag{3.2}$$

For these experiments, we take $\lambda = 1$. We ran experiments with several other values of $\lambda < 1$ (not shown) and found it did not have a significant impact on the results.

3.1.1 Numerical Covariance Propagation

We use two methods of propagation to illustrate the impact of numerical schemes on the discrete covariance propagation Eq. (1.3) and to leverage insights from continuum covariance analysis in the form of the polar decomposition. The covariance matrix is propagated discretely using either of two methods:

(1) Traditional propagation: the covariance is propagated as in Eq. (1.3), where the matrix $\boldsymbol{M}_{k,k-1}$ is the finite difference discretization of the fundamental solution equation 2.4.

(2) Polar decomposition propagation: the polar decomposition of the fundamental solution operator Eq. (2.24) is discretized and used in place of the matrix $\boldsymbol{M}_{k,k-1}$ in Eq. (1.3).

In the polar decomposition propagation, Eq. (2.24) is discretized as follows. The operator $\boldsymbol{\mathcal{D}}_t$ of Eq. (2.23) is a self-adjoint multiplication operator, therefore when $t = t_k$ its corresponding discretization is the diagonal matrix $\boldsymbol{D}_{k,0}$ whose diagonal elements are $(\boldsymbol{D}_{k,0})_{ii} = d(x_i, t_k)$. Here $d(x_i, t_k)$ is the solution to Eq. (2.16) evaluated on the discrete spatial grid at time t_k, which we generate using the exact solution to Eq. (2.16), as discussed in Appen. B.1. The discretization of the operator $\boldsymbol{\mathcal{U}}_t$ of Eq. (2.19) is done via finite differences to generate the propagation matrix

$\boldsymbol{U}_{k,k-1}$ corresponding to Eq. (2.17). The matrix $\boldsymbol{U}_{k,k-1}$, which propagates the solution from time t_{k-1} to t_k, is independent of time by virtue of Eq. (3.1), therefore we will denote $\boldsymbol{U}_{k,k-1} = \boldsymbol{U}$ for simplicity. To compute the covariance matrix at time t_k, $\boldsymbol{P}_k$, we do not construct the matrix $\boldsymbol{M}_{k,k-1}$ explicitly using the polar decomposition, but instead compute the covariance as follows:

$$\begin{aligned}\boldsymbol{P}_k &= \boldsymbol{D}_{k,0}\boldsymbol{U}_{k,k-1}\boldsymbol{U}_{k-1,k-2}...\boldsymbol{U}_{1,0}\boldsymbol{P}_0\boldsymbol{U}_{1,0}^{\mathrm{T}}...\boldsymbol{U}_{k,k-1}^{\mathrm{T}}\boldsymbol{D}_{k,0} \\ &= \boldsymbol{D}_{k,0}\boldsymbol{U}^k\boldsymbol{P}_0(\boldsymbol{U}^{\mathrm{T}})^k\boldsymbol{D}_{k,0}. \end{aligned} \tag{3.3}$$

The Lax-Wendroff (Lax and Wendroff, 1960) and Crank-Nicolson (Crank and Nicolson, 1947) finite difference schemes are used to generate the matrices $\boldsymbol{U}_{k,k-1}$ and $\boldsymbol{M}_{k,k-1}$ corresponding to their respective PDEs in Eqs. (2.17) and (2.4). Like $\boldsymbol{U}_{k,k-1}$, $\boldsymbol{M}_{k,k-1}$ is independent of time, therefore through the rest of the chapter we will denote $\boldsymbol{M}_{k,k-1}$ simply as $\boldsymbol{M}$. We choose these two simple finite difference schemes to illustrate their contrasting behaviors, particularly when generating the matrix $\boldsymbol{U}$. According to the continuum analysis, the operator $\boldsymbol{\mathcal{U}}_t$ is unitary. The Crank-Nicolson discretization preserves the unitary property of $\boldsymbol{\mathcal{U}}_t$ because the scheme is quadratically conservative, but the Lax-Wendroff scheme does not. In the numerical results, Secs. 3.2 and 3.3, propagated covariances are labeled by the finite difference scheme used to construct the matrices $\boldsymbol{M}$ and $\boldsymbol{U}$ (Lax-Wendroff or Crank-Nicolson) followed by the method of propagation (traditional or polar decomposition) as described in the beginning of this section.

3.1.2 Initial Covariances

We generate four types of initial covariances from the following two different correlation kernel functions: the Gaspari-Cohn (GC) compactly-supported, piecewise rational function (Gaspari and Cohn, 1999, Eq. 4.10), and the first order autoregressive function (FOAR) (Gaspari et al., 2006, their Eq. 23). Using each correlation function, we construct the initial covariance matrix with either a constant initial variance or spatially-varying initial variance. Initial covariances with constant initial variance take the initial variance to be one, while initial covariances with the spatially-varying

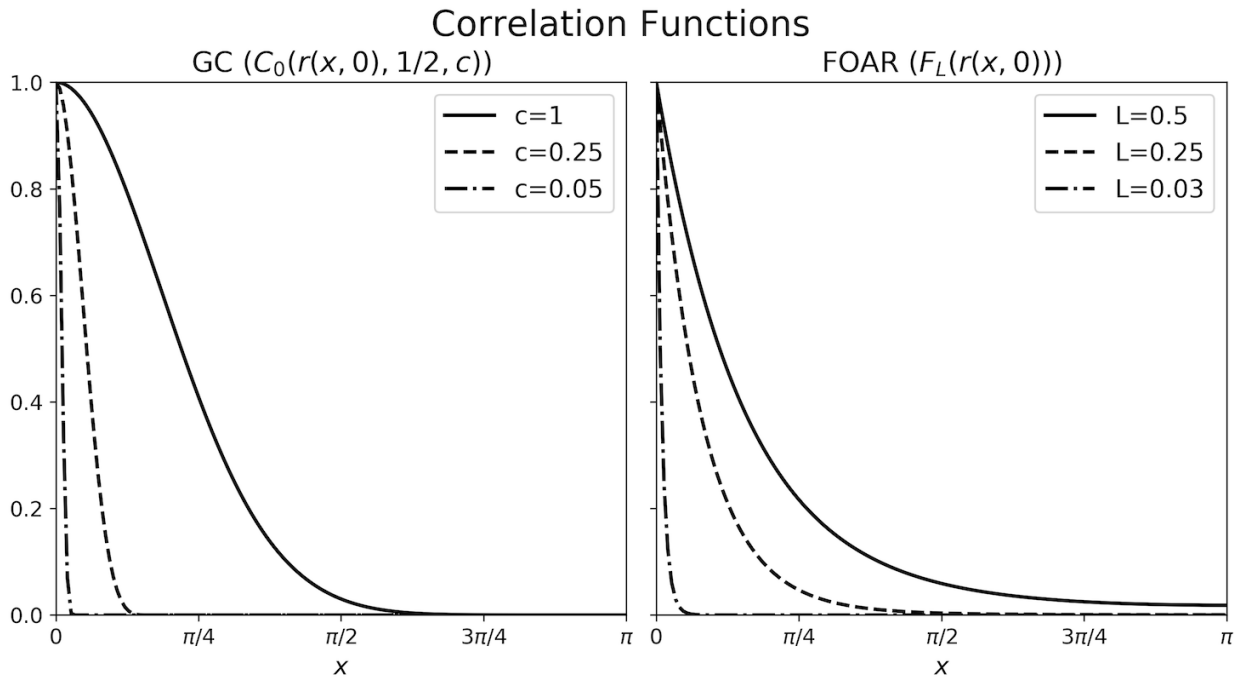

Figure 3.1: Examples of the correlation functions used to generate the initial covariances for numerical propagation. The GC correlation functions (left) are functions of compact support length parameter c, where the FOAR correlation functions (right) are functions of length scale L.

initial variance take the initial variance as the square of the standard deviation

$$\sigma_0(x) = \frac{\sin(3x)}{3} + 1. \tag{3.4}$$

The GC correlation function $C_0(r(x_i, x_j), 1/2, c)$ is a compactly supported approximation to a Gaussian function, supported on the interval $0 \leq r(x_i, x_j) \leq 2c$, where

$$r(x_i, x_j) = 2\sin(|x_i - x_j|/2) \tag{3.5}$$

is the chordal distance between x_i and x_j on $\mathbb{S}^1_1$. On the uniform spatial grid of 200 grid points for these experiments, values of $c = 1$, 0.25, and 0.05 correspond to 100, 16, and 3 grid lengths (Δx), respectively, from the peak of the correlation function to where it becomes zero, and 33, 8, and just 1 grid length, respectively, from the peak value of 1 to values less than 0.2; see Fig. 3.1 for these examples.

The FOAR correlation function given by

$$C_{FOAR}(x_i, x_j) = e^{-r(x_i,x_j)/\mathsf{L}} \tag{3.6}$$

is continuous but non-differentiable at the origin because of its cusp-like behavior (see Fig. 3.1). As with the GC correlation function, the chordal distance Eq. (3.5) is used to reflect periodicity of the domain. The FOAR correlation functions are nonzero on the full spatial domain. On the spatial grid, $\mathsf{L} = 0.5$, 0.25, and 0.03 correspond to 26, 12, and just 1 grid length, respectively, from the peak of the correlation to where it becomes less than 0.2.

3.2 Experimental Results

From the continuum analysis, one might expect the diagonal of the covariance matrix $\boldsymbol{P}_k$ to behave according to the variance equation for covariances with nonzero initial correlation lengths and according to the continuous spectrum equation for covariances with zero initial correlation length. To establish a baseline, Fig. 3.2 illustrates the solutions to the variance equation (1D version of Eq. (1.8) for $b = v_x$) and continuous spectrum equation (1D version of Eq. (1.9) for $b = v_x$) for

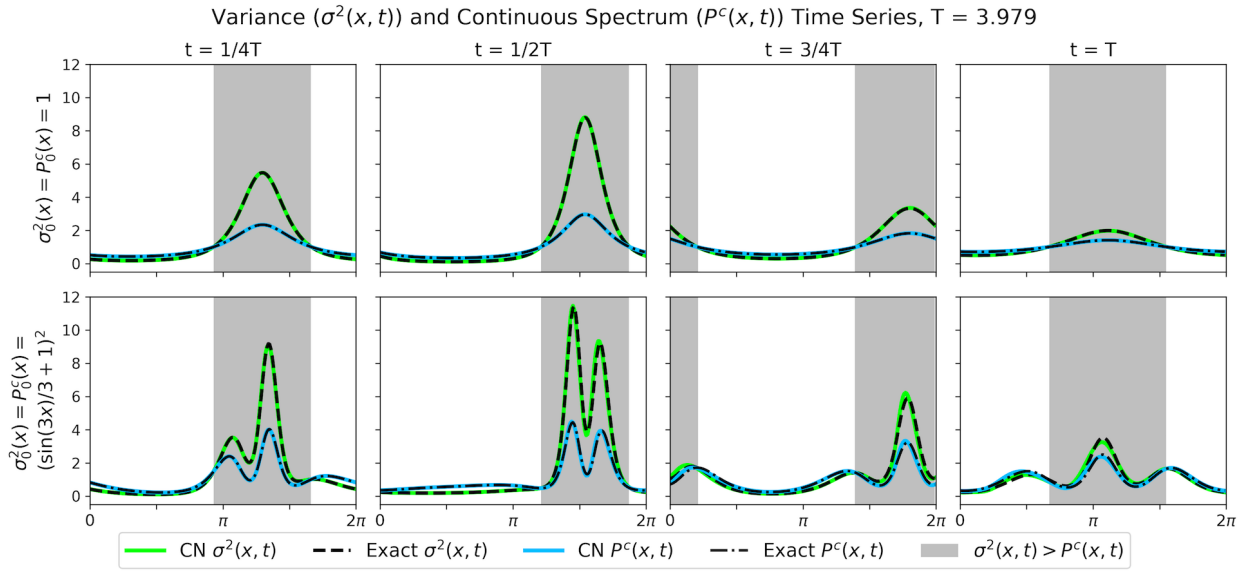

Figure 3.2: Top: Solutions to the one-dimensional version of variance equation Eq. (1.8) and continuous spectrum equation Eq. (1.9) for $b = v_x$ at various times with velocity field Eq. (3.1), for unit initial condition (top row) and spatially-varying initial condition (bottom row) taken to be the square of Eq. (3.4). Time $T = 3.979$ corresponds to slightly after a full period. Exact solutions for each case are given in black dashed (variance) and black dot-dashed lines (continuous spectrum). Green curves denote the solution to the variance equation computed with Crank-Nicolson (abbreviated CN); blue curves are solutions to the continuous spectrum equation computed with Crank-Nicolson. Regions highlighted in grey correspond to regions where the exact solution to the variance equation is greater than the exact solution to the continuous spectrum equation, i.e. $\sigma^2 > P^c$; unhighlighted regions correspond to regions where $\sigma^2 < P^c$.

unit initial condition and for the spatially-varying initial condition given by the square of Eq. (3.4). Considering the exact solutions first (black), we see that the dynamics of the variance solution and continuous spectrum solution are quite different due to the spatially-varying velocity field Eq. (3.1). We also see that solving either the variance equation or continuous spectrum equation directly using the Crank-Nicolson scheme (colored) produces solutions that are nearly indistinguishable from the exact solutions. Solutions to the variance and continuous spectrum equations computed using Lax-Wendroff differ very slightly from Crank-Nicolson and are not shown.

Figure 3.2 shows that propagating the diagonal of the covariance matrix $\boldsymbol{P}_k$ numerically, independently of the rest of the matrix, using the known dynamics of either the variance equation or the continuous spectrum equation, produces minimal discretization error. Figure 3.3 illustrates further that, at least for covariance matrices with relatively long initial correlation lengths, the numerically propagated full covariance itself also contains only minor discretization errors typically expected from finite difference approximations. Figure 3.3 shows results for the GC initial correlation supported on the full spatial domain ($c = 1$). The dissipative behavior of both Lax-Wendroff schemes are clear in the normalized spectra (left panels), whereas both the traditional Crank-Nicolson propagation and polar decomposition propagation using the Crank-Nicolson $\boldsymbol{U}$ (hereafter referred to as Crank-Nicolson polar decomposition) capture the normalized spectra quite well. The small amount of variance loss and gain seen in both the constant initial variance and spatially-varying initial variance cases (middle panels of Fig. 3.3) are consistent with dissipation and phase errors expected from finite differences. The polar decomposition methods (dashed) reduce the errors in the diagonals of the covariance matrices only slightly compared to the traditional methods (solid). The correlations at row 150 (right panels), which corresponds to where the variance reaches a maximum as the velocity field is at a minimum, are nearly identical to the exact correlations, suggesting minimal errors in correlation propagation.

The accuracy of numerically propagated full covariances is much worse for short initial correlation lengths, increasingly so as they approach zero. We monitor the total amount of variance lost or gained over time through the trace of the covariance matrix, $Tr(\boldsymbol{P}_k)$. Figure 3.4 shows the trace

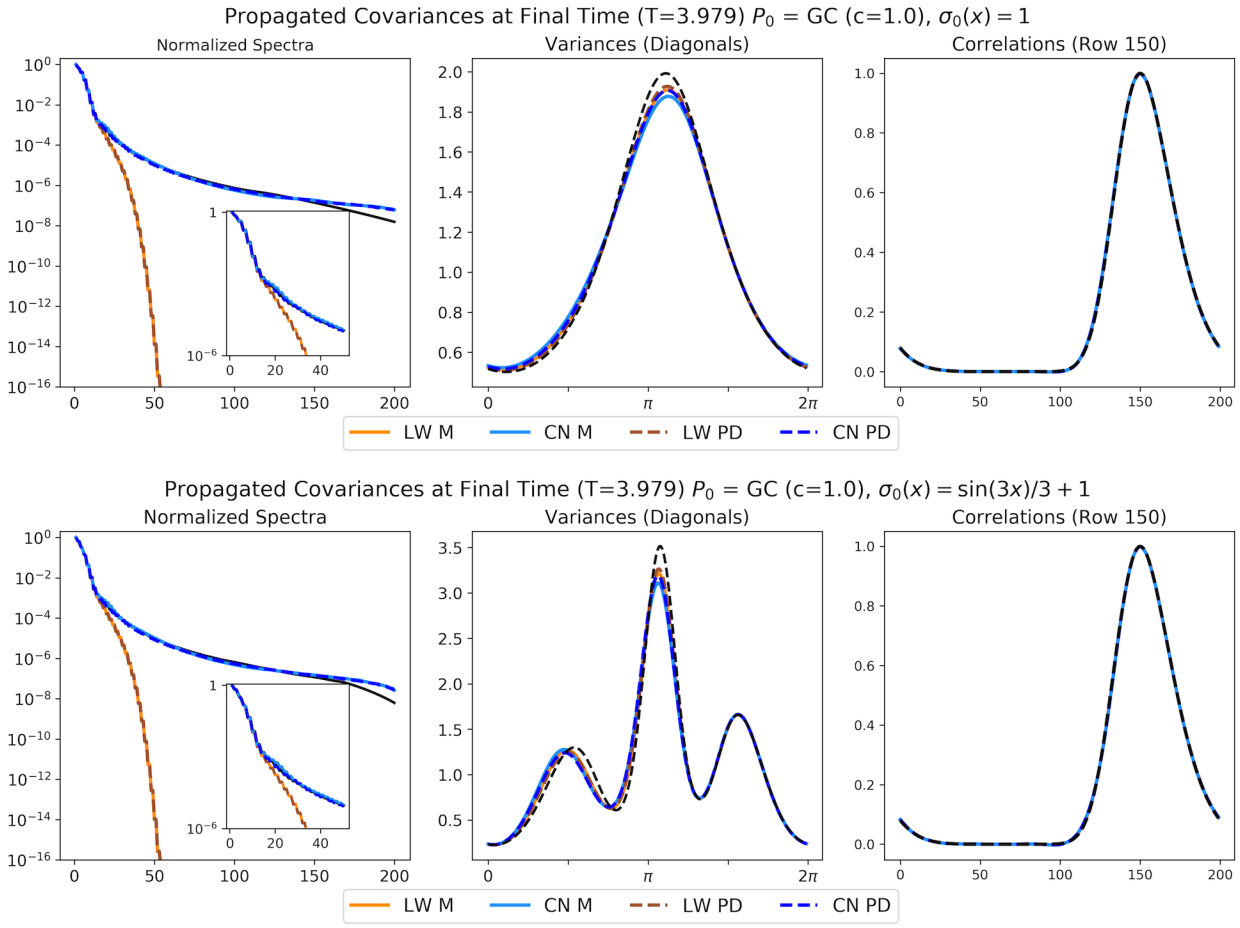

Figure 3.3: Propagated covariances at the final time T (slightly after a full period) for GC initial correlations with $c = 1$ (supported on the full domain) for all four propagation methods. Top row: constant initial variance. Bottom row: spatially-varying initial variance. The left panels correspond to the normalized spectra (relative to largest eigenvalue), the middle panels show the discrete variances (covariance matrix diagonals), and the right panels show the correlations at row 150 (location of maximum in variance in space and time). The exact normalized spectra are given in solid black, and the exact variances and correlations are given in black dashed. Crank-Nicolson (CN) and Lax-Wendroff (LW) M refer to traditional propagation using Crank-Nicolson or Lax-Wendroff (solid). CN and LW PD refer to propagation using the polar decomposition (dashed) with the matrix $\boldsymbol{U}$ constructed via the Crank-Nicolson or Lax-Wendroff scheme, respectively.

time series for both the GC and FOAR cases with spatially-varying initial variance as initial correlation lengths tend to zero (the constant initial variance case results are similar to Fig. 3.4 and not shown). The GC and FOAR cases in Fig. 3.4 exhibit similar behaviors in the trace over time, even though their initial correlation structures are quite different. As the initial correlation lengths tend to zero, the amount of variance lost during propagation increases strikingly in the Lax-Wendroff schemes. The polar decomposition propagation schemes (dashed) are an improvement over traditional propagation in some cases but are worse in others, and generally suffer similar amounts of variance loss. We also observe that as initial correlation lengths tend to zero, the numerical schemes gradually approach their own limiting behavior at $c = \mathsf{L} = 0$, rather than a discontinuous change in dynamics as seen in the continuum analysis.

Had we not performed the continuum analysis in Ch. 2, one might assume that the variance loss in Fig. 3.4 is caused simply by dissipation. However, the Crank-Nicolson scheme is not dissipative and yet produces significant variance loss. In fact, we see for the traditional Crank-Nicolson propagation (solid light blue) that there are regions of both variance loss and gain. We also see that for short, nonzero initial correlation lengths ($c = 0.05$, $\mathsf{L} = 0.03$ in particular), the numerical schemes better approximate the limiting case of $c = \mathsf{L} = 0$ than the correct behavior for $c,\ \mathsf{L} > 0$. This suggests that inaccurate discrete diagonal propagation is particularly pronounced for short correlation lengths.

The behavior of the trace time series in Fig. 3.4 indicates that covariance propagation itself can be a source of spurious loss and gain of variance, however it does not indicate where exactly this manifests itself. To gain a better understanding of the source of variance loss and gain, we examine various aspects of the propagated covariance matrix for different types of initial covariances as we did in Fig. 3.3, but now for covariances with shorter initial correlation lengths.

Figures 3.5 and 3.6 are final time snapshots (in the same format as shown in Fig. 3.3) of propagated covariances specified using the GC correlation function with $c = 0.25$ and FOAR correlation function with $\mathsf{L} = 0.25$, respectively, which are initially well-resolved as described in Sec. 3.1.2 and correspond to the mildest variance loss and gain cases shown in leftmost panels

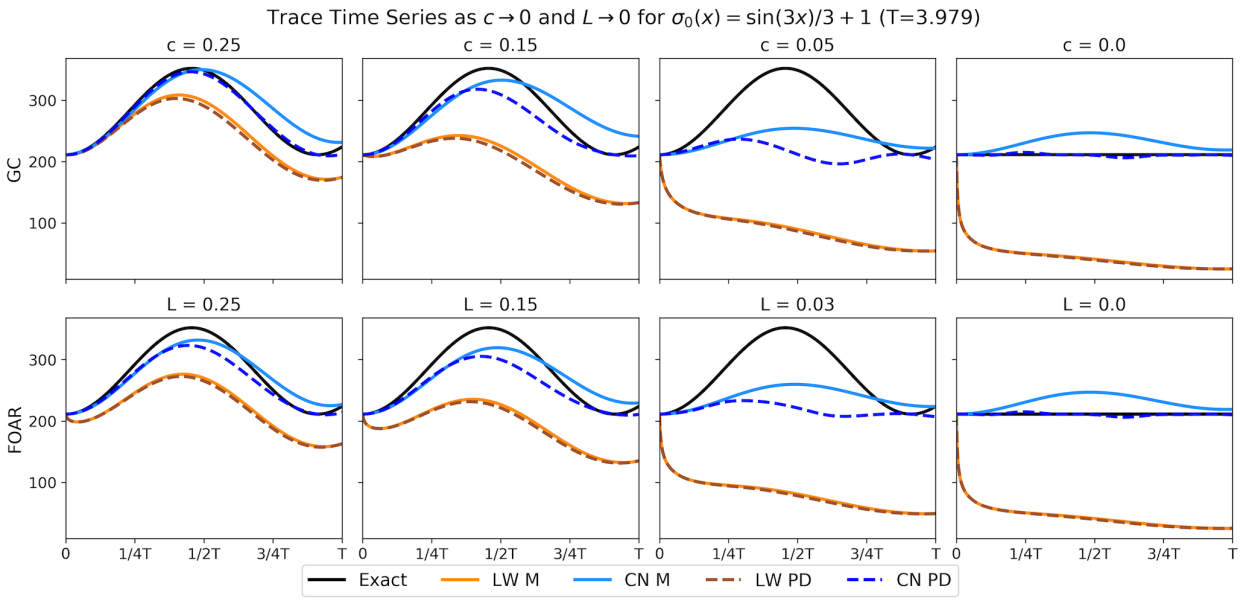

Figure 3.4: Trace time series for GC (top row) and FOAR (bottom row) for the spatially-varying initial variance (the square of Eq. 3.4). Each panel corresponds to different values of c and L, decreasing from left to right towards $c = \mathsf{L} = 0$. Refer to Fig. 3.3 for description of the curves. For the cases when $c = \mathsf{L} = 0$ (rightmost panels), the exact solutions (solid black) are constant in time due to the fact that continuous spectrum solution P^c satisfies the continuity equation (Eq. 1.9 with $b = v_x$), hence its integral over space is constant.

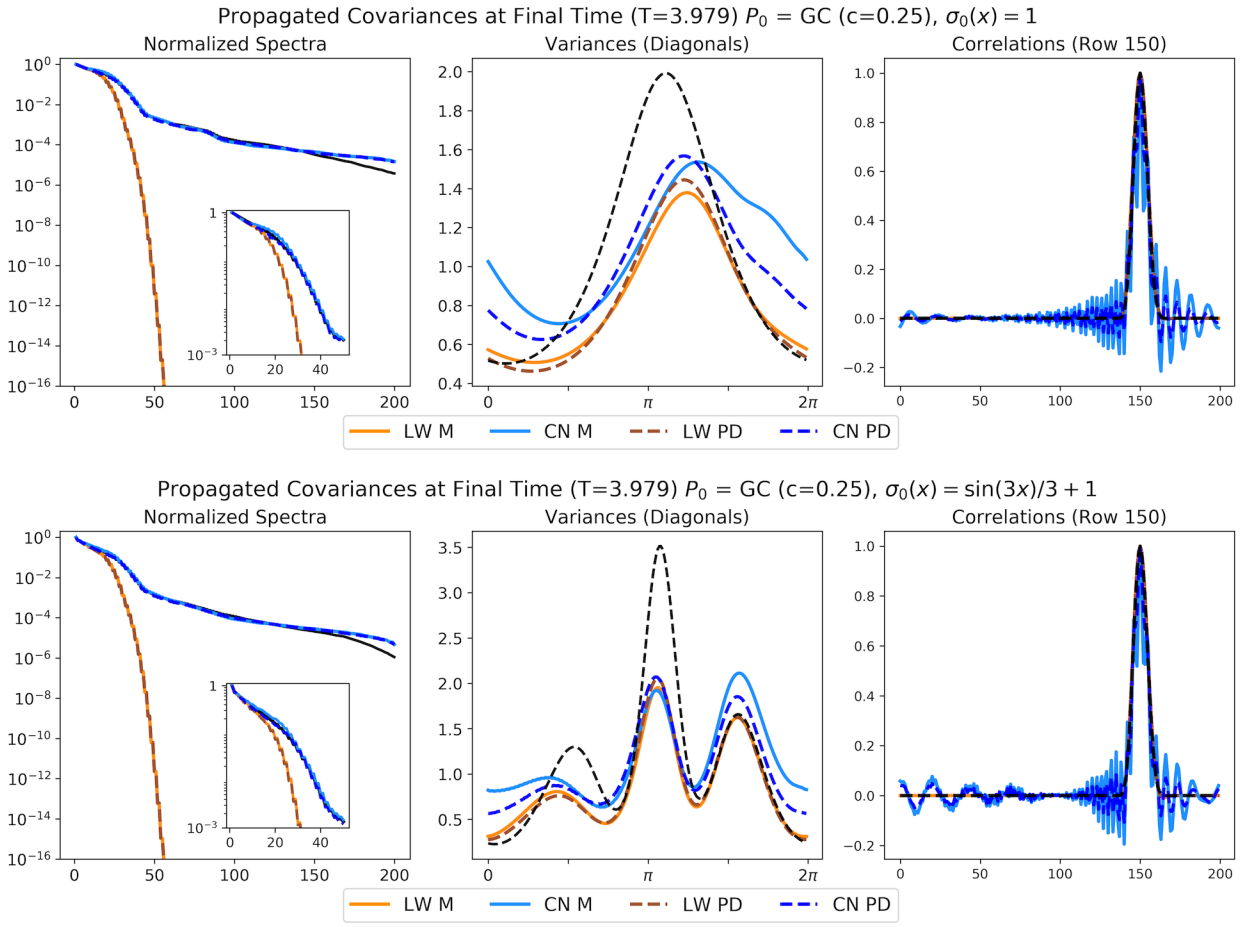

Figure 3.5: Same as Fig. 3.3 for $c = 0.25$.

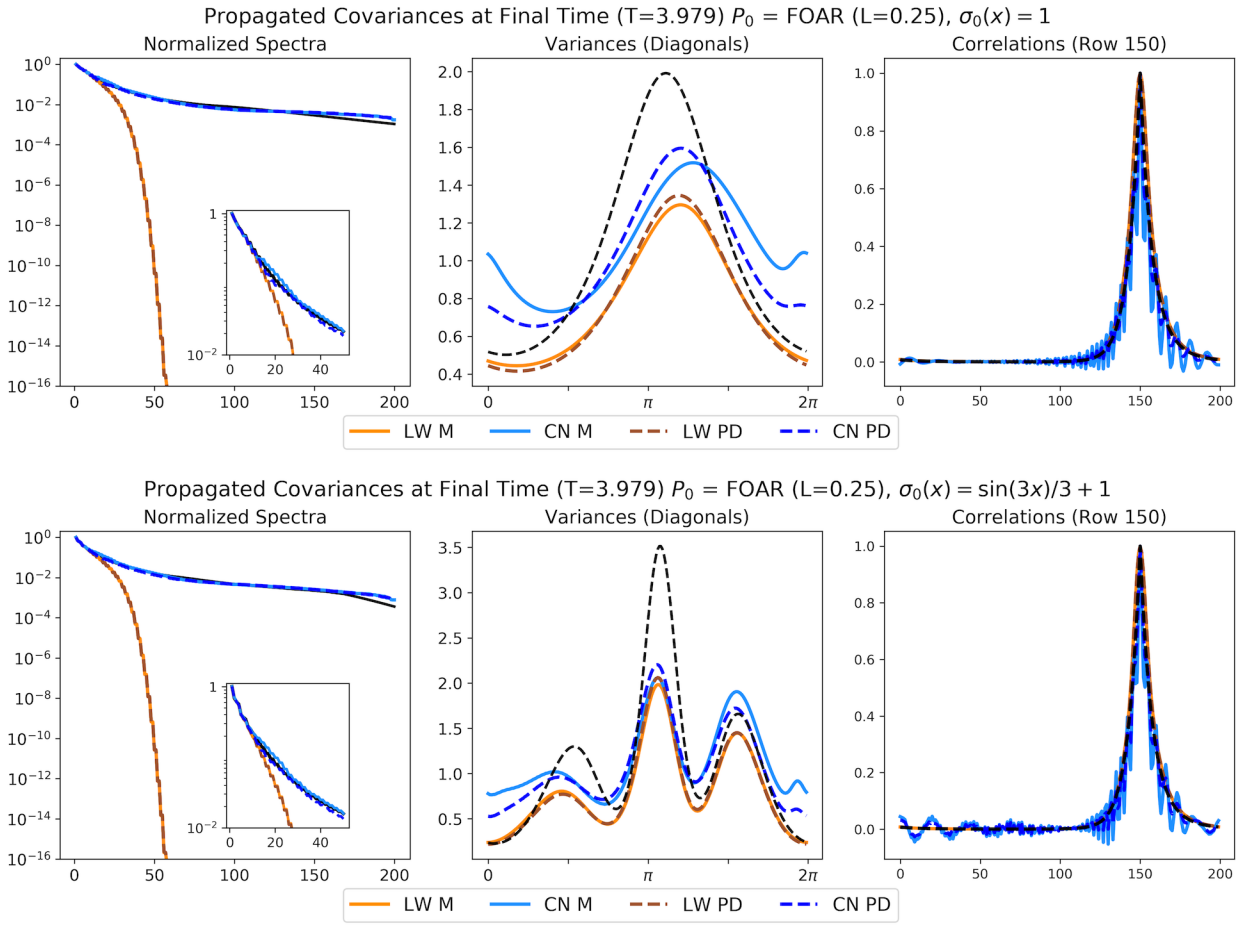

Figure 3.6: Same as Fig. 3.5 for the FOAR correlation function with $\mathsf{L} = 0.25$.

of Fig. 3.4. The normalized spectra in both of these cases are similar to the spectra seen in Fig. 3.3, where the Lax-Wendroff schemes are severely dissipative and both Crank-Nicolson methods approximate the exact spectrum moderately well.

The diagonals extracted from the numerically propagated covariances in Figs. 3.5 and 3.6 are strikingly different from the exact solution. Though covariances with these values of c and L are well-resolved initially, the extracted diagonals are smooth but wholly inaccurate. Across all four methods of propagation we see regions of both variance loss and variance gain, clearly illustrating that inaccurate discrete diagonal propagation is the problem more so than dissipation. The diagonal propagation becomes worse when the initial covariance has a spatially-varying variance (bottom rows of Figs. 3.5 and 3.6), which is a more realistic situation in practice. The errors we observe in the diagonals of the covariance matrices, interestingly, are not reflected in the normalized spectra; considering the normalized spectra alone would not even hint at the problems occurring in the discrete diagonal propagation. The correlations, as expected, show dispersion off the diagonal. The oscillatory behavior of both Crank-Nicolson schemes due to numerical dispersion is expected for this finite difference scheme (Beylkin and Keiser, 1997, p. 46).

In the limiting case when the initial correlation lengths become zero, we see two contrasting behaviors in the numerically propagated covariances depending on the initial variance; see Fig. 3.7. When the initial variance is constant (i.e. the initial covariance is the identity matrix), the Crank-Nicolson polar decomposition is the only scheme that correctly captures the behavior of the exact covariance (top row of Fig. 3.7). This is expected from the definition of the Crank-Nicolson polar decomposition in this case. Since the initial covariance is the identity matrix and the matrix $\boldsymbol{U}$ constructed using the Crank-Nicolson scheme is unitary, the Crank-Nicolson polar decomposition propagation, Eq. (3.3), reduces to $\boldsymbol{P}_k = \boldsymbol{D}_{k,0}\boldsymbol{U}^k I(\boldsymbol{U}^k)^{\mathrm{T}}\boldsymbol{D}_{k,0} = \boldsymbol{D}_{k,0}\boldsymbol{U}^k(\boldsymbol{U}^k)^{\mathrm{T}}\boldsymbol{D}_{k,0} = \boldsymbol{D}_{k,0}^2$, which is exact since the diagonal of $\boldsymbol{D}_{k,0}$ is evaluated analytically. When the initial variance varies spatially (bottom row of Fig. 3.2), the Crank-Nicolson polar decomposition propagation instead behaves more similarly to the Crank-Nicolson traditional propagation and both have regions of variance loss and variance gain. Carefully comparing the diagonals extracted from the traditional and

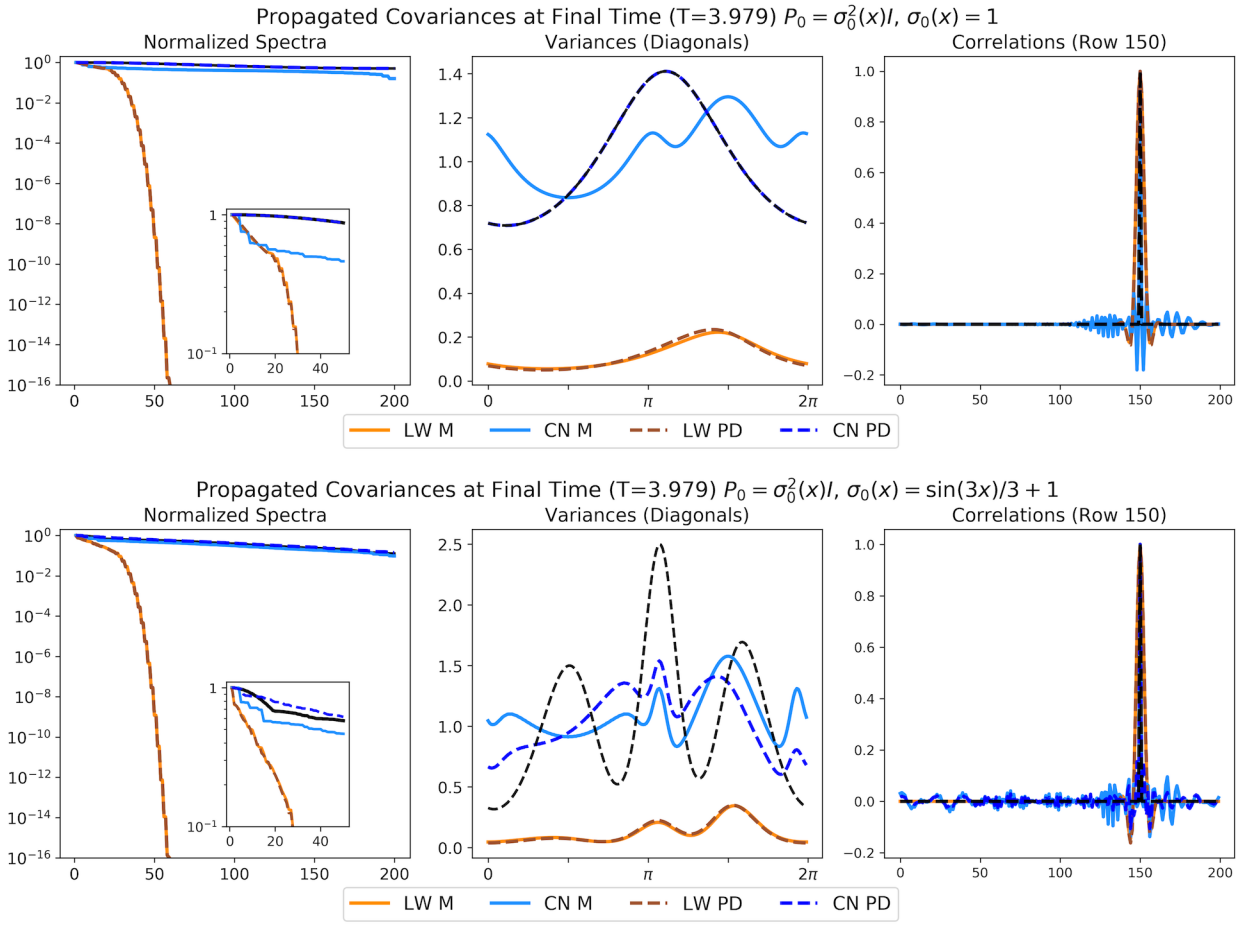

Figure 3.7: Propagated covariances at the final time T (slightly after a full period) for initial covariances with zero initial correlation lengths ($c = \mathsf{L} = 0$) for all four propagation methods. Top row: constant initial variance (identity matrix). Bottom row: spatially-varying initial variance. The left panels correspond to the normalized spectra (relative to largest eigenvalue), the middle panels show the discrete variances (covariance matrix diagonals), and the right panels show the correlations at row 150 (location of maximum in variance in space and time). The exact normalized spectra are given in solid black, the exact diagonals (solutions to the continuous spectrum equation) and correlations are given in black dashed. See Fig. 3.3 caption for description of the colored curves.

polar decomposition propagated covariances, the Crank-Nicolson polar decomposition propagation is a slight improvement over the Crank-Nicolson traditional propagation, but these differences are relatively minor compared to their absolute errors. The Lax-Wendroff schemes are substantially dissipative in all cases. We also observe that as the values of c and L decrease towards zero, the normalized spectra in Figs. 3.3, 3.5 – 3.7 decay more slowly and become relatively flat. This suggests that low-rank approximations would have difficulty capturing these covariances as correlation lengths shrink.

Comparing the diagonals of the covariance matrices at the final time across a series of initial correlation lengths in Figs. 3.8 and 3.9 demonstrates the severity of the variance loss and gain caused by inaccurate discrete diagonal propagation and provides a closer look at the approach to a limiting behavior seen in the trace time series. Without prior knowledge of the discontinuous change in dynamics as the initial correlation length tends to zero, one might surmise from Figs. 3.8 and 3.9 that the observed discrete diagonal behavior is caused by dissipation. Knowing the continuum behavior, however, makes it clear that we are observing inaccurate discrete diagonal propagation associated with the discontinuous change in continuum dynamics. The Crank-Nicolson polar decomposition propagation for a constant initial variance is the only scheme that captures the correct diagonal behavior when the initial covariance is the identity, and gradually approaches this behavior as the initial correlation length decreases. For all other cases in Figs. 3.8 and 3.9, the resulting covariance matrix diagonals are smooth, but grossly incorrect, and gradually approach their own limiting behavior at $c = \mathsf{L} = 0$ rather than changing abruptly as in the continuum case. The Lax-Wendroff schemes are severely dissipative, hardly resembling the correct dynamics, and are not shown in these and subsequent figures.

The errors in the discrete diagonal propagation are not limited to the final time; errors start to accumulate early on in the propagation cycle. We show this for the GC case in Fig. 3.10, where the FOAR results are similar but not shown. We can see clearly that rather than approximating the variance for $c > 0$ (black dashed), the Crank-Nicolson schemes (blue) tend to approximate the continuous spectrum ($c = 0$, brown), which is not the correct diagonal behavior for this case. Along

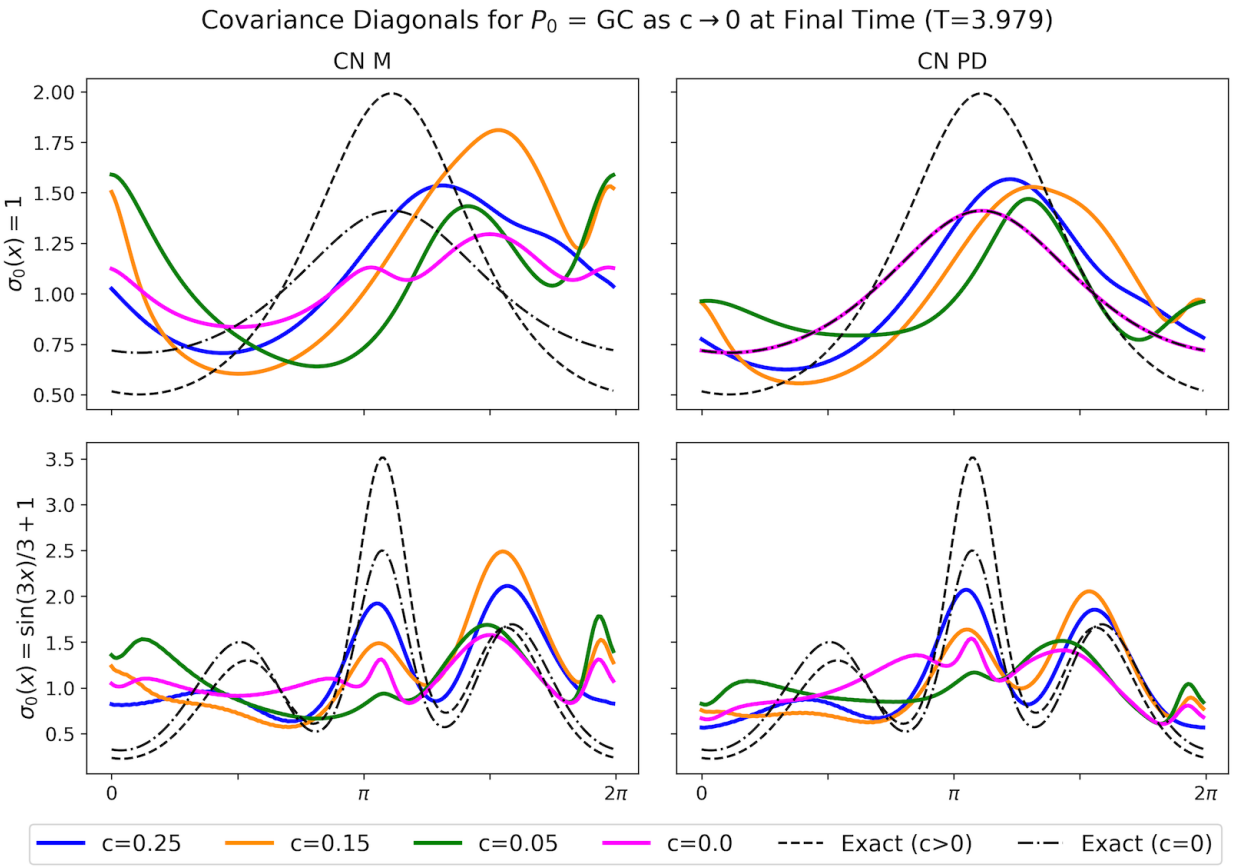

Figure 3.8: Covariance diagonals extracted from Crank-Nicolson traditional and polar decomposition propagation methods at the final time T for GC initial correlations as c approaches zero. Top row: constant initial variance. Bottom row: spatially-varying initial variance. Left column: propagation via traditional Crank-Nicolson (CN M). Right column: propagation via Crank-Nicolson polar decomposition (CN PD). Black curves are the exact diagonals, dashed for covariances with nonzero initial correlation lengths (variance solution) and dot-dashed for covariances with zero initial correlation length (continuous spectrum solution). These exact curves here labeled as Exact ($c > 0$) and Exact ($c = 0$) are the same exact curves labeled as Exact $\sigma^2(x,t)$ and Exact $P^c(x,t)$ in Fig. 3.2. In the top right panel, the exact curve for $c = 0$ (black dot-dashed) and the CN PD curve for $c = 0$ (magenta) identically overlap.

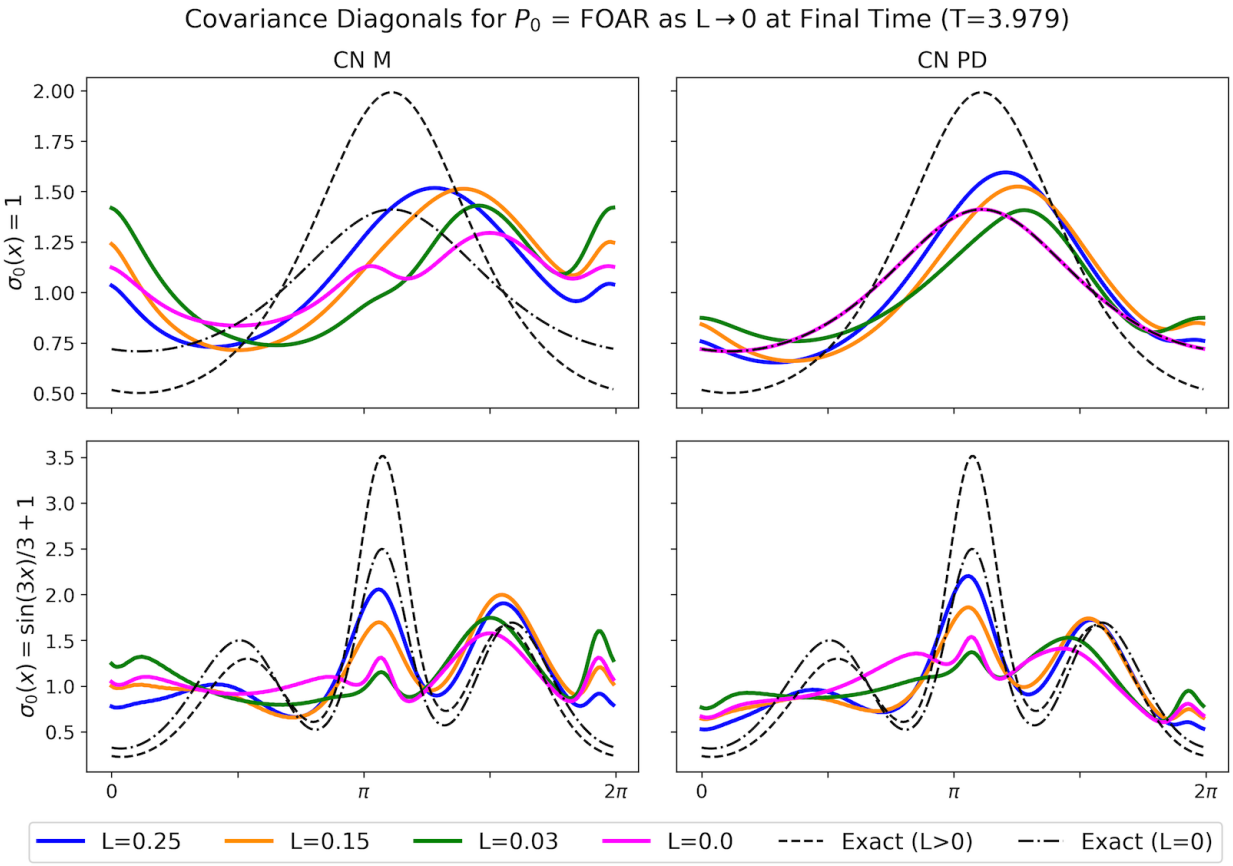

Figure 3.9: Same as Fig. 3.8 for FOAR correlation function initial covariances.

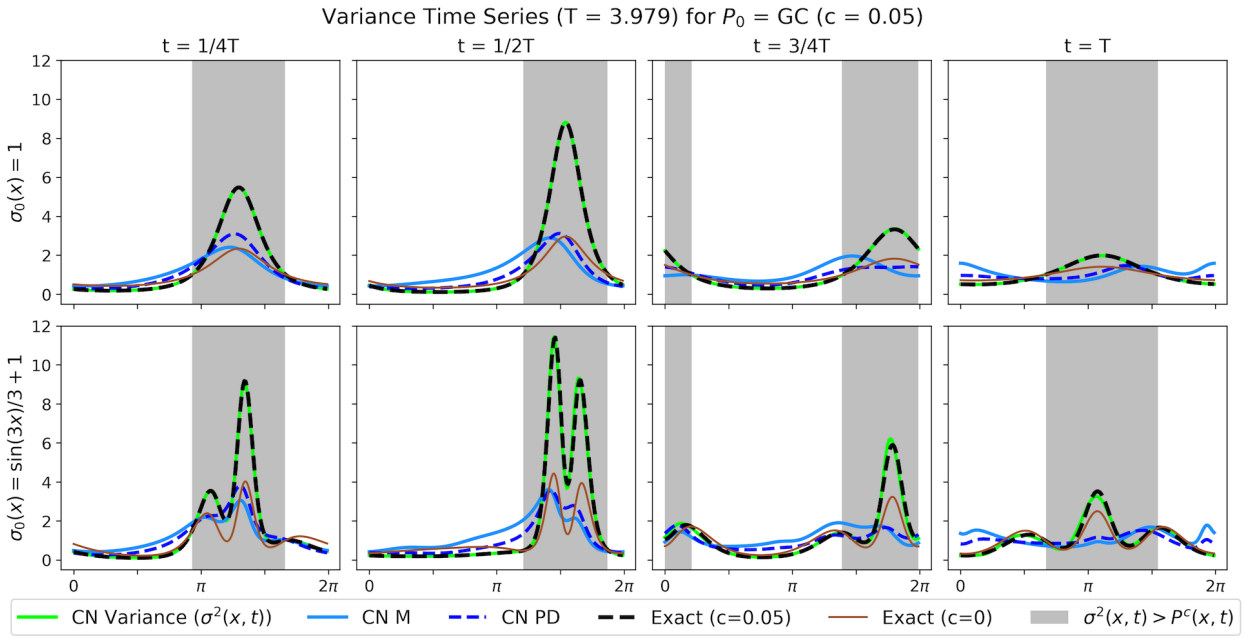

Figure 3.10: Variance time series for GC case with $c = 0.05$. Top row: constant initial variance. Bottom row: spatially-varying initial variance. Diagonals extracted from the Crank-Nicolson traditional and polar decomposition propagation are shown (solid and dashed blue) as well as propagating the variance by solving the variance equation (one-dimensional version of Eq. 1.8 for $b = v_x$) independently using Crank-Nicolson (green). The exact solutions to the variance and continuous spectrum equations are shown in black dashed and solid brown and denoted as Exact ($c = 0.05$) and Exact ($c = 0$), respectively. Regions where the exact variance is larger (smaller) than the exact continuous spectrum are highlighted in grey (white).

with the diagonals extracted from the propagated covariances in Fig. 3.10, we include the variance propagated independently by solving the one-dimensional version of Eq. (1.8) with $b = v_x$ using the Crank-Nicolson scheme, as was shown in Fig. 3.2. Including the variance solution computed using the Crank-Nicolson scheme in Fig. 3.10 emphasizes that propagating the covariance diagonal independently using the known diagonal dynamics significantly reduces the errors in the variance compared to the diagonal extracted from the propagated covariance matrix.

3.3 Interpretation of the Experimental Results

Grey regions in Fig. 3.10 correspond to where the exact variance solution is larger than the exact continuous spectrum solution, and these regions tend to correspond to where the numerically propagated diagonal for the short, nonzero initial correlation length (blue) exhibits variance loss. Further insight into this behavior can be gained by returning to the generalized continuum problem presented in Ch. 2. Assuming $P_0^d(\boldsymbol{x}, \boldsymbol{x}) = P_0^c(\boldsymbol{x})$, as is the case in Fig. 3.10, by multiplying the variance equation Eq. (1.8) by P^c and the continuous spectrum equation Eq. (1.9) by σ^2, taking the difference and dividing by $(P^c)^2$, we find from the quotient rule that the ratio

$$\frac{\sigma^2(\boldsymbol{x}, t)}{P^c(\boldsymbol{x}, t)} = m(\boldsymbol{x}, t) \tag{3.7}$$

satisfies the continuity equation with unit initial condition,

$$\begin{aligned} m_t + \boldsymbol{v} \cdot \boldsymbol{\nabla} m + (\boldsymbol{\nabla} \cdot \boldsymbol{v}) m = 0, \\ m(\boldsymbol{x}, t_0) = 1, \end{aligned} \tag{3.8}$$

for all $\boldsymbol{x} \in \Omega$ and $t \geq t_0$. From Eq. (3.8), the ratio σ^2/P^c must be conserved, and this holds for the generalized variance and continuous spectrum equations Eqs. (1.8) and (1.9), not just those presented in the numerical experiments. In regions where $m > 1$ we have $\sigma^2 > P^c$, and conversely in regions where $m < 1$ we have $\sigma^2 < P^c$. In fact, the function m of Eq. (3.8) on $\mathbb{S}_1^1$ with velocity field Eq. (3.1) can be expressed explicitly using Eq. (B.7), taking one as the initial condition. Solving $m = 1$, or equivalently $v(s(x,t)) - v(x) = 0$ (following the notation of Appen. B.1), determines the boundaries between regions where m is less than or greater than one in $\mathbb{S}_1^1$ at every time $t > t_0$.

For short initial correlation lengths, if the numerical schemes are better approximating the continuous spectrum solution P^c, regions where $m > 1$ should correspond to variance loss and regions where $m < 1$ should correspond to variance gain. Therefore, we refer to the ratio m as the amplitude index, which can be used to interpret the variance loss and gain observed in our numerical results. For example, regions where the index m is greater than one, i.e. $\sigma^2 > P^c$, highlighted in grey in Fig. 3.10, generally do coincide with regions where we see variance loss in the Crank-Nicolson traditional and polar decomposition schemes. Conversely, in the unshaded regions of Fig. 3.10 where the index m is less than one, i.e. $\sigma^2 < P^c$, the Crank-Nicolson traditional and polar decomposition schemes generally exhibit variance gain. Hence, we can exploit the amplitude index m to indicate regions of variance loss and gain.

Since the amplitude index m must be conserved according to Eq. (3.8), for a typical velocity field $\boldsymbol{v}$ there will always be regions of the spatial domain where m is greater than one and regions where m is less than one at any given time t. This implies that we should see both loss and gain of variance at individual times; whether there will be global variance loss or variance gain depends on the details of the velocity field $\boldsymbol{v}$. In all the numerical results, Figs. 3.3 – 3.10, we see both spurious variance loss and gain as a result of inaccurate discrete diagonal propagation. Data assimilation literature tends to focus on loss of variance because of its known negative impact on data assimilation schemes (Jazwinski, 1970, Sec. 8.8; Maybeck, 1982, Sec. 9.2; Anderson and Anderson, 1999; Houtekamer and Mitchell, 2000; Lyster et al., 2004). However, we see here for a general advective system that both loss and gain occur as reflected by the conservation law Eq. (3.8), cumulatively causing wholly inaccurate discrete diagonal propagation. Variance inflation, a tool often used in data assimilation practice to combat variance loss by rescaling the variance, typically by a multiplicative or additive factor, could perhaps be optimally adjusted in this context. Distinguishing between regions where the index m is greater than or less than one in general isolates the regions of variance loss and gain, respectively, thus the inflation factor could be tuned to only inflate where we expect variance loss. This would avoid using a single scale-factor that inflates the whole variance function and prevent variance inflation in regions where we already see variance

gain.

3.4 Conclusions and Discussion

Comparison of the numerical results with the continuum analysis in this chapter clearly illustrates the problems caused along the covariance diagonal during full-rank covariance propagation via Eq. (1.3). Standard methods of covariance propagation, like those considered in this chapter, produce covariance matrix diagonals that can be severely inaccurate. The result is not only variance loss, but also variance gain due to the inaccurate propagation along the diagonal.

It is important to recognize that the inaccurate diagonal propagation observed in these numerical experiments is a result of the discontinuous change in continuum covariance dynamics discussed in Ch. 2. Even when propagating covariance matrices using a fully Lagrangian scheme, as done in Lyster et al. (2004), propagated covariances still suffer from spurious loss of variance. This loss of variance is not due to the numerical scheme, but rather is due to the discontinuous change in dynamics identified in Ch. 2. Data assimilation schemes that do not propagate the covariance explicitly may experience errors similar to what we observe here because the underlying cause of these errors is the continuum covariance dynamics, not the numerical scheme. Studying full-rank covariance propagation as in Eq. (1.3) isolates the spurious loss of variance as an issue with covariance dynamics and implies that approximations of Eq. (1.3), such as in ensemble Kalman filters, can suffer variance loss in a similar manner. The errors caused during the covariance propagation may be a neglected source of the model error or "system error" observed in data assimilation schemes (Houtekamer and Mitchell, 2005, p. 3285).

The continuum analysis of Ch. 2 and numerical experiments presented in this chapter bring to light a fundamental issue associated with current approaches to discrete covariance propagation. The source of the problem is not numerical dissipation, but rather is rooted to more fundamental aspects of the propagation itself. Whether this is a problem that can be addressed by choosing an appropriate numerical scheme will be addressed in the next chapter.

Chapter 4

Numerical Demonstration Part II: The Limitations of Full-Rank Covariance Propagation

The numerical results in Ch. 3 demonstrate that standard methods of variance propagation, namely that corresponding to Eq. (1.3), are inaccurate, producing covariance matrix diagonals that hardly resemble the exact variance. One may ask, then, whether the inaccurate diagonal propagation observed in Ch. 3 can be mitigated either by choosing an appropriate numerical scheme or increasing the spatial resolution. In this chapter, we will address this first question by considering a slightly different example than that considered in Ch. 3. This example will demonstrate that the inaccurate diagonal propagation observed in Ch. 3 is not easily mitigated by the choice of the numerical scheme and is fundamentally a problem caused by the continuum dynamics.

4.1 Continuum Problem

In this example, we will consider the case where $b = \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{v}$ in Eq. (1.4), where for $\boldsymbol{x} \in \mathbb{S}^2_r$,

$$q_t + \boldsymbol{v} \cdot \boldsymbol{\nabla} q + \frac{1}{2}(\boldsymbol{\nabla} \cdot \boldsymbol{v})q = 0,$$
$$q(\boldsymbol{x}, t_0) = q_0(\boldsymbol{x}). \tag{4.1}$$

This case can arise when the state of a system is a tracer density, for example. If the tracer density satisfies the continuity equation ($b = \boldsymbol{\nabla} \cdot \boldsymbol{v}$ in Eq. 1.4), then its square root satisfies Eq. (4.1).

We can express Eq. (4.1) in conservative form,

$$q_t + \boldsymbol{\mathcal{S}} q = 0,$$
$$q(\boldsymbol{x}, t_0) = q_0(\boldsymbol{x}). \tag{4.2}$$

The operator $\boldsymbol{\mathcal{S}}$ is skew-symmetric on $L^2(\mathbb{S}_r^2)$,

$$\boldsymbol{\mathcal{S}} q = \frac{1}{2}\,\boldsymbol{\nabla\cdot}\,(\boldsymbol{v}q) + \frac{1}{2}\boldsymbol{v}\cdot\boldsymbol{\nabla} q, \tag{4.3}$$

meaning that for $\boldsymbol{v} \in C^1(\mathbb{S}_r^2)$ and $q \in H^1(\mathbb{S}_r^2)$, where $H^1(\mathbb{S}_r^2)$ is the Sobolev space where the weak first derivatives of q are in $L^2(\mathbb{S}_r^2)$, we have $(q, \boldsymbol{\mathcal{S}} q)_2 = 0$, where $(\cdot,\cdot)_2$ is the inner product on $L^2(\mathbb{S}_r^2)$. Using Eq. (4.2) and the skew-symmetry of $\boldsymbol{\mathcal{S}}$, it follows that q has a quadratically conserved energy functional,

$$\frac{d}{dt}\frac{1}{2}||q||_2^2 = \frac{d}{dt}\frac{1}{2}\int_{\mathbb{S}_r^2} q^2(\boldsymbol{x}, t)\, d\boldsymbol{x} = 0. \tag{4.4}$$

We can also define the associated continuum covariance equation using Eq. (1.6) and show (in a similar manner as done for the state q) that the covariance P satisfies its own conservation property,

$$\frac{d}{dt}\int_{\mathbb{S}_r^2}\int_{\mathbb{S}_r^2} P^2(\boldsymbol{x}_1, \boldsymbol{x}_2, t)\, d\boldsymbol{x}_1\, d\boldsymbol{x}_2 = 0. \tag{4.5}$$

The variance in this case satisfies the continuity equation, which implies that its L^1 norm on $\mathbb{S}_r^2$ is conserved,

$$\frac{d}{dt}||\sigma^2||_1 = \frac{d}{dt}\int_{\mathbb{S}_r^2} \sigma^2(\boldsymbol{x}, t)\, d\boldsymbol{x} = 0. \tag{4.6}$$

For initial states q_0 that are spatially correlated, the total variance, which is defined by Eq. (4.6), is conserved. Note that the trace of the operator $\boldsymbol{\mathcal{P}}_t^d$ (i.e., the total variance) is defined, therefore the covariance operator $\boldsymbol{\mathcal{P}}_t^d$ is trace class (Reed and Simon, 1972, pp. 206–208). For spatially uncorrelated initial states q_0, the covariance P remains diagonal for all time (see Ch. 2.3) and the dynamics along the diagonal satisfy the continuous spectrum equation, Eq. (1.9). When $b = \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{v}$, the continuous spectrum solution P^c satisfies the advection equation and is therefore constant along the characteristics. In contrast with the discrete spectrum covariance operator $\boldsymbol{\mathcal{P}}_t^d$, the continuous spectrum covariance operator $\boldsymbol{\mathcal{P}}_t^c$ in Eq. (2.30) is not trace class, and the L^1-norm

of the continuous spectrum solution P^c on $\mathbb{S}^2_r$ is not constant in time unless the initial condition P^c_0 is constant.

4.2 Numerical Demonstration and Discussion

When propagating the covariance discretely in Eq. (1.3), it is often best to choose a numerical scheme that will preserve the continuum conservation properties, described in this case by Eqs. (4.4) – (4.6). The Crank-Nicolson scheme implemented in Ch. 3, in fact, is well suited to preserve the continuum properties of the state, variance, and covariance. To demonstrate this, we will consider the one-dimensional version of Eq. (4.2) on the unit circle, $\mathbb{S}^1_1$, with the same experiment setup as described in Ch. 3: the velocity $v = v(x)$ is defined by Eq. (3.1) and the unit circle is discretized uniformly into $N = 200$ grid points.

By applying the Crank-Nicolson scheme to the one-dimensional version of Eq. (4.2), the (explicit) propagation matrix $\mathbf{M}$ (which does not depend on time since the velocity v is independent of time) is

$$\mathbf{M} = \left(\mathbf{I} + \frac{\Delta t}{2}\mathbf{S}\right)^{-1}\left(\mathbf{I} - \frac{\Delta t}{2}\mathbf{S}\right). \tag{4.7}$$

The matrix $\mathbf{I}$ is the $N \times N$ identity matrix and Δt is the time step associated with the time discretization. The matrix $\mathbf{S}$ is a discretization of the continuum operator $\boldsymbol{\mathcal{S}}$ according to the Crank-Nicolson scheme. This discretization ensures that $\mathbf{S}$ is skew-symmetric, which is consistent with the operator $\boldsymbol{\mathcal{S}}$. Since the matrix $\mathbf{S}$ is skew-symmetric, the propagation matrix $\mathbf{M}$ is unitary. Thus, the Crank-Nicolson scheme applied to the state dynamics in Eq. (4.2) is quadratically conservative, preserving the discrete counterpart of the continuum energy conservation defined in Eq. (4.4).

The unitary property of the propagation matrix $\mathbf{M}$ has important implications for the discrete covariance propagation, Eq. (1.3), as well. The Schatten p-norms, which are defined for any $N \times N$ matrix $\mathbf{A}$ as

$$||\mathbf{A}|| = \left(\sum_{i=1}^{N} |\lambda_i|^p\right)^{1/p}, \tag{4.8}$$

where λ_i are the singular values of the matrix $\mathbf{A}$, are unitarily invariant (Horn and Johnson, 1985,

pp. 440–441). Therefore, all Schatten p-norms of the covariance matrix $\mathbf{P}$ for $p < \infty$ are conserved when the propagation matrix $\mathbf{M}$ is defined by the Crank-Nicolson scheme, Eq. (4.7). For $p = 2$, Eq. (4.8) is the Frobenius norm, which is conserved and is the discrete counterpart of Eq. (4.5). When $p = 1$, Eq. (4.8) is equivalent to the trace of the covariance matrix $\mathbf{P}$, which is also conserved for all time. This is consistent with the variance conservation defined by Eq. (4.6). Interestingly, when the initial covariance matrix $\mathbf{P}_0$ is diagonal, the trace of the covariance matrix will still be conserved during the discrete propagation. According to the continuum dynamics, the trace of a diagonal covariance should not be conserved unless the initial covariance is constant along the diagonal. Thus, inaccurate propagation along the diagonal should be expected in this case.

Without prior knowledge of the discontinuous change in the continuum covariance dynamics, the Crank-Nicolson scheme seems well suited for discrete covariance propagation. One might expect little or no variance loss during propagation since the Crank-Nicolson scheme is not dissipative and will conserve total variance. We see, however, in Fig. 4.1, that this is not the case. Figure 4.1 is similar to Fig. 3.8 in Ch. 3.2, where covariance matrices with different initial cut-off lengths (proportional to their correlation length) are propagated up to a fixed time T (slightly after a full time period) using the Crank-Nicolson scheme and the matrix diagonals are extracted and plotted along the unit circle.

As seen in Ch. 3.2, the covariance diagonal extracted from the Crank-Nicolson propagation in Fig. 4.1 are smooth, but wholly inaccurate. First, consider the case where $c = 1$ (orange), in which the initial covariances are supported on the full domain. Correlation lengths for $c = 1$ are quite long, therefore it is expected that full-rank propagation will capture the variance well. We see, however, that the extracted covariance diagonal for $c = 1$ exhibits significant loss and gain of variance when compared to the exact variance (black), which is particularly pronounced when the initial variance is spatially-varying (panel b).

For all cases in Fig. 4.1, the trace of the numerically propagated covariance is conserved, though the diagonals themselves hardly resemble the exact variance solution (black). For initial covariances with $c > 0$, the total amount of variance loss is exactly equal to the amount of variance

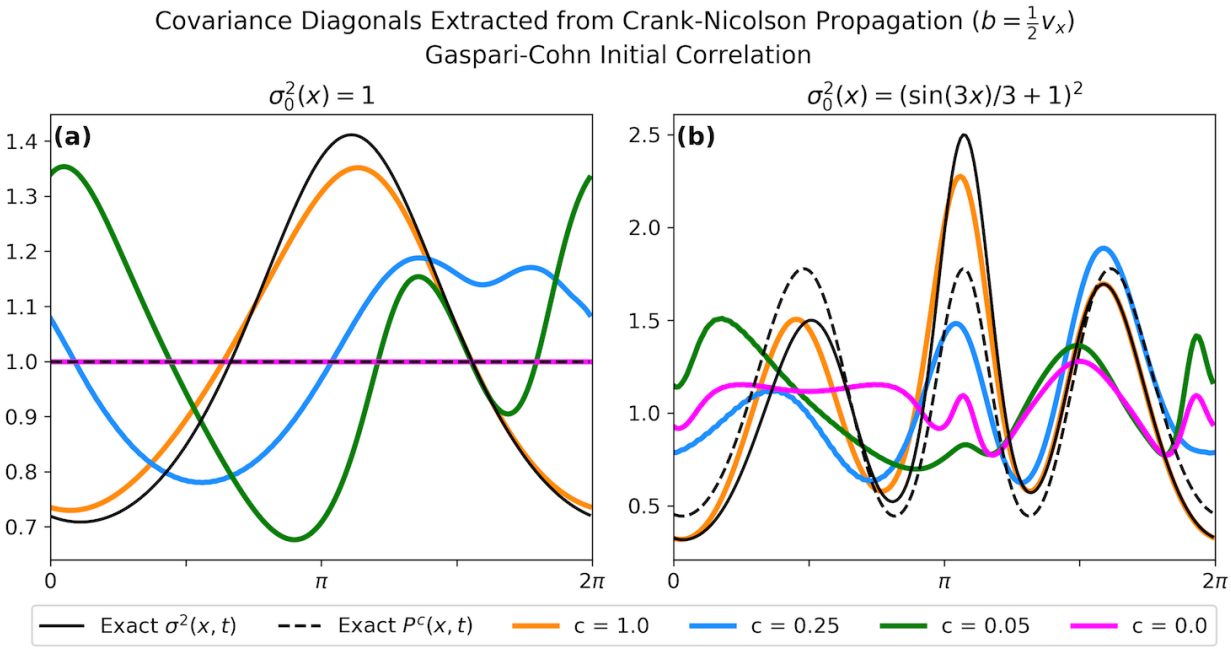

Figure 4.1: Covariance diagonals extracted at time $T = 3.98$ (slightly after a full time period) from the Crank-Nicolson covariance propagation defined for states with $b = \frac{1}{2}\, v_x$ (one-dimensional version of Eq. 4.1). In both panels, the colored curves denote matrix diagonals extracted from propagated initial covariances constructed with the Gaspari-Cohn (GC) correlation function with different initial cut-off lengths c (denoted in the legend). Panel (a) corresponds to initial covariances with unit initial variance, while panel (b) corresponds to a spatially-varying initial variance denoted in the panel title. The exact solution to the variance equation (solid black) and continuous spectrum equation (dashed black) are shown in both panels for reference, where in panel (a) both the $c = 0$ case (magenta) and continuous spectrum solution identically overlap.

gain due to the conservation imposed by the Crank-Nicolson scheme. Simply taking the trace of the covariance matrix as a diagnostic would not show the extent of inaccurate diagonal propagation in this case, as the Crank-Nicolson scheme will ensure that the trace is preserved regardless of how the scheme is actually approximating the diagonal.

The only case where the Crank-Nicolson scheme captures the correct diagonal dynamics is when the initial covariance is the identity matrix, shown in the panel (a) (magenta curve). Since the continuous spectrum solution satisfies the advection equation, a constant initial condition will remain constant for all time. The Crank-Nicolson scheme captures this behavior exactly when the initial covariance is the identity (or a scalar multiple of the identity matrix) since the propagation matrix $\mathbf{M}$ is unitary, as the propagation reduces to $\mathbf{MIM}^{\mathrm{T}} = \mathbf{I}$ for all time. When there is any spatial variability along the diagonal, as shown for $c = 0$ in panel (b), the discrete propagation does a very poor job of capturing the correct dynamics.

The main takeaway from Fig. 4.1 is that even though the numerical scheme, here the Crank-Nicolson scheme, preserves important conservation properties and is not numerically dissipative, we still observe considerably inaccurate diagonal propagation. The cause of this inaccurate diagonal propagation, as we will see explicitly in the next chapter and Appen. C.3, is the discontinuous change in the continuum dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane as correlation lengths tend to zero.

Chapter 5

Error Analysis of Full-Rank Covariance Propagation

Up to this point, we have observed that the discontinuous change in the continuum covariance dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane identified in Ch. 2 results in inaccurate variance propagation in discrete space. Chapters 3 and 4 demonstrate the severity of the inaccurate variance propagation caused by discrete covariance propagation, Eq. (1.3). These numerical results motivate an investigation into the continuum dynamics being approximated along the covariance diagonal during discrete covariance propagation. As mentioned in Ch. 1, discrete covariance propagation defined by Eq. (1.3) indicates that its structure will be problematic because of the discontinuous change in covariance dynamics along $\boldsymbol{x}_1 = \boldsymbol{x}_2$. The goal of this chapter is to investigate Eq. (1.3) further by determining the continuum dynamics being approximated along the diagonal. Through careful error analysis of Eq. (1.3), we will uncover the explicit cause of the observed inaccurate variance propagation.

5.1 Introduction

Before we derive the continuum dynamics being approximated along the covariance diagonal during discrete covariance propagation (i.e., Eq. 1.3), let's return to its standard formulation to gain an intuitive sense for how covariances are evolved in discrete space. To illustrate this in simplest terms, consider a discrete stochastic model state $N-$vector $\mathbf{q}$ that is propagated forward in time (forecasted) from time t_k to t_{k+1} (equivalent to Eq. 1.1 with slightly different notation),

$$\mathbf{q}^{k+1} = \mathbf{M}^{k+1,k}\mathbf{q}^k, \tag{5.1}$$

where $\mathbf{M}^{k+1,k}$ is the deterministic $N \times N$ propagation matrix representing the discretized model dynamics. For simplicity, we will assume the model dynamics are linear with no forcing, random or otherwise. Unlike Eq. (1.1) in Ch. 1, the time indices in Eq. (5.1) are superscripts, rather than subscripts, since subscripts will be used later in this chapter to denote the spatial discretization.

The corresponding $N \times N$ covariance matrix at time t_k, $\mathbf{P}^k$, is defined in terms of $\mathbf{q}^k$ using the expectation operator $\mathbb{E}[\cdot]$,

$$\mathbf{P}^k = \mathbb{E}\big[(\mathbf{q}^k - \overline{\mathbf{q}}^k)(\mathbf{q}^k - \overline{\mathbf{q}}^k)^{\mathrm{T}}\big], \tag{5.2}$$

where $\overline{\mathbf{q}}^k = \mathbb{E}[\mathbf{q}^k]$ is the mean state and superscript T denotes the transpose. The covariance matrix at time t_{k+1} can be derived from Eqs. (5.1) – (5.2),

$$\mathbf{P}^{k+1} = \mathbf{M}^{k+1,k}(\mathbf{M}^{k+1,k}\mathbf{P}^k)^{\mathrm{T}}. \tag{5.3}$$

Equation (5.3) is the explicit, full-rank propagation of the covariance matrix derived from the state dynamics (e.g. Kalman, 1960); this propagation is approximated at low rank using ensemble statistics in ensemble-based schemes (e.g. Leith, 1974; Evensen, 1994).

We can gain an intuitive sense of how inaccurate variance propagation can occur during covariance propagation by making a few observations from Eq. (5.3). Covariance propagation in Eq. (5.3) can be interpreted as a sequence of two operations on the covariance matrix $\mathbf{P}^k$. The first operation inside the parentheses, $\mathbf{M}^{k+1,k}\mathbf{P}^k$, is the propagation along the columns of the covariance matrix. This is followed by the propagation along the rows after the transpose is computed.[1] Since the sequence of operations in Eq. (5.3) occurs along the columns and rows of $\mathbf{P}^k$, diagonal elements of the covariance matrix, i.e. the variances, are necessarily approximated by off-diagonal elements. Recall that off-diagonal elements $\mathbf{P}^k_{ij}$, $i \neq j$, are smaller than the corresponding diagonal elements, $(\mathbf{P}^k_{ij})^2 \leq \mathbf{P}^k_{ii}\mathbf{P}^k_{jj}$, because the covariance matrix is positive semi-definite. Thus, when correlation lengths are small or, more generally, when there are sharp gradients near the diagonal, for instance, one would expect Eq. (5.3) to approximate the covariance diagonal poorly. In this way, the spurious

[1] Sometimes this interpretation of covariance propagation is termed "operator splitting" in the data assimilation literature (e.g., Ménard et al., 2021, Sec. 2.1.1). We do not to use this terminology because Eq. (5.3) is derived from first principles and without operator splitting as defined in the traditional sense (e.g. Iserles, 2009, pp. 378-380).

loss of variance observed during Eq. (5.3) is not necessarily a problem of numerical dissipation, but can also be caused by the structure of discrete covariance propagation itself.

Approximating diagonal elements with off-diagonal elements during covariance propagation, as we will see in this chapter, is particularly problematic for states governed advective dynamics, Eq. (1.4). Recall from Ch. 2 the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$ is a characteristic surface for the corresponding covariance evolution equation, Eq. (1.6). As a consequence, in the limit as correlation lengths tend to zero, there is a discontinuous change in the continuum covariance dynamics along $\boldsymbol{x}_1 = \boldsymbol{x}_2$. Approximating diagonal elements of the covariance with off-diagonal elements in Eq. (5.3) is therefore problematic, as the approximation is made across the discontinuity. For standard discretizations used to evaluate Eq. (5.3), we will find that Eq. (5.3) fails to approximate the variance σ^2, or even the continuous spectrum P^c, when correlation lengths are small enough. The result, therefore, is inaccurate variance propagation, as Eq. (5.3) is approximating the incorrect dynamics along the covariance diagonal.

This chapter demonstrates that traditional methods of variance propagation, namely that associated with Eq. (5.3), are inherently inaccurate for advective dynamics because of the discontinuous change in the continuum covariance dynamics along the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$. The loss and gain of variance observed during covariance propagation is therefore a manifestation of the inaccurate variance propagation caused by Eq. (5.3) and the continuum covariance dynamics. To show this, we will perform an error analysis of the diagonal propagation defined by Eq. (5.3) in one space dimension for two standard numerical schemes, one that is dissipative and one that is not. The error analysis reveals that the continuum dynamics being approximated along the covariance diagonal change as the ratio of the grid length to the correlation length becomes significant. When the ratio of the grid length to the correlation length approaches unity, the approximated continuum dynamics hardly resemble the continuum variance equation, Eq. (1.8), as several error terms that are powers of this ratio become significant. Both discretizations we analyze here are identical up to a highly dissipative-like term, indicating that the inaccurate variance propagation observed during discrete covariance propagation is driven by the discontinuous change in the continuum covariance

dynamics rather than the particular numerical scheme.

The goal of this chapter is to uncover the fundamental cause inaccurate variance propagation observed during full-rank covariance propagation in advective systems. This chapter does not present any explicit means to mitigate the inaccurate variance propagation, though we discuss the potential of correcting Eq. (5.3) in Ch. 6. The intention with this chapter is to carefully present the problem caused by standard methods of covariance propagation, identify the source of inaccurate variance propagation in advective systems, and thereby open the door to potential solutions. The results of this analysis have a direct impact in geophysical systems in which advective dynamics defined by Eq. (1.4) are an important component, such as general circulation models and chemical tracer transport models, for example.

This chapter is organized as follows. Section 5.2 establishes the preliminaries for the error analysis presented in Sec. 5.3. We will consider two, standard spatial discretizations of the state equation, and from these discretizations we will derive the associated discretizations for the covariance and its diagonal in Secs. 5.2.1 and 5.2.2. Section 5.3 derives the continuum dynamics the two numerical scheme are approximating along the covariance diagonal. Two cases are considered. Section 5.3.1 presents the case for covariances that are continuously differentiable across its diagonal, followed by the case when the covariance is not continuously differentiable across its diagonal in Sec. 5.3.2. The results for both cases are similar, but require slightly different approaches in their analysis due to their different differentiability assumptions. Section 5.4 discusses recent work by Pannekoucke et al. (2021), which is similar to, but distinct from what is presented here. Conclusions are given in Sec. 5.5. To supplement this chapter are Appen. C.1, which provides additional derivations needed for the error analysis, Appen. C.2, which presents the error analysis for the state equation in advection form to contrast with the main text which considers the state equation in flux form. Appendix C.3 performs the error analysis for the state dynamics in Ch. 4, following the analysis presented in this chapter.

5.2 Derivation of Covariance Diagonal Discretizations

Covariance propagation as defined in Eq. (5.3) represents a discretization in both time (index k) and space (represented by vectors and matrices in boldface). In this chapter, we will find that it is the spatial discretization that causes inaccurate variance propagation, independent of the time discretization. Therefore, we will study the semi-discretization of the diagonal of the covariance, where we only discretize in space and leave time as a continuous variable. This yields a system of ordinary differential equations to which several different time-integration schemes can be applied (and thus producing different numerical PDE schemes). Thus, the results presented here are quite general.

For this analysis, take the spatial domain to be the unit circle, $\mathbb{S}^1_1$, and consider two standard spatial discretizations for the state that will define the subsequent semi-discretizations for the covariance. The first is a first-order upwind spatial discretization, which is commonly applied to advective dynamics (Iserles, 2009, p. 398). The second scheme is a second-order centered difference scheme, which was the spatial discretization used for the experiments in Chs. 3 and 4. For this analysis, we will consider the one-dimensional version of Eq. (1.4) on $\mathbb{S}^1_1$ in flux form,

$$\begin{aligned} q_t + (vq)_x + (b - v_x)q &= 0, \\ q(x, t_0) &= q_0(x). \end{aligned} \tag{5.4}$$

Discretizing the state equation in flux form will help to preserve any conservative properties that may arise depending on the choice of b (e.g. mass conservation if $b = v_x$ in the context of the continuity equation, see Hesthaven, 2018, Ch. 4) and is often the approach taken when developing dynamical cores for weather and climate models (e.g. Lin and Rood, 1996). Appendix C.2 discusses the semi-discretization in advection form, by way of contrast.

The unit circle $\mathbb{S}^1_1$ is discretized into a uniform grid, $x_i = i\Delta x$ for $i = 1, 2, \ldots, N$, $\Delta x = 2\pi/N$, and assume periodicity. Quantities evaluated on the spatial grid are denoted by the subscript i, e.g. $q(x_i, t) = q_i(t)$, where the range $i = 1, 2, \ldots, N$ is included where needed for clarity. Sections 5.2.1 and 5.2.2 derive the semi-discretizations for the covariances using the upwind and centered difference

schemes, respectively, before proceeding to the error analysis in Sec. 5.3.

5.2.1 Upwind Spatial Discretization

The upwind spatial discretization is applied to the spatial derivative $(vq)_x$ in Eq. (5.4) and results in the following semi-discretization of the state,

$$\frac{d}{dt}q_i(t) = \frac{1}{\Delta x}\big[v_{i-1}(t)q_{i-1}(t) - v_i(t)q_i(t)\big] - \big[b_i(t) - (v_x)_i(t)\big]q_i(t), \quad i = 1, 2, \ldots, N. \tag{5.5}$$

It is assumed that $v \geq 0$ throughout the domain, so that Eq. (5.5) does represent upwind differencing. The derivative of the velocity $(v_x)_i(t)$ in the zeroth-order term of Eq. (5.5) is not discretized; leaving this term exact will become important for the error analysis in Sec. 5.3.

To derive the covariance semi-discretization associated with Eq. (5.5), we first define the error

$$\varepsilon(x,t) = q(x,t) - \overline{q}(x,t), \tag{5.6}$$

where $\overline{q} = \mathbb{E}[q]$ denotes the mean state. Using the definition of the error in Eq. (5.6) and the state semi-discretization in Eq. (5.5), we can derive the semi-discretization for the error,

$$\frac{d}{dt}\varepsilon_i(t) = \frac{1}{\Delta x}\big[v_{i-1}(t)\varepsilon_{i-1}(t) - v_i(t)\varepsilon_i(t)\big] - \big[b_i(t) - (v_x)_i(t)\big]\varepsilon_i(t), \quad i = 1, 2, \ldots, N. \tag{5.7}$$

This semi-discretization for the error is identical to the state semi-discretization in Eq. (5.5) since the discretization in Eq. (5.5) is linear and both b and v are deterministic. The covariance semi-discretization follows from the definition of the covariance using the error $\varepsilon(x,t)$,

$$P(x_1, x_2, t) = \mathbb{E}[\varepsilon(x_1,t)\varepsilon(x_2,t)], \tag{5.8}$$

which on the discretized spatial grid can be expressed as

$$P(x_i, x_j, t) = P_{i,j}(t) = \mathbb{E}[\varepsilon_i(t)\varepsilon_j(t)], \quad i, j = 1, 2, \ldots, N. \tag{5.9}$$

From Eq. (5.7), this yields the semi-discretization for the covariance,

$$\begin{aligned}\frac{d}{dt}P_{i,j}(t) = \frac{1}{\Delta x}\Big\{v_{i-1}(t)P_{i-1,j}(t) + v_{j-1}(t)P_{i,j-1}(t) - \big[v_i(t) + v_j(t)\big]P_{i,j}(t)\Big\} \\ - \left[b_i(t) + b_j(t) - (v_x)_i(t) - (v_x)_j(t)\right]P_{i,j}(t), \quad i, j = 1, 2, \ldots, N.\end{aligned} \tag{5.10}$$

We can see that the covariance semi-discretization reflects the action of the spatial discretization in Eq. (5.5) in the i and j directions, coinciding with actions along the columns and rows described by Eq. (5.3).

Since we are interested in the propagation along the diagonal of the covariance, we take $i = j$ in Eq. (5.10) to extract the semi-discretization for the diagonal of the covariance,

$$\frac{d}{dt}P_{i,i}(t) = \frac{1}{\Delta x}\Big\{v_{i-1}(t)\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big] - 2v_i(t)P_{i,i}(t)\Big\} - 2\big[b_i(t) - (v_x)_i(t)\big]P_{i,i}(t), \quad i = 1, 2, \ldots, N. \tag{5.11}$$

There are two important aspects of Eq. (5.11) that arise from this derivation. First are the factors of two that appear in front of $v_i(t)P_{i,i}(t)$ and the zeroth-order terms; the importance of these factors of two will be discussed in Sec. 5.3. Second is the appearance of the term $P_{i-1,i}(t) + P_{i,i-1}(t)$, which represents an average across the diagonal. This averaging term is a second-order accurate (centered) approximation of the diagonal element on the half-grid, $2P(x_{i-1/2}, x_{i-1/2}, t)$, which we will discuss further in Sec. 5.3.

5.2.2 Centered Difference Spatial Discretization

Centered difference spatial discretizations are common in numerical solvers for differential equations. Here, we apply a second-order centered difference scheme to the $(vq)_x$ term in Eq. (5.4). Though this spatial discretization scheme is different from the upwind scheme, we will observe similar behaviors in the resulting semi-discretization of the covariance diagonal.

As with the upwind spatial discretization, we begin by applying the centered difference spatial discretization scheme to the term $(vq)_x$ in the state equation to produce its semi-discretization,

$$\frac{d}{dt}q_i(t) = \frac{1}{2\Delta x}\big[v_{i-1}(t)q_{i-1}(t) - v_{i+1}(t)q_{i+1}(t)\big] - \big[b_i(t) - (v_x)_i(t)\big]q_i(t), \quad i = 1, 2, \ldots, N, \tag{5.12}$$

where we keep the spatial derivatives of the velocity $(v_x)_i(t)$ exact as done for the upwind case. To derive the covariance semi-discretization associated with Eq. (5.12), we follow the same procedure presented in Sec. 5.2.1. The error $\varepsilon_i(t)$ satisfies the same semi-discretization as the state, Eq. (5.12),

due to linearity, and using the procedure applied in Sec. 5.2.1, we have the centered difference semi-discretization for the covariance,

$$\frac{d}{dt}P_{i,j}(t) = \frac{1}{2\Delta x}\big[v_{j-1}(t)P_{i,j-1}(t) - v_{j+1}(t)P_{i,j+1}(t) + v_{i-1}(t)P_{i-1,j}(t) - v_{i+1}(t)P_{i+1,j}(t)\big] \\ - \big[b_i(t) + b_j(t) - (v_x)_i(t) - (v_x)_j(t)\big]P_{i,j}(t), \quad i,j = 1,2,\ldots,N. \tag{5.13}$$

Taking $i = j$ yields the centered difference semi-discretization for the diagonal of the covariance,

$$\frac{d}{dt}P_{i,i}(t) = \frac{1}{2\Delta x}\Big\{v_{i-1}(t)\big[P_{i,i-1}(t) + P_{i-1,i}(t)\big] - v_{i+1}(t)\big[P_{i,i+1}(t) + P_{i+1,i}(t)\big]\Big\} \\ - 2\big[b_i(t) - (v_x)_i(t)\big]P_{i,i}(t), \quad i = 1,2,\ldots,N \tag{5.14}$$

Like the upwind semi-discretization for the covariance diagonal in Eq. (5.11), a factor of two appears in front of the zeroth-order terms, but now there are two averaging terms, $P_{i,i-1}(t)+P_{i-1,i}(t)$ and $P_{i,i+1}(t)+P_{i+1,i}(t)$, instead of just one. This pair of averaging terms resulting from the centered difference scheme will be seen to play a key role in determining the continuum dynamics being approximated along the diagonal.

5.3 Analysis of the Covariance Diagonal Propagation

The semi-discretizations of the covariance diagonal for the upwind and centered difference schemes in Eqs. (5.11) and (5.14), respectively, both contain terms that average across the diagonal. These averaging terms influence the continuum dynamics that the numerical schemes are approximating, and we will show in this section how these terms manifest themselves during discrete variance propagation.

For the error analysis, we will use Taylor series expansions, which is a common technique used to study the continuum dynamics being approximated by numerical schemes. Since Taylor series expansions require the existence and continuity of a certain number of derivatives, we will consider two cases regarding the differentiability of covariances. The first case we consider are covariances $P(x_1, x_2, t)$ that are at least four times continuously differentiable across $x_1 = x_2$ in Sec. 5.3.1. This is followed by Sec. 5.3.2, where we assume the covariance is not continuously differentiable

across $x_1 = x_2$. Such discontinuities across $x_1 = x_2$ can occur when gradients are steep near the diagonal or as correlation lengths become small, for example. A covariance model with a cusp along $x_1 = x_2$, such as the First-Order Autoregressive (FOAR) correlation function, Eq. (3.6), is another such example. Though these two cases differ in their assumptions on differentiability, they yield similar results that describe the behavior observed along the diagonal during discrete covariance propagation.

5.3.1 Continuously Differentiable Across $x_1 = x_2$

We begin with the case that the covariance $P(x_1, x_2, t)$ is four times continuously differentiable in space, i.e. $P(x_1, x_2, t) \in C^4(\mathbb{S}^1_1 \times \mathbb{S}^1_1)$. We will also assume the velocity field $v(x, t)$ is four times continuously differentiable in space, $v(x, t) \in C^4(\mathbb{S}^1_1)$. Recall that the covariance satisfies $P_{i,i-1}(t) = P(x_i, x_{i-1}, t)$ and $P_{i,i+1}(t) = P(x_i, x_{i+1}, t)$ and is spatially symmetric, e.g. $P(x_i, x_{i-1}, t) = P(x_{i-1}, x_i, t)$.

Upwind Spatial Discretization

The term we are interested in applying the Taylor expansions to is $v_{i-1}(t)\big[P_{i-1,i}(t)+P_{i,i-1}(t)\big]$ in Eq. (5.11). The covariance terms are averaging across the diagonal of the covariance to approximate the covariance at the half-grid point $P(x_{i-1/2}, x_{i-1/2}, t)$. Therefore, we will expand the terms in $v_{i-1}(t)\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big]$ about $x_{i-1/2},\ x_{i-1/2}$.

Since the covariance is continuously differentiable and symmetric, the Taylor series expansion of the averaging term only contains even order derivatives as the odd order derivatives of the covariance vanish due to symmetry. When expanding both covariance terms about $x_{i-1/2},\ x_{i-1/2}$ and also expanding $v_{i-1}(t)$ about $x_{i-1/2}$, we have the following,

$$\begin{aligned} v_{i-1}(t)\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big] &= 2v(x_{i-1/2}, t)P(x_{i-1/2}, x_{i-1/2}, t) - \Delta x v_x(x_{i-1/2}, t)P(x_{i-1/2}, x_{i-1/2}, t) \\ &\quad + \left(\frac{\Delta x}{2}\right)^2 v(x_{i-1/2}, t)P_2(x_{i-1/2}, t) - \left(\frac{\Delta x}{2}\right)^3 v_x(x_{i-1/2}, t)P_2(x_{i-1/2}, t) \\ &\quad + J^-(x_{i-1/2}, t) + K(x_{i-1/2}, t) \end{aligned} \tag{5.15}$$

The terms $J^-(x_{i-1/2}, t)$ and $K(x_{i-1/2}, t)$ are error terms defined in Eqs. (C.1) and (C.2), respec-

tively. The term $J^-(x_{i-1/2},t)$ contains products of the covariance evaluated along the diagonal and spatial derivatives of the velocity and is $\mathcal{O}(\Delta x^2)$. In contrast, the term $K(x_{i-1/2},t)$ contains product of the velocity and its derivatives with derivatives of the covariance across the diagonal and is $\mathcal{O}(\Delta x^4)$. We use the Landau notation, $\mathcal{O}(\Delta x^4)$ for example, here to denote that the terms are bounded by $M\Delta x^4$ for some nonzero constant M due to the assumptions of differentiability. Though $K(x_{i-1/2},t)$ is $\mathcal{O}(\Delta x^4)$ for correlation lengths bounded away from zero, we will see later that these terms do become significant as correlation lengths become small for fixed Δx.

In Eq. (5.15), we introduce the notation $P_2(x_{i-1/2},t)$, which is the following linear combination of second spatial derivatives,

$$P_2(x,t) = \Big[P_{x_1x_1}(x_1,x_2,t) - 2P_{x_1x_2}(x_1,x_2,t) + P_{x_2x_2}(x_1,x_2,t)\Big]_{x_1=x_2=x}. \tag{5.16}$$

By substituting the expansion in Eq. (5.15) into the upwind covariance diagonal semi-discretization in Eq. (5.11) and rearranging slightly, we have the following,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) &= \frac{1}{\Delta x}\big[2v(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) - 2v(x_i,t)P(x_i,x_i,t)\big] \\ &\quad - 2\big[b(x_i,t) - v_x(x_i,t)\big]P(x_i,x_i,t) - v_x(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) \\ &\quad + \frac{\Delta x}{4}v(x_{i-1/2},t)P_2(x_{i-1/2},t) - \frac{\Delta x^2}{8}v_x(x_{i-1/2},t)P_2(x_{i-1/2},t) \\ &\quad + \frac{1}{\Delta x}\left[J^-(x_{i-1/2},t) + K(x_{i-1/2},t)\right].\end{aligned} \tag{5.17}$$

Equation (5.17) contains terms on the grid points x_i and $x_{i-1/2}$, therefore the final step is to expand all terms to the common grid point x_i. For example, the first group of terms on the right-hand side of Eq. (5.17) represents a first-order upwind spatial discretization of $-(vP)_x$ evaluated at $x_1 = x_2 = x_i$, where the factors of two are a result of this approximation being made with the points x_i and $x_{i-1/2}$. When rewriting Eq. (5.17) so that all terms are on the grid point x_i (see Appen. C.1.1 for details), we arrive at the continuum dynamics the numerical scheme is

approximating,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i} - \big[2b(x_i,t) - 2v_x(x_i,t)\big]P(x_i,x_i,t)\\ &- v_x(x_i,t)P(x_i,x_i,t) - \frac{\Delta x^2}{8}\big(vP_2\big)_x\big|_{x=x_i} - \frac{\Delta x^2}{8}v_x(x_i,t)P_2(x_i,t)\\ &+ \frac{\Delta x}{4}v(x_i,t)P_2(x_i,t) + G_u(x_i,t) + H_u(x_i,t).\end{aligned} \tag{5.18}$$

We have kept in Eq. (5.18) terms that are significant and defined new error terms $G_u(x_i,t)$ and $H_u(x_i,t)$. The quantity $G_u(x_i,t)$ contains terms involving the covariance and its derivatives along the diagonal and is $\mathcal{O}(\Delta x)$, while $H_u(x_i,t)$ contains the derivatives of the covariance across the diagonal and is $\mathcal{O}(\Delta x^3)$. See Appen. C.1.1 for further discussion and formulation of G_u and H_u. Since the error terms in Eq. (5.18) are at least $\mathcal{O}(\Delta x)$, the scheme is first order, which is the same as the order of the upwind semi-discretization.

Due to the similar behaviors of Eq. (5.18) and the centered difference case derived in the next section, we save the discussion of this equation for Sec. 5.3.1.

Centered Difference Spatial Discretization

The error analysis of the centered difference semi-discretizaion of the covariance diagonal in Eq. (5.14) is similar to the analysis presented for the upwind scheme in the previous section. In this case, there are two sets of averaging terms to which we will apply Taylor series expansions. The first averaging term $v_{i-1}(t)\big[P_{i-1,i}(t)+P_{i,i-1}(t)\big]$ is expanded about $x_{i-1/2},\ x_{i-1/2}$, as was done for the upwind scheme. The expansion for the other averaging term $v_{i+1}(t)\big[P_{i+1,i}(t)+P_{i,i+1}(t)\big]$ is done in a similar manner about $x_{i+1/2},\ x_{i+1/2}$,

$$\begin{aligned}v_{i+1}(t)\big[P_{i+1,i}(t)+P_{i,i+1}(t)\big] = &\ 2v(x_{i+1/2},t)P(x_{i+1/2},x_{i+1/2},t) + \Delta x v_x(x_{i+1/2},t)P(x_{i+1/2},x_{i+1/2},t)\\ &+ \left(\frac{\Delta x}{2}\right)^2 v(x_{i+1/2},t)P_2(x_{i+1/2},t) + \left(\frac{\Delta x}{2}\right)^3 v_x(x_{i+1/2},t)P_2(x_{i+1/2},t)]\\ &+ J^+(x_{i+1/2},t) + K(x_{i+1/2},t).\end{aligned} \tag{5.19}$$

As before, J^+ and K are error terms defined in Eqs. (C.1) and (C.2) and are $\mathcal{O}(\Delta x^2)$ and $\mathcal{O}(\Delta x^4)$, respectively. Substituting Eqs. (5.15) and (5.19) into the semi-discretization in Eq. (5.14) is a bit more involved than the upwind scheme; the details are given in Appen. C.1.2. Due to the expansions

about $x_{i-1/2},\ x_{i-1/2}$ and $x_{i+1/2},\ x_{i+1/2}$, the centered difference semi-discretization produces two types of terms. The odd-order terms in Δx are centered difference approximation of first derivatives at x_i, while the even-order terms in Δx are average approximations of those quantities at x_i. See Appen. C.1.2 for explicit examples.

By rewriting the pairs of odd and even terms obtained during the Taylor series expansion into derivative and average approximations, respectively, we obtain the approximated diagonal dynamics for the centered difference numerical scheme,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i} - \big[2b(x_i,t) - 2v_x(x_i,t)\big]P(x_i,x_i,t)\\ &- v_x(x_i,t)P(x_i,x_i,t) - \frac{\Delta x^2}{8}\big(vP_2\big)_x\big|_{x=x_i} - \frac{\Delta x^2}{8}v_x(x_i,t)P_2(x_i,t)\\ &+ G_c(x_i,t) + H_c(x_i,t).\end{aligned} \tag{5.20}$$

Similar to the upwind scheme, $G_c(x_i,t)$ contains derivatives of the covariance along the diagonal, while $H_c(x_i,t)$ contains the derivatives of the covariance across the diagonal. In this case, however, $G_c(x_i,t)$ is $\mathcal{O}(\Delta x^2)$ and $H_c(x_i,t)$ is $\mathcal{O}(\Delta x^4)$ as a result of the centered difference scheme; see Appen. C.1 for their explicit formulation. Since the error terms are at least second order, Eq. (5.20) is the same order of accuracy as the centered difference spatial discretization applied initially.

Approximated Dynamics and Their Relation to Correlation Length

The upwind and centered difference semi-discretizations produce similar approximated dynamics along the diagonal in Eqs. (5.18) and (5.20). The first two terms on the right-hand side in both of these equations captures what looks like the continuous spectrum equation in flux form,

$$\begin{aligned}P^c_t + (vP^c)_x + \big(2b - 2v_x\big)P^c = 0,\\ P^c(x,t_0) = P^c_0(x)\end{aligned} \tag{5.21}$$

rather than the variance equation,

$$\begin{aligned}\sigma^2_t + (v\sigma^2)_x + \big(2b - v_x\big)\sigma^2 = 0,\\ \sigma^2(x,t_0) = \sigma^2_0(x).\end{aligned} \tag{5.22}$$

If we take a closer look at the second and third terms on the right-hand sides of Eqs. (5.18) and (5.20), the additional zeroth-order term $-v_x(x_i,t)P(x_i,x_i,t)$ appearing next is what allows the schemes to approximate the variance equation. By canceling out one factor of $v_x(x_i,t)P(x_i,x_i,t)$ in the second term, this third term allows both schemes to recover what looks like the variance equation in Eq. (5.22).

The additional terms on the right-hand sides of Eqs. (5.18) and (5.20) include terms involving P_2 and its spatial derivative, and as we will see, these can alter the dynamics along the diagonal. To illustrate this, we can rewrite P_2 by decomposing the covariance $P(x_1,x_2,t)$ into the product of the diagonal quantity[2] $P(x,x,t)$ and correlation $C(x_1,x_2,t)$,

$$P(x_1,x_2,t) = \sqrt{P(x_1,x_1,t)}C(x_1,x_2,t)\sqrt{P(x_2,x_2,t)}. \tag{5.23}$$

After differentiating according to Eq. (5.16), we have

$$P_2(x_i,t) = P(x_i,x_i,t)\big\{\log[P(x,x,t)]\big\}_{xx}\big|_{x=x_i} + P(x_i,x_i,t)C_2(x_i,t), \tag{5.24}$$

where $C_2(x,t)$ is the same linear combination of derivatives used to define $P_2(x,t)$ in Eq. (5.16) applied to the correlation $C(x_1,x_2,t)$.

The term P_2 directly relates to the correlation length through C_2, where we can define the correlation length $L(x,t)$ as follows,

$$L^2(x,t) = -\frac{1}{C_2(x,t)}, \tag{5.25}$$

as defined in Eq. (4.33) of Cohn (1993). Therefore, terms involving P_2 in Eqs. (5.18) and (5.20)

[2] We choose to use a general diagonal element $P(x,x,t)$ rather than the standard deviation or variance in this decomposition because of the discontinuous change in dynamics as correlation lengths tend to zero, as described in Ch. 2. When states are initially uncorrelated, the dynamics along the diagonal are governed by the continuous spectrum equation, Eq. (1.9), not the variance equation, Eq. (1.8). Therefore, decomposing the covariance in Eq. (5.23) using the standard deviation would not be correct for spatially uncorrelated initial states.

can be rewritten in terms of the diagonal element and correlation length,

Upwind:

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i}\left[1-\frac{\Delta x^2}{8L^2(x_i,t)}\right]-\big[2b(x_i,t)-2v_x(x_i,t)\big]P(x_i,x_i,t)\\ &-v_x(x_i,t)P(x_i,x_i,t)\left[1-\frac{\Delta x^2}{8L^2(x_i,t)}\right]\\ &-\frac{\Delta x}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)+\tilde{G}_u(x_i,t)+\tilde{H}_u(x_i,t),\end{aligned} \tag{5.26}$$

Centered Difference:

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i}\left[1-\frac{\Delta x^2}{8L^2(x_i,t)}\right]-\left[2b(x_i,t)-2v_x(x_i,t)\right]P(x_i,x_i,t)\\ &-v_x(x_i,t)P(x_i,x_i,t)\left[1-\frac{\Delta x^2}{8L^2(x_i,t)}\right]+\tilde{G}_c(x_i,t)+\tilde{H}_c(x_i,t).\end{aligned} \tag{5.27}$$

The error terms G_u, H_u, G_c, and H_c have been updated after replacing P_2 with its expansion in Eq. (5.24) to define $\tilde{G}_u, \tilde{H}_u, \tilde{G}_c$, and $\tilde{H}_c$, which are given in Appen. C.1.

Both the flux term $-\big(vP\big)_x$ and zeroth-order term $-v_x(x_i,t)P(x_i,x_i,t)$ have a coefficient in Eqs. (5.26) – (5.27) that now depends on the ratio of the grid length Δx to the correlation length L. For correlation lengths that are sufficiently large, $\frac{\Delta x^2}{L^2(x,t)}$ is small and tends to zero as Δx tends to zero for this large, fixed L. This case yields an extra factor of $-v_x(x_i,t)P(x_i,x_i,t)$ that combines with the other zeroth-order terms in Eqs. (5.26) – (5.27) to recover the variance equation, Eq. (5.22). When correlation lengths become small enough such that the ratio $\frac{\Delta x^2}{8L^2(x,t)} \approx 1$, both the flux term $-(vP)_x$ and the extra $-v_x(x_i,t)P(x_i,x_i,t)$ term vanish, reducing the partial differential equation to an ordinary differential equation in time only.

We can write the behavior observed in Eqs. (5.26) – (5.27) concisely as

$$p_t + \lambda(x,t)(vp)_x + (2b-2v_x)p + \lambda(x,t)v_x p = 0, \tag{5.28}$$

where $p = p(x,t) = P(x,x,t)$ represents the general diagonal element of the covariance, and

$$\lambda(x,t) = 1 - \frac{\Delta x^2}{8L^2(x,t)}. \tag{5.29}$$

The coefficient $\lambda(x,t)$ plays a primary role in the dynamics the numerical schemes are approximating along the diagonal of the covariance as the correlation length $L(x,t)$ evolves over space and time.

As derived in Cohn (1993, Sec. 4b), the correlation length $L(x,t)$ for states governed by Eq. (1.4) satisfies its own differential equation,

$$L_t + vL_x - v_x L = 0,$$
$$L(x,t_0) = L_0(x). \tag{5.30}$$

For spatially-varying velocity fields ($v_x \neq 0$) the correlation length L will evolve over space and time. In particular, from Eq. (5.30), $1/L$ satisfies the continuity equation. This implies that regions of convergence (where $(v/L)_x < 0$) will cause correlation lengths to shrink, balanced by regions of divergence (where $(v/L)_x > 0$) where correlation lengths grow since $1/L$ integrated over $\mathbb{S}^1_1$ is conserved. Small correlation lengths can be expected to occur dynamically, therefore λ can become significant enough to alter the dynamics being approximated along the diagonal of the covariance.

The difference between the upwind and centered difference approximated dynamics occurs with the additional first-order error term, $-\frac{\Delta x}{4L^2(x_i,t)} v(x_i,t) P(x_i,x_i,t)$, in Eq. (5.26). This term does not appear in the centered difference scheme, as it becomes a second-order error term when performing the analysis (see Appen. C.1). The term $-\frac{\Delta x}{4L^2(x_i,t)} v(x_i,t) P(x_i,x_i,t)$ is a zeroth-order dissipative term, which can be even more problematic than the second-order dissipative terms that are commonly encountered during upwind differencing. Even for modest correlation lengths, this dissipative term can be quite significant, changing the approximated dynamics along the diagonal considerably. This zeroth-order dissipative term is a result of the discontinuous change in continuum dynamics, and is different from the numerical dissipation often discussed in the context of covariance propagation (i.e., a second-order dissipative term). Additional discussion on this in the context of other data assimilation literature is given in Secs. 5.4 and 6.2.

The magnitude of λ depends on the correlation length L, which we know from its governing dynamics in Eq. (5.30), can shrink for a spatially-varying velocity field. To gain an idea for how large $|\lambda|$ can become, we consider the following one-dimensional example in which the correlation lengths can be computed exactly. We let states q satisfy the one-dimensional version of Eq. (1.4) with $b = v_x$, and take the velocity $v = v(x) = \sin(x) + 2$ to vary in space, but not time. The

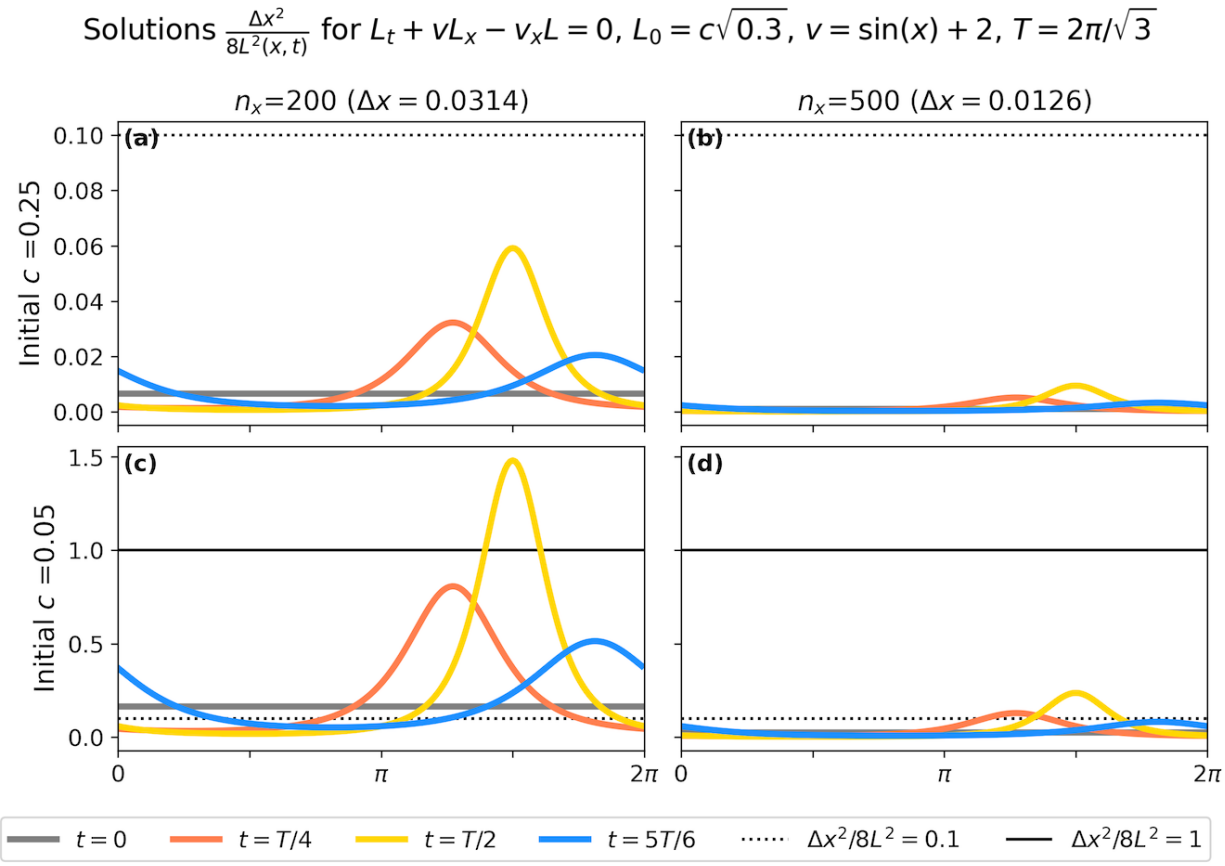

Figure 5.1: Time snapshots (colors) of the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ as functions of space for different initial correlation lengths L_0 and number of grid points on the discretized unit circle. Correlation lengths $L(x,t)$ are exact solutions to Eq. (5.30) with constant initial correlation lengths L_0, where L_0 corresponds to the correlation length of the Gaspari-Cohn correlation function introduced in Ch. 3. Each row corresponds to two different cut-off length parameters c, where $c = 0.25$ corresponds to the initial correlation function being compactly-supported on a quarter of the domain, for example. Each column plots solutions for two uniformally-spaced spatial grids, with 200 grid points on the left and 500 grid points on the right. Solid black lines indicated where the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ is equal to one, and black dotted when the ratio is one-tenth.

solution to the state q in this case, as well as the correlation length L, can be solved exactly and are given in Appen. B.2. For different initial correlation length and fixed grid lengths Δx, we can solve for the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ as a function of space and time. Figure 5.1 illustrates these results. Each panel plots the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ on the unit circle at different times (colored lines) by solving for the correlation length L from Eq. (5.30). Each row corresponds to a different initial correlation length, which is constant and taken to be the correlation length of the Gaspari-Cohn correlation function introduced in Ch. 3. Each column corresponds to two different uniform spatial discretizations, equivalently, two different values of Δx.

We can see in Fig. 5.1 that the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ evolves over time and can become quite large. For the larger initial correlation length (top row), the values of the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ are small, but in panel (a) can be significant enough to alter the dynamics along the diagonal in Eq. (5.28). This becomes more pronounced for the short initial correlation length (bottom row), where values in panel (c) vary significantly and become quite large. Values of $\frac{\Delta x^2}{8L^2(x,t)}$ that are greater than one reverse the direction of propagation, which we see can occur for the values shown in panel (c). Even for a relatively dense spatial grid with $n_x = 500$ in panel (d), the short initial correlation length case produces values of $\frac{\Delta x^2}{8L^2(x,t)}$ that can be significant, larger than the values shown in panel (a). One could surmise here that increasing spatial resolution may not be an adequate solution to combat inaccurate evolution along the covariance diagonal. Additionally, large-scale data assimilation schemes are already computationally expensive, and increasing spatial resolution such that the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ is small enough may be too costly.

From Fig. 5.1, we can also deduce the size of the dissipative first-order error term (fourth term on the right-hand side) in Eq. (5.26). Since the ratio $\frac{\Delta x^2}{8L^2(x,t)}$ can become large, this implies that the first-order error term $\frac{\Delta x}{4L^2(x,t)}$ may also become quite large. If this term is large enough, it can cause significant dissipation that manifests itself in what looks like variance loss during covariance propagation. From Eq. (5.26) we see that along with this dissipative term, the approximated continuum dynamics are changing as well, indicating that the underlying problem is inaccurate propagation.

The error terms $\tilde{H}_u$ and $\tilde{H}_c$ in Eqs. (5.26) and (5.27), which contain derivatives across the covariance diagonal, also depend on powers of the ratio $\Delta x/L(x,t)$. Though we do not include the explicit formulation of these error terms in Eqs. (5.26) and (5.27), they too can become large as the ratio $\Delta x/L(x,t)$ approaches unity. Hence, Fig. 5.1 can provide insight into how these error terms can behave and further alter the approximated diagonal dynamics when correlation lengths become small for fixed Δx.

Figure 5.1 quantitatively demonstrates that λ can significantly alter the dynamics along the diagonal of the covariance even for modest initial correlation lengths and dense spatial discretizations. Therefore, the $\mathcal{O}(\Delta x^2)$ coefficient in the first and third terms on the right-hand sides of Eqs. (5.26) and (5.27) is not a negligible error term as it directly influences the continuum dynamics along the diagonal of the covariance. The inaccurate variance propagation caused by the P_2 error term is a consequence of averaging across the diagonal.

5.3.2 Discontinuous Across $x_1 = x_2$

Differentiability of covariances across $x_1 = x_2$ is not always guaranteed. For example, gradients near the diagonal can become sharp enough that differentiability across $x_1 = x_2$ is lost relative to the size of the discrete grid, in which case the analysis of the previous section does not apply. In this section, we assume that the covariance is not differentiable across $x_1 = x_2$ and instead take advantage of the symmetry of the covariance function to derive results similar to those derived in the continuously differentiable case.

We will now assume that the covariance is four times continuously differentiable away from $x_1 = x_2$, but not across $x_1 = x_2$. The spatial derivatives from the left and the right of $x_1 = x_2$ are assumed to exist and are defined as follows,

$$P^{\pm}_{x_1}(x_1,x_2,t) = \lim_{x\to x_1^{\pm}} \frac{\partial P}{\partial x_1}(x,x_2,t); \quad P^{\pm}_{x_2}(x_1,x_2,t) = \lim_{x\to x_2^{\pm}} \frac{\partial P}{\partial x_2}(x_1,x,t), \tag{5.31}$$

where the superscript $+$ denotes derivative from the right and $-$ the derivative from the left. Using this notation, we will redefine P_2 of Eq. (5.16), and introduce two new terms P_1 and P_3, which will

now appear in the expansions,

$$P_1^{\pm}(x,t) = \Big[P_{x_1}^{\pm}(x_1,x_2,t) - P_{x_2}^{\mp}(x_1,x_2,t)\Big]_{x_1=x_2=x}, \tag{5.32}$$

$$P_2^{\pm}(x,t) = \Big[P_{x_1x_1}^{\pm\pm}(x_1,x_2,t) - 2P_{x_1x_2}^{\pm\mp}(x_1,x_2,t) + P_{x_2x_2}^{\mp\mp}(x_1,x_2,t)\Big]_{x_1=x_2=x}, \tag{5.33}$$

$$P_3^{\pm}(x,t) = \Big[P_{x_1x_1x_1}^{\pm\pm\pm}(x_1,x_2,t) - 3P_{x_1x_1x_2}^{\pm\pm\mp}(x_1,x_2,t) + 3P_{x_1x_2x_2}^{\pm\mp\mp}(x_1,x_2,t) - P_{x_2x_2x_2}^{\mp\mp\mp}(x_1,x_2,t)\Big]_{x_1=x_2=x}. \tag{5.34}$$

By symmetry of the covariance, we have that

$$P_1^{-}(x,t) = -P_1^{+}(x,t), \quad P_2^{-}(x,t) = P_2^{+}(x,t), \quad P_3^{-}(x,t) = -P_3^{+}(x,t), \tag{5.35}$$

for all x and t. During the expansions, derivatives will either be taken from the left or the right depending on where $x_{i-1/2}$ and $x_{i+1/2}$ lie on the spatial grid relative to the other spatial variables. We still assume that the velocity field is four times continuously differentiable in its spatial argument.

We begin with the upwind spatial discretization. As before, we will expand the average term $v_{i-1}(t)\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big]$ about $x_{i-1/2},\ x_{i-1/2}$, but this time taking partial derivatives from the left for spatial variable x_i and from the right for spatial variable x_{i-1}. After computing these expansions and invoking the symmetry property in Eq. (5.35) to rewrite derivatives in terms of the right partials only (see Eq. C.20 in Appen. C.1.3, for example), we have the following semi-discretization for the upwind scheme,

$$\begin{aligned}
\frac{d}{dt}P(x_i,x_i,t) = {} & \frac{1}{\Delta x}\big[2v(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) - 2v(x_i,t)P(x_i,x_i,t)\big] \\
& - 2\big[b(x_i,t) - v_x(x_i,t)\big]P(x_i,x_i,t) \\
& - v_x(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) + v(x_{i-1/2},t)P_1^{+}(x_{i-1/2},t) \\
& + \frac{\Delta x}{4}\Big[v(x_{i-1/2},t)P_2^{+}(x_{i-1/2},t) - 2v_x(x_{i-1/2},t)P_1^{+}(x_{i-1/2},t)\Big] \\
& - \frac{\Delta x^2}{8}v_x(x_{i-1/2},t)P_2^{+}(x_{i-1/2},t) + \frac{1}{\Delta x}\Big[J^{-}(x_{i-1/2},t) + \tilde{K}^{-}(x_{i-1/2},t)\Big].
\end{aligned} \tag{5.36}$$

Like in Eq. (5.17), we have two error terms $J^{-}(x_{i-1/2},t)$ and $\tilde{K}^{-}(x_{i-1/2},t)$, which are defined in Eqs. (C.1) and (C.19), respectively. Here, J^{-} is $\mathcal{O}(\Delta x^2)$ and contains covariances evaluated

along the diagonal, while $\tilde{K}^-$ is $\mathcal{O}(\Delta x^4)$ and contains terms that are derivatives (from the right) of the covariance across the diagonal. We observe here that Eq. (5.36) contains terms involving P_1^+ and P_3^+ that did not appear in Eq. (5.17), including in $\tilde{K}^-(x_{i-1/2}, t)$. When the covariance has continuous derivatives across $x_1 = x_2$, P_1^+, P_2^+, and P_3^+ become full derivatives and P_1^+ and P_3^+ vanish, so that Eq. (5.17) and Eq. (5.36) are equivalent.

Like the continuously differentiable case, the first group of terms on the right-hand side of Eq. (5.36) is a first-order upwind spatial discretization of $-(vP)_x$ evaluated at $x_1 = x_2 = x_i$, and there are additional terms evaluated on the half-grid that must be expanded to the grid point x_i. When expanding terms about x_i in Eq. (5.36), the approximated dynamics along the diagonal of the covariance for the upwind semi-discretization is the following,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i} - \big[2b(x_i,t) - 2v_x(x_i,t)\big]P(x_i,x_i,t)\\ &- v_x(x_i,t)P(x_i,x_i,t) - v_x(x_i,t)\left[\Delta x P_1^+(x_i,t) + \frac{\Delta x^2}{4}P_2^+(x_i,t)\right]\\ &+ v(x_i,t)\left[P_1^+(x_i,t) + \frac{\Delta x}{4}P_2^+(x_i,t)\right]\\ &- \frac{\Delta x}{2}v(x_i,t)\big(P_1^+\big)_x\big|_{x=x_i} - \frac{\Delta x^2}{8}v(x_i,t)\big(P_2^+\big)_x\big|_{x=x_i} + \hat{G}_u(x_i,t) + \hat{H}_u(x_i,t).\end{aligned} \tag{5.37}$$

The error terms $\hat{G}_u$ and $\hat{H}_u$ are of a similar definition to those in Sec. 5.3.1, with $\hat{G}_u$ being $\mathcal{O}(\Delta x)$ and $\hat{H}_u$ being $\mathcal{O}(\Delta x^2)$ due to the presence of P_1^+. See Appen. C.1.3 for further details.

There are important similarities and differences between the approximated dynamics in Eq. (5.37) compared to its counterpart when the covariance is continuously differentiable, Eq. (5.18). When the covariance is four times continuously differentiable, P_1^+ and P_3^+ vanish, and Eq. (5.37) would reduce to Eq. (5.18). When P_1^+ and P_3^+ are nonzero, additional error terms are introduced to Eq. (5.18) that are problematic. The zeroth-order term $v(x_i,t)P_1^+(x_i,t)$ on the right-hand side of Eq. (5.37) is particularly problematic, as the scheme that governs the diagonal propagation is no longer even first-order accurate because of this error term.

We see similar behaviors in the approximated dynamics for the centered difference spatial discretization. As in the continuously differentiable case, there are two averaging terms that need

to be expanded. The term $v_{i-1}(t)\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big]$ was already expanded for the upwind case, therefore we are left to expand the other averaging term, $v_{i+1}(t)\big[P_{i+1,i}(t) + P_{i,i+1}(t)\big]$, about $x_{i+1/2},\ x_{i+1/2}$. In this case, derivatives are taken from the right for spatial variables x_{i+1} and from the left for spatial variables x_{i-1}. The details of the expansions are lengthy and presented in Appen. C.1.3 for reference. The expansions produce P_1^+ and P_3^+ terms that appear due to the discontinuity in the derivatives across $x_1 = x_2$ and would vanish when the covariance has at least four continuous derivatives across $x_1 = x_2$. As a consequence, we retain odd-order error terms, unlike the continuously differentiable case. The approximated dynamics for the diagonal of the covariance from the centered difference spatial discretization scheme is as follows,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i} - \big[2b(x_i,t) - 2v_x(x_i,t)\big]P(x_i,x_i,t)\\ &- v_x(x_i,t)P(x_i,x_i,t) - v_x(x_i,t)\left[\Delta x P_1^+(x_i,t) + \frac{\Delta x^2}{4}P_2^+(x_i,t)\right]\\ &- \frac{\Delta x}{2}v(x_i,t)\big(P_1^+\big)_x\big|_{x=x_i} - \frac{\Delta x^2}{8}v(x_i,t)\big(P_2^+\big)_x\big|_{x=x_i} + \hat{G}_c(x_i,t) + \hat{H}_c(x_i,t),\end{aligned} \tag{5.38}$$

with the error terms $\hat{G}_c$ and $\hat{H}_c$ defined in Appen. C.1.3.

As in the upwind case, there are now additional error terms involving P_1^+ and P_3^+ as a consequence of the covariance having discontinuous derivatives across $x_1 = x_2$. The presence of the first-order term $-\Delta x v_x(x_i,t)P_1^+(x_i,t)$ in Eq. (5.38) renders the order of the diagonal propagation to be first-order accurate, reducing the order of accuracy that would be expected from a centered-difference scheme. The results for both the upwind and centered difference cases indicate that for covariances that are not continuously differentiable across $x_1 = x_2$, neither scheme can approximate the variance dynamics accurately, regardless of the magnitude of P_1^+. Similar to the continuously differentiable case, the upwind scheme contains additional terms in Eq. (5.37) that are not present in the centered difference scheme in Eq. (5.38), as they move to higher order.

Since the covariance is not continuously differentiable across $x_1 = x_2$, rewriting Eqs. (5.37) and (5.38) in terms of the correlation length defined in Eq. (5.25) cannot be done. However, we can gain a better idea of the behavior of the approximated dynamics in this case by rewriting P_1^+ and P_2^+ in terms of the diagonal element $P(x,x,t)$ and the correlations C_1^+ and C_2^+, which are defined

as in Eqs. (5.32) and (5.33) instead for the correlation C. First, we expand P_1^+ and P_2^+ as follows,

$$P_1^+(x,t) = P(x,x,t)C_1^+(x,t), \tag{5.39}$$

$$P_2^+(x,t) = P(x,x,t)\log(P)_{xx}\big|_{x_1=x_2=x} + P(x,x,t)C_2^+(x,t). \tag{5.40}$$

Thus, the approximated dynamics for the upwind and centered difference schemes can be rewritten in terms of correlations,

Upwind:

$$\begin{aligned}
\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i}\left[1+\frac{\Delta x}{2}C_1^+(x_i,t)+\frac{\Delta x^2}{8}C_2^+(x_i,t)\right]\\
&-\big[2b(x_i,t)-2v_x(x_i,t)\big]P(x_i,x_i,t)\\
&-v_x(x_i,t)P(x_i,x_i,t)\left[1+\frac{\Delta x}{2}C_1^+(x_i,t)+\frac{\Delta x^2}{8}C_2^+(x_i,t)\right]\\
&-v(x_i,t)P(x_i,x_i,t)\Bigg[-C_1^+(x_i,t)-\frac{\Delta x}{4}C_2^+(x_i,t)\\
&\qquad\qquad+\frac{\Delta x}{2}\big(C_1^+\big)_x\big|_{x=x_i}+\frac{\Delta x^2}{8}\big(C_2^+\big)_x\big|_{x=x_i}\Bigg]\\
&+\hat{G}_u(x_i,t)+\frac{\Delta x}{4}v(x_i,t)P(x_i,x_i,t)\log(P)_{xx}\big|_{x_1=x_2=x_i}\\
&+\hat{H}_u(x_i,t)-\frac{\Delta x^2}{8}\big[vP\log(P)_{xx}\big]_x\big|_{x_1=x_2=x_i}.
\end{aligned} \tag{5.41}$$

Centered Difference:

$$\begin{aligned}
\frac{d}{dt}P(x_i,x_i,t) = &-(vP)_x\big|_{x_1=x_2=x_i}\left[1+\frac{\Delta x}{2}C_1^+(x_i,t)+\frac{\Delta x^2}{8}C_2^+(x_i,t)\right]\\
&-\big[2b(x_i,t)-2v_x(x_i,t)\big]P(x_i,x_i,t)\\
&-v_x(x_i,t)P(x_i,x_i,t)\left[1+\frac{\Delta x}{2}C_1^+(x_i,t)+\frac{\Delta x^2}{8}C_2^+(x_i,t)\right]\\
&-v(x_i,t)P(x_i,x_i,t)\left[\frac{\Delta x}{2}\big(C_1^+\big)_x\big|_{x=x_i}+\frac{\Delta x^2}{8}\big(C_2^+\big)_x\big|_{x=x_i}\right]\\
&+\hat{G}_c(x_i,t)-\frac{\Delta x^2}{8}\big[vP\log(P)_{xx}\big]_x\big|_{x_1=x_2=x_i}+\hat{H}_c(x_i,t).
\end{aligned} \tag{5.42}$$

The impact of averaging across the diagonal in Eqs. (5.41) and (5.42) is quite similar to that of the continuously differentiable case discussed in Sec. 5.3.1 when the approximated dynamics are written in terms of derivatives of the correlations. In fact, if the covariance is continuously

differentiable across $x_1 = x_2$, then this collapses exactly to the equations in Sec. 5.3.1 as the correlation length can then be defined. The presence of C_1^+ and its derivative introduces additional error terms that change the order of accuracy of the schemes and further alter the dynamics.

Though the correlation length L of Eq. (5.25) is not well-defined in this case, one could assign a correlation length for C_1^+ and C_2^+. For example, consider the First-Order Autoregressive (FOAR) correlation function, first introduced in Eq. (3.6) and revisited here with slightly different notation,

$$C_{FOAR}(x_1, x_2) = e^{-||x_1 - x_2||/\mathsf{L}}, \qquad \text{(Eq. (3.6), revisited)}$$

where $\mathsf{L} > 0$ is the length-scale parameter and $||\cdot||$ is a norm (e.g., chordal distance on the unit sphere, Euclidean distance on the plane). The FOAR correlation function is not continuously differentiable across $x_1 = x_2$ due to the cusp along $x_1 = x_2$. Unlike correlation functions that are at least twice continuously differentiable across $x_1 = x_2$, the second derivative from the left and the right of the FOAR correlation function is positive. Therefore, C_1^+ and C_2^+ for the FOAR correlation function can be rewritten in terms of the length scale parameter L.

Generalizing the definition of a correlation length for C_2^+, however, is not straightforward. For twice-continuously differentiable correlation functions, the second derivative of the correlation must be negative along $x_1 = x_2$ since this is where the correlation is a maximum. However, when the correlation function is not continuously differentiable across $x_1 = x_2$, the sign of C_2^+ does not matter. The FOAR correlation function, for example, has its second derivatives C_2^+ and C_2^- both positive. Therefore, defining a correlation length analogous to the continuously differentiable case is not obvious. The correlation length scale associated with nonzero C_1^+, however, can be generalized. Since $x_1 = x_2$ is a local maximum for correlation functions by definition, we can define a correlation length whose sign will remain consistent for different correlation functions (similar to the first derivative test for determining local extrema). We can define correlation length scale L_1 to be

$$L_1(x, t) = -\frac{1}{C_1^+(x, t)}, \tag{5.43}$$

and express P_1^+ in terms of this correlation length,

$$P_1^+(x,t) = P(x,x,t)C_1^+(x,t) = -\frac{P(x,x,t)}{L_1(x,t)}. \tag{5.44}$$

$P_1^-(x,t)$ can be defined in terms C_1^-, where the relation $C_1^+ = -C_1^-$ will yield the a similar definition of $L_1(x,t)$ as given in Eq. (5.44) so that $L_1(x,t)$ remains positive.

5.4 Discussion of Pannekoucke et al. (2021)

Recent work by Pannekoucke et al. (2021) (hereafter referred to as P21 in this section) seeks to quantify the errors caused by numerical discretization during covariance propagation. The results of P21 are similar to, but distinct from, the error analysis of this chapter in the motivation, approach, and interpretation. In this section, we will discuss the results of P21 in the context of this thesis.

The motivation for P21 is to identify the contribution of numerical discretization of the state dynamics to the model-error covariance matrix, often denoted as $\mathbf{Q}$, which is added to the discrete propagation in Eq. (1.3). The authors quantify the errors caused by numerical discretization by determining the approximated variance dynamics. The authors compute the approximated variance dynamics as follows: first, the authors discretize the state equation (in both space and time) according to a chosen numerical scheme, then they Taylor-expand the state to recover the continuum state equation that is being approximated to higher order. Finally, the authors then derive the approximated dynamics for the variance from this approximated state equation. Though the error terms P21 recover in their approximated variance equation are identical to those derived in this chapter, their method of recovering these terms differs from this chapter's analysis and falls short of identifying the fundamental cause for these error terms, as described below.

We begin this section by showing that the error terms recovered by P21 are equivalent to those derived in the error analysis of the state equation in advection form, Appen. C.2. With this established, we conclude this section with further discussion of P21.

5.4.1 Analysis of Their Approximated Variance Dynamics

The state dynamics in P21 satisfy the one-dimensional advection equation, Eq. (1.4) for $b = 0$. The authors apply a first order semi-Lagrangian scheme with linear interpolation to compute the full discretization for the state. Here, we will apply the discretization in space only, which is a first-order upwind scheme, to recover the corresponding semi-discretization using the notation established in this chapter,

$$\frac{d}{dt}q_i(t) = \frac{v_i(t)}{\Delta x}\left[q_{i-1}(t) - q_i(t)\right]. \tag{5.45}$$

In line with the analysis applied in P21, we expand the state $q_{i-1}(t)$ about x_i to second order in space, then substitute this expansion back into Eq. (5.45),

$$\frac{d}{dt}q(x_i,t) = -v(x_i,t)q_x(x_i,t) + \frac{1}{2}\Delta x v(x_i,t)q_{xx}(x_i,t) + \mathcal{O}(\Delta x^2). \tag{5.46}$$

Equation (5.46) is what we will now use to determine the approximated dynamics for the variance. In P21, the authors skip Eq. (5.46) and instead present the "modified equation," which is the PDE the numerical scheme is approximating to higher order. To remain consistent with error analysis presented in this chapter, we will not do this and instead work directly with Eq. (5.46).

To proceed to the approximated dynamics for the variance, we will consider the error $\varepsilon(x,t)$ defined in Eq. (5.6). Due to linearity and v being deterministic, the error $\varepsilon(x_i,t)$ also satisfies Eq. (5.46). Next, we can determine the associated equation for the product $\varepsilon(x_i,t)\varepsilon(x_j,t)$,

$$\begin{aligned}\frac{d}{dt}\varepsilon(x_i,t)\varepsilon(x_j,t) =& - v(x_i,t)\varepsilon_x(x_i,t)\varepsilon(x_j,t) - v(x_j,t)\varepsilon_x(x_j,t)\varepsilon(x_i,t) \\ &+ \frac{1}{2}\Delta x\big[v(x_i,t)\varepsilon_{xx}(x_i,t)\varepsilon(x_j,t) + v(x_j,t)\varepsilon_{xx}(x_j,t)\varepsilon(x_i,t)\big].\end{aligned} \tag{5.47}$$

To deviate slightly from P21 so that the derivation of the approximated variance dynamics is clear, we will now adopt the notation used in this chapter. Observe the following relation for the sum of the second partial derivatives of the covariance P,

$$\big[P_{x_1x_1}(x_1,x_2,t) + P_{x_2x_2}(x_1,x_2,t)\big]_{x_1=x_2=x} = \frac{1}{2}\big[P_{xx}(x,x,t) + P_2(x,t)\big]. \tag{5.48}$$

Since we have that $P_{x_i x_i}(x_i, x_j, t) = \mathbb{E}\left[\varepsilon_{xx}(x_i, t)\varepsilon(x_j, t)\right]$, we can rewrite Eq. (5.47) after taking the expectation and evaluating along $i = j$ as follows,

$$\frac{d}{dt}P(x_i, x_i, t) = -v(x_i, t)P_x(x_i, x_i, t) + \frac{\Delta x}{4}v(x_i, t)P_{xx}(x_i, x_i, t) + \frac{\Delta x}{4}v(x_i, t)P_2(x_i, t). \tag{5.49}$$

By rewriting P_2 in terms of the diagonal element and correlation length L, Eq. (C.8), and expanding the derivative of $\log(P)_{xx}$, we have the following approximated dynamics along the covariance diagonal,

$$\begin{aligned}\frac{d}{dt}P(x_i, x_i, t) =& - v(x_i, t)P_x(x_i, x_i, t) + \frac{\Delta x}{2}v(x_i, t)P_{xx}(x_i, x_i, t) \\ &- \frac{\Delta x}{4}v(x_i, t)\frac{(P_x)^2(x_i, x_i, t)}{P(x_i, x_i, t)} - \frac{\Delta x}{4L^2(x_i, t)}v(x_i, t)P(x_i, x_i, t).\end{aligned} \tag{5.50}$$

Let's now compare Eq. (5.50) to Eq. (40a) of P21, which is their approximated dynamics for the variance. Equation (40a) of P21 in their notation is given as

$$\partial_t V^p + U\partial_x V^p = -\frac{V^p \kappa}{\nu^p} + \kappa\partial_x^2 V^p - \frac{\kappa(\partial_x V^p)^2}{2V^p}, \tag{5.51}$$

Now, remapping to the notation used in this chapter, noting that P21 defines an anisotropic tensor ν^p in their Eq. (25) that relates to the correlation length L,

$$\begin{aligned} V^p &\mapsto P(x, x, t) \\ U &\mapsto v(x, t) - \frac{\Delta t}{2}v_t(x, t) + \frac{\Delta t}{2}v(x, t)v_t(x, t) \;\Rightarrow v(x, t), \\ \kappa &\mapsto \frac{v(x, t)}{2}[\Delta x - v(x, t)\Delta t] \;\Rightarrow \frac{\Delta x}{2}v(x, t), \\ \nu^p &\mapsto \frac{1}{2}L^2(x, t). \end{aligned}$$

Since we only discretize in space, we can drop the time-dependent errors. We can now rewrite Eq. (5.51) to compare with Eq. (5.50),

$$\begin{aligned}P_t(x, x, t) + v(x, t)P_x(x, x, t) =& -\frac{\Delta x}{4L^2(x, t)}v(x, t)P(x, x, t) + \frac{\Delta x}{2}v(x, t)P_{xx}(x, x, t) \\ &- \frac{\Delta x}{4}v(x_i, t)\frac{(P_x)^2(x, x, t)}{P(x, x, t)}.\end{aligned} \tag{5.52}$$

The analysis presented in P21, therefore, recovers the same error terms (up to first order only) that are derived from the error analysis in advection form, Appen. C.2.

Retaining only the first-order error terms is where P21 stops in their analysis, claiming that the variance loss observed during covariance propagation is a result of numerical diffusion (see P21 p. 8). However, the error analysis in this chapter demonstrates that the second-order error terms are equally as important because they can change the approximated dynamics entirely. We can actually recover the second-order error terms in a similar process by expanding the state up to third order in space,

$$\frac{d}{dt}q(x_i,t) = -v(x_i,t)q_x(x_i,t) + \frac{1}{2}\Delta x v(x_i,t)q_{xx}(x_i,t) - \frac{1}{6}\Delta x^2 v(x_i,t)q_{xxx}(x_i,t) + \mathcal{O}(\Delta x^3). \quad (5.53)$$

As before, the error $\varepsilon(x_i,t)$ satisfies the same equation as the state shown above. The equation for the product $\varepsilon(x_i,t)\varepsilon(x_j,t)$ is almost identical to the previous case but with an additional term of order Δx^2,

$$\begin{aligned}\frac{d}{dt}\varepsilon(x_i,t)\varepsilon(x_j,t) = &- v(x_i,t)\varepsilon_x(x_i,t)\varepsilon(x_j,t) - v(x_j,t)\varepsilon_x(x_j,t)\varepsilon(x_i,t) \\ &+ \frac{\Delta x}{2}\left[v(x_i,t)\varepsilon_{xx}(x_i,t)\varepsilon(x_j,t) + v(x_j,t)\varepsilon_{xx}(x_j,t)\varepsilon(x_i,t)\right] \\ &- \frac{\Delta x^2}{6}\left[v(x_i,t)\varepsilon_{xxx}(x_i,t)\varepsilon(x_j,t) + v(x_j,t)\varepsilon_{xxx}(x_j,t)\varepsilon(x_i,t)\right]\end{aligned} \quad (5.54)$$

With a bit of algebra, one can derive the following relation,

$$\left[P_{x_1x_1x_1}(x_1,x_2,t) + P_{x_2x_2x_2}(x_1,x_2,t)\right]_{x_1=x_2=x} = \frac{1}{4}\left[P_{xxx}(x,x,t) + 3(P_2)_x(x,t)\right] \quad (5.55)$$

Therefore, we can rewrite the new term in Eq. (5.54) as

$$\begin{aligned}\mathbb{E}\Big\{ -\frac{\Delta x^2}{6}\left[v(x_i,t)\varepsilon_{xxx}(x_i,t)\varepsilon(x_j,t) + v(x_j,t)\varepsilon_{xxx}(x_j,t)\varepsilon(x_i,t)\right]_{x_i=x_j}\Big\} = \\ -\frac{\Delta x^2}{24}v(x_i,t)P_{xxx}(x_i,x_i,t) - \frac{\Delta x^2}{8}v(x_i,t)(P_2)_x(x_i,t).\end{aligned}$$

we have now recovered the problematic P_2 term originally derived in Sec. 5.3.1. In addition to the first-order error terms recovered already, we have derived the same approximated dynamics as done in the advection form case presented in Appen. C.2.

5.4.2 Discussion

The approach presented in P21 for deriving the approximated variance dynamics contrasts with the error analysis presented earlier in this chapter in two fundamental ways. The first is the analysis itself. The approach taken in P21 to Taylor-expand the state, rather than the covariance, is a logical choice from a data assimilation perspective since covariance propagation is defined from the discretized state propagation. As demonstrated in Ch. 2, however, to fully understand covariance propagation for advective dynamics requires careful analysis of the covariance itself since the hyperplane $x_1 = x_2$ is a characteristic surface. In addition, the discontinuous change in the continuum dynamics along $x_1 = x_2$ is only uncovered by carefully analyzing the continuum covariance dynamics. Therefore, deriving the approximated dynamics along the covariance diagonal from the covariance semi-discretization itself, as done in this chapter, is a more discerning approach in light of the continuum behavior.

Second, there is the interpretation of the error terms. The method for deriving the error terms that alter the variance dynamics in P21 bypasses the underlying issues that cause the observed errors. Though the same error terms noted by P21 are recovered through the analysis performed in this chapter, our approach shows that these terms are a result of using off-diagonal elements to approximate diagonal elements, which is particularly problematic due to the discontinuous change in the continuum dynamics along $x_1 = x_2$. Therefore, the errors we observe are not typical numerical discretization errors; they are errors that are caused by the continuum dynamics and inherent behavior of Eq. (1.3), independent of the numerical discretization. Including the second-order error terms along the diagonal is essential to see that inaccurate variance propagation is the problem, not dissipation. Only retaining the first-order error terms can be misleading, suggesting that numerical diffusion is the source of the problem, as is claimed by P21. Though numerical diffusion may dominate for first-order schemes, this is not the case in higher-order schemes.

The fact that P21 recovers the same first-order error terms along the covariance matrix diagonal further confirms that discrete covariance propagation causes variance loss in advective systems.

However, P21 does not address the underlying issue causing this loss of variance, namely the inaccurate variance propagation. To understand the full scope of the problem requires knowledge of the continuum covariance dynamics and careful error analysis of discrete propagation.

5.5 Conclusions and Discussion

In this chapter, we see that standard methods of covariance propagation, i.e. Eq. (5.3), are a direct cause of inaccurate variance propagation in advective systems. This inaccurate behavior is not due to numerical dissipation, but rather is due to the discontinuous change in the continuum covariance dynamics along $x_1 = x_2$ coupled with approximating the diagonal of the covariance with off-diagonal elements. The result is that covariance propagation approximates the incorrect continuum dynamics along the covariance diagonal as the correlation length approaches grid scale. Hence, the spurious loss (and gain) of variance observed during covariance propagation in Ménard et al. (2000); Ménard and Chang (2000); Lyster et al. (2004); Ménard et al. (2021), for example, is a manifestation of the inaccurate variance propagation explicitly derived in this chapter. Since ensemble-based schemes implement low-rank approximations of the full-rank propagation, the issues presented here will still be present in addition to the problems that arise with low-rank approximations.

The terms in the approximated continuum dynamics that differentiate across the covariance diagonal (P_2 for example) are the driving cause of the inaccurate variance propagation. We see in Eqs. (5.18), (5.20), (5.37), and (5.38) how these derivative terms directly influence the approximated continuum dynamics as they become large for fixed grid length. It's important to note that the error terms $H_u, H_c, \hat{H}_u$, and $\hat{H}_c$ in these equations contain higher order derivatives that differentiate across the covariance diagonal. As these derivatives become large for fixed grid length, they too can be come significant enough to alter the approximated continuum dynamics even further.

Numerical schemes rely on differentiability for stability, convergence to the exact solution, and to determine the scheme's order of accuracy. As gradients along the covariance diagonal become sharp, for example when correlation lengths become small, differentiability across the covariance

diagonal is lost relative to the grid resolution. This loss of differentiability results in large (and eventually unbounded) derivatives across the diagonal, causing the error terms derived in Eqs. (5.18), (5.20), (5.37), and (5.38) to become large and change the approximated continuum dynamics. We see in this chapter the consequences of propagating covariances according to Eq. (1.3) when gradients near the covariance diagonal become sharp. The question is whether this behavior can be corrected now that we have an explicit equation for the approximated dynamics; this is the subject of the next chapter.

Chapter 6

Developing a Correction to Full-Rank Propagation

The original motivation for the error analysis in Ch. 5 was to derive a correction to full-rank covariance propagation in an effort to mitigate the observed inaccurate variance propagation. The advantage of correcting full-rank covariance propagation is that it may be more readily adapted to current data assimilation algorithms. Low-rank approximations of Eq. (1.3) obtained by ensemble-based schemes are commonplace in today's data assimilation schemes; therefore correcting the full-rank propagation can in turn mitigate the variance loss caused by inaccurate variance propagation in the low-rank approximations.

The approach for determining a correction to full-rank covariance propagation presented in this chapter differs from current approaches in that it is derived directly from the continuum covariance dynamics. Recall that error analysis in Ch. 5 resulted in a several error terms that involve various derivatives of the covariance across the diagonal $x_1 = x_2$. These error terms are ideal candidates to add back to the diagonal of the covariance during full-rank propagation and (potentially) mitigate the inaccurate variance propagation. Since these error terms are common to both the upwind and centered difference numerical schemes and when the differentiability of the covariance across $x_1 = x_2$ changes, developing a correction using these error terms could likely be adapted to other numerical schemes. In this chapter, we will explore the idea of building a correction from the error terms accumulated from the error analysis in Ch. 5.

6.1 Correcting Full-Rank Covariance Propagation

We observed in Ch. 5.3 that averaging across the diagonal can change the dynamics the numerical schemes are approximating along the covariance diagonal, depending on the magnitude of correlation lengths or partial derivatives. Both the upwind and centered difference spatial discretizations, though different discretizations, produce similar error terms in their approximated dynamics. This suggests that the inaccurate variance propagation caused by Eq. (5.3) is fundamentally an issue of dynamics and may not be readily fixed by increasing spatial resolution or changing the numerical scheme. The generality of these results are also emphasized by the fact that they are derived from the semi-discretizations of the generalized advection equation, Eq. (1.4).

Both numerical schemes include averaging terms that are second-order (first-order in the discontinuous case) approximations of diagonal elements evaluated on the half-grid. Taylor series expansions we have computed in Ch. 5, in fact, reveal higher-order approximations of these diagonal elements, given sufficient differentiability. For example, consider the case where the covariance is not differentiable across $x_1 = x_2$ but is at least three times continuously differentiable away from $x_1 = x_2$. We can expand the average approximation across the diagonal about the diagonal element on the half-grid,

$$P(x_i, x_{i+1}, t)+P(x_{i+1}, x_i, t) = 2P(x_{i+1/2}, x_{i+1/2}, t)+\Delta x P_1^+(x_{i+1/2}, t)+\frac{\Delta x^2}{4}P_2^+(x_{i+1/2}, t)+\mathcal{O}(\Delta x^3). \tag{6.1}$$

We can rewrite the above equation to approximate the diagonal quantity $2P(x_{i+1/2}, x_{i+1/2}, t)$,

$$2P(x_{i+1/2}, x_{i+1/2}, t) = P(x_i, x_{i+1}, t)+P(x_{i+1}, x_i, t)-\Delta x P_1^+(x_{i+1/2}, t)-\frac{\Delta x^2}{4}P_2^+(x_{i+1/2}, t)+\mathcal{O}(\Delta x^3). \tag{6.2}$$

If the average term $P(x_i, x_{i+1}, t) + P(x_{i+1}, x_i, t)$ poorly approximates the diagonal quantity $2P(x_{i+1/2}, x_{i+1/2}, t)$ when correlation lengths become relatively small, including higher-order terms like P_1^+ and P_2^+ may be a viable option to try and correct the diagonal dynamics. For the state dynamics in Eq. (1.4), we can derive explicit PDEs for P_1^+ and P_2^+. This was first discussed in Sec. 4 of Cohn (1993); here, the main points from this work will be presented here in the context

this thesis.

Through a change of coordinates, we can more concisely define P_1^+ and P_2^+ as follows. First, we define

$$x = \frac{1}{2}(x_1 + x_2), \quad \xi = \frac{1}{2}(x_1 - x_2), \tag{6.3}$$

with

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}, \quad \frac{\partial}{\partial \xi} = \frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}. \tag{6.4}$$

Now define a function $\hat{P}$ that satisfies the following relation with the covariance P in the original space,

$$\hat{P}(x, \xi, t) = P(x + \xi, x - \xi, t). \tag{6.5}$$

The dynamics along $x_1 = x_2$ fall naturally out of this coordinate system, where $x_1 = x_2$ implies $x = x_1 = x_2$ and $\xi = 0$. We can then express P_1^+ and P_2^+ in terms of the partial derivatives of $\hat{P}$ with respect to ξ evaluated at $\xi = 0$, i.e. along $x_1 = x_2$,

$$P_n(x, t) = \left. \frac{\partial^n \hat{P}(x, \xi, t)}{\partial \xi^n} \right|_{\xi=0}, \tag{6.6}$$

$$P_n^-(x, t) = \lim_{\xi \to 0^-} \left. \frac{\partial^n \hat{P}(x, \xi, t)}{\partial \xi^n} \right|_{\xi=0}, \tag{6.7}$$

$$P_n^+(x, t) = \lim_{\xi \to 0^+} \left. \frac{\partial^n \hat{P}(x, \xi, t)}{\partial \xi^n} \right|_{\xi=0}, \quad n = 1, 2, \ldots \tag{6.8}$$

The derivation of the PDEs that govern P_1^+ and P_2^+ are given in Sec. 4d, 4a of Cohn (1993), respectively, and we suggest the reader refer to this work for further details. For the state dynamics defined in Eq. (1.4), the PDEs for $P_1^+ = P_1$ and $P_2^+ = P_2$ are given below.

$$\begin{gathered} (P_1)_t + v(P_1)_x + (v_x + 2b)P_1 = 0, \\ P_1(x, t_0) = (P_1)_0(x) \end{gathered} \tag{6.9}$$

$$\begin{gathered} (P_2)_t + v(P_2)_x + 2(v_x + b)P_2 + v_{xxx}\sigma_x^2 + 2b_{xx}\sigma^2 = 0, \\ P_2(x, t_0) = (P_2)_0(x) \end{gathered} \tag{6.10}$$

The equation for P_2 is coupled with the variance equation, but remains a PDE in the same number of space dimensions as the state.

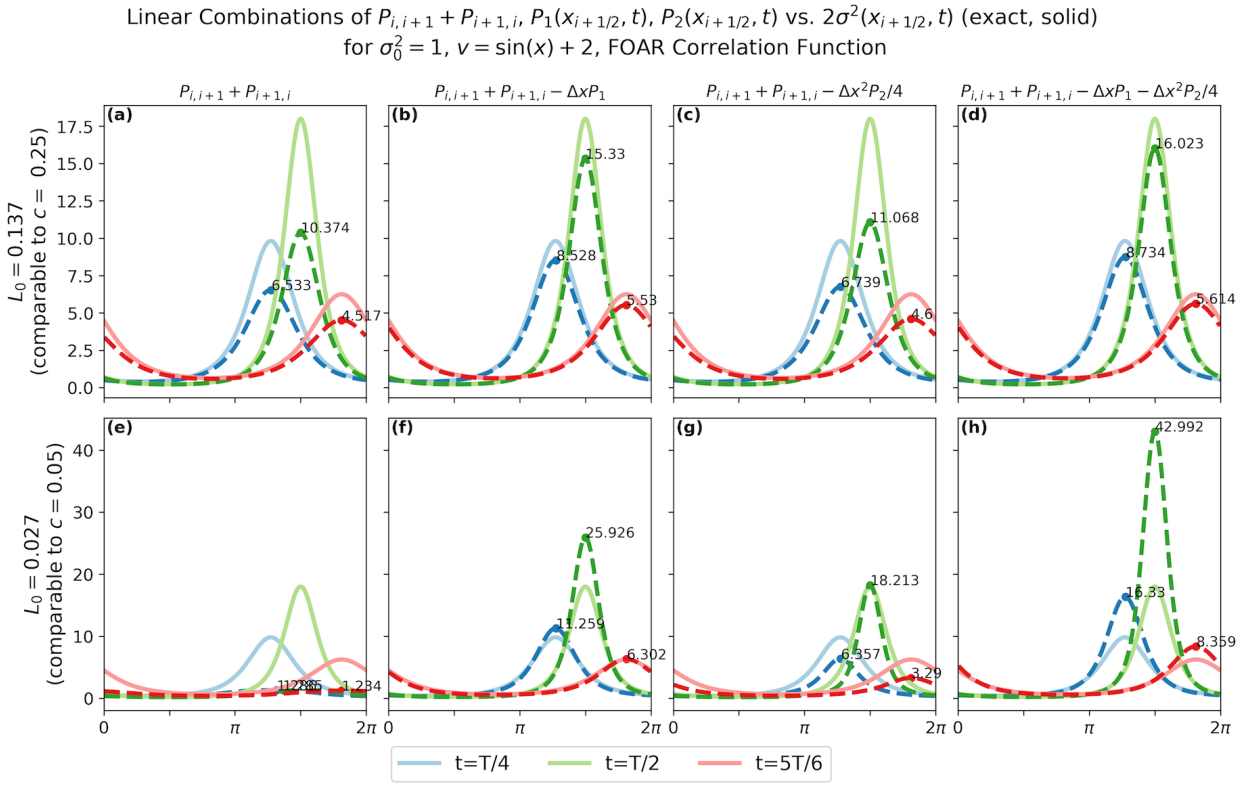

Figure 6.1: Comparisons of different approximations of $2P(x_{i+1/2}, x_{i+1/2}, t)$ based on Eq. (6.2). The solid lines in each panel are the exact solutions $2P(x_{i+1/2}, x_{i+1/2}, t) = 2\sigma^2(x_{i+1/2}, t)$ at three different times and are the same in each panel. Dashed curves indicated the different approximations at each time as indicated in the title of each column. Numerical values of the maximums are marked for reference. Each row corresponds to two different initial correlation lengths used to construct the initial covariances, which use the FOAR function defined in Eq. (3.6). Values of L_0 are chosen to be comparable to values of cut-off length c shown in Fig. 5.1.

With explicit PDEs for P_1 and P_2 in Eqs. (6.9) and (6.10), these terms can be computed, either analytically or numerically, and used in Eq. (6.2) to increase the order of the approximation the diagonal element. We can gauge how P_1 and P_2 can be used to improve the approximation of $2P(x_{i+1/2}, x_{i+1/2}, t)$ by considering the example in Fig. 6.1. In Fig. 6.1, we compare $2P(x_{i+1/2}, x_{i+1/2}, t) = 2\sigma^2(x_{i+1/2}, t)$ with several variations of approximations in Eq. (6.2) for states governed by the one-dimensional continuity equation (Eq. B.9 with $\alpha = 1$). All quantities in Fig. 6.1 are exact solutions (see Appen. B.2 for explicit solutions), therefore these results represent the best-case scenario of approximating the diagonal elements. Initial covariances are constructed with a constant initial variance of one and correlations are constructed using the First-Order Autoregressive (FOAR) correlation function, Eq. (3.6). We choose the FOAR correlation function because its derivatives are discontinuous across $x_1 = x_2$ due to the cusp along $x_1 = x_2$ so that its P_1^+ is nonzero.

Considering the first row of Fig. 6.1, which is the case for a modest initial correlation length, adding a correction either with just $-\Delta x P_1$ in (b), $-\frac{\Delta x^2}{4} P_2$ in (c), or both in (d), improves the approximation of twice the variance (solid lines) seen in panel (a), which is what the upwind and centered difference schemes are currently implementing. It is important to note here that the values used to compute the dashed-line approximations are exact solutions because these solutions are available. Therefore, these results in Fig. 6.1 represent the best-case scenario of the proposed corrections.

The impact of the various correction terms change drastically in the bottom row of Fig. 6.1, which considers an initial covariance with a short initial correlation length. Panel (e), which is the averaging term without any correction, is a very poor approximation of twice the variance. Interestingly, panels (f) and (h) which include the $-\Delta x P_1$ correction term, are over-approximations, with the most severe over-approximation in panel (h). Panel (g), which only contains the $-\frac{\Delta x^2}{4} P_2$ correction term, however, is a better correction in terms of not over-approximating twice the variance, but the shapes are not ideal.

The results in Fig. 6.1 suggest that using the P_1 and P_2 terms that appear during the Taylor

series expansions as corrections to the diagonal propagation may be helpful in some cases, but would be of limited use in others. Determining the type of correction that needs to be implemented may depend on space, time, grid length, the dynamics of the system, and the correlation structure of the covariance. Recall that the error terms $H_u, H_c, \hat{H}_u$, and $\hat{H}_c$ also contain derivatives of the covariance across the diagonal. These error terms can become significant as well, making a correction to Eq. (5.3) in this way even more difficult. In addition, developing an associated correction to the full covariance propagation defined in Eq. (5.3), is not straightforward, particularly to ensure positive semi-definiteness is maintained and correlation structures are not negatively impacted. The correction proposed here is not the only method of correcting the full-rank covariance propagation, though we suspect developing a correction may be quite complex.

6.2 Discussion of Ménard et al. (2021)

The results of Sec. 6.1 suggest that correcting the covariance diagonal during full-rank propagation is not straightforward due to the structure of the error terms. There may be ways to correct the covariance diagonal other than that presented in Sec. 6.1, though it is likely that any would be difficult to formulate because of the discontinuous change in the continuum covariance dynamics identified in Ch. 2.

A correction to full-rank propagation, however, has been identified by Ménard et al. (2021) (hereafter referred to as M21 in this section), resulting in successful mitigation of variance loss in their experiments. Closer inspection in the context of this work, however, suggests that the success of the correction proposed in M21 is limited to their problem and not readily adapted to other numerical schemes or dynamics. In this section, we will consider the correction presented in M21, first by reprising the derivation of their correction and then discussing this correction in the context of this thesis.

The state dynamics in M21 satisfy the advection equation ($b = 0$ in Eq. 1.4). In their Sec. 2.2, the authors discretize the state in both space and time using a first-order semi-Lagrangian discretization scheme with linear interpolation. From the full discretization of the state, the au-

thors determine the associated discretization for the covariance diagonal (the equation immediately following their Eq. 22). Their equation can be derived from Eq. (C.23) by applying a first-order Euler time-integration scheme to the first-order upwind semi-discretization.

The analysis and derivation of a correction following the second equation after Eq. (22) in M21 deviates from that presented in this work as follows. Rather than determining the approximated continuum dynamics along the covariance diagonal through Taylor expansion, the authors add and subtract the diagonal element $\alpha_i^n P_{i-1,i-1}^n$, where $\alpha_i^n = v_i^n \Delta t/\Delta x$, the superscripts n denote the time index, and the subscripts i the spatial index. As a consequence, the authors obtain in their Eq. (23): a first-order semi-Lagrangian discretization scheme with linear interpolation for the covariance diagonal, plus additional terms the authors denote as "HOT." Their Eq. (23) is reprised here using the notation of Ch. 5, with the boxed terms indicating the added and subtracted terms that are not a part of the original diagonal propagation:

$$P_{i,i}^{n+1} = (1-\alpha_i^n)P_{i,i}^n \boxed{+\alpha_i^n P_{i-1,i-1}^n} \tag{6.11}$$

$$\underbrace{-\alpha_i^n P_{i,i}^n + (\alpha_i^n)^2 P_{i,i}^n + \alpha_i^n(1-\alpha_i^n)\big(P_{i,i-1}^n + P_{i-1,i}^n\big)\boxed{-\alpha_i^n P_{i-1,i-1}^n} + (\alpha_i^n)^2 P_{i-1,i-1}^n}_{HOT}. \tag{6.12}$$

Therefore, their Eq. (23) recovers a numerical scheme that is a first-order approximation of the continuum variance dynamics, plus remaining terms they denote as "HOT". From these "HOT" the authors derive a correction, which will be addressed later.

There are a few observations to be made about Eqs. (6.11) – (6.12). By adding and subtracting the boxed term $\alpha_i^n P_{i-1,i-1}^n$, a first-order semi-Lagrangian scheme is recovered along the covariance diagonal. The continuum variance in this case does satisfy the advection equation, therefore the first-order semi-Lagrangian scheme approximates the continuum variance dynamics (up to first order). The "HOT" in Eq. (6.12), therefore, are not necessarily higher-order terms in the sense that they contain first-order terms in Δx and Δt, which are the same order as the discretization scheme. However, isolation of these "HOT" terms allows for recovery of the first order semi-Lagrangian scheme for the variance along the covariance matrix diagonal.

The correction proposed in Sec. 6 of M21 is defined by adding the "HOT" terms back to the

covariance diagonal during full-rank propagation, defined by their Eq. (36). Their results Sec. 6.1 – 6.3 suggest that this correction mitigates the majority of the variance loss caused by the discrete covariance propagation. The proposed correction in their Eq. (36) effectively enforces a first-order approximation of the variance along the diagonal. Therefore, it is expected to see first-order accurate propagation of the variance. The correction to the diagonal also impacts an important property of the correlations due to the particular dynamics for their case. For states governed by the advection equation, both the correlation C and covariance P satisfy the two-dimensional version of the advection equation (i.e. Eq. 1.6 with $b_1 = b_2 = 0$). Thus, the correlation C satisfies the same full discretization as the covariance, and its diagonal the same discretization defined in Eqs. (6.11) – (6.12). It follows that in the context of the correlation propagation, their correction term (i.e., adding back their "HOT" defined by their Eq. 35), will ensure that one is preserved along the diagonal of the correlation during propagation.

In the context of the analysis presented in Ch. 5 and Appen. C, the correction term proposed by M21 to the first-order scheme mitigates the dissipation errors caused by the upwind discretization. In Ch. 5, we see this in the $\Delta x/L^2$ term that appears in the upwind scheme. This error term, however, is not present in the centered difference scheme, as the centered difference scheme is not dissipative. Therefore, a correction derived using the approach presented in M21 will not address the errors observed for the centered difference scheme. Recall that the error terms we derive in Ch. 5 and Appen. C involve powers of the ratio $\Delta x/L$, and as this ratio approaches unity, there are several terms that become significantly problematic, completely changing the approximated continuum dynamics (while also introducing a zeroth-order dissipative term in the upwind case). Therefore, the correction of M21 may work for first-order schemes by mitigating dissipative error term, but will not be able to generalize to higher-order schemes, as the analysis presented here demonstrates.

6.3 Conclusion and Discussion

Current data assimilation schemes implement covariance propagation defined by Eq. (1.3), whether it be explicitly, implicitly, or approximated at low rank as in ensemble-based schemes. Therefore, correcting full-rank propagation to mitigate the inaccurate variance propagation observed in advective systems is a logical choice. We see in this chapter, however, that correcting full-rank propagation may be possible, but is likely limited due to the structure of the error terms. The results in Fig. 6.1 illustrate in a simple example that using higher-order derivatives of the covariance across its diagonal, which are derived from the error analysis, do not necessarily produce more accurate covariance diagonals. Additional terms from the Taylor expansion may be necessary, and the resulting correction may need to depend on both space and time. Though these error terms are derived directly from the continuum, it is unlikely that formulating a correction using these terms would be tractable in practice.

The approach taken in this chapter to derive a correction is just one method of correcting the inaccurate variance propagation caused by Eq. (1.3). However, the continuum analysis of Ch. 2 and error analysis of Ch. 5, together with this chapter, question whether standard methods of covariance propagation can accurately propagate the variance in a manner suited for data assimilation applications. Recall that data assimilation is typically practiced in high-dimensional geophysical applications. The high dimensionality of data assimilation limits the choices for numerical schemes simply due to computational costs. Thus, developing a correction to full-rank propagation that is practical in a data assimilation framework adds an additional challenge.

In the next chapter, we will consider an alternative method of covariance propagation that bypasses Eq. (1.3) altogether that is computationally tractable and derived directly from the continuum covariance dynamics.

Chapter 7

An Alternative Approach to Standard Covariance Propagation

The discontinuous change in the continuum covariance dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane as correlation lengths tend to zero is the cause of inaccurate variance propagation associated with Eq. (1.3) for advective dynamics. We see the severity of the inaccurate variance propagation in the simple numerical experiments of Chs. 3 and 4, and derive the explicit cause of the inaccurate propagation through the error analysis in Ch. 5. Correcting full-rank covariance propagation, as discussed in Ch. 6, may be possible in certain situations, but in general, it can be quite complicated due to the structure of the error terms. Current methods for propagating covariances have not changed significantly since the 1990s, therefore deriving an alternative method of covariance propagation that bypasses Eq. (1.3) entirely is relevant and timely.

Section 4 of Cohn (1993) introduces the theoretical foundation for an alternative method of covariance propagation that can directly address inaccurate variance propagation. First introduced as "local covariance evolution" (LCE), for spatial variables $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$, this method derives a system of PDEs in n space dimensions from the governing dynamics that locally approximates the covariance through the variance and correlation, rather than directly solving the covariance evolution equation in $2n$ space dimensions. Atmospheric, ocean, and other large-scale data assimilation algorithms already solve PDEs in n space dimensions when evolving the model state, therefore LCE is tractable in high-dimensional contexts. In particular, this method mitigates the inaccurate variance propagation caused by Eq. (1.3) by evolving the variance directly.

To implement LCE in a data assimilation schemes requires reconstructing the full covariance

P from the evolved local quantities, as the covariance P provides critical information for the measurement update. This is where my thesis extends the theory introduced by Cohn (1993), in which parametric correlation functions are used to approximate the full correlation fields. By relating the local correlation quantities governed by the derived system of PDEs in n space dimensions to the parameters of the correlation function, we can then approximate the correlation field. Rescaling the approximated correlations with the evolved variance then reconstructs the full covariance P.

Local covariance evolution differs from standard methods of covariance propagation (i.e., Eq. 1.3 and its reduced-rank approximations) in that it is derived from the continuum covariance dynamics. In this chapter, we will derive the LCE equations for the one-dimensional version of Eq. (1.4) for three types of parametric correlation functions: continuously differentiable at $x_1 = x_2$, discontinuous derivatives at $x_1 = x_2$, and the Hadamard product of two correlation functions. The first two cases are discussed in Sec. 4 of Cohn (1993), where the Hadamard product case is novel to this work. In this chapter, we will reprise and expand upon that presented in Sec. 4 of Cohn (1993) in the context of this thesis.

7.1 Derivation of the LCE Equations in One Space Dimension

Consider states q that satisfy the generalized advective dynamics of Eq. (1.4) on the circle of radius $r > 0$, $\mathbb{S}^1_r$. Its corresponding covariance P satisfies Eq. (1.6) for $x_1, x_2 \in \mathbb{S}^1_r$ and can be decomposed into the product of the standard deviation σ and correlation C,

$$P(x_1, x_2, t) = \sigma(x_1, t) C(x_1, x_2, t) \sigma(x_2, t). \tag{7.1}$$

Using this decomposition and the covariance PDE, Eq. (1.6), the standard deviation σ satisfies the same dynamics as the state q (i.e., Eq. 1.4 for σ). Thus, the variance σ^2 satisfies Eq. (1.8), which serves as the first equation needed for LCE.

The correlation function C satisfies the advection equation two space dimensions,

$$\begin{aligned} &C_t + v_1 C_{x_1} + v_2 C_{x_2} = 0, \\ &C(x_1, x_2, t_0) = C_0(x_1, x_2). \end{aligned} \tag{7.2}$$

The goal now is to derive a system of PDEs, each in one space dimension in this case, that locally approximates the correlation C. This local approximation is achieved by expanding the correlation C about zero separation, $x_1 = x_2$, similar to the Taylor expansions performed in Ch. 5. To streamline this expansion, we will use the following change of coordinates (which was first introduced in Ch. 6),

$$x = \frac{1}{2}(x_1 + x_2), \quad \xi = \frac{1}{2}(x_1 - x_2). \tag{6.3 revisited}$$

Using this coordinate transformation, we can define a new function $\tilde{C}$ that relates to C as follows,

$$\tilde{C}(x, \xi, t) = \tilde{C}(x_1(x, \xi), x_2(x, \xi), t) = C(x + \xi, x - \xi, t). \tag{7.3}$$

From Eq. (7.2) and (7.3), we have a PDE for $\tilde{C}$,

$$\begin{aligned} \tilde{C}_t + \frac{1}{2}\big[v(x + \xi, t) + v(x - \xi, t)\big]\tilde{C}_x + \frac{1}{2}\big[v(x + \xi, t) - v(x + \xi, t)\big]\tilde{C}_\xi = 0, \\ \tilde{C}(x, \xi, t_0) = \tilde{C}_0(x, \xi) = C_0(x + \xi, x - \xi). \end{aligned} \tag{7.4}$$

To obtain the system of PDEs that locally approximates the correlation C, we expand $\tilde{C}$ about zero separation, $\xi = 0$. Like the error analysis in Ch. 5, we will consider two cases for the differentiability of $\tilde{C}$ at zero separation. First assume that $\tilde{C}$ is at least four times continuously differentiable at $\xi = 0$. Then, consider the case where $\tilde{C}$ is not continuously differentiable at $\xi = 0$, but is four times continuously differentiable away from $\xi = 0$. In both cases, we will assume the velocity field v is four times continuously differentiable in its spatial argument.

Remark 2. *The derivation of the LCE system of PDEs can be done in the original space by expanding the correlation C about $x_1 = x_2$. This approach, however, is quite messy, as seen in the Taylor expansions derived in Ch. 5, but can be shown to be equivalent to the derivation presented in the next two sections.*

7.1.1 Continuously Differentiable at Zero Separation

Suppose the correlation C is at least four times continuously differentiable in both spatial variables x_1 and x_2. Therefore, $\tilde{C}$ is also at least four times continuously differentiable in its spatial

variables x and ξ. We can define the following derivatives of $\tilde{C}$ with respect to ξ which we will use during the Taylor expansions,

$$C_k(x,t) = \frac{\partial^k \tilde{C}}{\partial \xi^k}(x,\xi,t)\bigg|_{\xi=0}, \quad k = 1,2,\ldots,K \geq 4. \tag{7.5}$$

Using Eq. (7.5), we can expand the transformed correlation $\tilde{C}$ about $\xi = 0$. The expansion of $\tilde{C}$ about $\xi = 0$ only contains even orders in ξ since $\tilde{C}$ is an even function (as a consequence of C being symmetric). By also expanding the velocities about $\xi = 0$ and substituting each of these expansions into Eq. (7.4), we can now collect terms in like powers of ξ. The zeroth-order term satisfies $\tilde{C}(x,0,t) = 1$ by definition of the correlation. The second-order terms in ξ yields the following PDE for C_2,

$$\begin{aligned} (C_2)_t + v(C_2)_x + 2v_x C_2 &= 0, \\ C_2(x,t_0) &= (C_2)_0(x). \end{aligned} \tag{7.6}$$

Observe that C_2 is the same quantity defined in Ch. 5 by rewriting the derivatives with respect to x and ξ in terms of x_1 and x_2. Equations for higher-order derivatives in $\tilde{C}$ can be derived as well (see Cohn, 1993, Sec. 4b for further discussion), yielding the system of PDEs in one space dimension that locally approximates the correlation C.

To reconstruct the full correlation fields (Ch. 9 will provide further discussion), C_2 in Eq. (7.6) may provide sufficient information for the chosen parametric correlation function. Alternatively, C_2 can be used to define the correlation length $L(x,t)$ as we saw in Ch. 5,

$$L^2(x,t) = -\frac{1}{C_2(x,t)}, \tag{5.25 revisited}$$

which satisfies its own PDE,

$$\begin{aligned} L_t + vL_x - v_x L &= 0, \\ L(x,t_0) &= L_0(x). \end{aligned} \tag{5.30 revisited}$$

The correlation length L is a parameter often used to describe correlation functions (e.g., Daley, 1991, p. 117), therefore the correlation length can be evolved instead of C_2. Chapter 9 will describe

how to use the evolved correlation length fields to reconstruct correlations using a parametric correlation function.

7.1.2 Discontinuous Derivatives At Zero Separation

In certain geophysical applications, correlations can have cusp-like structures or sharp gradients near $x_1 = x_2$, and are therefore no longer continuously differentiable at zero separation (e.g., Ménard et al., 2000; Ménard and Chang, 2000; Waller et al., 2016). Thus, a new system of PDEs must be derived to approximate the correlations locally for LCE. This derivation combines the approach presented in the previous section with aspects of the error analysis in Ch. 5 for the case where derivatives are discontinuous at $x_1 = x_2$.

Now suppose the correlation C is not continuously differentiable at $x_1 = x_2$, but is at least four times continuously differentiable away from $x_1 = x_2$. Though the quantities C_k defined by Eq. (7.5) do not exist, we can define their corresponding left and right derivatives,

$$C_k^{\pm}(x,t) = \lim_{\xi \to 0^{\pm}} \frac{\partial^k \tilde{C}}{\partial \xi^k}(x,\xi,t), \quad k = 1,2,\ldots,K \geq 4. \tag{7.7}$$

From the definition of $\tilde{C}$, we have the following relationships for the even and odd derivatives,

$$C_{2k}^{+}(x,t) = C_{2k}^{-}(x,t), \quad C_{2k+1}^{+}(x,t) = -C_{2k+1}^{-}(x,t), \quad k = 1,2,\ldots,K. \tag{7.8}$$

Using Eqs. (7.7) and (7.8), we can proceed with the expansion of $\tilde{C}$. Due to the symmetry property in Eq. (7.8), we can combine the expansions from the left and right of $\xi = 0$ into a single expression,

$$\tilde{C}(x,0,t) = 1 + |\xi| C_1(x,t) + \frac{1}{2}|\xi|^2 C_2(x,t) + \ldots, \tag{7.9}$$

where $C_k = C_k^{+}$ for $k = 1,2,\ldots,K$. Unlike the continuously differentiable case, the odd-order derivatives of $\tilde{C}$ remain in its Taylor expansion.

Following the same procedure described in Sec. 7.1.1 using Eq. (7.9) and collecting like powers of ξ, we obtain a new system of PDEs that locally approximates the correlation. The PDEs for the even-order derivatives remain the same as in Sec. 7.1.1. Since the odd-order derivatives of the

correlation $C_k^{\pm}$ are nonzero, they satisfy PDEs as well. The PDE for $C_1 = C_1^+ = -C_1^-$, for example, is

$$(C_1)_t + v(C_1)_x + v_x C_1 = 0,$$
$$C_1(x, t_0) = (C_1)_0(x). \tag{7.10}$$

Equations for the higher-order odd derivatives can also be derived (see Sec. 4d of Cohn, 1993, applied to Eq. 7.2), where the PDEs for odd-order derivatives greater than one are coupled with the previous odd-order derivative and are uncoupled from the even-order derivatives.

7.1.3 Hadamard Product

The theory of the previous two sections can be extended to the case where the initial correlation C_0 is the product of two correlation functions. In discrete space, this corresponds to the Hadamard product of two correlation matrices. Constructing correlations from the Hadamard product of two correlation functions is useful since the resulting product remains positive semi-definite and inherits features from both functions. In this section, we will derive the equations to implement LCE for this case, which is novel to this work.

Suppose that the initial correlation C_0 is a product of two correlation functions,

$$C_0(x_1, x_2) = F_0(x_1, x_2) G_0(x_1, x_2). \tag{7.11}$$

From the correlation dynamics in Eq. (7.2), correlations are simply advected along the characteristics. Thus, solutions to Eq. (7.2) are of the form

$$C(x_1, x_2, t) = C_0\big(s(x_1, t), s(x_2, t)\big), \tag{7.12}$$

following the notation of Appen. B.1. If the initial correlation is defined by Eq. (7.11), we have

$$C(x_1, x_2, t) = F_0\big(s(x_1, t), s(x_2, t)\big) G_0\big(s(x_1, t), s(x_2, t)\big) = F(x_1, x_2, t) G(x_1, x_2, t), \tag{7.13}$$

and the product is preserved with space and time. Equation (7.13) implies that both F and G satisfy the advection equation in two space dimensions, Eq. (7.2).

Since the correlation C remains the product of two correlations F and G, this implies that F and G can be evolved independently, then at a fixed time t their solutions can be used to reconstruct the correlation C according to Eq. (7.13). For LCE, the equations derived in Secs. 7.1.1 and 7.1.2 define the dynamics that can be used to approximate the solutions F and G depending on the differentiability of F_0 and G_0 at $x_1 = x_2$.

7.2 Reconstructing Correlations with Parametric Correlation Functions: Introduction

The final step of LCE is to reconstruct the full covariance P, first by reconstructing the correlations from the evolved local quantities such as C_1 or C_2, then rescaling by the evolved variance. We will accomplish this by using parametric correlation functions. In general, parametric correlation functions are functions of spatial variables x_1 and x_2 (often as a normed difference $||x_1 - x_2||$) and a finite number of parameters that influence the shape and behavior of the correlations. The implementation of LCE will relate the evolved local quantities derived in Secs. 7.1.1 – 7.1.2 to the parameters of the correlation function chosen to the approximate the full correlation field.

The specific implementation of LCE depends on the particular parametric correlation function. Since we are evolving local correlation quantities in space and time and then relating these quantities to the parameters of the correlation function, we need to choose parametric correlation functions that will allow its parameters to vary. Gaspari et al. (2006) gives several examples of correlation functions that can work in this context. Chapter 9 will discuss how to implement LCE with Eq. (23) of Gaspari et al. (2006).

One parametric correlation function commonly used in data assimilation is the Gaspari-Cohn (1999) compactly-supported approximation to a Gaussian; in fact this function was used to construct the initial correlations for the numerical experiments in Chs. 3 and 4 (abbreviated in those chapters as GC). Gaspari and Cohn (1999) introduce a family of compactly-supported, piecewise rational correlation functions in their Sec. 4(c) that depend on two parameters a and c. The parameter c, as we have seen before, is the cut-off length that determines the region of

compact support. The parameter a influences the shape of the correlation, where $a = 1/2$ yields the compactly-supported approximation to a Gaussian. For fixed parameters a and c, the Gaspari-Cohn correlation function has as a single, fixed correlation length L that depends on both a and c. To use this function for LCE, however, requires a Gaspari-Cohn-like correlation function that allows the correlation length L to vary. This motivates the next chapter, in which we will derive a generalization of the Gaspari-Cohn correlation function to allow for variable correlation length fields with the intention to use this correlation function later for LCE.

Remark 3. *The Parametric Kalman Filter (PKF), first introduced by Pannekoucke et al. (2016), is similar to, but distinct from, local covariance evolution (LCE). In general, the PKF is a data assimilation filter that updates and evolves covariances that are parameterized by a variance and a local anisotropic tensor (referred to as VLATcov in Pannekoucke et al., 2021). The dynamics for these parameters are derived from the discrete state problem and often make particular assumptions about the variance and structure of the correlations (e.g., "local homogeneity" of the variance and Gaussian-like correlations, see Pannekoucke et al., 2016, 2018, 2021). Local covariance evolution is different from the PKF in that the dynamics that approximate the covariance for LCE are derived in the continuum, directly from the continuum covariance dynamics, with no assumptions on the variance or correlations aside from differentiability. In addition, LCE applies to general classes of correlation functions, where we do not need to make any assumptions on the correlation structure. Though the PKF and LCE are related, LCE is arguably the mathematical foundation of the forecast step of the PKF.*

Chapter 8

A Generalized Gaspari-Cohn Correlation Function

In this chapter, we will derive a new, generalized version of the compactly-supported, piecewise rational correlation function introduced by Gaspari and Cohn (1999) and its subsequent generalization by Gaspari et al. (2006). The motivation for this new function is so that it to be used for local covariance evolution, introduced in the previous chapter. However, this new correlation function stands on its own as a unique contribution to both the data assimilation and spatial statistics communities. This work has been published in Gilpin et al. (2023).

8.1 Introduction

Parametric correlation functions play an essential role in data assimilation, where they are used to model covariances given a set of tunable parameters or applied as tapering functions to localize sample covariances in ensemble-based schemes (Daley, 1991; Gaspari and Cohn, 1999; Gneiting, 1999; Houtekamer and Mitchell, 1998, 2001; Hamill et al., 2001; Gaspari et al., 2006). Compactly-supported correlation functions are particularly useful for modeling covariances wherein correlations can be assumed identically zero after some distance while maintaining positive semi-definiteness. In addition, correlation functions that do not have compact support can gain compact support through the Hadamard (element-wise) product with compactly-supported correlation functions, which preserves positive semi-definiteness by the Schur product theorem (Horn and Johnson, 1985, p. 458). Compactly-supported correlation functions are particularly useful in high-dimensional contexts, such as data assimilation and spatial statistics, as the resulting covariances are computationally

tractable due to their sparsity.

One of the most commonly used compactly-supported correlation functions in data assimilation is the Gaspari and Cohn (1999) (hereafter referred to as GC99) piecewise rational, compactly-supported approximation to a Gaussian (their Eq. 4.10), which we will refer to as the GC99 function. This parametric correlation function depends on only two parameters: the tunable cut-off length parameter c, and the parameter a that when taken to be one-half produces correlations with Gaussian-like behavior. For covariance tapering applications such as localization, the GC99 function has been used frequently since it was first applied by Houtekamer and Mitchell (2001) because of its Gaussian-like shape and compact support controlled through the cut-off length parameter c. Though the GC99 function has several attractive features, it is limited by its homogeneity (invariance under translations, GC99, p. 729) and isotropy (invariance under rotations, GC99, p. 729), unless constructed using coordinate stretching (e.g., GC99, pp. 725, 732; Ménard et al., 2016, pp. 882 – 883). The fixed cut-off length parameter c of the GC99 function typically requires special tuning per application, and it has been shown that the optimal cut-off length, sometimes termed cut-off radius or localization radius, depends on the observation type, spatial location, dynamics, and whether localization is applied in model space or observation space (Evensen, 2009, Ch. 15 and references therein).

There is thus a need for compactly-supported correlation functions with more flexible features that can adjust to the situation. Several compactly-supported, parametric correlation functions have been introduced that vary in their flexibility as correlation functions (e.g., Wu, 1995; Wendland, 1995; Gaspari and Cohn, 1999; Gneiting, 1999; Buhmann, 2000; Gneiting, 2002; Gaspari et al., 2006; Porcu et al., 2013; Kleiber and Porcu, 2015; Liang and Marcotte, 2016; Stanley et al., 2021). In particular, the follow-on work to GC99, namely Gaspari et al. (2006) (hereafter referred to as G06), presents a generalization of the GC99 function in which the parameter a can vary while keeping c fixed (see their Eq. 33). This generalization introduces inhomogeneity and anisotropy through variable a. However, like the GC99 function, the G06 function is limited by its fixed cut-off length c.

In this chapter, we further generalize the compactly-supported, piecewise rational correlation functions introduced in GC99 and G06 by allowing *both* the parameters a and c to vary, as functions, over the spatial domain. This correlation function, which we refer to as the Generalized Gaspari-Cohn (GenGC) correlation function, is a compactly-supported parametric correlation function where inhomogeneity and anisotropy are introduced as both a and c vary over space. By now allowing the cut-off length c to vary over the domain, the GenGC correlation function gains much more flexibility in its construction, and in particular is no longer restricted to support regions of fixed spheres. The parameters a and c can be estimated, tuned manually, or derived from dynamics. The added flexibility of GenGC while maintaining positive semi-definiteness and compact support can be useful for covariance modeling and tapering in data assimilation and spatial statistics applications.

We can see the impact of allowing both parameters a and c to vary over some spatial index k by considering the correlation length L_k of GenGC,

$$L_k = c_k \left(\frac{3(22a_k^2 + 3a_k + 1)}{40(8a_k^2 - 2a_k + 1)} \right)^{1/2}. \tag{8.1}$$

The correlation length L_k of GenGC varies spatially as c_k and a_k vary over space, and its linear dependence on variable c_k allows for direct control of the correlation length L_k. In contrast, the correlation length of the GC99 function where $a = 1/2$ and c is fixed, remains constant at $L = c\sqrt{0.3}$. For the G06 function, the correlation length L_k does vary over space as a_k varies, but its dependence on a_k is such that L_k/c_k in Eq. (8.1) is tightly bounded, $0.2297 \leq L_k/c_k \leq 0.5485$, as discussed in Sec. 8.3.2. Therefore, varying c_k yields more control and influence on the correlation length, introducing flexibility that the GC99 and G06 functions lack.

This flexibility gained by allowing the correlation length L_k to vary through both a_k and c_k can be particularly useful when reconstructing correlations from correlation lengths L_k that are evolved dynamically. For example, Cohn (1993, pp. 3136 – 3144) derives explicit PDEs for the correlation length field associated with state dynamics governed by the continuity equation and related PDEs. By evolving the correlation length field according to its governing PDE and

with additional information regarding c_k or a_k, GenGC and Eq. (8.1) can be used to reconstruct full correlation functions, as we discuss in Sec. 8.3.2. Alternatively, correlation lengths L_k can be obtained by estimating the local correlation tensor from an ensemble of states (Michel et al., 2016, and references therein) and in a similar manner used to construct correlations with GenGC.

The layout of this chapter is as follows. In Sec. 8.2, we derive the GenGC correlation function, beginning with definitions and theory in Sec. 8.2.1, followed by its construction in Sec. 8.2.2 using convolutions and brief discussion of its evaluation. Section 8.3 discusses important consequences of allowing both a and c to vary over space. Since the correlations modeled and tapered during data assimilation are generally continuous, it is natural to consider the conditions on a and c such that GenGC is continuous, which is done in Sec. 8.3.1. This is followed by Sec. 8.3.2, which defines the correlation length L_k of GenGC and discusses methods for implementing GenGC given L_k. In Sec. 8.4, we present two sets of examples, with one-dimensional examples in Sec. 8.4.1 and two-dimensional examples in Sec. 8.4.2, to highlight the flexibility gained by varying a and c. Concluding remarks are given in Sec. 8.5, followed by two appendices, Appen. D.1 and D.2, which contain a proof and an example to supplement Sec. 8.2, and the formulas for GenGC, respectively.

8.2 Derivation of the GenGC Correlation Function

The GC99 function (Eq. 4.10 of GC99) is constructed through self-convolution of a compactly-supported, radially symmetric function over Euclidean 3-space, $\mathbb{R}^3$. The G06 function (Eq. 33 of G06) and the Generalized Gaspari-Cohn (GenGC) correlation function in this chapter generalize this approach as follows. Rather than constructing a correlation function using *one* compactly-supported, radially symmetric function h, we build a correlation function by convolving functions from a whole *collection* of m compactly-supported, radially symmetric functions $h_k,\ k = 1, 2, ..., m$ on $\mathbb{R}^3$. Our particular choice of functions h_k is given in Eq. (8.18), though this approach can be generalized to other collections of radially symmetric functions, which we discuss at the beginning of Sec. 8.2.2.

The derivation of the GenGC correlation function is presented in two parts. In Sec. 8.2.1,

we reprise definitions and theory from G06 that establish the theoretical framework in which we can build scalar covariance functions with inhomogeneous and anisotropic features. Using this framework, in Sec. 8.2.2 we construct scalar covariance functions by convolving functions from a collection of radially-symmetric functions $h_k,\ k = 1, 2, ..., m$. With the general approach established at the beginning of Sec. 8.2.2, we then specify the functions h_k and construct the GenGC correlation function, followed by a brief discussion on its evaluation.

8.2.1 Definitions and Theoretical Framework

We begin by revisiting some of the theory established in G06, where particular definitions and discussions from G06 serve as the starting point for this chapter. We consider covariance functions defined on a domain $\Omega \subseteq \mathbb{R}^N$, which we take to be an open, connected set, with smooth boundary $\partial\Omega$ in the case that Ω is not all of $\mathbb{R}^N$, where the general theory applies. In later discussions, we will restrict ourselves to $\mathbb{R}^3$ or to the surface of the sphere with radius $r > 0$ embedded in $\mathbb{R}^3$, which we denote as $\mathbb{S}^2_r$, as needed by applications.

We start by defining covariance functions of *vector* random fields, which will serve as the foundation for building inhomogeneous and anisotropic covariance functions.

Definition 1. *(Def. 2.1 of G06) The multi-level covariance function of a vector random field*

$$\mathbf{X}(\boldsymbol{r}) := \big[X_1(\boldsymbol{r}), X_2(\boldsymbol{r}), ..., X_m(\boldsymbol{r})\big], \quad \boldsymbol{r} \in \Omega, \tag{8.2}$$

is the $m \times m$ matrix function

$$\mathbf{B}(\boldsymbol{r}, \boldsymbol{s}) := \big\langle \mathbf{Y}(\boldsymbol{r})\mathbf{Y}(\boldsymbol{s})^T \big\rangle, \quad \boldsymbol{r}, \boldsymbol{s} \in \Omega, \tag{8.3}$$

where

$$\mathbf{Y}(\boldsymbol{r}) := \big[Y_1(\boldsymbol{r}), Y_2(\boldsymbol{r}), ..., Y_m(\boldsymbol{r})\big]^T,\ Y_k(\boldsymbol{r}) := X_k(\boldsymbol{r}) - \big\langle X_k(\boldsymbol{r}) \big\rangle, \tag{8.4}$$

where $X_k(\boldsymbol{r}),\ k = 1, 2, ..., m$, are real-valued random fields on Ω, and where $\langle\cdot\rangle$ denotes mathematical expectation.

The covariance defined in Eq. (8.3) is a *matrix* function, where we use boldface, e.g., $\mathbf{B}$, to denote matrices. We use the term "multi-level," as done in G06, to give an intuitive sense of how Eq. (8.3) will be used to generalize single-level, scalar covariance functions. As discussed in G06 (p. 1816), this terminology does not capture the full generality of this and the following definition, but it is useful for our purposes.

The matrix structure of covariance functions defined in Eq. (8.3) becomes more explicit in the following definition, which is equivalent to Def. 1 in the case of continuous covariance functions.

Definition 2. *(Def. 2.2 of G06) The $m \times m$ matrix of functions*

$$\mathbf{B}(\boldsymbol{r}, \boldsymbol{s}) = \{B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})\}, \quad \boldsymbol{r}, \boldsymbol{s} \in \Omega, \quad k, \ell = 1, 2, ..., m, \tag{8.5}$$

is a multi-level covariance function on Ω if, for each positive integer n, and for each choice of points $\boldsymbol{r}_1, ..., \boldsymbol{r}_n \in \Omega$, the $mn \times mn$ matrix

$$\{B_{k\ell}(\boldsymbol{r}_i, \boldsymbol{r}_j)\}, \quad i, j = 1, 2, ..., n, \quad k, \ell = 1, 2, ..., m, \tag{8.6}$$

is positive semi-definite, i.e., for arbitrary scalars c_{ik} with $i = 1, 2, ..., n$ and $k = 1, 2, ..., m$,

$$\sum_{k,\ell=1}^{m} \sum_{i,j=1}^{n} c_{ik} c_{j\ell} B_{k\ell}(\boldsymbol{r}_i, \boldsymbol{r}_j) \geq 0. \tag{8.7}$$

From any given matrix covariance function $\mathbf{B}(\boldsymbol{r}, \boldsymbol{s}) = \{B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})\}$, we now want to extract a *scalar* (i.e., single-level) covariance function to evaluate in practice. To build such a scalar covariance function, we choose the indices $k, \ell = 1, 2, ..., m$ to index m *subdomains*, or *subregions* Ω_k, $\Omega_\ell \subset \Omega$, which are open, connected sets with smooth boundaries $\partial\Omega_k$, $\partial\Omega_\ell$. These subregions partition the domain Ω into m non-overlapping, non-empty subdomains, i.e.,

$$\bigcup_{k=1}^{m} \overline{\Omega}_k = \overline{\Omega}, \quad \Omega_k \cap \Omega_\ell = \emptyset \quad \forall k \neq \ell, \tag{8.8}$$

where the overbar denotes the closure, e.g., $\overline{\Omega}_k = \Omega_k \cup \partial\Omega_k$. Now, for every $\boldsymbol{r} \in \Omega_k$ and $\boldsymbol{s} \in \Omega_\ell$, and for each fixed $k, \ell = 1, 2, ..., m$, we define the following scalar function B,

$$B(\boldsymbol{r}, \boldsymbol{s}) := B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}). \tag{8.9}$$

Equation (8.9) is a *piecewise-defined, scalar* (single-level) covariance function on Ω, where the non-boldface B denotes a scalar function. The proof that Eq. (8.9) is a covariance function is given in Appen. D.1.

From the matrix covariance function in Eq. (8.5), we have extracted particular entries that collapse the multi-level covariance function into the single-level covariance function[1] defined in Eq. (8.9). For $\boldsymbol{r}, \boldsymbol{s} \in \Omega_k$, $B(\boldsymbol{r}, \boldsymbol{s}) = B_{kk}(\boldsymbol{r}, \boldsymbol{s})$ is an auto-covariance, while if $\boldsymbol{r} \in \Omega_k$, $\boldsymbol{s} \in \Omega_\ell$, and $k \neq \ell$, then $B(\boldsymbol{r}, \boldsymbol{s}) = B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ is a cross-covariance. The piecewise nature of Eq. (8.9) is inherited from the partitioning of Ω: different functions $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ are evaluated depending on the subregions $\boldsymbol{r}$ and $\boldsymbol{s}$ belong to, piecing together a scalar covariance function that is defined over the full domain Ω. In Appen. D.1, we illustrate how to construct Eq. (8.9) in a simple example, which makes the piecewise nature of Eq. (8.9) more explicit.

The advantage of scalar covariance functions defined via Eq. (8.9) is that the subregions Ω_k can have different properties, such as length-scale or cut-off length. Therefore, as these properties vary over the different subregions Ω_k of the full domain Ω, inhomogeneous and anisotropic features are introduced to the resulting scalar covariance function. The choice of subregions Ω_k can be quite general and can adapt to the needs of different applications. For example, $\Omega_1, \Omega_2, ..., \Omega_m$ can represent different vertical levels of a global circulation model, which motivated the terminology "multi-level" in G06. Each vertical level can just as well be partitioned again into individual grid cells, so that we represent the subregions Ω_k with multi-index $k = (k_{lat}, k_{lon}, k_{vert})$, for instance.

8.2.2 Construction

The construction of scalar covariance functions from matrix covariance functions described in the previous section provides the framework in which we can introduce inhomogeneous and anisotropic features by varying quantities such as length-scale or cut-off length over the different subregions $\Omega_1, \Omega_2, ..., \Omega_m$. There are several different ways to use Eq. (8.9) to build inhomogeneous

[1] A matrix covariance function, Eq. (8.3) for example, can be interpreted as "multi-level" in the sense that each k corresponds to a "level" in which a single-level covariance function is defined over the full domain Ω with m different levels as $k = 1, 2, ..., m$. Equation (8.9) extracts a portion from each of these k levels to collapse the multi-level covariance function into a single-level, piecewise-defined, scalar covariance function on Ω.

and anisotropic covariance functions, as illustrated in Exs. 2.3, 2.5, and 2.6 of G06, for instance. The focus of the present work is on the construction of scalar covariance functions from a collection of radially symmetric, compactly-supported functions. We will choose the functions $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ in Eq. (8.9) to be the convolution of such functions, applying theory established in GC99 to the framework introduced here in Sec. 8.2.1. We begin by describing the construction of scalar covariance functions from convolutions first on $\mathbb{R}^N$ for general N, then restricted to $N = 3$. This general approach will lay the groundwork for the construction of GenGC on $\mathbb{R}^3$ and on the surface of the sphere $\mathbb{S}^2_r$, which we describe at the end of this section.

Suppose we have a collection of functions $h_k \colon \mathbb{R}^N \mapsto \mathbb{R}$, $k = 1, 2, ..., m$, with each $h_k \in L^1(\mathbb{R}^N) \cap L^2(\mathbb{R}^N)$. Assume each h_k is radially symmetric on $\mathbb{R}^N$, meaning that there exists an even function $h_k^0 \colon \mathbb{R} \mapsto \mathbb{R}$ such that for any $\boldsymbol{r}, \boldsymbol{s} \in \mathbb{R}^N$,

$$h_k(\boldsymbol{r} - \boldsymbol{s}) = h_k^0(||\boldsymbol{r} - \boldsymbol{s}||), \tag{8.10}$$

where each function h_k^0 depends only on the distance $||\cdot||$, which here denotes the Euclidean norm on $\mathbb{R}^N$. Other norms can be used in Eq. (8.10), however the Euclidean norm is the natural choice in this context. Now, for each fixed $k, \ell = 1, 2, ..., m$, and all $\boldsymbol{r}, \boldsymbol{s} \in \mathbb{R}^N$, define $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ to be the following convolution over $\mathbb{R}^N$ of functions from this collection,

$$B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = (h_k * h_\ell)(\boldsymbol{r} - \boldsymbol{s}) = \int_{\mathbb{R}^N} h_k(\boldsymbol{v}) h_\ell(\boldsymbol{r} - \boldsymbol{s} - \boldsymbol{v}) d\boldsymbol{v}. \tag{8.11}$$

As shown in G06 (see Remarks on p. 1820), the resulting matrix of convolutions $\mathbf{B}(\boldsymbol{r}, \boldsymbol{s}) = \{B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})\} = \{(h_k * h_\ell)(\boldsymbol{r} - \boldsymbol{s})\}$ for $k, \ell = 1, 2, ..., m$ defines a multi-level covariance function. To extract the scalar covariance function from this matrix of convolutions, we partition the given domain $\Omega \subseteq \mathbb{R}^N$ into subregions Ω_k, $k = 1, 2, ..., m$, where these indices k correspond to the indices of the radially symmetric functions h_k. Equation (8.9) then defines the scalar covariance function $B(\boldsymbol{r}, \boldsymbol{s})$ extracted from the matrix covariance function $\mathbf{B}(\boldsymbol{r}, \boldsymbol{s}) = \{(h_k * h_\ell)(\boldsymbol{r} - \boldsymbol{s})\}$ of Eq. (8.11). That is, the scalar covariance function $B(\boldsymbol{r}, \boldsymbol{s})$ is obtained by evaluating Eq. (8.11) only for $\boldsymbol{r} \in \Omega_k$ and $\boldsymbol{s} \in \Omega_\ell$, for each fixed $k, \ell = 1, 2, ..., m$.

For high-dimensional applications, we are interested in constructing covariance functions in which long-range correlations are identically zero. We can accomplish this by requiring the radially symmetric functions h_k to have compact support, i.e., each of these functions go to zero at some fixed distance. For the rest of the section, we will require the functions h_k to be supported on spheres of fixed radius $c_k > 0$, where c_k is referred to as the *cut-off length*.

With our scalar covariance function $B(\boldsymbol{r}, \boldsymbol{s})$ defined through Eq. (8.9) by the convolutions in Eq. (8.11), all that remains is to evaluate the convolution integrals. For evaluation, we will restrict ourselves to $\mathbb{R}^3$, which is sufficient for most applications and will make explicit computation of Eq. (8.11) feasible. To simplify the calculation of Eq. (8.11), observe that for each pair k, ℓ, the function $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = (h_k * h_\ell)(\boldsymbol{r} - \boldsymbol{s})$ is radially symmetric, which is a property inherited from the radial symmetry of h_k and h_ℓ (see Thm. 3.a.1 in GC99). Therefore, for fixed k, ℓ, and for $\boldsymbol{r} \in \Omega_k$ and $\boldsymbol{s} \in \Omega_\ell$, we can express $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = B^0_{k\ell}(||\boldsymbol{r} - \boldsymbol{s}||)$ for some even function $B^0_{k\ell}$ defined on the real axis, which simplifies the evaluation of Eq. (8.11) significantly and is summarized in the following theorem from GC99.

Theorem 1. *(Theorem 3.c.1 of GC99) Suppose that $h_k \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ is radially symmetric and supported on a sphere of radius c_k, $0 < c_k \leq \infty$, for $k = 1, 2, ..., m$. Let h^0_k and $B^0_{k\ell}$ denote the even functions on $\mathbb{R}$ given by*

$$h^0_k(||\boldsymbol{r}||) := h_k(\boldsymbol{r}), \tag{8.12}$$

$$B^0_{k\ell}(||\boldsymbol{r}||) := (h_k * h_\ell)(\boldsymbol{r}) = \int_{\mathbb{R}^3} h_k(\boldsymbol{v}) h_\ell(\boldsymbol{r} - \boldsymbol{v}) d\boldsymbol{v}, \tag{8.13}$$

for $\boldsymbol{r} \in \mathbb{R}^3$ and $k, \ell = 1, 2, ..., m$, where $||\cdot||$ denotes Euclidean distance. If $c_k \leq c_\ell$, then

$$B^0_{k\ell}(z) = \frac{2\pi}{z} \int_0^{c_k} r h^0_k(r) \int_{|r-z|}^{r+z} s h^0_\ell(s)\, ds\, dr, \quad z > 0, \tag{8.14}$$

$$B^0_{k\ell}(0) = 4\pi \int_0^{c_k} r^2 h^0_k(r) h^0_\ell(r)\, dr. \tag{8.15}$$

Therefore, for $\boldsymbol{r} \in \Omega_k$, $\boldsymbol{s} \in \Omega_\ell$, and $c_k \leq c_\ell$, we evaluate the scalar covariance function $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = B^0_{k\ell}(||\boldsymbol{r} - \boldsymbol{s}||)$ by evaluating Eqs. (8.14) – (8.15) with $z = ||\boldsymbol{r} - \boldsymbol{s}||$.

The correlation function $C_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ can be computed by decomposing the covariance $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ for $\boldsymbol{r} \in \Omega_k$ and $\boldsymbol{s} \in \Omega_\ell$ as follows,

$$B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = \sqrt{B_{kk}(\boldsymbol{r}, \boldsymbol{r})} C_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) \sqrt{B_{\ell\ell}(\boldsymbol{s}, \boldsymbol{s})}, \tag{8.16}$$

where $B_{kk}(\boldsymbol{r}, \boldsymbol{r})$ and $B_{\ell\ell}(\boldsymbol{s}, \boldsymbol{s})$ are the variances at $\boldsymbol{r} \in \Omega_k$ and $\boldsymbol{s} \in \Omega_\ell$, respectively. Using Thm. 1, we can therefore evaluate the correlation function explicitly for each k, ℓ and $\boldsymbol{r} \in \Omega_k,\ \boldsymbol{s} \in \Omega_\ell$ as follows,

$$C_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = \frac{B^0_{k\ell}(||\boldsymbol{r} - \boldsymbol{s}||)}{\sqrt{B^0_{kk}(0) B^0_{\ell\ell}(0)}}. \tag{8.17}$$

Up to this point, we have described the construction of scalar covariance functions from convolutions of arbitrary compactly-supported, radially symmetric functions h_k. To construct the Generalized Gaspari-Cohn (GenGC) correlation function, we specify the functions h_k as follows. Given scalars $a_k \in \mathbb{R}$ and $c_k > 0$, for $k = 1, 2, ..., m$, define $h_k \colon \mathbb{R}^3 \mapsto \mathbb{R}$ as

$$h_k(\boldsymbol{r}; a_k, c_k) = \begin{cases} \big(2(a_k - 1)||\boldsymbol{r}||/c_k + 1\big) n_k, & 0 \leq ||\boldsymbol{r}|| \leq c_k/2, \\ 2a_k n_k (1 - ||\boldsymbol{r}||/c_k), & c_k/2 \leq ||\boldsymbol{r}|| \leq c_k, \\ 0, & c_k \leq ||\boldsymbol{r}||, \end{cases} \tag{8.18}$$

where $n_k = (44a_k^2 + 6a_k + 2)^{-1/2}$. The functions $h_k(\boldsymbol{r}; a_k, c_k)$ are piecewise linear, compactly-supported, radially symmetric functions that are supported on a sphere of radius c_k. Each $h_k(\boldsymbol{r}; a_k, c_k)$ is parametrized by the scalar a_k and the cut-off length c_k. Equation (8.18) generalizes the piecewise linear functions defined in Sec. 4(c) of GC99 and Eq. (32) of G06 by letting both parameters a_k and c_k vary over the index k. Examples of Eq. (8.18) are given in Fig. 8.1 for different values of a_k. The parameter a_k determines the value of $h_k(\boldsymbol{r}; a_k, c_k)$ at $||\boldsymbol{r}|| = c_k/2$, subsequently determining the slopes of each line in its piecewise definition.

To construct the GenGC correlation function, we partition the domain $\Omega \subseteq \mathbb{R}^3$ into subregions Ω_k, $k = 1, 2, ..., m$, where each subregion Ω_k has an associated cut-off length c_k and parameter a_k used to define h_k. The functions $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ for $\boldsymbol{r} \in \Omega_k$, $\boldsymbol{s} \in \Omega_\ell$ are defined using Eq. (8.11) on $\mathbb{R}^3$ for h_k, h_ℓ given in Eq. (8.18). For fixed k, ℓ, we can use Thm. 1 to evaluate $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$. Equation (8.17)

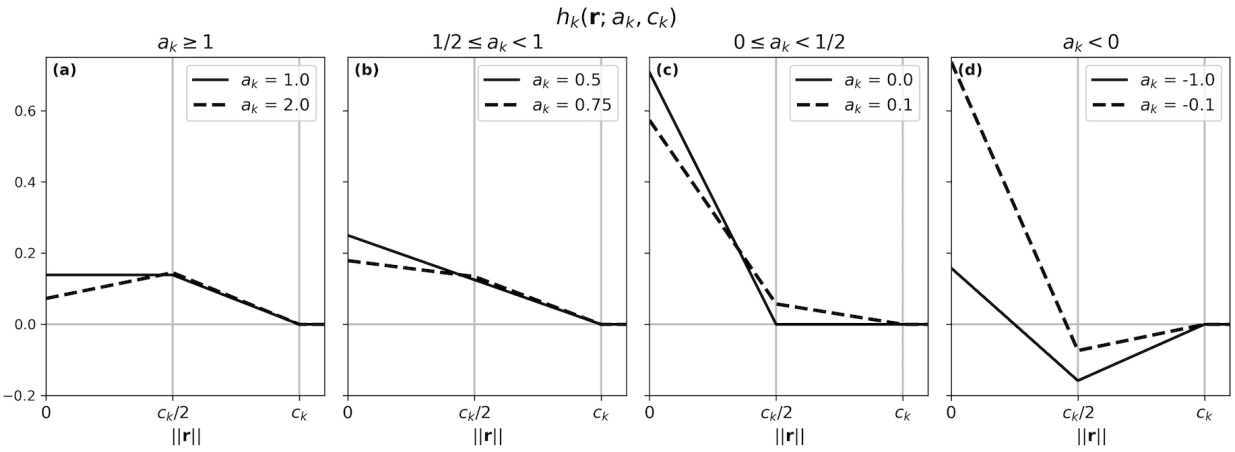

Figure 8.1: Examples of Eq. (8.18) for various values of a_k, plotted as a function of $||\boldsymbol{r}||$. In panel (a), $h_k(\boldsymbol{r}; 1, c_k)$ is constant on $0 \leq ||\boldsymbol{r}|| \leq c_k/2$. For $a_k = 1/2$ in panel (b), the function $h_k(\boldsymbol{r}; 1/2, c_k)$ collapses to a single line on the full domain $0 \leq ||\boldsymbol{r}|| \leq c_k$. In the case that $a_k = 0$, as in panel (c), $h_k(\boldsymbol{r}; 0, c_k)$ is zero for $c_k/2 \leq ||\boldsymbol{r}|| \leq c_k$. Values of $a_k < 0$ as shown in panel (d) are the only cases in which the function $h_k(\boldsymbol{r}; a_k, c_k)$ can become negative.

then defines the GenGC correlation function. Like the GC99 and G06 functions, GenGC is sixth-order piecewise rational (fifth-order near the origin)[2] and three-times continuously differentiable in space for fixed a_k, a_ℓ, c_k, c_ℓ. In the case that $c_k = c$ for $k = 1, 2, ..., m$ is fixed while a_k and a_ℓ vary, GenGC is the G06 function; similarly if also $a_k = 1/2$ for $k = 1, 2, ..., m$, GenGC is exactly the GC99 function. For general choices of cut-off lengths c_k and c_ℓ, GenGC is not supported on a sphere of uniform radius. Rather, its support depends on the values of c_k and c_ℓ. This, in combination with the parameters a_k and a_ℓ, allows for correlation functions that can be highly inhomogeneous and anisotropic.

GenGC is a piecewise-defined scalar correlation function that is first split into six functions depending on the relationship of c_k and c_ℓ, and subsequently split again into twelve different functions that depend on how $||\boldsymbol{r} - \boldsymbol{s}||$ for $\boldsymbol{r} \in \Omega_k$ and $\boldsymbol{s} \in \Omega_\ell$ relates to the particular c_k and c_ℓ. Appendix D.2 describes the construction of GenGC in further detail and gives the explicit formulae. We verified that this function is symmetric positive definite numerically by generating a variety of different matrices and checking for symmetry and positive eigenvalues. Note that these formulae reduce to those of the G06 function (their Eq. 33 and Appendix C) in the case that $c_k = c$ for $k = 1, 2, ..., m$.

Though the explicit formulae for the GenGC correlation function are lengthy (see Appen D.2), its evaluation can be implemented efficiently in practice, for instance as an operator acting on a vector (Gilpin, 2023). The evaluation of the GenGC correlation function can be formulated as a two step process: first, there are the two sets of conditional statements that determine the correct piecewise function to evaluate based on c_k, c_ℓ, and $||\boldsymbol{r} - \boldsymbol{s}||$, then a small number of scalar operations are done to evaluate the chosen function. These two steps are not computationally expensive, which can be optimized per coding language, and simply extend existing implementations of the GC99 and G06 functions.

[2] The GenGC, GC99, and G06 correlation functions are sixth-order piecewise rational functions in that the highest-degree terms are of the form z^6/z (near the origin, the highest-degree terms are z^5 since the $1/z$ terms vanish and are therefore fifth-order). We therefore correct the mis-stated order of the GC99 function in Sec. 4(c) of GC99 and the G06 function in G06, p. 1821.

Remark 4. *As the GenGC correlation function is constructed via convolution over the Euclidean space* $\mathbb{R}^3$*, for* $\boldsymbol{r},\ \boldsymbol{s} \in \mathbb{R}^3$ *the norm* $||\boldsymbol{r}-\boldsymbol{s}||$ *is taken to be the Euclidean norm. Restricting* $\boldsymbol{r}, \boldsymbol{s} \in \mathbb{R}^3$ *to* $\mathbb{S}_r^2$ *for* $r > 0$ *embedded in* $\mathbb{R}^3$ *requires restricting the Euclidean norm to* $\mathbb{S}_r^2$*, which is chordal distance,*

$$||\boldsymbol{r}-\boldsymbol{s}||_{cd} := 2r\sin(\theta/2), \quad 0 \leq \theta \leq \pi, \quad \boldsymbol{r}, \boldsymbol{s} \in \mathbb{S}_r^2 \tag{8.19}$$

where θ *is the angle between the vectors* $\boldsymbol{r}$ *and* $\boldsymbol{s}$*,*

$$r^2\cos(\theta) = \boldsymbol{r}^T\boldsymbol{s}, \quad \boldsymbol{r}, \boldsymbol{s} \in \mathbb{S}_r^2, \tag{8.20}$$

(GC99, pp. 730 – 732, 737). This implies that the cut-off length c_k *should also be specified in terms of chordal distance.*

8.3 Continuity Properties Involving a and c

With the GenGC correlation function derived in the previous section, we now discuss two of its properties resulting from a_k and c_k varying over the domain. Since correlation modeling and tapering applications in data assimilation typically require continuity, we first study the influence of a_k and c_k on the continuity behavior of GenGC. We show that GenGC is not, in general, continuous at the boundaries $\partial\Omega_k$ of the subregions Ω_k. If, however, a_k and c_k are defined as discretizations of continuous functions a and c on Ω, then continuity of GenGC, in an appropriate sense, ensues. We conclude this section with a discussion of the correlation length of GenGC, first introduced in Eq. (8.1), and of implementation of GenGC when the correlation length is available over the domain.

8.3.1 Continuity of a and c as Functions on Ω

The GenGC correlation function is defined piecewise on the domain Ω. The piecewise nature is inherited from the partitioning of Ω into subregions $\Omega_1, \Omega_2, ..., \Omega_m$, with each subregion assigned constants a_k and c_k used to define the functions h_k for $k = 1, 2, ..., m$ in Eq. (8.11). For fixed k, ℓ, the scalar function $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ in Eq. (8.11) used to define GenGC in Eq. (8.17) is continuous

for $\boldsymbol{r} \in \Omega_k$, $\boldsymbol{s} \in \Omega_\ell$, since the functions h_k and h_ℓ are continuous in their arguments. However, continuity of $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ over the full domain Ω (i.e., as k, ℓ vary), and hence that of the GenGC correlation function, is not always guaranteed. Rather, the continuity of $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ over Ω depends on the behavior of $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ across the boundaries of the subregions $\Omega_1, \Omega_2, ..., \Omega_m$.

To see this, fix k, ℓ with $k \neq \ell$ and let $\boldsymbol{r}_* \in \partial\Omega_k$ and $\boldsymbol{s}_* \in \partial\Omega_\ell$ with $\boldsymbol{r}_* = \boldsymbol{s}_*$ a point in common to the two boundaries. We define $B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*)$ and $B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*)$ in terms of limits,

$$B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) := \lim_{\boldsymbol{r} \to \boldsymbol{r}_*,\ \boldsymbol{r} \in \Omega_k} B_{kk}(\boldsymbol{r}, \boldsymbol{r}) = \int_{\mathbb{R}^3} h_k^2(\boldsymbol{v}; a_k, c_k)\, d\boldsymbol{v}, \tag{8.21}$$

$$B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*) := \lim_{\boldsymbol{s} \to \boldsymbol{s}_*,\ \boldsymbol{s} \in \Omega_\ell} B_{\ell\ell}(\boldsymbol{s}, \boldsymbol{s}) = \int_{\mathbb{R}^3} h_\ell^2(\boldsymbol{v}; a_\ell, c_\ell)\, d\boldsymbol{v}, \tag{8.22}$$

where the last equality in each case comes from the convolutions defined in Eq. (8.11). We also define $B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*)$ in a similar manner,

$$B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*) := \lim_{\boldsymbol{r} \to \boldsymbol{r}_*,\ \boldsymbol{r} \in \Omega_k} \lim_{\boldsymbol{s} \to \boldsymbol{s}_*,\ \boldsymbol{s} \in \Omega_\ell} B_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = \int_{\mathbb{R}^3} h_k(\boldsymbol{v}; a_k, c_k) h_\ell(\boldsymbol{v}; a_\ell, c_\ell)\, d\boldsymbol{v}. \tag{8.23}$$

The limits in Eqs. (8.21) – (8.23) exist due to the continuity of $B_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ for $\boldsymbol{r} \in \Omega_k, \boldsymbol{s} \in \Omega_\ell$ and fixed k, ℓ.

For the covariance function to be well-defined over all of Ω, including at the boundaries between subregions, and in particular for the variance to exist at each common boundary point $\boldsymbol{r}_* = \boldsymbol{s}_*$, we must have that $B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) = B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*) = B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*)$. It is clear, however, from Eqs. (8.21) and (8.22) that $B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) \neq B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*)$ in general, since h_k and h_ℓ under the integral signs are different functions if $a_k \neq a_\ell$ or $c_k \neq c_\ell$. We also have

$$|B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*)| \leq |B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*)|^{1/2}, \tag{8.24}$$

by the Cauchy-Schwartz inequality. Though the quantity $B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*)$ is defined at the common boundary point $\boldsymbol{r}_* = \boldsymbol{s}_*$ when $a_k \neq a_\ell$ or $c_k \neq c_\ell$, it is not the variance, and all we are guaranteed is the bound given in Eq. (8.24).

We illustrate the consequences of specifying a_k and c_k arbitrarily over the different subregions in the top row of Fig. 8.2. For this example, we consider the GenGC correlation function on the

domain $(0,1)$, which is discretized into 201 equally-spaced grid points (including both endpoints 0 and 1). We then partition the one-dimensional line segment $(0,1)$ into two subregions $\Omega_1 = (0,1/2)$ and $\Omega_2 = (1/2,1)$, where $a_1 \neq a_2$ and $c_1 \neq c_2$ (values shown in the top left panel). The GenGC correlation matrix produced for this case, shown in the top middle panel, is continuous within the four blocks (which correspond to fixed $k,\ell = 1,2$), but is discontinuous where the boundaries of the subregions meet at $r_* = s_* = 1/2$. This is more explicit in the one-dimensional correlations shown in the top right panel, where we see away from the boundary at $1/2$, the correlations are continuous, but each correlation has a jump discontinuity at $1/2$.

Though $a_1, a_2, ..., a_m$ and $c_1, c_2, ..., c_m$ can be arbitrarily specified on the different subregions $\Omega_1, \Omega_2, ..., \Omega_m$, the covariance function, and hence the GenGC correlation function, will not necessarily be continuous over the full domain Ω. Since data assimilation applications typically require continuous covariance functions, we must impose conditions on a_k, a_ℓ and c_k, c_ℓ so that the covariance function is well-defined and continuous where the boundaries of the subregions Ω_k and Ω_ℓ meet. This requires defining a_k, a_ℓ and c_k, c_ℓ from continuous functions on Ω, which we describe in the following theorem.

Theorem 2. *Let $a\colon \Omega \mapsto \mathbb{R}$ and $c\colon \Omega \mapsto (0,\infty)$ be continuous functions on Ω, and for each $k = 1,2,...,m$, define*

$$a_k := \frac{\int_{\Omega_k} a(\boldsymbol{r})\, d\boldsymbol{r}}{\int_{\Omega_k} 1 \cdot d\boldsymbol{r}}, \quad c_k := \frac{\int_{\Omega_k} c(\boldsymbol{r})\, d\boldsymbol{r}}{\int_{\Omega_k} 1 \cdot d\boldsymbol{r}}. \tag{8.25}$$

Fix $k \neq \ell$ and suppose $\boldsymbol{r}_* \in \partial\Omega_k$, $\boldsymbol{s}_* \in \partial\Omega_\ell$ with $\boldsymbol{r}_* = \boldsymbol{s}_*$. In the limit as Ω_k and Ω_ℓ shrink towards the common boundary point $\boldsymbol{r}_* = \boldsymbol{s}_*$, then $B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) = B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*) = B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*)$.

Proof. In the limit as Ω_k and Ω_ℓ shrink towards their boundary points $\boldsymbol{r}_*$ and $\boldsymbol{s}_*$, respectively, it follows from Eq. (8.25) that $a_k \to a(\boldsymbol{r}_*)$, $c_k \to c(\boldsymbol{r}_*)$, $a_\ell \to a(\boldsymbol{s}_*)$, and $c_\ell \to c(\boldsymbol{s}_*)$. Since the boundary points satisfy $\boldsymbol{r}_* = \boldsymbol{s}_*$, we have that $a_k = a_\ell$ and $c_k = c_\ell$ in the limiting case. Therefore, as Ω_k and Ω_ℓ shrink to their common boundary point $\boldsymbol{r}_* = \boldsymbol{s}_*$, it follows from Eqs. (8.21) – (8.22) that $B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) = B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*)$, i.e., the variances are equal and well-defined. In addition, it follows from Eqs. (8.21) – (8.23) that $B_{k\ell}(\boldsymbol{r}_*, \boldsymbol{s}_*) = B_{kk}(\boldsymbol{r}_*, \boldsymbol{r}_*) = B_{\ell\ell}(\boldsymbol{s}_*, \boldsymbol{s}_*)$. □

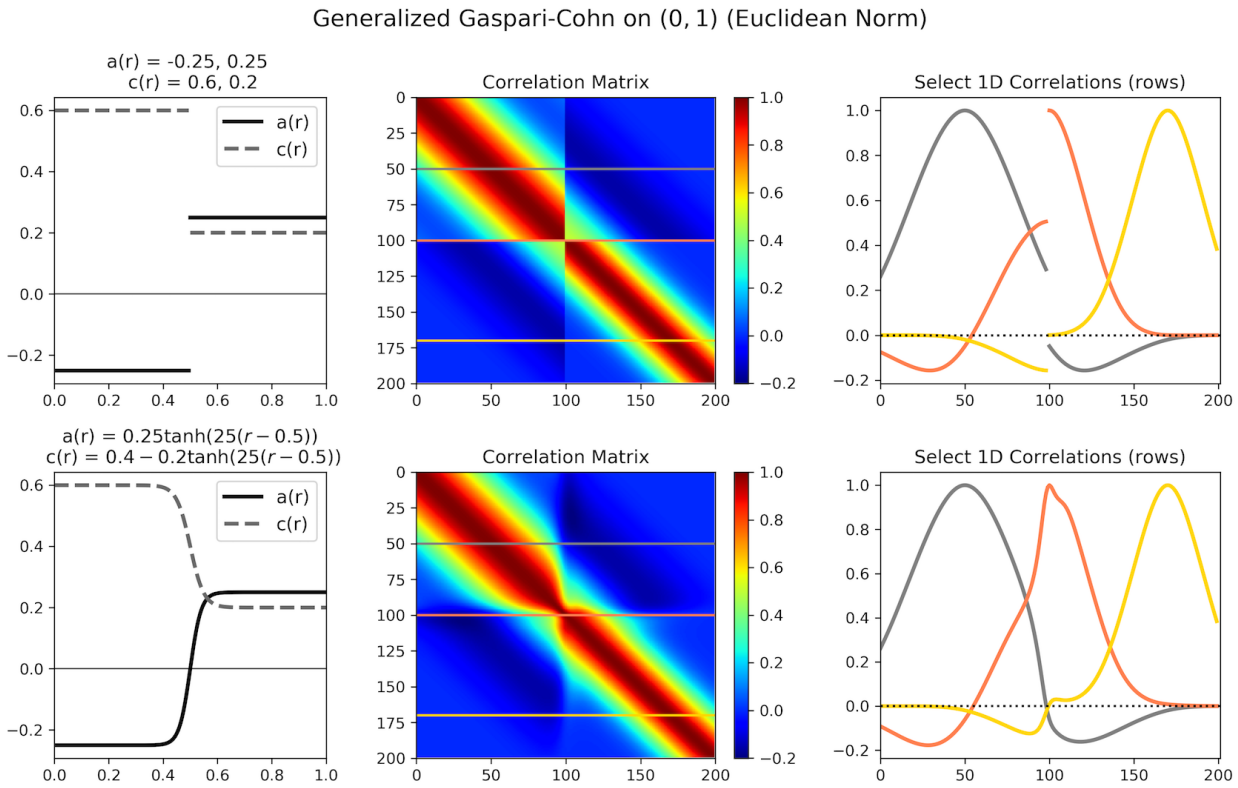

Figure 8.2: One-dimensional examples of the GenGC correlation function on $(0, 1)$ for different choices of functions for a and c (and subsequently partitions of $(0, 1)$). In both cases, the domain $(0, 1)$ is discretized into 201 equally-spaced grid points (including both endpoints 0 and 1). Top row: the domain is partitioned into two subregions, $\Omega_1 = (0, 1/2)$ and $\Omega_2 = (1/2, 1)$, where values of a and c are constant over these two domains and chosen to mimic the constant values of the continuous case, shown on the bottom left panel. The middle panel illustrates the corresponding GenGC correlation matrix, with horizontal lines indicating the selected one-dimensional correlations shown on the right at rows 50 (grey), 100 (orange), and 170 (yellow). Bottom row: both a and c vary continuously over $(0, 1)$ according to the functions specified in the subplot title. The corresponding GenGC correlation matrix is in the middle (see text for the description of its construction), with the right panel illustrating rows 50, 100, and 170 from the correlation matrix.

Therefore, the variances are continuous throughout Ω when a_k and c_k are chosen as appropriate discretizations of continuous functions, in the sense described in Thm. 2, and this implies continuity of the GenGC correlation function in the same sense. In practice, the integrals in Eq. (8.25) can be approximated numerically, such as with the midpoint rule, for example.

Theorem 2 provides the framework to construct continuous GenGC correlation functions on Ω. Suppose the domain Ω is discretized into m grid cells. Take each grid cell to be a subregion Ω_k for $k = 1, 2, ..., m$, then define a_k and c_k according to Eq. (8.25) given continuous functions a and c on Ω. The bottom row of Fig. 8.2 constructs GenGC on $(0, 1)$ in this framework. The bottom left panel plots the functions $a(r)$ and $c(r)$ for $r \in (0, 1)$, both constructed using a hyperbolic-tangent function to allow for a smooth transition between two constant values on either side of $1/2$. The domain $(0, 1)$ is discretized into 201 equally-spaced grid points (including both endpoints 0 and 1), and the subregions Ω_k for $k = 1, 2, .., m = 200$ are the grid intervals of width $1/200$ centered halfway between each grid point. Values of a_k and c_k in Eq. (8.25) are evaluated using the midpoint rule with four subintervals within each subregion Ω_k. The resulting GenGC correlation matrix is shown in the bottom center panel, along with selected rows of the correlation matrix on the right. We can see that when a and c are continuous functions and each subregion is taken to be a grid cell on the discrete grid, the GenGC correlation matrix does not have the jump discontinuities seen in the top row, and instead has a smooth transition across $1/2$ corresponding to the smooth behavior in a and c shown on the left. In fact, by choosing hyperbolic tangent to construct a and c, we have created a continuous version of the discontinuous case shown in the top row of Fig. 8.2, as is clear under mesh-refinement (not shown).

We emphasize here that the choice of a and c and the partitioning of the domain have a significant impact not only on the inhomogeneity and anisotropy of the GenGC correlation function, but also on its continuity over space, which is not readily available from convolution theory alone. For applications where continuity of the GenGC correlation function is required, choosing a and c to be continuous functions over the spatial domain will ensure that continuity holds. In cases where a and c are not continuous functions, such as the piecewise-constants shown in the top row

of Fig. 8.2, the GenGC correlation function will not be continuous across the boundaries of the subregions. There are applications where jump discontinuities like those shown in Fig. 8.2 may be suitable, for instance at the interface between different fluids, such as the atmosphere and ocean (e.g., Stanley et al., 2021). Care should be taken to ensure that the GenGC correlation function constructed satisfies the requirements of the particular application.

8.3.2 Correlation Length

Evaluation of the GenGC correlation function requires specifying $a_1, a_2, ..., a_m$ and $c_1, c_2, ..., c_m$ over the different subregions of the domain. The GenGC correlation function does not impose restrictions on how $a_1, a_2, ..., a_m$ and $c_1, c_2, ..., c_m$ are specified, therefore these quantities can be specified arbitrarily. If GenGC must be continuous over the full domain, the only restriction is that a_k and c_k for $k = 1, 2, .., m$ are defined from continuous functions, as described in the previous section, but these functions can also be arbitrary.

In some applications, the cut-off length values c_k for $k = 1, 2, ..., m$ may be readily specified from physical or computational considerations, while specifying each a_k may be less straightforward and might instead be computed from other known quantities. The desired correlation length values L_k, for instance, may be available while holding the cut-off length constant. For example, for state dynamics governed by the continuity equation and related hyperbolic partial differential equations (PDEs), an explicit PDE can be derived for the correlation length field (see Cohn, 1993, pp. 3137–3138). Therefore, if the PDE for the correlation length field is solved and the cut-off field is known or specified, correlations can be reconstructed using GenGC, provided that we can relate the correlation length L_k and cut-off length c_k to a_k. We derive an expression here that does so, and we discuss methods of reconstructing correlations with GenGC in this context.

For fixed k, ℓ and $\boldsymbol{r} \in \Omega_k$, $\boldsymbol{s} \in \Omega_\ell$, the GenGC correlation function, Eq. (8.17), can be expressed as $C_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = C^0_{k\ell}(z)$ for an even function $C^0_{k\ell}$ defined on the real axis with $z = ||\boldsymbol{r} - \boldsymbol{s}||$. Therefore, we can define the correlation length for GenGC in terms of its second derivative with respect to z for $k = \ell$ and at $z = 0$ (Daley, 1991, p. 110; Michel et al., 2016, special case of their

Eq. 2),

$$L_k = \left(-\frac{d^2}{dz^2} C^0_{kk}(z) \Big|_{z=0} \right)^{-1/2}, \quad k = 1, 2, ..., m. \tag{8.26}$$

From the explicit formula for GenGC (see Appen. D.2), we derive an expression for the correlation length L_k as a function of a_k and c_k,

$$L_k = c_k \left(\frac{3(22a_k^2 + 3a_k + 1)}{40(8a_k^2 - 2a_k + 1)} \right)^{1/2}. \tag{8.27}$$

As the quantities a_k and c_k vary over the different subregions Ω_k for $k = 1, 2, ..., m$, the correlation length L_k also varies. In particular, if a_k and c_k are defined from continuous functions on Ω, for example by Eq. (8.25), then it follows from Eq. (8.27) that the correlation length L_k is also. Observe that while L_k depends linearly on c_k for fixed a_k, L_k depends only weakly on a_k away from $a_k = 0$, but strongly near $a_k = 0$, as shown in the left panel of Fig. 8.3.

In applications where the cut-off lengths c_k and correlation lengths L_k are available, to construct GenGC one would need to recover a_k. From Eq. (8.27), we can solve for a_k as a function of L_k and c_k, which yields a quadratic equation with the following two solutions,

$$a_k^{\pm} = \frac{80 + 9\kappa_k^2 \pm \sqrt{(80 + 9\kappa_k^2)^2 - 4(320 - 66\kappa_k^2)(40 - 3\kappa_k^2)}}{640 - 132\kappa_k^2}, \tag{8.28}$$

where we define $\kappa_k = c_k/L_k$, and $a_k^+(\kappa_k)$ ($a_k^-(\kappa_k)$) refers to the solution where the square root is added (subtracted).

Since the values of a_k in Eq. (8.28) are solutions to a quadratic equation, there are restrictions on c_k and L_k so that the solutions remain real, as well as restrictions on the choice of a_k^+ and a_k^- that should be considered. The right panel of Fig. 8.3 plots a_k^+ and a_k^- as functions of $\kappa_k = c_k/L_k$ and illustrates these restrictions. For the solutions in Eq. (8.28) to remain real and bounded, we require that the discriminant remain positive and the denominator to be non-zero. It follows that two real roots a_k^+ and a_k^- do not exist for all values of κ_k. Rather, we obtain two sets of restrictions on values of a_k^+ and a_k^- in Eq. (8.28) and the corresponding values of κ_k, which we will denote as

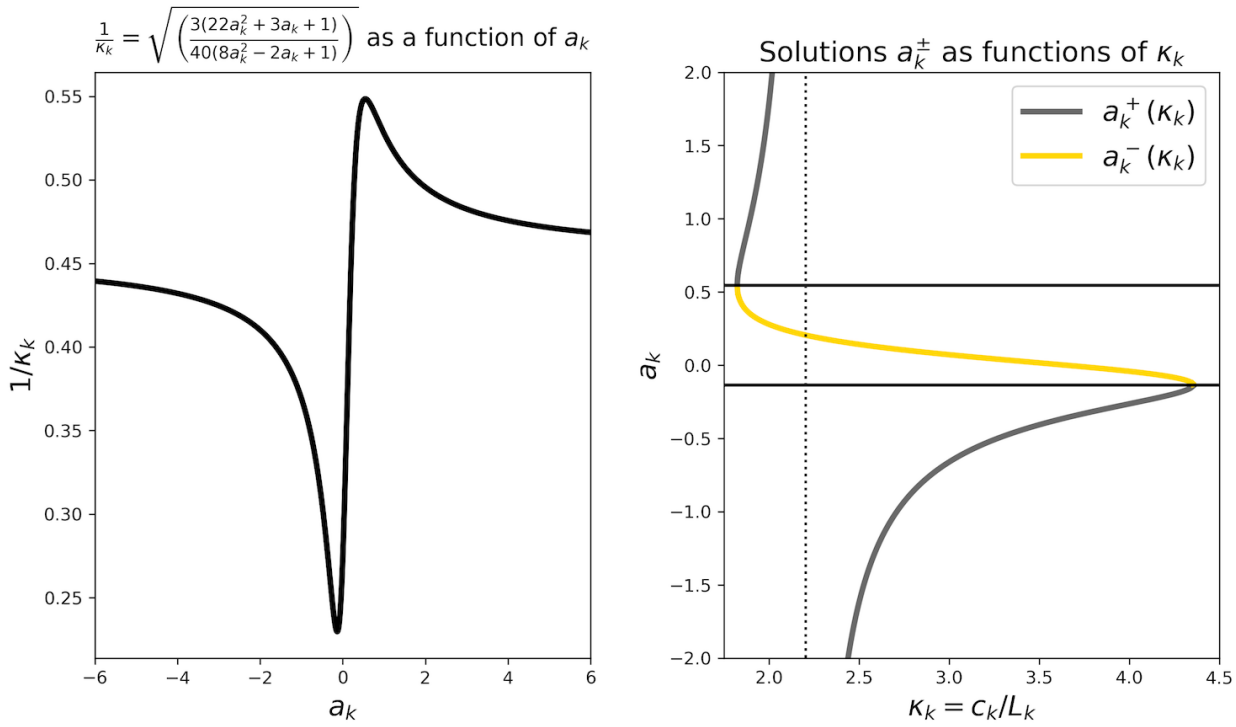

Figure 8.3: Left: Equation (8.27) solved for $1/\kappa_k = L_k/c_k$ plotted as a function of a_k. Right: Solutions a_k^+ (grey) and a_k^- (orange) to Eq. (8.28) as functions of $\kappa_k = c_k/L_k$. The vertical dotted line marks the vertical asymptote at $\kappa_k = \sqrt{160/33}$ for a_k^+. The horizontal black lines mark the transitions between the roots a_k^+ and a_k^-.

κ_k^+ and κ_k^- for a_k^+ and a_k^-, respectively,

$$a_k^+ \in (-\infty, -0.1348) \cup (0.5463, \infty) \text{ for } \kappa_k^+ \in (1.8233, 2.2019) \cup (2.2019, 4.3536) \tag{8.29}$$

$$a_k^- \in (-0.1348, 0.5463) \text{ for } \kappa_k^- \in (1.8233, 4.3536) \tag{8.30}$$

The root a_k^+ (grey curves in the right panel of Fig. 8.3) is asymmetric about the vertical asymptote at $\kappa_k^+ = \sqrt{160/33} \cong 2.2019$ (dotted line in the right panel of Fig. 8.3). Therefore, values of a_k^+ that cross the vertical asymptote at κ_k^+ will produce a GenGC correlation function that is not continuous. The root a_k^- (orange curve), however, remains continuous for all values κ_k^- specified in Eq. (8.30), but the values of a_k^- that can be achieved are limited, bounded by the black horizontal lines in the right panel of Fig. 8.3. As κ_k varies between the limits expressed in Eqs. (8.29) – (8.30), $1/\kappa_k = L_k/c_k$ is restricted to vary only between $1/4.3536 \cong 0.2297$ and $1/1.8233 \cong 0.5485$.

Solving for a_k directly from Eq. (8.28) given κ_k is possible, but has the limitation expressed by Eqs. (8.29) – (8.30) and illustrated in the right panel of Fig. 8.3. For applications where a_k may not vary substantially, an alternative to solving for a_k directly would be to fix $\kappa_k = \kappa$ as constant, which is equivalent to fixing $a_k = a$. Given a set of correlation lengths L_k, one could then solve for the cut-off lengths by evaluating $c_k = \kappa L_k$, making sure to observe the restrictions presented in Eqs. (8.29) – (8.30) when choosing κ. For example, correlation lengths L_k may be obtained dynamically, as in Cohn (1993, Sec. 4b), and fixing $\kappa = 1.8233$, the lower bound of κ^- and κ^+, we recover the cut-off lengths c_k and fix a. This choice of κ allows for recovery of the smallest possible cut-off lengths c_k given the correlation lengths L_k, which may be advantageous to reduce computational expense. Alternatively, one could simply take $a = 1/2$ fixed as in the GC99 correlation function, which according to the right panel of Fig. 8.3, would be nearly the same as choosing κ at its minimum $\kappa = 1.8233$.

8.4 Examples

The advantage of GenGC is that both a_k and c_k can vary over the subregions Ω_k, producing correlations that can be markedly inhomogeneous and anisotropic over the full domain Ω. This

section presents simple examples that highlight the anisotropy of GenGC on the sphere by allowing these quantities to vary, first through a series of one-dimensional examples, followed by a two-dimensional example on the surface of the Earth.

8.4.1 One-Dimensional Examples

We begin with one-dimensional examples on the discretized unit circle $\mathbb{S}_1^1$ with $m = 200$ equally-spaced grid points. The subregions Ω_k for $k = 1, 2, ..., m = 200$ are the grid intervals of width $\pi/100$ centered at the grid points. We choose $c(r)$ and $a(r)$ to be continuous functions on the unit circle and evaluate $c(r)$ and $a(r)$ at each grid point to define c_k and a_k, respectively (which is equivalent to evaluating Eq. 8.25 using the midpoint rule with one subinterval equal to the subregion). The corresponding GenGC correlation matrices are constructed using the chordal distance, Eq. (8.19), for different choices of functions $a(r)$ and $c(r)$. We give three examples that illustrate the results of allowing a and c to vary as continuous functions over the spatial domain.

Example 1

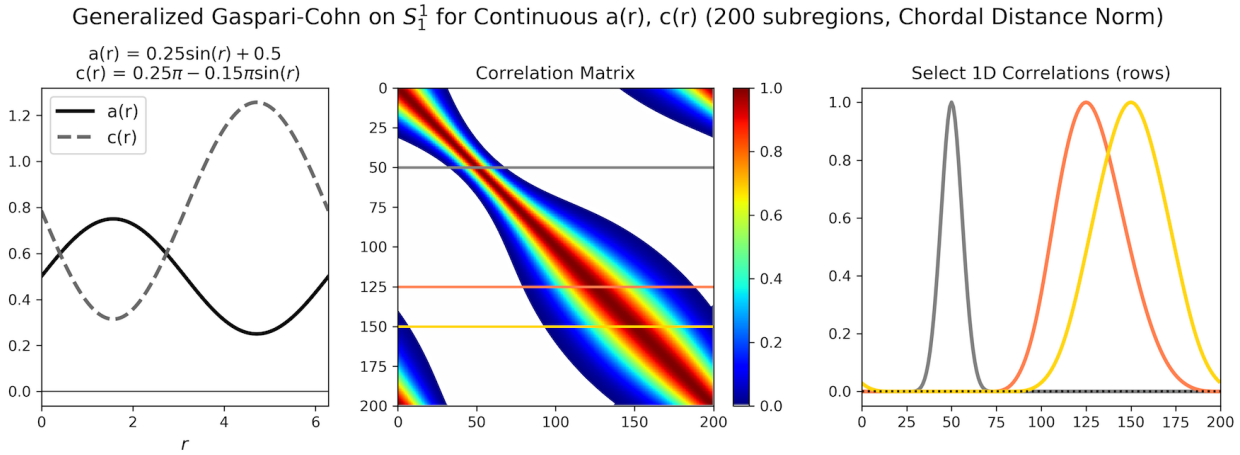

Figure 8.4: Example of the GenGC correlation function on the unit circle $\mathbb{S}_1^1$ constructed using continuous functions for a (solid black) and c (dashed grey) shown in the left-panel. The correlation matrix is shown in the center, constructed using chordal distance defined in Eq. (8.19). The white region corresponds to correlations between -0.003 and 0.003, highlighting the regions where the correlations become zero. Horizontal lines on the correlation matrix mark the row 50 (grey), 125 (orange), and 150 (yellow) shown in the right panel, with the dotted black line marking zero.

The first example, shown in Fig. 8.4, constructs the GenGC correlation matrix with a and c defined using simple sine functions illustrated in the left panel. The resulting correlation matrix in the middle panel of Fig. 8.4 has a non-uniform region where correlations vanish (white), with larger non-zero regions in the right half of the domain and smaller non-zero regions on the left. This reflects the choice of function for c and its impact on the region of compact support. The specific one-dimensional correlations, as we see in the right panel, also illustrate the impact of the spatial variability of the cut-off parameter c, where the correlations are compactly supported and asymmetric about zero separation. Though a also varies over space in this example, these values of a do not have a significant impact on the shape of the resulting correlations.

Correlation functions like the GenGC correlation function shown in Fig. 8.4 may be useful for covariance tapering, particularly for localization in ensemble-based data assimilation schemes. Current data assimilation schemes often implement the GC99 function for localization, which is isotropic on the sphere since c is fixed and $a = 1/2$ is held constant. Alternate means of localization have been proposed, such as adaptive localization (Bishop and Hodyss, 2007), spatially-dependent localization (Buehner, 2012), scale-dependent localization (Buehner and Shlyaeva, 2015), empirical localization functions (Anderson and Lei, 2013), and cross-localization for coupled data assimilation (Stanley et al., 2021)[3] . The GenGC correlation function removes the restriction of a single cut-off length and produces correlations that can be anisotropic and asymmetric about zero separation, as illustrated in Fig. 8.4.

Example 2

We can see the impact of allowing a to vary over space on the anisotropy of the GenGC correlation function more clearly in Fig. 8.5. In this example, the function a varies significantly on the unit circle and attains values less than zero. As a result, the GenGC correlation matrix shown in the middle panel of Fig. 8.5 is highly anisotropic and has negative correlations, all while remaining compactly supported (white region). The one-dimensional correlations for this case are

[3] The mulitvariate Gaspari-Cohn localization function introduced by Stanley et al. (2021) is in fact a rescaled version of the GenGC correlation function for two subregions with different values of c on the two subregions while $a = 1/2$ is fixed.

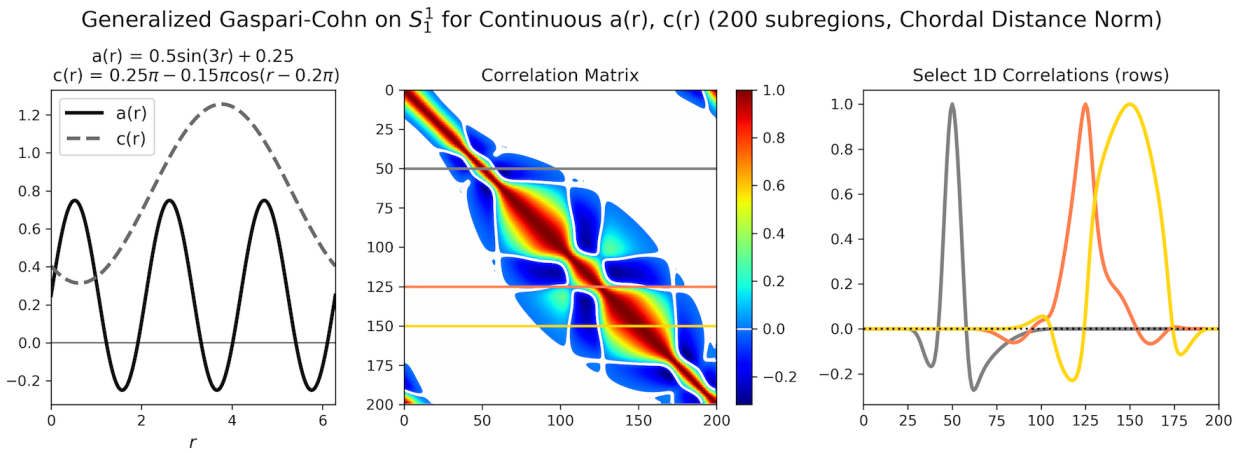

Figure 8.5: Same as Fig. 8.4 for different functions a and c shown in the left panel.

much more asymmetric about zero separation (right panel of Fig. 8.5) and vary significantly in their behavior. The grey correlation of row 50 is quite narrow, as it is in part of the domain with the smallest values of c, while the correlation at row 150 (yellow) is very broad close to zero separation.

The GenGC correlation function becomes negative for values of a less than zero, due to the behavior of the piecewise linear functions defined in Eq. (8.18). Correlations that become negative over certain regions occur naturally in several contexts, such as for geostrophically balanced wind fields (Daley, 1991, pp. 160 – 161, 163), for instance. Allowing a to vary across $a = 0$ significantly changes the interior shape of the correlations and can be tuned to suit different applications.

Example 3 A common application of correlation functions with compact support is to

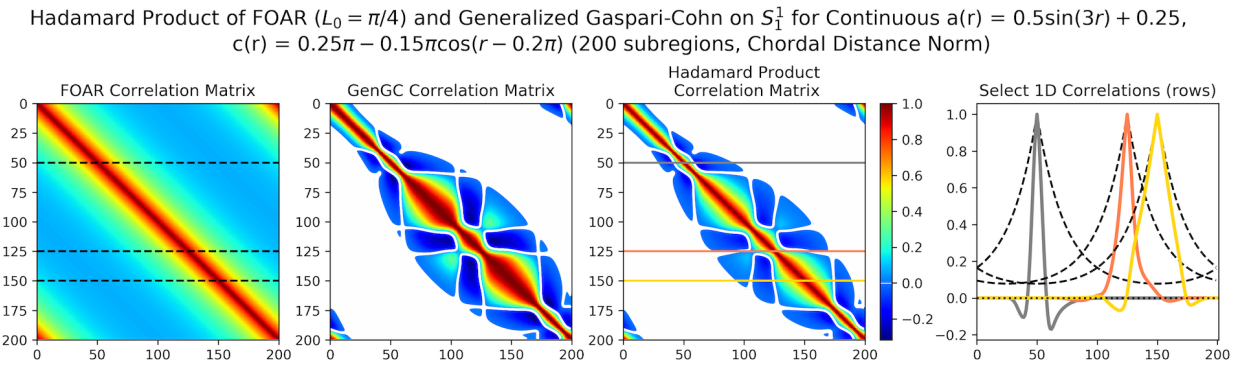

Figure 8.6: Hadamard product of the First-Order Autoregressive (FOAR) correlation function, Eq. (8.31), with correlations computed in Fig. 8.5. Left: original FOAR correlation matrix with length-scale $\mathsf{L}_0 = \pi/4$. Left-middle: GenGC correlation matrix, same as middle panel of Fig. 8.5. Right-middle: correlation matrix corresponding to the Hadamard product of the FOAR correlation matrix (left) and GenGC correlation matrix (left-middle). Right: colored lines are correlations corresponding to horizontal colored lines in right-middle panel; dashed-black lines correspond to same rows in the FOAR correlation matrix (rows 50, 125, and 150). The dashed lines are identical since the FOAR correlation function used here is isotropic.

compute the Hadamard (element-wise) product with a correlation function that does not have compact support. Figure 8.6 is an example of such a case, showing the Hadamard product of the GenGC correlation function in Fig. 8.5 with the First Order Autoregressive (FOAR) correlation function. The FOAR correlation function is defined as

$$C_{FOAR}(\boldsymbol{r}, \boldsymbol{s}; \mathsf{L}_0) = e^{-||\boldsymbol{r}-\boldsymbol{s}||/\mathsf{L}_0} \tag{8.31}$$

where L_0 is commonly referred to as the length-scale (not to be confused with the correlation length L defined in Sec. 8.3.2). The FOAR correlation function is not differentiable due to the cusp at zero separation and never becomes zero. We see in Fig. 8.6 that taking the Hadamard product of the correlations of Fig. 8.5 and the FOAR correlation with $\mathsf{L}_0 = \pi/4$ produces correlations that have compact support and a mixture of features from both functions. The cusp-like behavior of the FOAR is preserved by the Hadamard product, but it is now asymmetric about zero separation, contains negative correlations, and is compactly supported. The Hadamard product of GenGC with the FOAR combines the unique features from each function into one correlation function, which can be a powerful tool for covariance modeling, particularly for applications in which cusp-like behavior can be expected (e.g., Ménard et al., 2000; Ménard and Chang, 2000; Waller et al., 2016).

8.4.2 Two-Dimensional Example

In addition to one-dimensional examples, we consider a simple two-dimensional case motivated by the example presented in Paciorek and Schervish (2006). We construct correlations over the state of Colorado, which we divide into two regions: a western mountainous region with short cut-off lengths, and an eastern plains region with long cut-off lengths. Figure 8.7 plots the corresponding a and c fields as functions of latitude and longitude, where both fields are constructed using hyperbolic tangent (as we did in Fig. 8.2) and are latitudinally invariant. These are relatively homogeneous fields in both the mountainous and plains regions, with a smooth, sharp but continuous, transition between the two.

We consider correlations at three different locations, indicated by the marked points in Fig. 8.7, and compare four different cases of GenGC depending on the choice of a and c, Figs. 8.8–8.10. Panels (a) of Figs. 8.8–8.10 plot the correlations constructed using the GC99 function (Eq. 4.10 of GC99), where GenGC is evaluated with c fixed at the value at the center location and a fixed at one-half. Panels (b) plot correlations constructed using the G06 function (Eq. 33 of G06), which is GenGC with c constant (as in panel a) but with a varying according to the background field in the left panel of Fig. 8.7. Panels (c) evaluate GenGC where a is fixed at one-half, and c is allowed to

vary according to the background field in the right panel of Fig. 8.7. Finally, panels (d) are the full GenGC correlation function which uses the full c and a background fields shown in Fig. 8.7. These examples highlight the anisotropy gained by letting a and c vary over the surface of the Earth and illustrate how this affects the correlation function when compared with the case where either one or both are fixed.

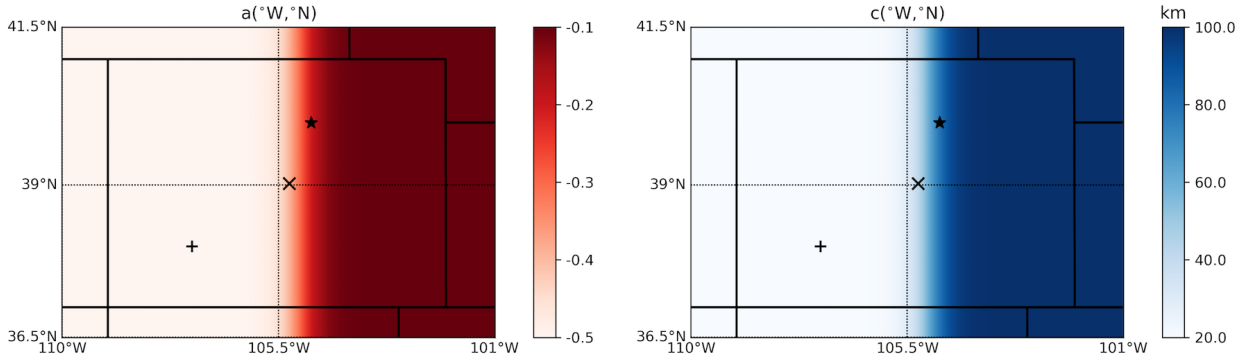

Figure 8.7: Fields a (left) and c (right) as functions of latitude and longitude over the state of Colorado, plotted on the surface of the Earth using a Mercator projection. Both fields are latitudinally invariant and are constructed in the longitudinal direction using hyperbolic tangent: $a(lon) = -0.2\tanh(10(lon - 105.05°W)) - 0.3$, $c(lon) = 40\tanh(10(lon - 105.05°W)) + 60$; c is expressed in km. Marked points indicate locations where the two-dimensional correlations are constructed: $+$ for Fig. 8.8, $\times$ for Fig. 8.9, and $\star$ for Fig. 8.10.

The western location in Fig. 8.7 (marked $+$) is in a region of relatively constant values of a and values of c that are small. As a result, the correlations in Fig. 8.8 are nearly identical, as the long-range behavior of a and c have little influence because of the very short local cut-off lengths. Panels (a) and (c) are nearly identical but differ slightly from panels (b) and (d) in their interior behavior because each pair uses different a values for their construction. The correlations in this case are all fairly isotropic and decay to zero relatively quickly beyond the 0.05 contour.

Moving eastward to the central marked point in Fig. 8.7 ($\times$ point), we obtain the correlations shown in Fig. 8.9. This location sits on the western edge of the steep gradients in a and c, which heavily influences the behavior of the correlations that allow for a and/or c to vary (panels b – d). The GC99 function in panel (a), as expected, is isotropic and only differs from panel (a) in Fig. 8.8 in how fast it decays to zero. Panels (b) and (d) have the most strikingly different interior behaviors

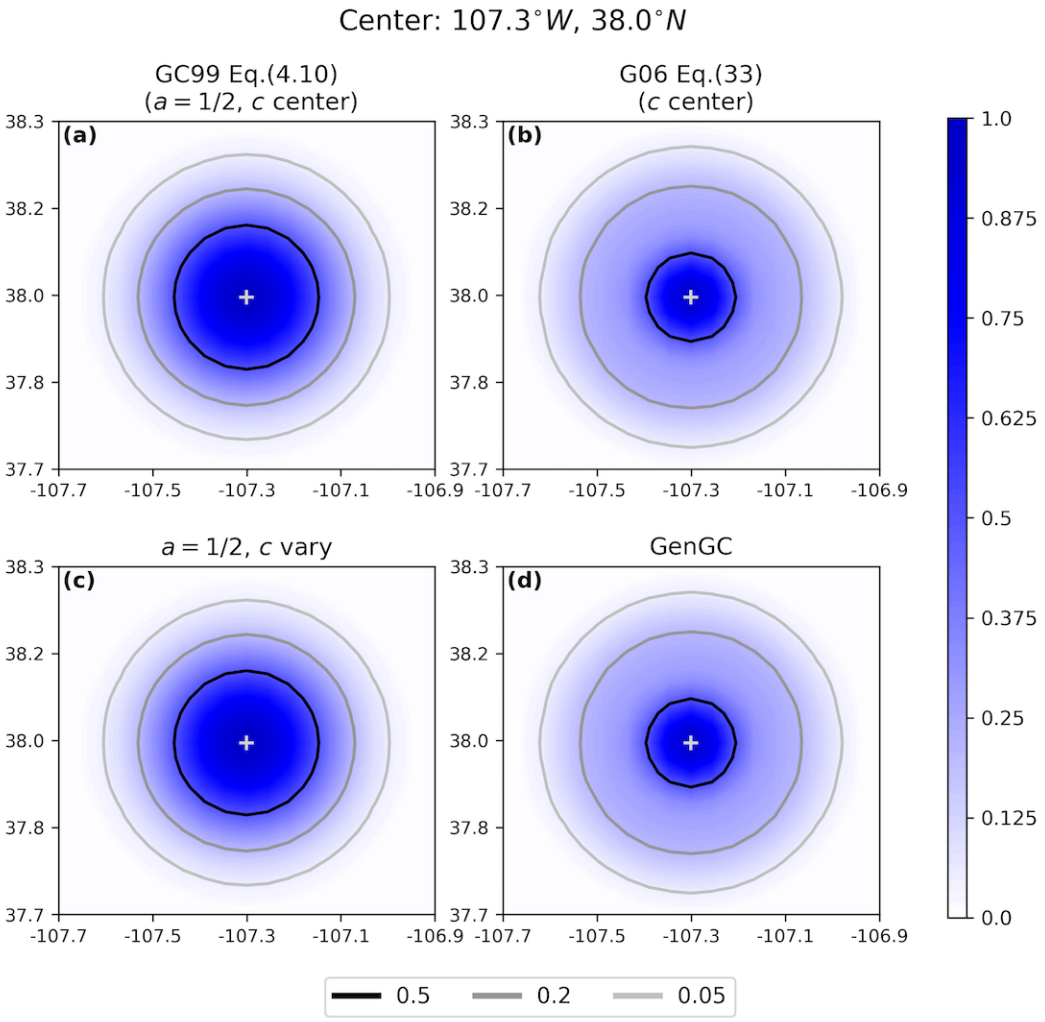

Figure 8.8: Two-dimensional correlations with respect to $(107.3°W, 38°N)$ (plus-sign + in Fig. 8.7) given background fields in Fig. 8.7. Panel (a) is the GC99 function with $a = 1/2$ fixed and the value of c taken to be the value at $(107.3°W, 38°N)$. Panel (b) is the G06 function using the a field of Fig. 8.7 and c at $(107.3°W, 38°N)$, fixed. Panel (c) is GenGC with $a = 1/2$ fixed and c varying according to Fig. 8.7, while panel (d) is the full GenGC using both a and c fields from Fig. 8.7. The colorbar indicates the value of the correlations, along with selected values marked by contour lines defined in the legend. Correlations are plotted on the surface of the Earth using a Mercator projection.

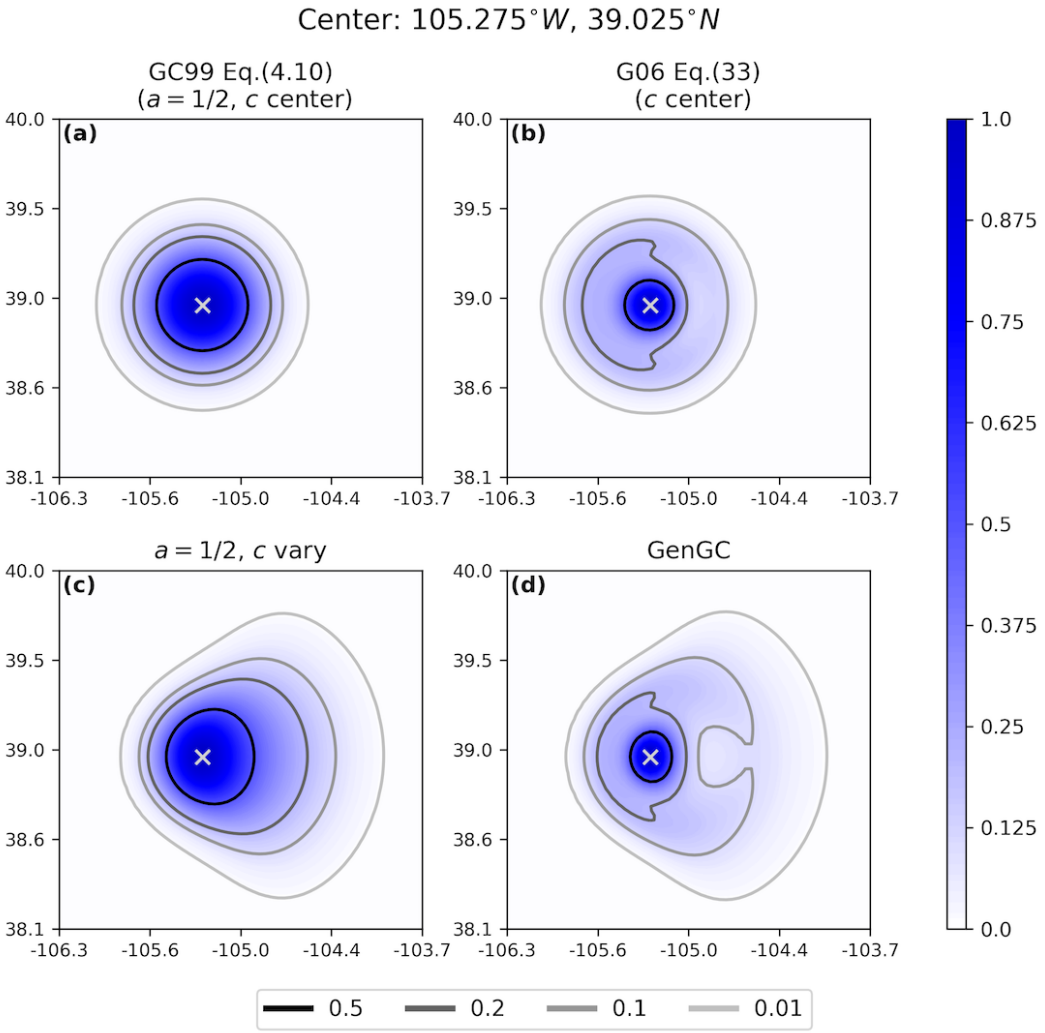

Figure 8.9: Same as Fig. 8.8, evaluated at $(105.275°W, 39.025°N)$ ($\times$ in Fig. 8.7).

due to the spatial dependence of a, where structures are no longer radially symmetric. Panels (c) and (d) capture the changes in c, noting that to the west of the center location cut-off lengths are small and increase when moving east. This is reflected in the eastward expansion of the correlations in panels (c) and (d), whereas the correlations in panels (a) and (b) go to zero within a fixed radius determined by the value of c at the center. Examining Fig. 8.9 in its totality shows the step-by-step progression of the GC99 function to its most general form of GenGC in panel (d).

Correlations in Fig. 8.10 are evaluated at the starred location in Fig. 8.7 ($\star$ point) on the eastern edge of the steep gradient in both the a and c fields. In this region, we obtain correlations that become negative for both the G06 function in panel (b) and GenGC in panel (d), which are not captured when a is fixed at one-half in panels (a) and (c). As in Fig. 8.9, we see the influence of the gradient in the cut-off length field c, where correlations that allow the cut-off length field to vary (i.e., panels c and d) exhibit expansion eastward towards larger cut-off lengths and shrinkage westward for shorter cut-off lengths, relative to the fixed cut-off lengths in panels (a) and (b).

8.5 Concluding Remarks

This chapter presents a generalized version of the Gaspari and Cohn (1999) compactly-supported, piecewise rational correlation function (the GC99 function) and its subsequent extension in Eq. (33) of Gaspari et al. (2006) (the G06 function). The function presented here, which we refer to as the Generalized Gaspari-Cohn (GenGC) correlation function, extends the GC99 and G06 functions to let the cut-off length c and parameter a vary over arbitrary subregions $\Omega_1, \Omega_2, ..., \Omega_m$, such as individual grid cells, that partition the domain Ω on which GenGC is defined. As a consequence, the GenGC correlation function is a piecewise-defined, scalar correlation function that can be inhomogeneous and anisotropic while remaining compactly-supported. In particular, by allowing the cut-off length c to vary over the different subregions, GenGC can be compactly-supported on regions other than spheres of fixed radius with more direct control of the correlation lengths, introducing features that the GC99 and G06 functions lack.

The general framework to construct GenGC and other scalar covariance functions that allows

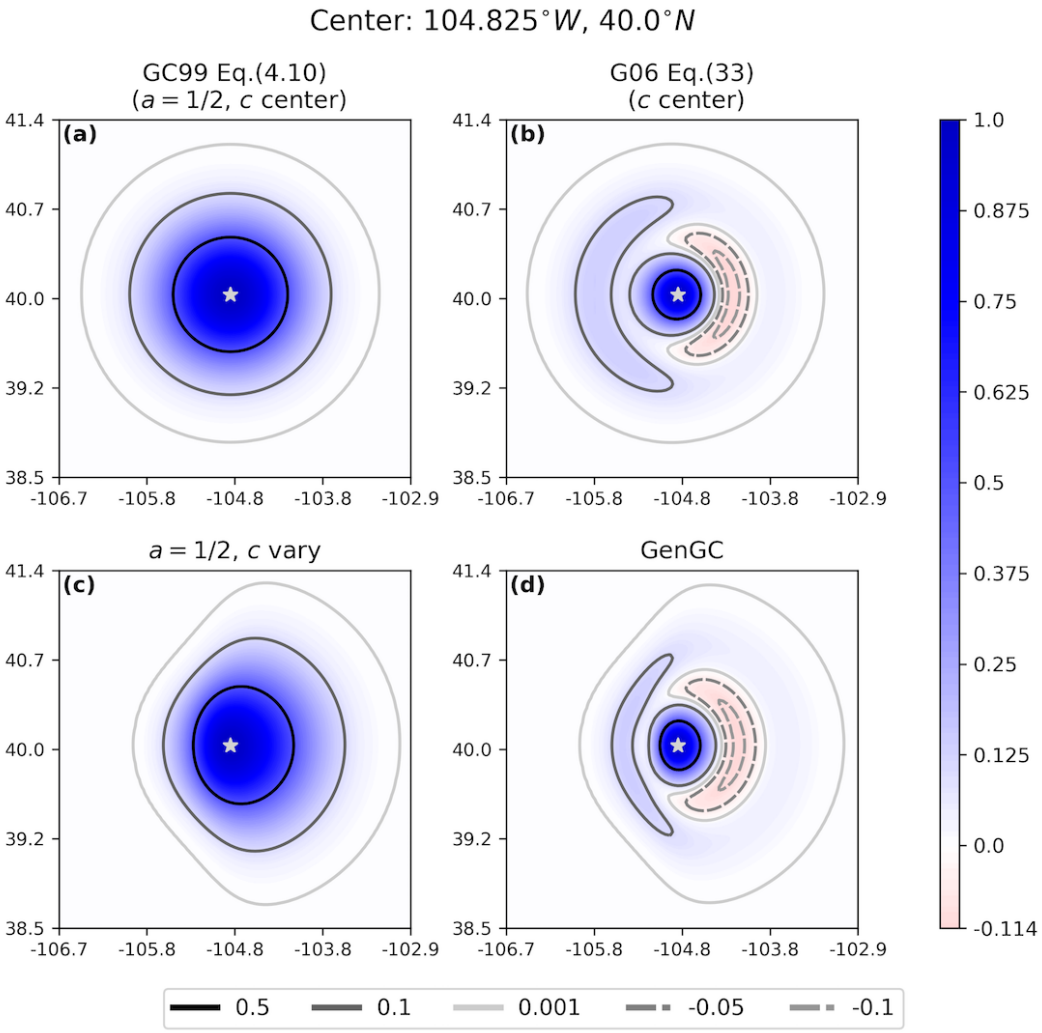

Figure 8.10: Same as Fig. 8.8, evaluated at $(104.825°W, 40°N)$ ($\star$ in Fig. 8.7).

parameters to vary over the domain is built from matrix covariance functions of vector random fields. As we describe in Sec. 8.2.1, matrix covariance functions provide the initial structure in which we can introduce inhomogeneity and anisotropy, and from these matrix functions we can extract the piecewise-defined scalar covariance function evaluated in practice. In this way, we build the GenGC correlation function from the convolution of compactly-supported, radially-symmetric functions, as done in Sec. 8.2.2.

Since GenGC is defined piecewise, continuity over the full domain is not always guaranteed, as discussed in Sec. 8.3.1. In practice, choosing the subregions $\Omega_1, \Omega_2, ..., \Omega_m$ to be the grid cells of the discretized domain and allowing a and c to vary continuously over the full domain will ensure the GenGC correlation function remains continuous, in an appropriate sense. The one- and two-dimensional examples in Sec. 8.4 are constructed in this manner. These examples are intended to illustrate the flexibility of the GenGC correlation function while maintaining its compact support. They suggest potential areas of application where this function may be useful, such as for localization or covariance modeling. In particular, the GenGC correlation function can play an important role in reconstructing covariances from correlation length fields and variance fields, which will be described in the next chapter.

The increased flexibility of the GenGC correlation function compared to the GC99 and G06 functions is reflected in its correlation length, Eq. (8.27), which varies over the domain Ω as a_k and c_k vary over the different subregions Ω_k for $k = 1, 2, ..., m$. This correlation length is a scalar value, independent of the dimensionality of Ω, due to the radial symmetry of the functions h_k in Eq. (8.18) used to construct GenGC. This implies that for Ω in two and three dimensions, GenGC is still *locally* isotropic, in the sense that in each subregion Ω_k, there is only one, scalar correlation length defined, independent of direction. To be fully anisotropic in two and three dimensions requires correlation lengths that depend on direction (for example: Cohn, 1993, Sec. 5b, Eqs. 5.34a,b; local correlation tensors defined in Michel et al., 2016, Sec. 2.4). Though the GenGC correlation function is a generalization of the GC99 and G06 functions, further work is needed to construct a version that is fully anisotropic in two and three dimensions.

Chapter 9

Implementation of Local Covariance Evolution and Preliminary Results

Chapter 7 introduces the theoretical framework for local covariance evolution (LCE), followed by Ch. 8, which derives a new parametric correlation function that can be used to reconstruct correlations from variable correlation length fields. In this chapter, we will now implement LCE in one space dimension. Since the implementation of LCE depends on the parametric correlation function, in this chapter we will develop the methodology for two parametric correlation functions: the Generalized Gaspari-Cohn (GenGC) correlation function defined in Ch. 8, and the First-Order Autoregressive (FOAR) correlation function. In addition, we will implement LCE for the Hadamard product of these two parametric correlations. This yields a total of three LCE methodologies, in which their implementation is described and preliminary results for each case are presented.

9.1 Methodology

We begin by discretizing the domain, taken to be the unit circle $\mathbb{S}^1_1$, into open intervals Ω_k for $k = 1, 2, \ldots, m = 200$ of uniform width of $\pi/100$. The grid point $x_k \in \mathbb{S}^1_1$ is taken to be at the midpoint of Ω_k for $k = 1, 2, \ldots, m$.

Reconstructing covariances using LCE consists of two parts, mirroring the decomposition of the covariance P into the standard deviation σ and correlation C in Eq. (7.1). The first part of LCE is the evolution of the variance, σ^2. This is done by solving the variance equation, Eq. (1.8), either analytically or numerically. The second part is the evolution and reconstruction of the correlations. The covariance is then recovered by rescaling the correlations by the variance.

Evolving and reconstructing the correlations for LCE is more involved than evolving the variance since it depends on the particular parametric correlation function. Therefore in the next three sections, we will see how to evolve and reconstruct the correlations for three different cases.

9.1.1 Generalized Gaspari-Cohn (GenGC)

Since the standard Gaspari-Cohn correlation function (referred to as GC99 in Ch. 8) is compactly supported and frequently used in the data assimilation community, implementing LCE for the Gaspari-Cohn correlation function is a natural choice. This motivates the derivation of the GenGC correlation function in Ch. 8. By allowing both parameters a_k and c_k to vary over the domain, the correlation length L_k of the GenGC correlation function has the full flexibility needed for LCE in one space dimension.

The correlation length L_k of GenGC defined in Eq. (8.27) depends on both parameters a_k and c_k. Therefore, the correlation length field alone is not sufficient to determine both a_k and c_k. Even if both the correlation length L_k and cut-off length c_k for $k = 1, 2, \ldots, m$ are known, the restrictions on the values of a_k for fixed L_k and c_k (as discussed in Ch. 8.3.2) make choosing an appropriate value of a_k difficult. One option would be to include the dynamics for a higher-order derivative of the correlation at zero separation, C_4 for example, in addition to solving for the correlation length. Solving for a_k with an additional equation for C_4, however, is algebraically complicated. For fixed L_k and c_k, a_k satisfies a quartic equation with even more restrictions due to the nonlinear relationship of a_k with the ratio L_k/c_k.

Rather than solving for both a_k and c_k over the full spatial domain, we will fix the values of $a_k = a$ to be constant for $k = 1, 2, ..., m$. By assuming the parameter $a_k = a$ is constant over space and time, we are now able to compute the cut-off length field c_k for $k = 1, 2, \ldots, m$ from the evolved correlation lengths. For cases where the shape of the correlations depends primarily on the cut-off length or when the values of a are in a regime such that the shape of GenGC does not vary (see Fig. 8.4, for example), this assumption is sufficient. We will see the impacts of this assumption in the examples presented in Sec. 9.2.

The implementation of LCE with the GenGC correlation function must be done using the framework described by Thm. 2 in Ch. 8.3.1 to ensure the resulting correlations are continuous. Since $a_k = a$ for $k = 1, 2, \ldots, m$ is constant over the domain, a can be directly evaluated in GenGC. The values of c_k, however, should be evaluated as the average value of the function $c(x, t)$ over the subregion Ω_k, Eq. (8.25). Since the grid points x_k are the midpoints of the subregions Ω_k, simply evaluating $c_k = c(x_k, t)$ is sufficient, as it is equivalent to computing Eq. (8.25) with the midpoint rule over one subinterval equal taken to be the full subregion Ω_k. Evaluating Eq. (8.25) to determine c_k using more than one subinterval in Ω_k could be done using appropriate mesh-refinement methods.

The general implementation of LCE for the GenGC correlation function with fixed $a_k = a$ is summarized by the following three steps:

(1) Define the initial correlation $C_0(x_1, x_2)$ using GenGC for initial cut-off values c_k^0 and $a_k^0 = a$ fixed for $k = 1, 2, \ldots, m$, where the superscript zero indicate the values of c_k and a_k at the initial time t_0. The initial correlation lengths L_k^0 are computed from $a_k = a$ and c_k^0 for $k = 1, 2, ..., m$ using Eq. (8.27).

(2) At time $t > t_0$, the correlation length field $L(x_k, t)$ is evaluated by solving Eq. (5.30) with the initial condition from step (1). With fixed $a_k^t = a$, $k = 1, 2, \ldots, m$ and correlation lengths $L(x_k, t)$, the cut-off length field $c(x_k, t)$ is computed using Eq. (8.27). The values c_k^t are then computed from $c(x_k, t)$ using Eq. (8.25), where the superscript t denotes the value of c_k at time t.

(3) GenGC is evaluated using the fixed value of $a_k^t = a$ and computed cut-off length field c_k^t for $k = 1, 2, \ldots, m$ at time t.

Remark 5. *In practice, one would likely not solve the correlation length PDE given by Eq. (5.30) directly. Instead, one would solve the PDE for* $1/L$ *since it satisfies the continuity equation and a numerical scheme that will preserve its continuum conservation properties can be implemented.*

9.1.2 First Order Autoregressive (FOAR)

Another parametric correlation function that can be used for LCE is the First-Order Autoregressive (FOAR) correlation function, first introduced in Eq. (3.6) in Ch. 3.1. Its definition is repeated here with slightly different notation,

$$C_{FOAR}(x_1, x_2) = e^{-||x_1 - x_2||/\mathsf{L}}, \tag{3.6 revisited}$$

where L is the length-scale parameter and $||x_1 - x_2||$ is the normed distance between points x_1 and x_2, which is taken to be chordal distance on the unit circle defined by Eq. (3.5). The FOAR correlation function has a cusp at $x_1 = x_2$, therefore its derivatives at $x_1 = x_2$ are discontinuous. The dynamics for LCE in this case are therefore defined in Ch. 7.1.2.

We can implement LCE using the FOAR correlation function by making use of the generalization in Eq. (23) of Gaspari et al. (2006) referred to as the "multi-level FOAR." The multi-level FOAR generalizes Eq. (3.6) to allow the length-scale parameter $\mathsf{L} = \mathsf{L}_k$ to vary over the spatial domain, similar to how a_k and c_k vary in the GenGC correlation function. For a given length-scale field L_k, Eq. (23) of Gaspari et al. (2006) can be used to reconstruct FOAR-like correlations that can be inhomogeneous and anisotropic.

At zero separation (i.e., $x_1 = x_2$), the multi-level FOAR correlation function reduces to Eq. (3.6). Therefore, we can compute C_1 from Eq. (3.6) and find it only depends on the length-scale parameter L,

$$C_1(x, t) = -\frac{1}{\mathsf{L}(x, t)}, \tag{9.1}$$

where we now allow the length-scale parameter L to vary over space and time. Using the relation between C_1 and the length-scale L in Eq. (9.1), we can derive an explicit PDE for the FOAR length-scale,

$$\begin{aligned} \mathsf{L}_t + v\mathsf{L}_x - v_x\mathsf{L} = 0, \\ \mathsf{L}(x, t_0) = \mathsf{L}_0(x), \end{aligned} \tag{9.2}$$

which happens to satisfy the same dynamics as the correlation length L in the continuously differentiable correlation case.

Like the GenGC correlation function, the multi-level FOAR correlation function is constructed by convolving functions from a collection of radially-symmetric functions on $\mathbb{R}^3$ (see the proof on pp. 1821 – 1822 of Gaspari et al., 2006). Therefore, to ensure that the resulting correlation is continuous, the values of L_k used in Eq. (23) of Gaspari et al. (2006) should be evaluated using the average value of $\mathsf{L}(x,t)$ over the subregion Ω_k. In fact, Thm. 2 can be generalized for arbitrary parameters and then applied to the multi-level FOAR correlation function to ensure continuity.

The implementation of LCE for the FOAR correlation function is more straightforward than the GenGC case since the length-scales $\mathsf{L}(x,t)$ can be solved directly. Given initial length-scales $\mathsf{L}_k^0 = \mathsf{L}_0(x_k)$, we can solve Eq. (9.2) for $\mathsf{L}(x,t)$ at some time $t > t_0$, then evaluate L_k^t using Eq. (8.25) for $\mathsf{L}(x,t)$ in place of $c(x,t)$. Finally, we can use Eq. (23) of Gaspari et al. (2006) to recover the full correlations.

9.1.3 Hadamard Product of GenGC and FOAR

Though the cusp-like behavior of the FOAR correlation function is useful for certain geophysical applications, the resulting correlation matrix is dense since the exponential is everywhere nonzero. Therefore, considering correlations built from the Hadamard product of the FOAR correlation function with a compactly-supported correlation function, such as GenGC, is computationally beneficial, particularly in high-dimensional settings. We can implement LCE for initial correlations C_0 that are the Hadamard product of the FOAR correlation and GenGC correlation following the analysis presented in Ch. 7.1.3.

Following the notation of Ch. 7.1.3, assume that F_0 in Eq. (7.11) is the FOAR correlation function (standard or multi-level) and G_0 is the GenGC correlation function. Since the correlation C remains the product of these two correlation functions for all time, $C(x_1, x_2, t) = F(x_1, x_2, t)G(x_1, x_2, t)$, it is sufficient to evolve the FOAR and GenGC correlations independently, then reconstruct the full correlation as their product. The implementation in LCE in this case

reduces to implementing Sec. 9.1.1 and Sec. 9.1.2 simultaneously to approximate G and F, respectively.

9.2 Preliminary Results

The primary goal of LCE is to evolve covariances in a way that mitigates the inaccurate variance propagation we observed during full-rank covariance propagation. This is accomplished by evolving the variance directly, which can be done to any desired accuracy given the dynamics (here by Eq. 1.8) and an appropriate numerical scheme. To determine if LCE is an effective means of covariance propagation then requires testing the its ability to approximate the full correlation fields. This section presents preliminary results that implement the evolution and reconstruction of full correlation fields as described in this chapter.

For these experiments, consider the same one-dimensional problem described in Ch. 3 for the discretization described at the beginning of Sec. 9.1. The exact solutions to the correlation C, correlation length L, and length-scale L are known and their solutions are given explicitly in Appen. B.2. For each of the three correlation function cases (GenGC, FOAR, and the Hadamard product of the two), the initial correlation C_0 is taken to be isotropic, meaning the initial correlation length $L_0(x,t)$ or length-scale $\mathsf{L}_0(x,t)$ is constant over the domain. Using the exact solutions to either the correlation length or length-scale PDE, we can then compute $L(x,t)$ and $\mathsf{L}(x,t)$ at time $t = T$, then reconstruct the correlations following the procedures outlined in Secs. 9.1.1 – 9.1.3.

In Ch. 3, we observe that full-rank propagation produces the most significant errors in both the variance and correlations when initial correlation lengths are small (see Figs. 3.5 – 3.6 for example). The first set of LCE experiments, therefore, targets these regimes, where we will consider cases where the initial correlation length and length-scale parameters are relatively small. Figure 9.1 contains examples of the correlation matrices obtained in the LCE experiments for each of the three parametric correlation function cases, where the discrete values $c_k^0 = 0.25$ and $\mathsf{L}_k^0 = 0.25$ for all $k = 1, 2, \ldots, m = 200$. Each row shows the exact correlation matrix (left), the LCE-approximated correlation matrix (middle), and selected rows from each to illustrate the explicit one-dimensional

correlations (right).

We see for the cases in Fig. 9.1 that the LCE-approximated correlations reconstruct the exact correlation fields accurately and eliminate the major problems observed in the full-rank propagation examples of Ch. 3. Recall that the full-rank propagation results in diffusion of information off the diagonal, or numerical dispersion in the Crank-Nicolson case, that is particularly problematic when correlation lengths shrink in regions of convergence (see Figs. 3.5 – 3.6, for example). In the region around row 150, where full-rank propagation performs quite poorly due to the extremely sharp gradient near the diagonal, LCE captures the correlations quite well. Overall, the LCE-approximated correlations are quite accurate and are able to capture anisotropy simply from the evolved correlation length or length-scale fields.

Though LCE is "local" in the sense that the evolved parameters are derived from a local approximation of the correlation, we see in Fig. 9.1 that correlations are well approximated away from the diagonal. The FOAR case (middle row) has the largest errors away from zero separation, but these errors are still relatively small. The GenGC and Hadamard product cases (top and bottom rows of Fig. 9.1, respectively), capture the full correlations very well, indicating that for these dynamics fixing a as constant is sufficient to recover the cut-off length fields from the evolved correlation lengths.

Though most geophysical applications will consider cases where the initial cut-off lengths c_k^0 for GenGC or length-scale L_k^0 for the FOAR are relatively small over the different subregions, it is worth considering other values of c_k^0 and L_k^0 to test the ability of LCE to reconstruct correlations. In Fig.9.2, we consider the case for $c_k^0 = 1$ and $\mathsf{L}_k^0 = 1$ for all $k = 1, 2, \ldots, 200$, in which the full-rank propagation examples in Ch. 3 have minimal errors in the variance and correlations.

The LCE-reconstructed correlations in Fig. 9.2 have more notable errors than those in Fig. 9.1 for smaller initial correlation lengths and length-scales. For these long initial correlation lengths and length-scales, it is the region around row 50 in which the LCE-approximated correlations have the largest errors. Interestingly, around row 150 where full-rank covariance propagation is generally poor due to the sharp gradient near the diagonal, the LCE-approximated correlations

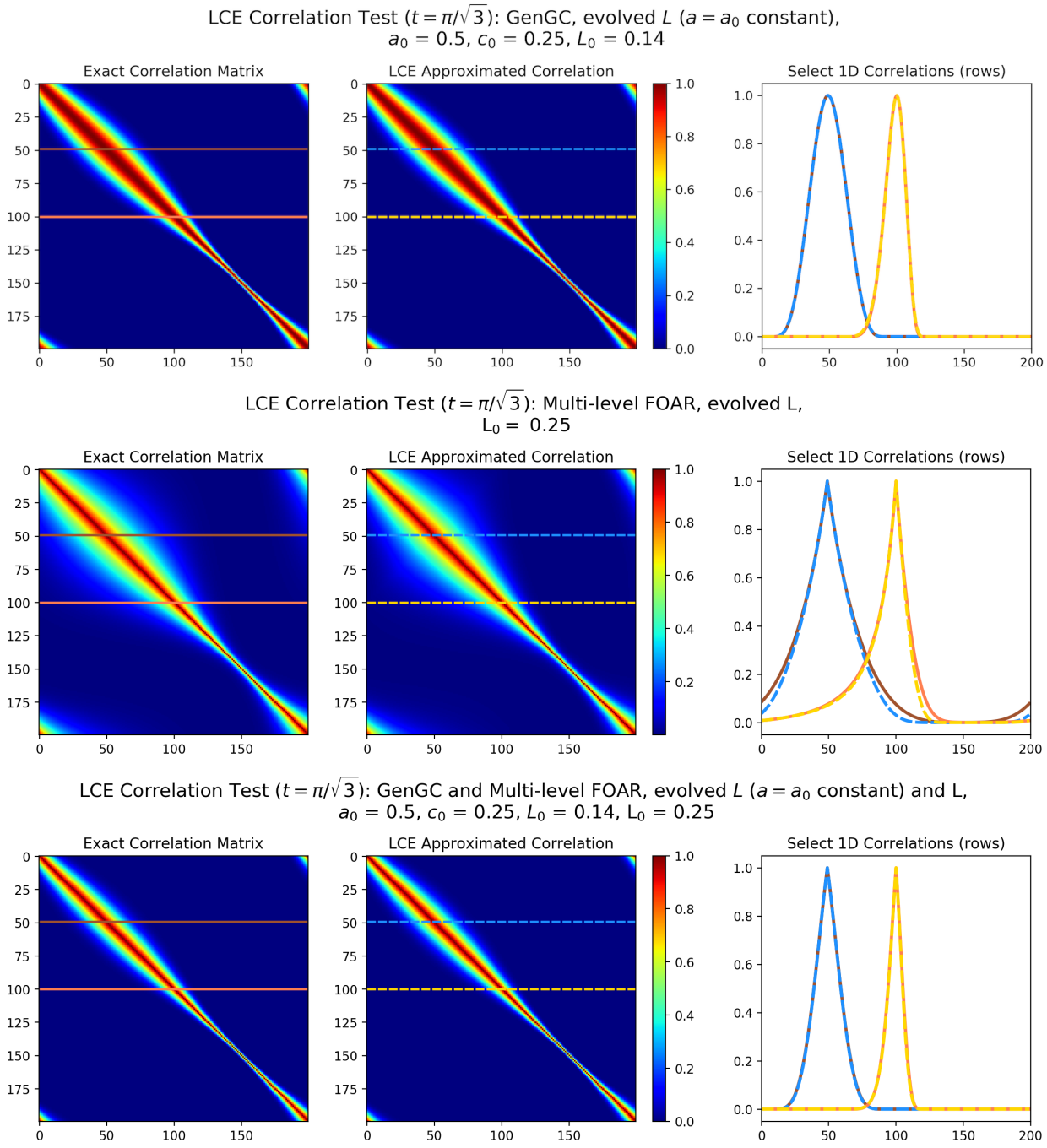

Figure 9.1: Preliminary results for the LCE experiments for three correlation functions: GenGC (top row), FOAR (middle row), Hadamard product of GenGC and FOAR (bottom row) for $c_k^0 = 0.25$ and $\mathsf{L}_k^0 = 0.25$ for all $k = 1, 2, \ldots, 200$. In each row title, italicized L denotes the correlation length for GenGC while roman L denotes the correlation length scale for the FOAR correlation function. In each row, the left panel is the exact correlation matrix, the middle is the LCE-approximated correlation matrix, and the right panel are rows 49 and 100 corresponding to the horizontal lines in each correlation matrix. All matrices are evaluated at time $t = \pi/\sqrt{3}$, corresponding to half a time period.

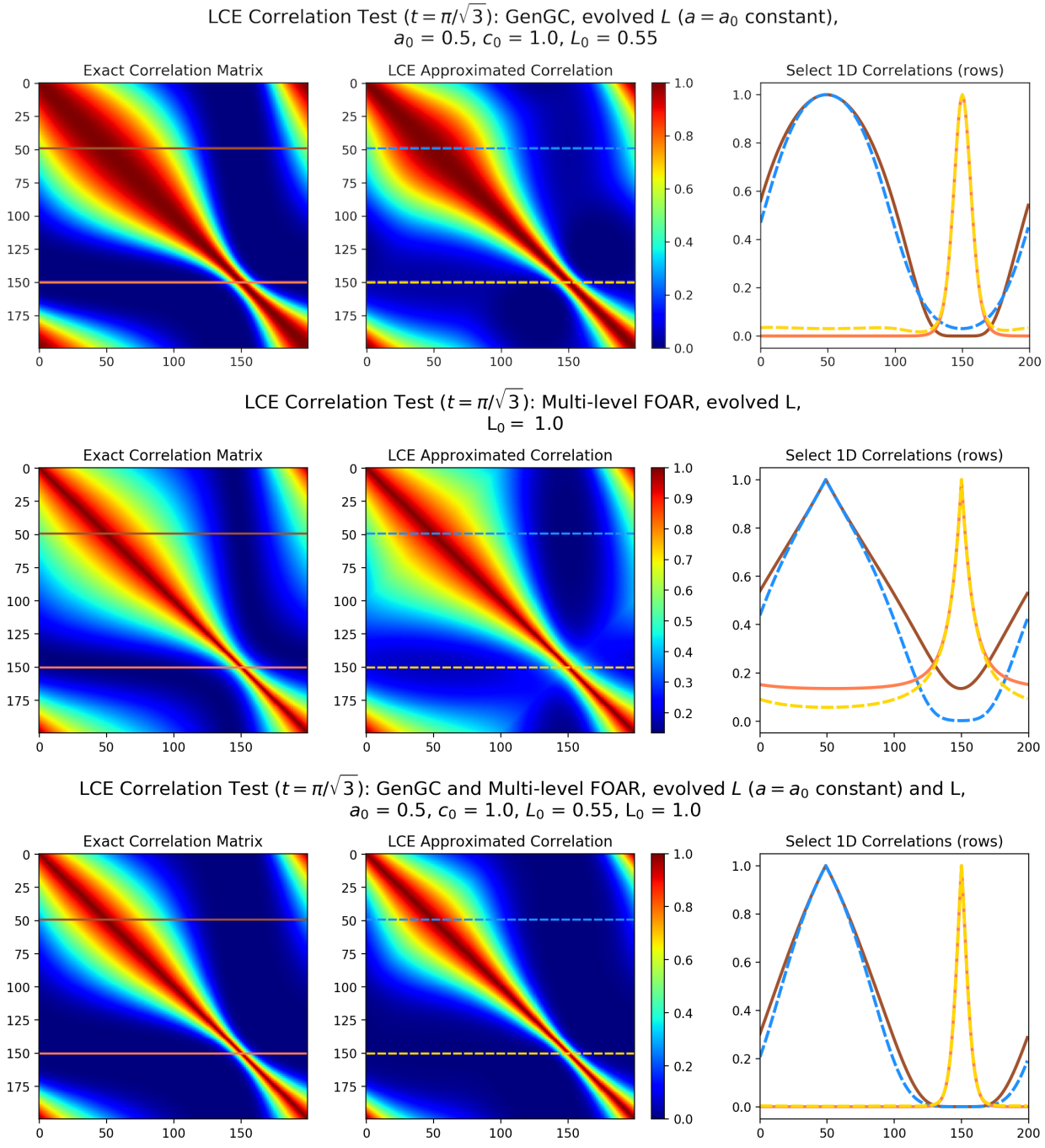

Figure 9.2: Same as Fig. 9.1 for $c_k^0 = 1$, $L_k^0 = 1$ for all $k = 1, 2, \ldots, 200$. The left panels show rows 49 and 150.

are still accurate. Near the correlation diagonal, the LCE-approximated correlations are accurate, though further from the diagonal the errors start to increase, indicating the extent of the "local" approximation. Further investigation into the LCE approximated correlations for larger initial correlation lengths and length-scales is the subject of future work.

Overall, the preliminary results for the LCE-approximated correlations suggest that LCE is a viable alternative to standard methods covariance propagation in advective systems, particularly when correlation lengths become small or sharp gradients form near the covariance diagonal. LCE largely eliminates the errors observed in the correlations produced during standard covariance propagation and is a significant cost reduction for evolving correlations relative to full-rank propagation and even to low-rank approximations. This work establishes the foundation for further investigation into LCE, which is addressed in more detail in Ch. 10.3.

Chapter 10

Conclusions and Future Work

10.1 Summary

In this thesis, the fundamental cause of inaccurate variance propagation associated with standard methods of covariance propagation has been investigated. Through careful analysis of the continuum covariance dynamics for advective systems, Ch. 2 identifies a discontinuous change in dynamics along the hyperplane $\boldsymbol{x}_1 = \boldsymbol{x}_2$ in the limit as correlation lengths tend to zero. The impacts of this discontinuity on discrete covariance propagations is then observed in the following chapters. Chapter 3 illustrates in a simple, one-dimensional example with a known exact solution, that the diagonals produced during full-rank covariance propagation are smooth, but wholly inaccurate. The errors observed result in both spurious loss and gain of variance, which are a manifestation of inaccurate variance propagation more so than numerical dissipation. One may surmise these errors can be mitigated by choosing an appropriate numerical scheme. However, Ch. 4 demonstrates that even when a numerical scheme is chosen to suit the continuum conservation properties, inaccurate variance propagation is still a considerable problem. Thus, inaccurate variance propagation is not simply a problem of numerics per se, but rather is due to the discontinuous change in the continuum covariance dynamics.

Motivated by the numerical results in Chs. 3 and 4, Ch. 5 performs an error analysis of full-rank covariance propagation in order to identify the explicit sources of inaccurate variance propagation. Although the original intent of Ch. 5 was to derive a correction to full-rank propagation, the results in Ch. 5 provide important insights into variance propagation and clearly explain

the behavior observed in Chs. 3 and 4. The development of a correction, motivated by the results of Ch. 5, is discussed in Ch. 6. Though Chs. 5 and 6 suggest that correcting full-rank covariance propagation may be quite challenging, the findings in Ch. 5 provide a new and important perspective on covariance propagation. These results are particularly relevant to the data assimilation applications that include advective dynamics, such chemical tracer transport models, for example.

Chapters 2 – 6 demonstrate that standard methods of variance propagation, namely that associated with Eq. (1.3), are inherently inaccurate. This motivates the investigation into an alternative approach to covariance propagation that bypasses Eq. (1.3) entirely. Chapters 7 and 9 present an alternative method of covariance propagation that can directly address inaccurate variance propagation. Termed "local covariance evolution" (LCE), this method is different from current approaches in that it is derived from the continuum covariance dynamics and evolves the variance directly. The preliminary results presented in Ch. 9 demonstrate that LCE can accurately reconstruct correlations in regions where full-rank covariance propagation behaves poorly. In addition, LCE is computationally efficient and therefore tractable in data assimilation applications. To implement LCE requires parametric correlation functions that allow for variable parameters, therefore Ch. 8 introduces a new correlation function to then use for reconstructing correlations in Ch. 9.

In its totality, my thesis provides a new perspective on covariance propagation, which is a mathematically significant problem that is highly relevant to data assimilation. My thesis shows that the spurious loss of variance often observed during covariance propagation in advective systems is more than just loss of variance. Rather, the underlying problem is inaccurate variance propagation, caused by the discontinuous behavior of the continuum dynamics coupled with the structure of discrete covariance propagation. To uncover the root of this problem requires returning to the underlying mathematics that govern both the continuum and discrete covariance propagation. From this perspective, my thesis provides new insights that help us better understand covariance propagation for data assimilation applications.

10.2 Main Contributions

Issues with variance propagation associated with standard methods of discrete covariance propagation are investigated in this thesis. In data assimilation practice, this problem arises as spurious loss of variance and is often interpreted and addressed in discrete terms. My thesis demonstrates the importance of returning to the continuum formulation of covariance propagation, in particular to understand the propagation of covariance matrices in space and time. My thesis shows that the spurious loss of variance associated with covariance propagation in advective systems is not simply caused by reduced-rank representations or the numerical method. Rather, the fundamental cause of this loss of variance is rooted within the continuum dynamics itself. Without returning to the continuum formulation of covariance propagation, the source of the observed spurious loss of variance in advective systems cannot be properly identified.

Below is a summary of the original contributions and new perspectives that address inaccurate variance propagation in advective systems:

- Chapter 2 identifies the discontinuous change in the continuum covariance dynamics along the $\boldsymbol{x}_1, \boldsymbol{x}_2$-hyperplane in the limit as correlation lengths tend to zero. This result is mathematically novel and highly relevant to geophysical data assimilation applications.

- Through a series of numerical experiments, Chs. 3 and 4 shows that inaccurate variance propagation is a considerable problem during full-rank covariance propagation in discrete space. The inaccurate variances are due to the tendency of discrete, full-rank covariance propagation to approximate the incorrect continuum dynamics.

- The error analysis in Ch. 5 explicitly demonstrates that inaccurate variance propagation observed during full-rank covariance propagation is not caused by numerical dissipation or the particular numerical method employed, but is due to the nature of the continuum covariance dynamics and structure of covariance propagation itself.

- Chapters 7 and 9 propose and implement an alternative method of covariance propagation

in one space dimension. This method is derived directly from the continuum covariance dynamics in an effort to mitigate inaccurate variance propagation and is computationally tractable when scaled to high-dimensional geophysical systems. Preliminary results suggest this method can eliminate the major problems associated with full-rank covariance propagation.

- Also, Ch. 8 introduces a new compactly-supported correlation function that generalizes the Gaspari-Cohn correlation function. This function is used in Ch. 9 for the alternative method of covariance propagation, and can be useful in other data assimilation applications, such as for covariance modeling and localization.

Covariance information plays a critical role in data assimilation schemes. My work uncovers the source of inaccurate variance propagation for advective systems, with the hope of providing a framework to develop tractable solutions for data assimilation applications.

10.3 Future Work

This thesis lays the groundwork for several different research directions that address inaccurate variance propagation in data assimilation schemes. Additional research directions that stem from this thesis are described below.

Continuum Covariance Propagation for Stochastic $\boldsymbol{v}$ and Systems of Hyperbolic PDEs The analysis presented in Ch. 2 considers a PDE for a single state variable q and a background velocity field $\boldsymbol{v}$ that is deterministic. There are two immediate directions in which this work can be extended. The first is to consider the case where $\boldsymbol{v}$ is stochastic. This will introduce two new covariance quantities: an auto-covariance for the velocity $\boldsymbol{v}$ and cross-covariance between the state q and velocity $\boldsymbol{v}$. One can perform an analysis of these PDEs for this case in a similar manner to that done in Ch. 2 in order to uncover the continuum covariance dynamics. Since stochastic $\boldsymbol{v}$ is often the case in physical applications, this analysis can provide further insights into behaviors observed in data assimilation applications that implement advective dynamics.

The second direction to consider is covariance propagation for systems of hyperbolic PDEs. Several questions arise in this case: what are the continuum covariance dynamics? What are the characteristic surfaces? What additional behaviors arise in the covariance dynamics that are not present in the single state variable case? There is a vast amount of analytical work that can be done here that builds upon the work established in Ch. 2. To begin, one can consider states that satisfy the standard wave equation on a periodic domain. Like the generalized advection equation, the wave equation is a hyperbolic PDE, and due to the two time derivatives, the wave equation can be rewritten as a system of two PDEs with one time derivative in each equation. Analysis of the covariance dynamics in this case is likely a tractable first step to understanding covariance propagation for systems of equations. On its own, analysis of the covariance dynamics associated with the wave equation is an important component of understanding covariance propagation for hyperbolic PDEs.

Continued Error Analysis of Full-Rank Covariance Propagation The error analysis of the full-rank covariance propagation in Ch. 5 considers two spatial discretization schemes, a first-order upwind spatial discretization and a second-order centered difference spatial discretization. Though both spatial discretizations approximate the incorrect continuum dynamics as correlation lengths become small, it is possible that there are other numerical schemes that are able to approximate the correct continuum dynamics along the covariance diagonal. Attempting to construct such schemes and performing error analysis for other numerical schemes is worth pursuing, particularly for schemes that are tractable in high-dimensional contexts. This additional error analysis has the potential to uncover numerical schemes that are better suited for covariance propagation or to further demonstrate that covariance propagation is unable to accurately propagate the variance in advective systems.

Two projects stem from the work presented in Ch. 5. The first is to investigate whether an alternative numerical scheme can be used for full-rank covariance propagation that can accurately propagate the variance when correlation lengths become small. Before determining such a scheme, one would first need to derive the conditions in which the correlation length L is bounded

from below. This information will be necessary to understand whether a numerical method can be developed that can "sense" when correlation lengths become small and adapt accordingly. Numerical methods to consider are multi-grid methods, for example, which can adapt in their spatial resolution and can yield sparse matrices.

The second project is to perform a convergence analysis on the approximated continuum diagonal dynamics derived in Ch. 5. This convergence analysis can help determine whether standard spatial discretizations for Eq. (1.3), under appropriate conditions on Δx, L, and velocity v, will converge to the correct continuum diagonal dynamics. The lower bound analysis for the correlation length L, as discussed above, will be useful here. Since the approximated dynamics presented in Ch. 5 depend on powers of the ratio $\Delta x/L$, careful analysis will be needed when considering the limiting cases for this ratio. This project will help to further our understanding of the inaccurate variance propagation caused by Eq. (1.3) and whether it can be mitigated.

Generalized Gaspari-Cohn in Two and Three Space Dimensions As mentioned at the end of Ch. 8, the Generalized Gaspari-Cohn (GenGC) correlation function is locally isotropic, meaning that in two and three space dimensions, GenGC has a single, scalar correlation length that is independent of direction and the dimension of the space. This is due to GenGC (and its predecessors in Gaspari and Cohn (1999) and Gaspari et al. (2006)) depending on the normed distance between grid points, enforcing radial symmetry that results in correlations being locally isotropic in more than one space dimension. The GenGC correlation function derived in Ch. 8 may be sufficient in certain applications. However, a fully anisotropic and inhomogeneous version in two and three space dimensions may also be needed. This generalization is particularly relevant for implementing local covariance evolution in two and three dimensions, in which the evolved correlation length fields will depend on direction and the dimension of the space.

To derive a fully generalized Gaspari-Cohn-like function for two and three space dimensions is likely more involved than the generalization presented in Ch. 8. The theory of Ch. 8.2.1, which builds scalar covariance functions from matrix covariance functions, holds for both two and three space dimension. The construction of the scalar covariance functions in Ch. 8.2, needs to be

modified so that the correlation length scales can depend on the dimension of the space. One first step would be to choose a different collection of functions for the convolutions in Eq. (8.11), for example, by altering the radial symmetry assumption on h_k. By generalizing the framework presented in Ch. 8.2.2, this project can result in the derivation of several different fully anisotropic correlation functions for two and three space dimensions, including new versions of the Gaspari-Cohn and FOAR correlation functions.

Further Development of LCE For Evolving Covariances The local covariance evolution (LCE)-approximated correlations presented in Ch. 9 are preliminary results. Further experimentation is necessary to fully test the ability of LCE to approximate correlations governed by the advection equation in one space dimension. This includes testing different initial conditions and implementing LCE for other parametric correlation functions that are relevant to the data assimilation community. In addition, more can be explored regarding the parameter a used to reconstruct correlations with the GenGC correlation function. For example, further analysis and experiments can be done to determine whether holding a constant is a sufficient assumption, or if one can derive equations that evolve a over space and time. It is possible, however, that developing LCE for $2D$ may provide better insights into developing LCE for a spatially dependent a.

For LCE to be applied in geophysical data assimilation applications, it is necessary to extend the LCE methodology to two and three space dimensions. Moving from $1D$ to $2D$ is non-trivial: correlation length fields require careful definition in $2D$, and it is possible that, unlike the $1D$ case, variance and correlation length information may not be sufficient to capture covariances in $2D$. Michel et al. (2016), for example, may be a useful starting point to define correlation length fields in two and three space dimensions. In addition, LCE in $2D$ (and in $3D$) requires a careful choice of parametric correlation functions that allow for correlation length fields that depend on direction. One approach would be to assume that the parametric correlation function is separable in the latitudinal and longitudinal components: in $2D$, for example, we can construct a correlation function in two space dimensions the product of two GenGC correlation functions, each in one space dimension. If this assumption is not sufficient, derivation of a fully anisotropic, non-separable

GenGC correlation function may be necessary.

LCE in $2D$ can then be tested for the advection equation on an isentropic surface of Earth's atmosphere using background wind fields from a numerical weather model or reanalysis data set. These experiments can then establish a LCE methodology that can be adapted for chemical tracer transport data assimilation applications, for instance. With the theory developed for $2D$, the extension to $3D$ may be more readily achieved. Once LCE for $2D$ and $3D$ is complete for advective dynamics, the next step is to consider other geophysical systems, for example the advection-diffusion equation.

A New Data Assimilation Filter Built Upon Local Covariance Evolution Local covariance evolution (LCE) establishes the mathematical foundation for a new, transformative approach to data assimilation by replacing full-rank covariance propagation and its low-rank approximation in the data assimilation cycle. To fully incorporate LCE in data assimilation, however, requires an associated measurement-update step, in which the evolved continuum quantities, such as the variance and correlation length, are updated. Current measurement updates are built to update the covariance matrix. Therefore, a new measurement update step is needed in order to update the quantities evolved from LCE. This problem will be formulated in the continuum, defining the appropriate functional spaces and necessary closure assumptions to derive the measurement update for a general class of observation operators.

10.4 Final Remark

Covariance propagation, namely the evolution of the covariance according to partial differential equations, is overlooked in both the mathematics and data assimilation communities. There is a gap in the theory of covariance propagation, particularly for the case where there is a discontinuity along the covariance diagonal. In an effort to address this gap, my thesis opens the door to a new area of mathematics where covariance propagation can be formally investigated.

There is also a gap in knowledge regarding covariance propagation in the data assimilation community that serves as the motivating application for this work. The fundamental, mathematical

nature of this thesis provides a new perspective on covariance propagation for the data assimilation community, bringing to light important issues that have gone mostly unnoticed. The results of this work are immediately impactful in chemical tracer data assimilation, whose dynamical models are governed by the advective dynamics considered here. By providing a new perspective on covariance propagation for data assimilation, this work can help to improve uncertainty quantification, which can ultimately help increase forecasting accuracy. This work, therefore, sets the stage to have a lasting impact in mathematics, data assimilation, and the broader community.

Bibliography

Anderson, J. and Lei, L. (2013). Empirical localization of observation impact in ensemble Kalman filters. Mon. Weath. Rev., 141:4140–4153.

Anderson, J. L. and Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Weather Rev., 127:2741–2758.

Barkmeijer, J., Gijzen, M. V., and Bouttier, F. (1998). Singular vectors and estimates of the analysis-error covariance metric. Q. J. Roy. Meteor. Soc., 124:1695–1713.

Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. Nature, 525:47–56.

Berner, J., Achatz, U., Batté, L., Bengtsson, L., Cámara, A. d. l., Christensen, H. M., Colangeli, M., Coleman, D. R. B., Crommelin, D., Dolaptchiev, S. I., Franzke, C. L. E., Friederichs, P., Imkeller, P., Järvinen, H., Juricke, S., Kitsios, V., Lott, F., Lucarini, V., Mahajan, S., Palmer, T. N., Penland, C., Sakradzija, M., von Storch, J.-S., Weisheimer, A., Weniger, M., Williams, P. D., and Yano, J.-I. (2017). Stochastic parameterization: Towards a new view of weather and climate models. B. Am. Meteorol. Soc, 98:565–588.

Beylkin, G. and Keiser, J. (1997). An adaptive pseudo-wavelet approach for solving nonlinear partial differential equations. In Multiscale Wavelet Methods for Partial Differential Equations, volume 6 of Wavelet Analysis and Applications. Academic Press.

Bishop, C. and Hodyss, D. (2007). Flow adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. Q. J. Roy. Meteor. Soc., 133:2021–2044.

Buehner, M. (2012). Evaluation of a spatial/spectral covariance localization approach for atmospheric data assimilation. Mon. Weath. Rev., 140:617–636.

Buehner, M. and Shlyaeva, A. (2015). Scale-dependent background-error covariance localisation. Tellus A, 67:1–17.

Buhmann, M. D. (2000). A new class of radial basis functions with compact support. Mathematics of Computation, 70(233):307–318.

Buizza, R. and Palmer, T. (1995). The singular-vector structure of the atmospheric global circulation. J. Atmos. Sci., 54(9):1434–1456.

Coddington, E. A. and Levinson, N. (1955). Theory of Ordinary Differential Equations. McGraw-Hill.

Cohn, S. E. (1993). Dynamics of short-term univariate forecast error covariances. Mon. Weather Rev., 121:3123–3149.

Cohn, S. E. (2010). The principle of energetic consistency in data assimilation. In Data Assimilation, chapter 7, pages 137–216. Springer Berlin Heidelberg.

Courant, R. and Hilbert, D. (1962). Methods of Mathematical Physics, volume II. Interscience Publishers.

Courtier, P. and Talagrand, O. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. Q. J. Roy. Meteor. Soc., 113:1329–1347.

Crank, J. and Nicolson, P. (1947). A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. Math. Proc. Cambridge, 43:50–67.

Daley, R. (1991). Atmospheric Data Analysis. Cambridge University Press.

Ehrendorfer, M. and Tribbia, J. (1997). Optimal prediction of forecast error covariances through singular vectors. J. Atmos. Sci., 57:286–313.

El Gharamti, M. (2018). Enhanced adaptive inflation algorithm for ensemble filters. Mon. Weather Rev., 146(2):623–640.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., 99:10143–10162.

Evensen, G. (2009). Data Assimilation: The Ensemble Kalman Filter. Springer, 2nd edition.

Furrer, R. and Bengtsson, T. (2007). Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. J. Multivariate Anal., 98:227–255.

Gaspari, G. and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. Q. J. Roy. Meteor. Soc., 125:723–757.

Gaspari, G., Cohn, S. E., Guo, J., and Pawson, S. (2006). Construction and application of covariance functions with variable length-fields. Q. J. Roy. Meteor. Soc., 132:1815–1838.

Gilpin, S. (2023). A generalized gaspari-cohn correlation function. https://doi.org/10.5281/zenodo.7859258.

Gilpin, S., Matsuo, T., and Cohn, S. E. (2022). Continuum covariance propagation for understanding variance loss in advective systems. SIAM/ASA J. Uncertainty Quantification, 10:886–914.

Gilpin, S., Matsuo, T., and Cohn, S. E. (2023). A generalized, compactly-supported correlation function for data assimilation applications. Accepted for Publication in Q. J. Roy. Meteor. Soc.

Gneiting, T. (1999). Correlation functions for atmospheric data analysis. Q. J. Roy. Meteor. Soc., 125:2449–2464.

Gneiting, T. (2002). Compactly supported correlation functions. Journal of Multivariate Analysis, 83:493–508.

Hamill, T. M., Whitaker, J., and Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon. Weather Rev., 129:2776–2790.

Hesthaven, J. S. (2018). Numerical Methods for Conservation Laws: From Analysis to Algorithms. SIAM.

Holton, J. and Hakim, G. J. (2013). An Introduction to Dynamic Meteorology. Academic Press, 5th edition.

Horn, R. and Johnson, C. R. (1985). Matrix Analysis. Cambridge University Press.

Houtekamer, P. L. and Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. Mon. Weather Rev., 126:769–811.

Houtekamer, P. L. and Mitchell, H. L. (2000). An adaptive ensemble Kalman filter. Mon. Weather Rev., 128:416–433.

Houtekamer, P. L. and Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Weather Rev., 128:123–137.

Houtekamer, P. L. and Mitchell, H. L. (2005). Ensemble Kalman filtering. Q. J. Roy. Meteor. Soc.., 131:3269–3289.

Hunter, J. K. and Nachtergaele, B. (2001). Applied Analysis. World Scientific Publishing Co.

Iserles, A. (2009). A First Course in the Numerical Analysis of Differential Equations. Cambridge University Press, 2nd edition.

Jazwinski, A. H. (1970). Stochastic Processes and Filter Theory. Academic Press.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82:35–45.

Kalnay, E. (2003). Atmospheric modeling, data assimilation and predictability. Cambridge University Press.

Kleiber, W. and Porcu, E. (2015). Nonstationary matrix covariances: compact support, long range dependence and quasi-arithmetic constructions. Stoch. Environ. Res. Risk Assess., 29:193–204.

Lax, P. and Wendroff, B. (1960). Systems of conservation laws. Communications on Pure and Applied Mathematics, 13:217–237.

Leith, C. E. (1974). Theoretical skill of Monte Carlo forecasts. Mon. Weather Rev., 102(6):409–418.

Liang, M. and Marcotte, D. (2016). A class of non-stationary covariance functions with compact support. Stoch. Environ. Res. Risk Assess., 30:973–987.

Lin, S. and Rood, R. B. (1996). Multidimensional Flux-Form Semi-Lagrangian Transport Schemes. Mon. Weath. Rev., 124:2046–2070.

Lorenz, E. (1965). A study of the predictability of a 28-variable atmospheric model. Tellus, 21:321–333.

Lyster, P. M., Cohn, S. E., Zhang, B., Chang, L.-P., Ménard, R., Olson, K., and Renka, R. (2004). A Lagrangian trajectory filter for constituent data assimilation. Q. J. Roy. Meteor. Soc., 130:2315–2334.

Maybeck, P. S. (1982). Stochastic models, estimation and control, volume 2. Academic Press.

Ménard, R. and Chang, L. P. (2000). Assimilation of stratospheric chemical tracer observations using a Kalman Filter. Part II: χ^2-validated results and analysis of variance and correlation dynamics. Mon. Weather Rev., 128:2672–2686.

Ménard, R., Cohn, S. E., Chang, L. P., and Lyster, P. M. (2000). Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part I: Formulation. Mon. Weather Rev., 128:2654–2671.

Ménard, R., Deshaies-Jacques, M., and Gasset, N. (2016). A comparison of correlation-length estimation methods for the objective analysis of surface pollutants at Environment and Climate Change Canada. Journal of the Air & Waste Management Association, 66(9):874–895.

Ménard, R., Skachko, S., and Pannekoucke, O. (2021). Numerical discretization causing error variance loss and the need for inflation. Q. J. Roy. Meteor. Soc., 147(740):3498–3520.

Michel, Y., Ménétrier, B., and Montmerle, T. (2016). Objective filtering of the local correlation tensor. Q. J. Roy. Meteor. Soc., 142:2314–2323.

Mitchell, H. L. and Houtekamer, P. L. (2000). An adaptive ensemble Kalman filter. Mon. Weather Rev., 128:416–433.

Paciorek, C. and Schervish, M. (2006). Spatial modelling using a new class of nonstationary covariance functions. Environmetrics, 17:483–506.

Pannekoucke, O., Bocquet, M., and Ménard, R. (2018). Parametric covariance dynamics for the nonlinear diffusive Burgers equataion. Nonlin. Processes Geophys., 25:481–495.

Pannekoucke, O., Ménard, R., Aabaribaoune, M. E., and Plu, M. (2021). A methodology to obtain model-error covariances due to the discretization scheme from the parametric Kalman filter perspective. Nonlin. Processes Geophys., 28:1–22.

Pannekoucke, O., Ricci, S., Barthelemy, S., Ménard, R., and Thual, O. (2016). Parametric Kalman Filter for chemical transport model. Tellus A, 68:31547.

Phillips, N. A. (1986). The spatial statistics of random geostrophic modes and first-guess errors. Tellus A, 38:314–332.

Porcu, E., Daley, D. J., Buhmann, M., and Bevilacqua, B. (2013). Radial basis functions with compact support for multivariate geostatistics. Stoch. Environ. Res. Risk Assess., 27:909–922.

Reed, M. and Simon, B. (1972). Methods of Modern Mathematical Physics I: Functional Analysis. Academic Press.

Shearer, M. and Levy, R. (2015). Partial Differential Equations: An Introduction to Theory and Applications. Princeton University Press.

Stanley, Z., Grooms, I., and Kleiber, W. (2021). Multivariate localization functions for strongly coupled data assimilation in the bivariate Lorenz '96 system. Nonlin. Processes Geophys., 28:656–683.

Talagrand, O. and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. Q. J. Roy. Meteor. Soc., 113:1311–1328.

Tippett, M., Anderson, J., Bishop, C., Hamill, T., and Whitaker, J. (2003). Ensemble square root filters. Mon. Weather Rev., 131:1485–1490.

Waller, J., Simonin, D., Dance, S., Nichols, N., and Ballard, S. (2016). Diagnosing Observation Error Correlations for Doppler Radar Radial Winds in the Met Office UKV Model Using Observation-Minus-Background and Observation-Minus-Analysis Statistics. Mon. Weather Rev., 144:3533–3551.

Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree. Advances in Computational Mathematics, 4:389–396.

Wolfram Research, Inc. (2022). Mathematica, Version 13.1. Champaign, IL, 2022.

Wu, Z. M. (1995). Compactly supported positive definite radial functions. Advances in Computational Mathematics, 4:283–292.

Zwillinger, D., editor (1996). CRC Standard Mathematical Tables and Formulae. CRC Press, Inc., 30th edition.

Appendix A

Supplement to Chapter 2

This appendix contains additional discussion of the continuous spectrum solution presented in Ch. 2. Appendix A.1 is published in Gilpin et al. (2022). Appendix A.2 provides an alternative, direct approach, to deriving the continuous spectrum solution in Ch. 2.

A.1 Proofs of Solutions with Dirac Delta Initial Conditions

We begin by verifying that $P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = d^2(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$, where d^2 satisfies Eq. (2.22), is the solution to the covariance evolution equation Eq. (1.6) for $P_0(\boldsymbol{x}_1, \boldsymbol{x}_2) = \delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$.

Proof. The function $P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = d^2(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is a weak (distribution) solution of Eq. (1.6) with $P_0(\boldsymbol{x}_1, \boldsymbol{x}_2) = \delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ if

$$(\mathcal{L}^*\phi, P) = 0 \tag{A.1}$$

for all test functions $\phi \in C^1(\Omega \times \Omega \times [t_0, T])$ with period $t = T$, where $\mathcal{L}^*$ is the adjoint differential operator

$$\mathcal{L}^* = -\partial_t - \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_1 - \boldsymbol{v}_2 \cdot \boldsymbol{\nabla}_2 + b_1 + b_2 - \boldsymbol{\nabla}_1 \cdot \boldsymbol{v}_1 - \boldsymbol{\nabla}_2 \cdot \boldsymbol{v}_2 \tag{A.2}$$

corresponding to the differential operator

$$\mathcal{L} = \partial_t + \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_1 + \boldsymbol{v}_2 \cdot \boldsymbol{\nabla}_2 + b_1 + b_2 \tag{A.3}$$

of Eq. (1.6). Here we denote $(\cdot, \cdot)$ as the inner product over $L^2(\Omega \times \Omega \times [t_0, T])$ and we will denote $(\cdot, \cdot)'$ as the inner product over $L^2(\Omega \times [t_0, T])$. Substituting $P = d^2(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ into the

expression for $(\mathcal{L}^*\phi, P)$, applying the Dirac delta, and expanding the result yields

$$(\mathcal{L}^*\phi, P) = (-\phi_t(\boldsymbol{x}_1, \boldsymbol{x}_1, t) - \boldsymbol{\nabla}_1 \cdot (\boldsymbol{v}_1\phi(\boldsymbol{x}_1, \boldsymbol{x}_1, t)) + (2b_1 - \boldsymbol{\nabla}_1 \cdot \boldsymbol{v}_1)\phi(\boldsymbol{x}_1, \boldsymbol{x}_1, t), d^2(\boldsymbol{x}_1, t))', \quad \text{(A.4)}$$

where we note that $\boldsymbol{\nabla}_2 \cdot \boldsymbol{v}_2 \,|_{\boldsymbol{x}_2=\boldsymbol{x}_1} = \boldsymbol{\nabla}_1 \cdot \boldsymbol{v}_1$ and
$\boldsymbol{\nabla}_1\phi(\boldsymbol{x}_1, \boldsymbol{x}_1, t) = \boldsymbol{\nabla}_1\phi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)|_{\boldsymbol{x}_2=\boldsymbol{x}_1} + \boldsymbol{\nabla}_2\phi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)|_{\boldsymbol{x}_2=\boldsymbol{x}_1}$. Using integration by parts to move derivatives off of the test function ϕ onto d^2 in Eq. (A.4) gives us

$$(\mathcal{L}^*\phi, P) = (\phi(\boldsymbol{x}_1, \boldsymbol{x}_1, t), d_t^2 + \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_1 \, d^2 + (2b_1 - \boldsymbol{\nabla}_1 \cdot \boldsymbol{v}_1)d^2)' = 0, \quad \text{(A.5)}$$

since d^2 satisfies Eq. (2.22). Thus, $P(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = d^2(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is a weak solution to Eq. (1.6).

□

To show that $\tilde{P} = \tilde{P}^c(\boldsymbol{x}_1, t)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the solution to 2.27 for $\tilde{P}_0 = P_0^c(\boldsymbol{x}_1)\delta(\boldsymbol{x}_1, \boldsymbol{x}_2)$ follows an identical proof as given above, where we take

$$b_i \mapsto \frac{1}{2}\boldsymbol{\nabla}_i \cdot \boldsymbol{v}_i, \;\; i = 1, 2 \quad \text{(A.6)}$$

$$d^2(\boldsymbol{x}_1, t) \mapsto \tilde{P}^c(\boldsymbol{x}_1, t), \quad \text{(A.7)}$$

where $\tilde{P}^c$ satisfies Eq. (2.29).

A.2 Direct Approach to Deriving the Continuous Spectrum Equation

There are several ways to derive the continuous spectrum solution, Eq. (1.9) and its associated covariance operator $\boldsymbol{\mathcal{P}}_t^c$. This section presents an alternative approach to the previous derivation that is a direct approach to deriving the continuous spectrum solution and operator $\boldsymbol{\mathcal{P}}_t^c$.

We begin with the operator $\boldsymbol{\mathcal{U}}_t$ and its adjoint $\boldsymbol{\mathcal{U}}_t^*$. From Eq. (2.17), we can express the kernel of the operator $\boldsymbol{\mathcal{U}}_t$, namely $u(\boldsymbol{x}, t; \boldsymbol{\xi})$, as follows,

$$u(\boldsymbol{x}, t; \boldsymbol{\xi}) = \delta(s(\boldsymbol{x}, t), \boldsymbol{\xi})r^{1/2}(\boldsymbol{x}, t). \quad \text{(A.8)}$$

Here, $s(\boldsymbol{x}, t)$ represents the inverse of the solution to the characteristic equation, Eq. (1.7), where $\boldsymbol{x}(t_0) = s(\boldsymbol{x}, t)$ (following the notation of Shearer and Levy, 2015, Ch. 3). In addition, we have

defined the function

$$r(\boldsymbol{x},t) = e^{-\int_{t_0}^{t} \boldsymbol{\nabla}\cdot\boldsymbol{v}(\boldsymbol{x}(\tau,s),\tau)d\tau}, \tag{A.9}$$

which follows directly by rewriting Eq. (2.17) into its corresponding ODE along the characteristics.

By defining the kernel u as in Eq. (A.8), we can derive an explicit expression for the adjoint operator $\boldsymbol{\mathcal{U}}_t^*$. We compute the adjoint over $L^2(\Omega)$ for $f, g \in L^2(\Omega)$ and $\Omega = \mathbb{S}_r^2$,

$$\begin{aligned}
(g, \boldsymbol{\mathcal{U}}_t f)_2 &= \int_\Omega \int_\Omega g(\boldsymbol{x}) f(\boldsymbol{\xi}) u(\boldsymbol{x}, t; \boldsymbol{\xi})\, d\boldsymbol{\xi}\, d\boldsymbol{x} \\
&= \int_\Omega \int_\Omega g(\boldsymbol{x}) f(\boldsymbol{\xi}) \delta(s(\boldsymbol{x},t), \boldsymbol{\xi}) r^{1/2}(\boldsymbol{x},t)\, d\boldsymbol{\xi}\, d\boldsymbol{x} \\
&= \int_\Omega g(\boldsymbol{x}) f(s(\boldsymbol{x},t)) r^{1/2}(\boldsymbol{x},t)\, d\boldsymbol{x} \\
&= \int_\Omega \int_\Omega g(\boldsymbol{x}) f(\boldsymbol{\xi}) r^{-1/2}(\boldsymbol{x},t) \delta(s^{-1}(\boldsymbol{\xi},t), \boldsymbol{x})\, d\boldsymbol{x}\, d\boldsymbol{\xi} = (\boldsymbol{\mathcal{U}}_t^* g, f)_2.
\end{aligned}$$

The derivation of the adjoint operator $\boldsymbol{\mathcal{U}}_t^*$ makes use of the change of variables $\boldsymbol{\xi} = s(\boldsymbol{x}, t)$. For this change of variables we have that $d\boldsymbol{\xi} = r(\boldsymbol{x}, t) d\boldsymbol{x}$ and $\boldsymbol{x} = s^{-1}(\boldsymbol{\xi}, t)$, where in particular, the solution r of Eq. (A.9) is the determinant of the Jacobian of the coordinate transformation (Coddington and Levinson, 1955, Ch. 1). Therefore, we can define the adjoint operator for $f \in L^2(\Omega)$

$$(\boldsymbol{\mathcal{U}}_t^* f)(\boldsymbol{\xi}) = \int_\Omega u^*(\boldsymbol{\xi}, t; \boldsymbol{x}) f(\boldsymbol{x})\, d\boldsymbol{x}, \tag{A.10}$$

with the kernel

$$u^*(\boldsymbol{\xi}, t; \boldsymbol{x}) r^{-1/2}(\boldsymbol{x}, t) \delta(s^{-1}(\boldsymbol{\xi}, t), \boldsymbol{x}). \tag{A.11}$$

We can now compute an explicit expression for $\boldsymbol{\mathcal{P}}_t^c$. For this computation, we will use the polar decomposition of $\boldsymbol{\mathcal{M}}_t$. Recall that the operator $\boldsymbol{\mathcal{D}}_t$ is a self-adjoint multiplication operator

defined in Eq. (2.23), which we will apply first, then proceed as follows:

$$
\begin{aligned}
(\boldsymbol{\mathcal{P}}_t^c f)(\boldsymbol{x}_1) &= (\boldsymbol{\mathcal{D}}_t \boldsymbol{\mathcal{U}}_t \boldsymbol{\mathcal{P}}_0^c \boldsymbol{\mathcal{U}}_t^* \boldsymbol{\mathcal{D}}_t f)(\boldsymbol{x}_1) \\
&= \int_\Omega f(\boldsymbol{x}_2) \int_\Omega \int_\Omega \Big[d(\boldsymbol{x}_1,t)\delta(s(\boldsymbol{x}_1,t),\boldsymbol{\xi}_1) r^{1/2}(\boldsymbol{x}_1,t) P_0^c(\boldsymbol{\xi}_1)\delta(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2) \\
&\qquad\qquad \cdot r^{-1/2}(\boldsymbol{x}_2,t)\delta(s^{-1}(\boldsymbol{\xi}_2,t),\boldsymbol{x}_2,) d(\boldsymbol{x}_2,t)\, d\boldsymbol{\xi}_1\, d\boldsymbol{\xi}_2 \Big]\, d\boldsymbol{x}_2 \\
&= \int_\Omega f(\boldsymbol{x}_2) d(\boldsymbol{x}_1,t) d(\boldsymbol{x}_2,t) r^{1/2}(\boldsymbol{x}_1,t) r^{-1/2}(\boldsymbol{x}_2,t) \int_\Omega \Big[P_0^c(s(\boldsymbol{x}_1,t))\delta(s(\boldsymbol{x}_1,t),\boldsymbol{\xi}_2) \\
&\qquad\qquad \cdot \delta(s^{-1}(\boldsymbol{\xi}_2,t),\boldsymbol{x}_2)\, d\boldsymbol{\xi}_2 \Big]\, d\boldsymbol{x}_2 \\
&= \int_\Omega f(\boldsymbol{x}_2) d(\boldsymbol{x}_1,t) d(\boldsymbol{x}_2,t) r^{1/2}(\boldsymbol{x}_1,t) r^{-1/2}(\boldsymbol{x}_2,t) P_0^c(s(\boldsymbol{x}_1,t))\delta(\boldsymbol{x}_1,\boldsymbol{x}_2)\, d\boldsymbol{x}_2 \\
&= f(\boldsymbol{x}_1) d^2(\boldsymbol{x}_1,t) P_0^c(s(\boldsymbol{x}_1,t))
\end{aligned}
$$

Thus, the continuous spectrum operator $\boldsymbol{\mathcal{P}}_t^c$ is a multiplication operator whose multiplying function $P^c(\boldsymbol{x},t) = d^2(\boldsymbol{x}_1,t) P_0^c(s(\boldsymbol{x}_1,t))$ satisfying the continuous spectrum equation, Eq. (1.9). This implies that for initial covariances of the form $P_0(\boldsymbol{x}_1,\boldsymbol{x}_2) = P_0^c(\boldsymbol{x}_1)\delta(\boldsymbol{x}_1,\boldsymbol{x}_2,)$, the corresponding covariance operator $\boldsymbol{\mathcal{P}}_t$ is the continuous spectrum operator, and the Dirac delta is preserved for all time.

Appendix B

Exact Solutions to One-Dimensional Equations

This appendix includes the exact solutions to the various one-dimensional equations used to generate the numerical results in Chs. 3, 6, and 9. Appendix B.1 is published in Gilpin et al. (2022).

B.1 Solution to the One-Dimensional Continuity Equation

We begin by solving the one-dimensional version of Eq. (1.4) for $b = v_x$ used in the numerical experiments,

$$\begin{aligned} q_t + vq_x + v'q = 0,\ x \in \mathbb{S}^1_1, \\ q(x, t_0) = q_0(x), \\ v = \sin(x) + 2. \end{aligned} \tag{B.1}$$

To solve Eq. (B.1), we start by rewriting the equation in terms of the Lagrangian/total derivative

$$\frac{D}{Dt} \equiv \partial_t + v\partial_x. \tag{B.2}$$

By doing so, Eq. (B.1) becomes the ODE,

$$\begin{aligned} \frac{Dq(x(t), t)}{Dt} = -v'(x(t))q(x(t), t), \\ q(x, t_0) = q_0(x). \end{aligned} \tag{B.3}$$

Associated with Eq. (B.3) are the Lagrangian trajectories, or characteristic equations, that determine how the spatial variable x evolves over time,

$$\begin{aligned} \frac{dx(t)}{dt} &= v(x(t)), \\ x(t_0) &= s. \end{aligned} \tag{B.4}$$

Here, we take s as a general initial parameter and define our solutions to Eq. (B.4) as $x(t;s)$, following the notation of (Shearer and Levy, 2015, Ch. 3) and discussion in Ch. 1.

Since the state equation is reduced to an ODE in Eq. (B.3), solutions are of the form

$$q(x,t) = q_0(x(t_0))e^{\int_{t_0}^{t} -v'(x(\tau))d\tau}. \tag{B.5}$$

We can solve the integral in the exponential of Eq. (B.5) using a simple $u-$substitution of $u = x(t)$ along with Eq. (B.4),

$$\int_{t_0}^{t} -v'(x(\tau))d\tau = \int_{s}^{x(t)} -\frac{v'(u)}{v(u)}du = \ln\left(\frac{v(s)}{v(x(t))}\right). \tag{B.6}$$

To rewrite Eq. (B.6) in terms of (x,t), we first solve for $x(t;s)$ using the characteristic equation Eq. (B.4), then invert to determine $s = s(x,t)$, which can be done since the velocity field v is continuously differentiable (Shearer and Levy, 2015, Ch. 3). Therefore, substituting Eq. (B.6) and $x(t_0) = s(x,t)$ into Eq. (B.5), we have the explicit solution to Eq. (B.1),

$$q(x,t) = q_0(s(x,t))\left(\frac{\sin(s(x,t))+2}{\sin(x)+2}\right), \tag{B.7}$$

$$s(x,t) = 2\tan^{-1}\left(\frac{\sqrt{3}}{2}\tan\left(\tan^{-1}\left(\frac{2\tan(x/2)+1}{\sqrt{3}}\right) - \frac{\sqrt{3}t}{2}\right) - \frac{1}{2}\right). \tag{B.8}$$

The function $\tan^{-1}$ refers to the principle branch of inverse tangent, where $y = \tan^{-1}(x),\ x \in \mathbb{R},\ y \in (-\pi/2, \pi/2)$, defined by solving $x = \tan(y)$ (Zwillinger, 1996, Sec. 6.3.3). Figure B.1 plots Eq. (B.7) for two different initial conditions, along with solving Eq. (B.1) using Lax-Wendroff and Crank-Nicolson schemes on a uniform grid of 200 grid points and unit Courant-Friedrichs-Lewy number.

The amplitude index m of Eq. (3.8) on $\mathbb{S}^1_1$ with velocity field Eq. (3.1) can be expressed explicitly using Eq. (B.7) taking one as the initial condition. The exact solution d of Eq. (2.16) on

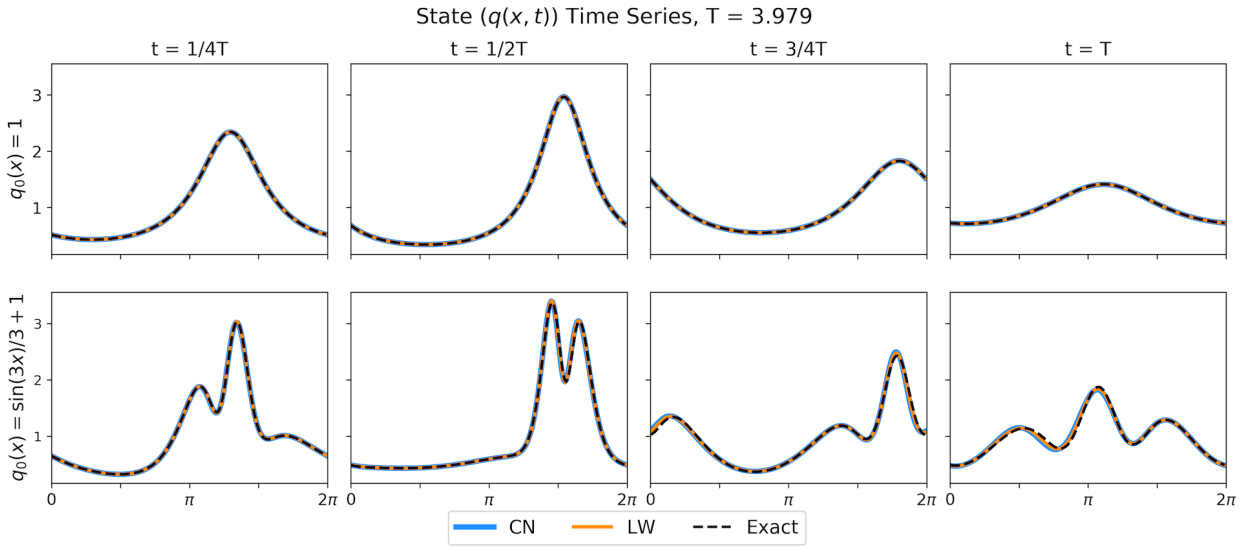

Figure B.1: Solutions to Eq. (B.1) for $q_0(x) = 1$ (top row) and $q_0(x) = \sin(3x)/3 + 1$ (bottom row) at four different times. The exact solution (black dashed) is shown along with finite-difference solutions using Crank-Nicolson (blue) and Lax-Wendroff (orange). The finite difference solutions are nearly indistinguishable from the exact solution.

$\mathbb{S}^1_1$ for $b = v_x$, as in the numerical experiments, can also be computed explicitly from Eq. (B.7) (as d^2 in this case satisfies Eq. B.1 with unit initial condition).

B.2 Exact Solutions to Additional One-Dimensional Equations

Figures 5.1–6.1 illustrates use of exact solutions to Eqs. (5.30), (6.9), and (6.10). This appendix details how to compute the exact solutions to these equations and gives explicit formulas.

For this appendix, we will consider the following PDE for the state $p = p(x,t)$ on the unit circle,

$$\begin{aligned} p_t + vp_x + \alpha v_x p &= 0, \\ p(x,t_0) &= p_0(x), \\ v(x) &= \sin(x) + 2, \end{aligned} \tag{B.9}$$

where $\alpha \in \mathbb{R}$ is a scalar that does not depend on space or time. We include the scalar α so that solutions to Eq. (B.9) can represent the different equations observed in this thesis; for $\alpha = -1$, Eq. (B.9) represent correlation length equation, Eq. (5.30), for example. Solutions to Eq. (B.9) can be computed explicitly for $v(x) = \sin(x) + 2$, though solutions to the PDE in this case can be derived for a general, time-independent velocity field $v(x)$. We will use the general velocity field $v(x)$ to derive the solutions to Eq. (B.9), with explicit solutions using $v(x) = \sin(x) + 2$ when necessary.

To solve Eq. (B.9), we first transform the equation using the following change of variables,

$$p(x,t) = g(x,t)^\alpha. \tag{B.10}$$

By the chain rule, the partial derivatives of with respect to time and space satisfy

$$p_t(x,t) = \alpha g(x,t)^{\alpha-1} g_t(x,t), \quad p_x(x,t) = \alpha g(x,t)^{\alpha-1} g_x(x,t). \tag{B.11}$$

Inserting Eqs. (B.10) – (B.11) into Eq. (B.9), then dividing by $\alpha g(x,t)^{\alpha-1}$, we have that $g = g(x,t)$

satisfies the continuity equation,

$$g_t + vg_x + v_x g = 0,$$
$$g(x, t_0) = g_0(x) = p_0^{1/\alpha}(x, t), . \tag{B.12}$$

Calculating the explicit solution to Eq. (B.12) is given in the previous section, where following its notation we have

$$g(x, t) = g_0(s(x, t)) \left(\frac{v(s(x, t))}{v(x)} \right), \tag{B.13}$$

where $s(x, t)$ are the inverted solutions to the characteristic equation, Eq. (B.4), given an initial parameter s. We can then recover p, where

$$p(x, t) = p_0(s(x, t)) \left(\frac{v(s(x, t))}{v(x)} \right)^{\alpha} \tag{B.14}$$

This approach is used to compute the correlation lengths in Fig. 5.1 (where $\alpha = -1$) and for P_1 in Fig. 6.1, where for $b = v_x$, the solution P_1^+ satisfies Eq. (B.9) with $\alpha = 3$.

In Fig. 6.1, P_2 is the exact solution the following equation,

$$(P_2)_t + v(P_2)_x + 4v_x P_2 + v_{xxx}\sigma_x^2 + 2v_{xx}\sigma^2 = 0,$$
$$P_2(x, t_0) = (P_2)_0(x), \tag{B.15}$$
$$v(x, t) = \sin(x) + 2.$$

Equation (B.15) is an inhomogeneous PDE that can be solved using Duhamel's principle (e.g., Appen. A Cohn, 1993; Shearer and Levy, 2015, Sec. 5.5). To do so, we first rewrite Eq. (B.15) as its corresponding ODE along the characteristics,

$$\frac{DP_2}{Dt} = -4v_x P_2 - v_{xxx}\sigma_x^2 - 2v_{xx}\sigma^2, \tag{B.16}$$

where the Lagrangian (total) derivative is defined in Eq. (B.2). To solve Eq. (B.16), we define the integrating factor

$$\Phi(t, \tau; s) = exp\left\{ -4 \int_{\tau}^{t} v(x(y; s))dy \right\} = \left(\frac{v(x(\tau; s))}{v(x(t; s))} \right)^4 \tag{B.17}$$

where s is the initial parameter defining the characteristic equation, Eq. (B.4). The second equality comes from evaluating the integral itself using the methodology presented in Sec. B.1.

By multiplying both sides of Eq. (B.16) by the integrating factor $\Phi(t,\tau;s)$ and integrating in τ from the initial time t_0 to t, we have the following,

$$P_2(x,t) = \Phi(t,t_0;s)(P_2)_0(s(x,t)) + \int_{t_0}^{t} \Bigg(\Phi(t,\tau;s) \cdot \left[-v''(x(\tau))\sigma_x^2(x(\tau),\tau) - 2v'''(x(\tau))\sigma^2(x(\tau),\tau) \right] \Bigg) d\tau, \tag{B.18}$$

where the spatial derivatives of the velocity v are replaced with primes $'$ since the velocity only depends on space. The first term in Eq. (B.18) is the homogeneous solution to Eq. (B.15), while the integral in Eq. (B.18) is the inhomogeneity that produces the particular solution. This integral can be solved analytically since the explicit solution to the variance equation can be evaluated using Eqs. (B.9) and (B.12) where $p = \sigma^2$ and $\alpha = 2$. Through careful factoring and observing that

$$\frac{ds}{dx} = \frac{v(s(x,t))}{v(x)}, \tag{B.19}$$

the integral in Eq. (B.18) can be evaluated an explicit solution for P_2 governed by Eq. (B.15) becomes

$$\begin{aligned} P_2(x,t) = (P_2)_0(s(x,t)) &\left(\frac{v(s(x,t))}{v(x} \right)^4 \\ &+ \frac{2v(s(x,t))^2}{v(x)^4} \left(\cos^2(s(x,t)) - \cos(s(x,t))\cos(x) - 2\sin(s(x,t)) + \sin(x) \right). \end{aligned} \tag{B.20}$$

The explicit formulation of $s(x,t)$ is given in Eq. (B.8).

Appendix C

Supplement to Chapter 5

In this appendix, we provide further discussion of the Taylor series expansions and their error terms to supplement what is described in the body of the text. There are three sections to this appendix. Appendix C.1 provides additional details on the error analysis conducted in Ch. 5, including more detailed derivations and formulations of the error terms. Appendix C.2 performs the error analysis for the state equation in advection form, in contrast to Ch. 5. The analysis of covariance propagation in advection form is referenced in both Ch. 5.4 and 6.2. We conclude with Appen. C.3, which performs an error analysis for the state dynamics considered in Ch. 4 in their conservative form, Eq. (4.2).

C.1 Taylor Series Expansions

First, we define the error terms common to both the upwind and centered difference scheme that appear in Eqs. (5.15) and (5.19). For general spatial variable x, these are

$$J^{\pm}(x,t) = \left(\frac{\Delta x}{2}\right)^2 P(x,x,t)\left[v_{xx}(x,t) \pm \frac{\Delta x}{6} v_{xxx}(x,t) + \frac{\Delta x^2}{48} v_{xxxx}(x,t)\right], \tag{C.1}$$

$$K(x,t) = \left(\frac{\Delta x}{2}\right)^4 \left[\frac{1}{12} v(x,t) P_4(x,t) + \frac{1}{2} v_{xx}(x,t) P_2(x,t)\right]. \tag{C.2}$$

Here, P_4 is defined by Eq. (6.6) for $n = 4$, which contains a linear combination of fourth derivatives of the covariance across its diagonal.

C.1.1 Upwind, Continuously Differentiable Across $x_1 = x_2$

To arrive at the approximated dynamics given in Eq. (5.18) requires expanding its precursor in Eq. (5.17) about x_i. Recall that we assume the velocity field is four times continuously differentiable in its spatial argument and the covariance is four times continuously differentiable in both spatial arguments.

For example, the first line of Eq. (5.17) represents a first-order upwind approximation of the first derivative,

$$\begin{aligned}\frac{1}{\Delta x}\Big[2v(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) - 2v(x_i,t)P(x_i,x_i,t)\Big] \\ = -\big(vP\big)_x\big|_{x_1=x_2=x_i} + \frac{\Delta x}{4}\big(vP\big)_{xx}\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{24}\big(vP\big)_{xxx}\big|_{x_1=x_2=x_i} + \mathcal{O}(\Delta x^3),\end{aligned} \tag{C.3}$$

where the first term is kept in the approximated dynamics while the last two terms are the resulting error terms. When expanding all terms involving the velocity v and the covariance about the grid point x_i (including $\frac{1}{\Delta x}J^-(x_{i-1/2},t)$) in Eq. (5.17), we arrive at Eq. (5.18). In Eq. (5.18), we define $G_u(x_i,t)$ as the error terms that involve only derivatives of the covariance along the diagonal,

$$\begin{aligned}G_u(x_i,t) = \frac{\Delta x}{4}\big(vP\big)_{xx}\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{24}\big(vP\big)_{xxx}\big|_{x_1=x_2=x_i} + \frac{\Delta x}{2}\big(v_xP\big)_x\big|_{x_1=x_2=x_i} \\ - \frac{\Delta x^2}{8}\big(v_xP\big)_{xx}\big|_{x_1=x_2=x_i} + \frac{\Delta x^3}{48}\big(v_xP\big)_{xxx}\big|_{x_1=x_2=x_i} + J_u(x_i,t).\end{aligned} \tag{C.4}$$

Here, $J_u(x_i,t)$ is the expansion of $\frac{1}{\Delta x}J^-(x_{i-1/2},t)$ about x_i and is $\mathcal{O}(\Delta x)$. Since we assume the both the covariance and velocity are four times continuously differentiable, $G_u(x_i,t)$ is bounded and tends to zero in the limit as Δx tends to zero for fixed, nonzero correlation length L.

Expanding the terms involving $P_2(x_{i-1/2},t)$ about x_i in Eq. (5.17) yield important terms to consider,

$$\frac{\Delta x}{4}v(x_{i-1/2},t)P_2(x_{i-1/2},t) = \frac{\Delta x}{4}v(x_i,t)P_2(x_i,t) - \frac{\Delta x^2}{8}\big(vP_2\big)_x\big|_{x=x_i} \tag{C.5}$$

$$-\frac{\Delta x^2}{8}v_x(x_{i-1/2},t)P_2(x_{i-1/2},t) = -\frac{\Delta x^2}{8}v_x(x_i,t)P_2(x_i,t) + \frac{\Delta x^3}{16}\big(v_xP_2\big)\big|_{x=x_i} \tag{C.6}$$

In Eq. (5.18), we keep both terms in Eq. (C.5) and only the first term in Eq. (C.6). We then define the error term $H_u(x_i,t)$ as the second term in Eq. (C.6) plus the expansion of the error term

$\frac{1}{\Delta x}K(x_{i-1/2},t)$ about x_i, which we will define as $K_u(x_i,t)$

$$H_u(x_i,t) = \frac{\Delta x^3}{16}\big(v_x P_2\big)\big|_{x=x_i} + K_u(x_i,t). \tag{C.7}$$

Here, $K_u(x_i,t)$ remains $\mathcal{O}(\Delta x^3)$. Thus, H_u contains all error terms that involve derivatives of the covariance across its diagonal.

Since we assume the covariance is four times continuously differentiable in both spatial arguments, we can redefine P_2 in terms of the correlation length according to Eqs. (5.24) and (5.25),

$$P_2(x,t) = P(x,x,t)\log(P)_{xx}|_{x_1=x_2=x} - \frac{P(x,x,t)}{L^2(x,t)}. \tag{C.8}$$

In doing so, the error terms $G_u(x_i,t)$ and $H_u(x_i,t)$ can be redefined to include new terms accumulated by expanding P_2 in this way. When doing so, we have the following,

$$\begin{aligned}\tilde{G}_u(x_i,t) = G_u(x_i,t) &- \frac{\Delta x^2}{8}v_x(x_i,t)P(x_i,x_i,t)\log\big(P\big)_{xx}\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{8}\big[vP\log\big(P\big)_{xx}\big]_x\big|_{x_1=x_2=x_i},\\ &+ \frac{\Delta x^3}{16}\big[v_x P\log(P)_{xx}\big]_x\big|_{x_1=x_2=x_i} \qquad\qquad (C.9)\\ \tilde{H}_u(x_i,t) = &-\frac{\Delta x^2}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)\log\big(L\big)_x\big|_{x=x_i} - \frac{\Delta x^3}{16L^2(x_i,t)}v_x(x_i,t)P_x(x_i,x_i,t)\\ &+ \frac{\Delta x^3}{8L^2(x_i,t)}v_x(x_i,t)P(x_i,x_i,t)\log(L)_x\big|_{x=x_i} + K_u(x_i,t). \qquad\qquad (C.10)\end{aligned}$$

Now, $\tilde{G}_u(x_i,t)$ contains terms evaluated on the diagonal, while $\tilde{H}_u(x_i,t)$ contains all terms involving the correlation length L and derivatives across the diagonal. Under the differentiability assumptions, $\tilde{G}_u(x_i,t)$ is bounded and tends to zero as Δx tends to zero for fixed, nonzero correlation length L. The term $\tilde{H}_u(x_i,t)$, however, can become large as correlation lengths L become small for fixed Δx.

C.1.2 Centered Difference, Continuously Differentiable Across $x_1 = x_2$

Section 5.3.1 derives the centered difference semi-discretization for the diagonal of the covariance. When expanding the pairs of terms that average across the covariance diagonal, one about $x_{i-1/2},\ x_{i-1/2}$ and the other about $x_{i+1/2},\ x_{i+1/2}$, as well as the corresponding velocities about

$x_{i-1/2}$ and $x_{i+1/2}$, we have the following,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) &= \frac{1}{\Delta x}\big[v(x_{i-1/2},t)P(x_{i+1/2},x_{i+1/2},t) - v(x_{i+1/2},t)P(x_{i-1/2},x_{i-1/2},t)\big] \\ &\quad - \frac{1}{2}\big[v_x(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) + v_x(x_{i+1/2},t)P(x_{i+1/2},x_{i+1/2},t)\big] \\ &\quad - \big[2b(x_i,t) - 2v_x(x_i,t)\big]P(x_i,x_i,t) \\ &\quad + \frac{\Delta x}{8}\Big[v(x_{i-1/2},t)P_2(x_{i-1/2},t) - v(x_{i+1/2},t)P_2(x_{i+1/2},t)\Big] \qquad \text{(C.11)} \\ &\quad - \frac{\Delta x^2}{16}\Big[v_x(x_{i-1/2},t)P_2(x_{i-1/2},t) + v_x(x_{i+1/2},t)P_2(x_{i+1/2},t)\Big] \\ &\quad + \frac{1}{2\Delta x}\Big\{\big[J^-(x_{i-1/2},t) - J^+(x_{i+1/2},t)\big] + \big[K(x_{i-1/2},t) - K(x_{i+1/2},t)\big]\Big\}, \qquad \text{(C.12)}\end{aligned}$$

with $J^{\pm}$ and K defined in Eqs. (C.1) and (C.2).

There is a pattern for reducing Eq. (C.11): odd-order terms in Δx are centered difference approximation of derivatives at x_i, while even-order terms in Δx are average approximations at of those quantities x_i. For example, the first-order terms in Δx (the fourth group of terms on the right-hand side) in Eq. (C.11) produce centered difference approximations of two derivatives,

$$\begin{aligned}&\frac{\Delta x}{8}\Big[v_{xx}(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) - v_{xx}(x_{i+1/2},t)P(x_{i+1/2},x_{i+1/2},t) \\ &\qquad + v(x_{i-1/2},t)P_2(x_{i-1/2},t) - v(x_{i+1/2},t)P_2(x_{i+1/2},t)\Big] \qquad \text{(C.13)} \\ &\qquad = -\frac{\Delta x^2}{8}(v_{xx}P)_x\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{8}(vP_2)_x\big|_{x=x_i} + \mathcal{O}(\Delta x^4).\end{aligned}$$

By approximating first derivatives, the odd-order terms in Δx increase in order by one and become even-order error terms of the next degree.

The even-order terms in Δx are average approximations of terms about x_i and remain at that even order. For example, consider the second-order terms in Δx,

$$\begin{aligned}&-\frac{\Delta x^2}{16}\Big[v_x(x_{i-1/2},t)P_2(x_{i-1/2},t) + v_x(x_{i+1/2},t)P_2(x_{i+1/2},t) \\ &\qquad + \frac{1}{3}v_{xxx}(x_{i-1/2},t)P(x_{i-1/2},,x_{i-1/2},t) + \frac{1}{3}v_{xxx}(x_{i+1/2},t)P(x_{i+1/2},,x_{i+1/2},t)\Big] \\ &\quad = -\frac{\Delta x^2}{8}v_x(x_i,t)P_2(x_i,t) - \frac{\Delta x^2}{24}v_{xxx}(x_i,t)P(x_i,x_i,t) + \mathcal{O}(\Delta x^4). \qquad \text{(C.14)}\end{aligned}$$

These average terms remain at their original order. By collapsing each line in Eq. (C.11), we recover only even-order error terms and subsequently the approximated dynamics in Eq. (5.20).

Like the upwind case, we can collect quantities we do not include in Eq. (5.20) in the term $G_c(x_i,t)$, and perform a similar analysis as discussed in the latter half of Appen. C.1.1). We define $G_c(x_i,t)$ as the error terms involving the derivatives of the covariance along its diagonal and $H_c(x_i,t)$ as the error terms involving derivatives across the covariance diagonal,

$$G_c(x_i,t) = -\frac{\Delta x^2}{24}\big(vP\big)_{xxx}\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{8}\big(v_xP\big)_{xx}\big|_{x_1=x_2=x_i} + J_c(x_i,t) \tag{C.15}$$

$$H_c(x_i,t) = K_c(x_i,t). \tag{C.16}$$

The term $J_c(x_i,t)$ is the expansion of $\frac{1}{2\Delta x}\left[J^-(x_{i-1/2},t) - J^+(x_{i+1/2},t)\right]$ about x_i and is $\mathcal{O}(\Delta x^2)$. Here, $H_c(x_i,t)$ is only defined by $K_c(x_i,t)$, which is the expansion of $\frac{1}{2\Delta x}\left[K(x_{i-1/2},t) - K(x_{i+1/2},t)\right]$ about x_i and is $\mathcal{O}(\Delta x^4)$.

When we rewrite P_2 in terms of the correlation length, we obtain $\tilde{H}_c(x_i,t)$ that contains terms involving the correlation length. To see this, we follow the same procedure as done in Appen. C.1.1 by expanding P_2, and we obtain the following new error terms,

$$\begin{aligned}\tilde{G}_c(x_i,t) = G_c(x_i,t) &- \frac{\Delta x^2}{8}v_x(x_i,t)P(x_i,x_i,t)\log(P)_{xx}\big|_{x_1=x_2=x_i} \\ &- \frac{\Delta x^2}{8}\big[vP\log(P)_{xx}\big]_x\big|_{x_1=x_2=x_i}\end{aligned} \tag{C.17}$$

$$\tilde{H}_c(x_i,t) = -\frac{\Delta x^2}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)\log(L)_x\big|_{x=x_i} + K_c(x_i,t). \tag{C.18}$$

The new term in $\tilde{H}_c$ is a consequence of expanding $-\frac{\Delta x^2}{8}(vP_2)_x\big|_{x=x_i}$ in Eq. (5.20).

C.1.3 Discontinuous Across $x_1 = x_2$

The Taylor expansions become a bit more complicated when the covariance is no longer continuously differentiable across $x_1 = x_2$, but the results are quite similar. We introduce a new

term

$$\begin{aligned}\tilde{K}^{\pm}(x,t) = &\left(\frac{\Delta x}{2}\right)^3 \left[v_{xx}(x,t)P_1^+(x,t) + \frac{1}{3}v(x,t)P_3^+(x,t)\right] \\ &+ \left(\frac{\Delta x}{2}\right)^4 \left[\pm\frac{1}{6}v_{xxx}(x,t)P_1^+(x,t) + \frac{1}{2}v_{xx}(x,t)P_2^+(x,t)\right. \\ &\left.\pm\frac{1}{3}v_x(x,t)P_3^+(x,t) + \frac{1}{12}v(x,t)P_4^+(x,t)\right].\end{aligned} \tag{C.19}$$

In the case that the covariance is four times continuously differentiable across its diagonal, then $\tilde{K}^{\pm}$ becomes K exactly as P_1^+ and P_3^+ become zero.

Partial derivatives of the covariance need to be taken from the left and the right as necessary, but can be simplified using the symmetry property in Eq. (5.35) to only include derivatives from the right. For example, if we expand $P(x_{i-1}, x_i, t)$ about $x_{i-1/2}, x_{i-1/2}$ we can rewrite partial derivatives from the left in terms of partial derivatives from the right,

$$\begin{aligned}P(x_{i-1},x_i,t) = P(x_{i-1/2},x_{i-1/2},t) &+ \frac{\Delta x}{2}\underbrace{\left(P_{x_2}^+(x_{i-1/2},x_{i-1/2},t) - P_{x_1}^-(x_{i-1/2},x_{i-1/2},t)\right)}_{=-P_1^-(x_{i-1/2},t)=P_1^+(x_{i-1/2},t)} \\ &+ \frac{1}{2}\left(\frac{\Delta x}{2}\right)^2 \underbrace{\left(P_{x_1x_1}^+ - 2P_{x_1x_2}^{+-} + P_{x_2x_2}^-\right)(x_{i-1/2},x_{i-1/2},t)}_{=P_2^-(x_{i-1/2},t)=P_2^+(x_{i-1/2},t)} \\ &+ \frac{1}{6}\left(\frac{\Delta x}{2}\right)^3 \underbrace{\left(-P_{x_1x_1x_1}^- + 3P_{x_1x_1x_2}^{--+} - 3P_{x_1x_2x_2}^{-++} + P_{x_2x_2x_2}^+\right)(x_{i-1/2},x_{i-1/2},t)}_{=-P_3^-(x_{i-1/2},t)=P_3^+(x_{i-1/2},t)} \\ &+ \mathcal{O}(\Delta x^4)\end{aligned} \tag{C.20}$$

By following this procedure for the quantities in the upwind semi-discretization, Eq. (5.11), we

arrive at the approximated dynamics given in Eq. (5.37) with the following error terms,

$$\begin{aligned}
\hat{G}_u(x_i,t) &= \frac{\Delta x}{4}\big(vP\big)_{xx}\big|_{x_1=x_2=x} - \frac{\Delta x^2}{24}\big(vP\big)_{xxx}\big|_{x_1=x_2=x} + \frac{\Delta x^4}{4!\cdot 16}(vP)_{xxxx}\big|_{x_1=x_2=x_i} \\
&\quad + \frac{\Delta x}{2}\big(v_xP\big)_{x}\big|_{x_1=x_2=x} - \frac{\Delta x^2}{8}\big(v_xP\big)_{xx}\big|_{x_1=x_2=x} + \frac{\Delta x^3}{48}(v_xP)_{xxx}\big|_{x_1=x_2=x_i} \\
&\quad + \hat{J}_u(x_i,t) \\
\hat{H}_u(x_i,t) &= \frac{\Delta x^2}{8}v_{xx}(x_i,t)P_1^+(x_i,t) + \frac{\Delta x^2}{8}(vP_1^+)_{xx}\big|_{x=x_i} - \frac{\Delta x^3}{48}(vP_1^+)_{xxx}\big|_{x=x_i} + \frac{\Delta x^3}{32}(vP_2^+)_{xx}\big|_{x=x_i} \\
&\quad - \frac{\Delta x^3}{16}(v_xP_1^+)_{xx}\big|_{x=x_i} + \frac{\Delta x^4}{96}(v_xP_1^+)_{xxx}\big|_{x=x_i} - \frac{\Delta x^3}{16}(v_xP_2^+)_{x}|_{x=x_i} + \frac{\Delta x^4}{64}(v_xP_2^+)_{xx}\big|_{x=x_i} \\
&\quad + \hat{K}_u(x_i,t)
\end{aligned}$$

Since P_1^+ and P_3^+ are nonzero, $\hat{G}_u$ and $\hat{H}_u$ contain more terms than their counterparts G_u and H_u. $\hat{J}_u(x_i,t)$ is the expansion of $\frac{1}{\Delta x}J^-(x_{i-1/2},t)$ about x_i and remains $\mathcal{O}(\Delta x)$, while $\hat{K}_u(x_i,t)$ is the expansion of $\frac{1}{\Delta x}\tilde{K}^-(x_{i-1/2},t)$ about x_i and is $\mathcal{O}(\Delta x^3)$.

The centered difference semi-discretization requires Taylor-expanding four different covariance quantities. Following the procedure shown in Eq. (C.20) for each of these terms and substituting these expansions into the centered difference semi-discretization, we have the following,

$$\begin{aligned}
\frac{d}{dt}P(x_i,x_i,t) &= \frac{1}{\Delta x}\Big[v(x_{i-1/2},t)P(x_{i+1/2},x_{i+1/2},t) - v(x_{i+1/2},t)P(x_{i-1/2},x_{i-1/2},t)\Big] \\
&\quad - \frac{1}{2}\Big[v_x(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) + v_x(x_{i+1/2},t)P(x_{i+1/2},x_{i+1/2},t)\Big] \\
&\quad + \frac{1}{2}\Big[v(x_{i-1/2},t)P_1^+(x_{i-1/2},t) - v(x_{i+1/2},t)P_1^+(x_{i+1/2},t)\Big] \\
&\quad + \frac{\Delta x}{8}\Big[v(x_{i-1/2},t)P_2(x_{i-1/2},t) - v(x_{i+1/2},t)P_2(x_{i+1/2},t) \\
&\qquad - 2v_x(x_{i-1/2},t)P_1^+(x_{i-1/2},t) - 2v_x(x_{i+1/2},t)P_1^+(x_{i+1/2},t)\Big] \\
&\quad + \frac{\Delta x^2}{16}\Big[-v_x(x_{i-1/2},t)P_2(x_{i-1/2},t) - v_x(x_{i+1/2},t)P_2(x_{i+1/2},t)\Big] \\
&\quad + \frac{1}{2\Delta x}\Big\{\big[J^-(x_{i-1/2},t) - J^+(x_{i-1/2},t)\big] + \Big[\tilde{K}^-(x_{i-1/2},t) - \tilde{K}^+(x_{i+1/2},t)\Big]\Big\},
\end{aligned} \tag{C.21}$$

with $J^\pm$ and $\tilde{K}^\pm$ defined in Eqs. (C.1) and (C.19), respectively.

Reducing Eq. (C.21) to its approximated continuum dynamics is a bit more involved than the continuously differentiable case. There are pairs of terms that are average approximations and

pairs of terms that approximate first derivatives, however, we will retain odd-order error terms due to the presence of P_1^+ and P_3^+. For example, the zeroth-order terms in Eq. (C.21) consist of an average term and derivative term,

$$\begin{aligned} -\frac{1}{2}\big[v_x(x_{i-1/2},t)P(x_{i-1/2},x_{i-1/2},t) + v_x(x_{i+1/2},t)P(x_{i+1/2},x_{i+1/2},t)\big] \\ + \frac{1}{2}\Big[v(x_{i-1/2},t)P_1^+(x_{i-1/2},t) - v(x_{i+1/2},t)P_1^+(x_{i+1/2},t)\Big] \\ = -v_x(x_i,t)P(x_i,x_i,t) + \frac{\Delta x}{2}\big(vP_1^+\big)_x\big|_{x=x_i} + \mathcal{O}(\Delta x^2) \end{aligned} \tag{C.22}$$

Therefore, we retain a first-order error term that involves P_1^+, which would vanish if the covariance was continuously differentiable across $x_1 = x_2$. The same approach used to derive Eq. (C.22) can be applied to the other orders in Eq. (C.21) to recover the approximated dynamics in Eq. (5.38) with the following error terms,

$$\begin{aligned} \hat{G}_c(x_i,t) &= -\frac{\Delta x^2}{24}(vP)_{xxx}\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{4}(v_xP)_{xx}\big|_{x_1=x_2=x_i} + \hat{J}_c(x_i,t) \\ \hat{H}_c(x_i,t) &= -\frac{\Delta x^3}{48}(vP_1^+)_{xx}\big|_{x=x_i} - \frac{\Delta x^2}{16}(v_xP_1^+)_{xx}\big|_{x=x_i} - \frac{\Delta x^4}{64}(v_xP_2^+)_{xx}\big|_{x=x_i} + \hat{K}_c(x_i,t) \end{aligned}$$

Like the continuously differentiable case, $\hat{J}_c(x_i,t)$ is the expansion of the second-to-last term on the right-hand side of Eq. (C.21) and $\hat{K}_c(x_i,t)$ the expansion of the last term on the right-hand side of Eq. (C.21). These error terms are $\mathcal{O}(\Delta x^2)$ and $\mathcal{O}(\Delta x^3)$, respectively, with $\hat{K}_c(x_i,t)$ being $\mathcal{O}(\Delta x^3)$ due to the presence of P_1^+.

C.2 Semi-Discretization in Advection Form

The derivation of the approximated dynamics along the diagonal in Sec. 5.3 is done when the state equation in flux form, Eq. (5.4). In this appendix, we present the results for discretizing the state in advection form. These results apply, for example, to the advection equation (i.e., $b = 0$) where discretizing in advection form is the more natural choice. We assume that the velocity v is four times continuously differentiable in its spatial argument and the covariance P is four times continuously differentiable in both spatial arguments. We leave the case where the covariance is not differentiable across $x_1 = x_2$ to the interested reader.

We begin with the first-order upwind spatial discretization applied to q_x in one-dimensional version of Eq. (1.4). This yields the following semi-discretization of the state,

$$\frac{d}{dt}q_i(t) = \frac{v_i(t)}{\Delta x}\big[q_{i-1}(t) - q_i(t)\big] - b_i(t)q_i(t). \tag{C.23}$$

The two differences between the state semi-discretization of the advection form, Eq. (C.23), and the flux form case, Eq. (5.5), are that the velocity remains on the grid point x_i and that there is no zeroth-order term involving the spatial derivative of the velocity, v_x.

Following the same procedure performed in Sec. 5.3.1, but here applied to Eq. (C.23), we arrive at the semi-discretization for the covariance diagonal,

$$\frac{d}{dt}P_{i,i}(t) = \frac{v_i(t)}{\Delta x}\Big\{\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big] - 2P_{i,i}(t)\Big\} - 2b_i(t)P_{i,i}(t). \tag{C.24}$$

Like the flux form case, the pair $P_{i-1,i}(t) + P_{i,i-1}(t)$ averages across the diagonal to approximate the covariance at the half grid point $x_{i-1/2}, x_{i-1/2}$. By substituting the Taylor expansions of these averaging terms about $x_{i-1/2}, x_{i-1/2}$ into Eq. (C.24), including the first error term involving P_2, and we obtain the following,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = {} & \frac{2v(x_i,t)}{\Delta x}\big[P(x_{i-1/2},x_{i-1/2},t) - P(x_i,x_i,t)\big] - 2b(x_i,t)P(x_i,x_i,t) \\ & + \frac{\Delta x}{4}v(x_i,t)P_2(x_{i-1/2},t) + \frac{v(x_i,t)}{\Delta x}D(x_{i-1/2},t).\end{aligned} \tag{C.25}$$

Here, $D(x_{i-1/2},t)$ contain the additional derivatives of the covariance across the diagonal when expanding the averaging term and is $\mathcal{O}(\Delta x^4)$.

All that remains is to expand terms on the half-grid back to the grid point x_i. In this case, the expansion back to x_i only applies to the first and third terms in Eq. (C.25). Concluding by rewriting P_2 in terms of correlation length yields the approximated dynamics along the diagonal,

$$\begin{aligned}\frac{d}{dt}P(x_i,x_i,t) = {} & -v(x_i,t)P_x(x_i,x_i,t)\left[1 - \frac{\Delta x^2}{8L^2(x_i,t)}\right] - 2b(x_i,t)P(x_i,x_i,t) \\ & - \frac{\Delta x}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t) + A_u(x_i,t) + B_u(x_i,t).\end{aligned} \tag{C.26}$$

The error terms A_u and B_u are of the form

$$A_u(x_i,t) = -\frac{\Delta x}{4}v(x_i,t)P_{xx}(x_i,x_i,t) + \frac{\Delta x}{4}v(x_i,t)P(x_i,x_i,t)\log(P)_{xx}\big|_{x_1=x_2=x_i} \tag{C.27}$$
$$- \frac{\Delta x^3}{24}v(x_i,t)P_{xxx}(x_i,x_i,t) - \frac{\Delta x^2}{8}v(x_i,t)\big[P\log(P)_{xx}\big]_x\big|_{x_1=x_2=x_i}$$
$$B_u(x_i,t) = -\frac{\Delta x^2}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)\log(L)_x\big|_{x=x_i} + D_u(x_i,t), \tag{C.28}$$

where $D_u(x_i,t)$ expands the last term on the right-hand side of Eq. (C.25) about x_i and is $\mathcal{O}(\Delta x^3)$.

The approximated dynamics resulting from discretizing the state in advection form yields the same coefficient in front of the spatial derivative vP_x that depends on correlation length, and the zeroth-order dissipative term with coefficient $-\frac{\Delta x}{4L^2}$. Equation (C.26) does not, however, have any zeroth-order terms involving the spatial derivative of the velocity, v_x.

The centered difference semi-discretization of the state in advection form is similar,

$$\frac{d}{dt}q_i(t) = \frac{v_i(t)}{2\Delta x}\big[q_{i-1}(t) - q_{i+1}(t)\big] - b_i(t)q_i(t), \tag{C.29}$$

which then yields the following semi-discretization for the covariance diagonal,

$$\frac{d}{dt}P_{i,i}(t) = \frac{v_i(t)}{2\Delta x}\Big\{\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big] - \big[P_{i+1,i}(t) + P_{i,i+1}(t)\big]\Big\} - 2b_i(t)P_{i,i}(t). \tag{C.30}$$

By expanding both pairs of averaging terms, as done in Sec. 5.3.1, we arrive at the approximated dynamics for the centered difference scheme,

$$\frac{d}{dt}P(x_i,x_i,t) = -v(x_i,t)P(x_i,x_i,t)\left[1 - \frac{\Delta x^2}{8L^2(x_i,t)}\right] - 2b(x_i,t)P(x_i,x_i,t)$$
$$+ A_c(x_i,t) + B_c(x_i,t), \tag{C.31}$$

with error terms

$$A_c(x_i,t) = -\frac{\Delta x^2}{8}v(x_i,t)P_{xxx}(x_i,x_i,t) - \frac{\Delta x^2}{8}v(x_i,t)\big[P\log(P)_{xx}\big]_x\big|_{x_1=x_2=x_i} \tag{C.32}$$
$$B_c(x_i,t) = -\frac{\Delta x^2}{L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)\log(L)_x\big|_{x=x_i} + D_c(x_i,t). \tag{C.33}$$

Similar to the flux form case, we define $D_c(x_i,t)$ as the $\mathcal{O}(\Delta x^4)$ term that contains derivatives of the covariance across the diagonal resulting from the expansion of the pair of averaging terms in

Eq. (C.30). Like the upwind case in Eq. (C.26), the approximated dynamics derived from states in advection form contain the same terms as that derived from flux form except for the zeroth-order terms involving v_x.

Though the approximated dynamics derived from flux form and advection form are similar, they are distinct when the velocity field is divergent, $v_x \neq 0$. When correlation lengths become small, the two sets of discretizations will approximate different dynamics. Though these approximated dynamics are different when $v_x \neq 0$, the similarity in the error terms obtained from both approaches does suggest that the underlying problem is the same, namely the discontinuous change in continuum dynamics.

C.3 Semi-Discretization for Ch. 4 State Dynamics

To supplement the results presented in Ch. 4, in this appendix we will derive the approximated continuum dynamics for states governed by Eq. (4.2). We begin with the error analysis for the centered difference spatial discretization, since that is the discretization applied in Ch. 4, followed by the error analysis for the upwind spatial discretization at the end of this section for completeness. We will only consider the case where the covariance and velocity are four times continuously differentiable in their spatial arguments and leave the case where the covariance is not continuously differentiable across $x_1 = x_2$ to the interested reader.

We begin by applying a second-order centered difference spatial discretization to the two spatial derivative terms the state in Eq. (4.2) to obtain the following semi-discretization,

$$\frac{d}{dt}q_i(t) = \frac{1}{2\Delta x}\left[v_{i-1}(t)q_{i-1}(t) - v_i(t)q_i(t)\right] + \frac{1}{2\Delta x}v_i(t)\left[q_{i-1}(t) - q_i(t)\right], \quad i = 1, 2, \ldots, N. \quad \text{(C.34)}$$

Following the same procedure outline in Ch. 5.2.2, we can derive the corresponding semi-discretization

for the covariance,

$$\frac{d}{dt}P_{i,j}(t) = \frac{1}{4\Delta x}\big[v_{i-1}(t)P_{i-1,j}(t) - v_{i+1}(t)P_{i+1,j}(t)\big] + \frac{1}{4\Delta x}v_i(t)\big[P_{i-1,j}(t) - P_{i+1,j}(t)\big] + \frac{1}{4\Delta x}\big[v_{j-1}(t)P_{i,j-1}(t) - v_{j+1}(t)P_{i,j+1}(t)\big] + \frac{1}{4\Delta x}v_j(t)\big[P_{i,j-1}(t) - P_{i,j+1}(t)\big], \quad i,j = 1,2,\ldots,N. \tag{C.35}$$

Now, let $i = j$ and we have the semi-discretization for the covariance diagonal,

$$\frac{d}{dt}P_{i,j}(t) = \frac{1}{4\Delta x}\Big\{v_{i-1}(t)\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big] - v_{i+1}(t)\big[P_{i+1,i}(t) + P_{i,i+1}(t)\big]\Big\} + \frac{1}{4\Delta x}v_i(t)\Big\{\big[P_{i-1,i}(t) + P_{i,i-1}(t)\big] - \big[P_{i+1,i}(t) + P_{i,i+1}(t)\big]\Big\}, \quad i = 1,2,\ldots,N. \tag{C.36}$$

To obtain the approximated continuums dynamics along the covariance diagonal requires expanding the pairs of averaging terms (grouped in square braces) in Eq. (C.36) about the half-grid points $x_{i-1/2}$ and $x_{i+1/2}$. This analysis mirrors that presented in Ch. 5.3.1 for the first group of terms in curly braces and Appen. C.2 for the second group of terms in curly braces. In both cases, we obtain second-order average approximations for the even-order terms in Δx, and centered difference approximations of first derivatives for the odd-order terms in Δx. Rewriting these average and derivative terms about the grid point x_i, we obtain the following approximated continuum dynamics along the covariance diagonal,

$$\frac{d}{dt}P(x_i,x_i,t) = -\big(vP\big)_x\big|_{x_1=x_2=x_i} - \frac{1}{2}\left(\frac{\Delta x}{2}\right)^2\big(vP_2\big)_x(x_i,t) + G_{s,c}(x_i,t) + H_{s,c}(x_i,t). \tag{C.37}$$

The error terms are defined as follows,

$$G_{s,c}(x_i,t) = -\frac{\Delta x^2}{48}(vP)_{xxx}\big|_{x_1=x_2=x_i} - \frac{\Delta x^2}{48}v(x_i,t)P_{xxx}(x_i,x_i,t) - \frac{\Delta x^2}{8}(v_xP)_{xx}\big|_{x_1=x_2=x_i} + \frac{1}{2}J_c(x_i,t) \tag{C.38}$$

$$H_{s,c}(x_i,t) = -\frac{\Delta x^4}{16\cdot 24}(vP_2)_{xxx}(x_i,t) - \frac{\Delta x^4}{16\cdot 8}(v_xP_2)_{xx}(x_i,t) + \frac{1}{2}K_c(x_i,t) + \frac{1}{2}D_c(x_i,t), \tag{C.39}$$

where J_c and K_c, are given in Appen. C.1.2 and D_c is given in Appen. C.2. Here, $G_{s,c}(x_i,t)$ is $\mathcal{O}(\Delta x^2)$ while $H_{s,c}(x_i,t)$ is $\mathcal{O}(\Delta x^4)$.

The final step is to rewrite P_2 in terms of the correlation length L,

$$\frac{d}{dt}P(x_i,x_i,t) = -\big(vP\big)_x\big|_{x_1=x_2=x_i}\left[1-\frac{\Delta x^2}{8L^2(x_i,t)}\right] + \tilde{G}_{s,c}(x_i,t) + \tilde{H}_{s,c}(x_i,t), \tag{C.40}$$

where

$$\tilde{G}_{s,c}(x_i,t) = G_{s,c}(x_i,t) - \frac{\Delta x^2}{8}\left[vP\log(P)_{xx}\right]_x\big|_{x_1=x_2=x_i}, \tag{C.41}$$

$$\tilde{H}_{s,c}(x_i,t) = H_{s,c}(x_i,t) - \frac{\Delta x^2}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)\log(L)_x\big|_{x=x_i}. \tag{C.42}$$

For the state dynamics defined by Eq. (4.2), the variance, σ^2, satisfies the continuity equation. Therefore, for fixed grid length Δx and long correlation lengths L, Eq. (C.40) does approximate the correct continuum variance dynamics. As the ratio of the grid length to the correlation length approaches unity, however, the dynamics along the covariance diagonal change completely, either becoming constant in time when the $\frac{\Delta x^2}{8L^2(x_i,t)}$ is exactly one, or reversing wave propagation when this ratio is greater than one. The continuous spectrum solution P^c associated with Eq. (4.2) satisfy the advection equation; the approximated dynamics in Eq. (C.40) will only approximate the continuous spectrum solution correctly when the initial covariance is the identity matrix (or scalar multiple of the identity matrix) and the ratio $\frac{\Delta x^2}{8L^2(x_i,t)}$ is one, as Eq. (C.40) in this case will ensure the covariance diagonal does not evolve over time. In general, however, Eq. (C.40) will not approximate the correct continuous spectrum dynamics.

The approximated dynamics defined in Eq. (C.40) and results in Fig. 5.1 provide additional insights into the covariance diagonals produced by the Crank-Nicholson covariance propagation in Fig. 4.1. Panels (a) and (c) of Fig. 5.1 plot the magnitude of the ratio $\frac{\Delta x^2}{8L^2(x_i,t)}$ at different times for the $c = 0.25$ and $c = 0.05$ cases; the impact of $\frac{\Delta x^2}{8L^2(x_i,t)}$ can then be seen in the diagonals extracted in panel (a) of Fig. 4.1 (blue and green curves, respectively). The magnitude of $\frac{\Delta x^2}{8L^2(x_i,t)}$ for the $c = 0.05$ case in panel (c) of Fig. 5.1 is large, in fact becomes larger than one, which we can see significantly alters on the diagonal propagation in panel (a) of Fig. 4.1. Even though the magnitude of $\frac{\Delta x^2}{8L^2(x_i,t)}$ of the $c = 0.25$ case (panel a of Fig. 5.1) is relatively small (no larger than about 0.075), it is large enough to alter the dynamics being approximated along the diagonal

during discrete covariance propagation. These results suggest that the discontinuous change in the continuum covariance dynamics, which is the source of the errors in Eq. (C.40), can cause considerably inaccurate variance propagation, even for modest correlation lengths.

For completeness, the error analysis for the upwind spatial discretization is included to compare with the centered difference spatial discretization. As before, we will begin with the upwind spatial discretization of the state equation defined by Eq. (4.2). The semi-discretization for the state in this case is as follows,

$$\frac{d}{dt}q_i(t) = \frac{1}{2\Delta x}\left[v_{i-1}(t)q_{i-1}(t) - v_i(t)q_i(t)\right] + \frac{1}{2\Delta x}v_i(t)\left[q_{i-1}(t) - q_i(t)\right], \quad i = 1, 2, \ldots, N. \tag{C.43}$$

Following the same procedure applied in Ch. 5.2.1, we have the semi-discretization for the full covariance,

$$\begin{aligned}\frac{d}{dt}P_{i,j}(t) = &\frac{1}{2\Delta x}\left[v_{i-1}(t)P_{i-1,j}(t) - v_i(t)P_{i,j}(t)\right] + \frac{1}{2\Delta x}v_i(t)\left[P_{i-1,j}(t) - P_{i,j}(t)\right] \\ &+ \frac{1}{2\Delta x}\left[v_{j-1}(t)P_{i,j-1}(t) - v_j(t)P_{i,j}(t)\right] + \frac{1}{2\Delta x}v_j(t)\left[P_{i,j-1}(t) - P_{i,j}(t)\right], \\ &i, j = 1, 2, \ldots, N.\end{aligned} \tag{C.44}$$

The semi-discretization for the covariance diagonal then follows,

$$\begin{aligned}\frac{d}{dt}P_{i,i}(t) = &\frac{1}{2\Delta x}\Big\{v_{i-1}(t)\left[P_{i-1,i}(t) + P_{i,i-1}(t)\right] - 2v_i(t)P_{i,i}(t)\Big\} \\ &+ \frac{1}{2\Delta x}v_i(t)\Big\{\left[P_{i-1,i}(t) + P_{i,i-1}(t)\right] - 2P_{i,i}(t)\Big\}, \quad i = 1, 2, \ldots, N.\end{aligned} \tag{C.45}$$

To obtain the approximated continuum dynamics along the covariance diagonal requires expanding the two groups of terms in the curly braces in Eq. (C.45) about the half grid point $x_{i-1/2}$. Since this is done in Ch. 5.2.1 and Appen. C.2 for the first and second groups of terms, respectively, we will proceed directly to the approximated dynamics,

$$\begin{aligned}\frac{d}{dt}P(x_i, x_i, t) = &-(vP)_x\big|_{x_1=x_2=x_i} + \frac{\Delta x}{4}v(x_i, t)P_2(x_i, t) - \frac{\Delta x^2}{8}(vP_2)_x(x_i, t) \\ &+ G_{s,u}(x_i, t) + H_{s,u}(x_i, t),\end{aligned} \tag{C.46}$$

where

$$G_{s,u}(x_i,t) = \Bigg[\frac{\Delta x}{8}(vP)_{xx} - \frac{\Delta x^2}{48}(vP)_{xxx} + \frac{\Delta x^3}{24\cdot 6}(vP)_{xxxx} + \frac{\Delta x}{4}(v_xP)_x - \frac{\Delta x^2}{16}(v_xP)_{xx} + \frac{\Delta x^3}{12\cdot 8}(v_xP)_{xxx}\Bigg]_{x_1=x_2=x_i} \tag{C.47}$$
$$+ \frac{\Delta x}{8}v(x_i,t)P_{xx}(x_i,x_i,t) - \frac{\Delta x^2}{48}v(x_i,t)P_{xxx}(x_i,x_i,t) + \frac{\Delta x^3}{16\cdot 24}v(x_i,t)P_{xxxx}(x_i,x_i,t),$$
$$+ \frac{1}{2}J_u(x_i,t)$$
$$H_{s,u}(x_i,t) = \frac{\Delta x^3}{32}(vP_2)_{xx}(x_i,t) + \frac{\Delta x^3}{32}(v_xP_2)_x(x_i,t) - \frac{\Delta x^4}{4\cdot 16}(v_xP_2)_x(x_i,t) + \frac{\Delta x^4}{4\cdot 16}v(P_2)_{xx}(x_i,t)$$
$$+ D_u(x_i,t). \tag{C.48}$$

The terms J_u and K_u are the same terms defined in Appen. (C.1.1) and are $\mathcal{O}(\Delta x)$ and $\mathcal{O}(\Delta x^3)$, respectively, while D_u is defined in Appen. C.2 and is also $\mathcal{O}(\Delta x^3)$.

Since we assume the covariance is at least four times continuously differentiable across $x_1 = x_2$, we can rewrite Eq. (C.46) in terms of correlation length,

$$\frac{d}{dt}P(x_i,x_i,t) = -(vP)_x\big|_{x_1=x_2=x_i}\left[1 - \frac{\Delta x^2}{8L^2(x_i,t)}\right] - \frac{\Delta x}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)$$
$$+ \hat{G}_{s,u}(x_i,t) + \hat{H}_{s,u}(x_i,t), \tag{C.49}$$

where

$$\hat{G}_{s,u}(x_i,t) = G_{s,u}(x_i,t) - \frac{\Delta x}{4}v(x_i,t)P(x_i,x_i,t)\log(P)_{xx}\big|_{x_1=x_2=x_i}$$
$$- \frac{\Delta x^2}{8}\left[vP\log(P)_{xx}\right]_x\big|_{x_1=x_2=x_i}, \tag{C.50}$$
$$\hat{H}_{s,u}(x_i,t) = H_{s,u}(x_i,t) - \frac{\Delta x^2}{4L^2(x_i,t)}v(x_i,t)P(x_i,x_i,t)\log(L)_x\big|_{x=x_i}. \tag{C.51}$$

As we saw in Ch. 5.3.1, the upwind spatial discretization results in coefficient in front of the flux term that depends on correlation length and a zeroth-order dissipative term (namely the second term on the right-hand side of Eq. C.49). Like the centered difference case, for fixed grid length Δx and long correlation lengths L, Eq. (C.49) will approximate the correct continuum variance dynamics. The dynamics along the covariance diagonal change, however, as correlation lengths become small for fixed Δx, and in particular will result in strongly dissipative behavior due to the second term on the right-hand side of Eq. (C.49).

Appendix D

Supplement to Chapter 8

This appendix includes an additional discussion on matrix and scalar covariance functions, followed by the computation of the Generalized Gaspari-Cohn correlation function to supplement Ch. 8. Both sections of this appendix are published in Gilpin et al. (2023).

D.1 Matrix and Scalar Covariance Functions

Section 8.2.1 describes the general methodology for extracting a scalar covariance function, Eq. (8.9), from a matrix covariance function. In this appendix, we prove that Eq. (8.9) is in fact a covariance function, followed by a concrete example of a matrix covariance function and the extracted scalar covariance function.

We begin by proving that Eq. (8.9) is a covariance function, elaborating the discussion on pp. 1818 – 1819 in G06.

Proof. Let $\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_n \in \Omega$ and let $\Omega_k \subset \Omega$ for $k = 1, 2, ..., m$ be non-overlaping, non-empty subregions that partition the domain Ω into m regions. Each subregion Ω_k contains n_k points, where $\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_{n_1} \in \Omega_1$, $\boldsymbol{r}_{n_1+1}, ..., \boldsymbol{r}_{n_1+n_2} \in \Omega_2$, and so on, where $n_1 + n_2 + ... + n_m = n$. Consider the matrix covariance function $\mathbf{B}(\boldsymbol{r}, \boldsymbol{s})$ defined in Eq. (8.5). It is positive semi-definite, in that for arbitrary scalars c_{ik} for $i = 1, 2, ..., n$ and $k = 1, 2, ..., m$, Eq. (8.7) is satisfied. In particular, if we take the coefficients $c_{i1} = 0$ for $i > n_1$, $c_{i2} = 0$ for $i < n_1 + 1, i > n_1 + n_2$, ... , $c_{im} = 0$ for

$i < n_1 + n_2 + ... + n_{m-1} + 1$ while keeping the rest arbitrary, then

$$\sum_{k,\ell=1}^{m} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{\ell-1}+1}^{N_\ell} c_{ik} c_{j\ell} B_{k\ell}(\boldsymbol{r}_i, \boldsymbol{r}_j) \geq 0, \quad N_0 = 0,\ N_k = n_1 + n_2 + ... + n_k,\ k = 1, 2, ..., m. \tag{D.1}$$

It follows that the function $B(\boldsymbol{r}, \boldsymbol{s})$ defined in Eq. (8.9) is positive semi-definite, i.e.,

$$\sum_{i,j=1}^{n} \hat{c}_i \hat{c}_j B(\boldsymbol{r}_i, \boldsymbol{r}_j) \geq 0, \tag{D.2}$$

where $\hat{c}_1 = c_{11}$, $\hat{c}_2 = c_{21}$, ... , $\hat{c}_{n_1} = c_{n_1 1}$, $\hat{c}_{n_1+1} = c_{n_1+12}$, ... , $\hat{c}_{n_1+n_2} = c_{n_1+n_2 2}$, ... , and so on, and is therefore a covariance function. □

In the following example, we will illustrate how to extract a scalar covariance function defined in Eq. (8.9) from a matrix covariance function.

Example 1. *Suppose we partition the domain Ω into $m = 2$ subregions Ω_1 and Ω_2. The matrix covariance function $\mathbf{B}(\boldsymbol{r}, \boldsymbol{s})$ for $\boldsymbol{r}, \boldsymbol{s} \in \Omega$ is the following 4×4 block matrix, whose arguments in each block we represent abstractly with the subregions Ω_1 and Ω_2,*

$$\mathbf{B}(\boldsymbol{r}, \boldsymbol{s}) = \begin{pmatrix} \boxed{B_{11}(\Omega_1, \Omega_1)} & B_{11}(\Omega_1, \Omega_2) & B_{12}(\Omega_1, \Omega_1) & \boxed{B_{12}(\Omega_1, \Omega_2)} \\ B_{11}(\Omega_2, \Omega_1) & B_{11}(\Omega_2, \Omega_2) & B_{12}(\Omega_2, \Omega_1) & B_{12}(\Omega_2, \Omega_2) \\ B_{21}(\Omega_1, \Omega_1) & B_{21}(\Omega_1, \Omega_2) & B_{22}(\Omega_1, \Omega_1) & B_{22}(\Omega_1, \Omega_2) \\ \boxed{B_{21}(\Omega_2, \Omega_1)} & B_{21}(\Omega_2, \Omega_2) & B_{22}(\Omega_2, \Omega_1) & \boxed{B_{22}(\Omega_2, \Omega_2)} \end{pmatrix}. \tag{D.3}$$

The boxed terms correspond to the entries we extract to build the piecewise-defined scalar covariance function in Eq. (8.9),

$$B(\boldsymbol{r}, \boldsymbol{s}) = \begin{pmatrix} B_{11}(\Omega_1, \Omega_1) & B_{12}(\Omega_1, \Omega_2) \\ B_{21}(\Omega_2, \Omega_1) & B_{22}(\Omega_2, \Omega_2) \end{pmatrix}. \tag{D.4}$$

D.2 Computation Example and Formulae

The explicit computation of GenGC is very lengthy. In this appendix, we give an example of how to compute one of the functions using Eq. (8.17), followed by the complete formulae for GenGC.

For these computations, assume that $c_k \leq c_\ell$. In the case that $c_\ell \leq c_k$, the indices in the functions below simply need to be switched. Computation of $C_{k\ell}(\boldsymbol{r}, \boldsymbol{s})$ for $\boldsymbol{r} \in \Omega_k$, $\boldsymbol{s} \in \Omega_\ell$ in Eq. (8.17) is first split into the following six cases depending on the relationship between c_k and c_ℓ, with twelve subcases each:

$$C_{k\ell}(\boldsymbol{r}, \boldsymbol{s}) = \begin{cases} f_{1j}(z), & 0 < c_k \leq \frac{c_\ell}{4}, \\ f_{2j}(z), & \frac{c_\ell}{4} \leq c_k \leq \frac{c_\ell}{3} \\ f_{3j}(z), & \frac{c_\ell}{3} \leq c_k \leq \frac{c_\ell}{2} \\ f_{4j}(z), & \frac{c_\ell}{2} \leq c_k \leq \frac{2c_\ell}{3} \\ f_{5j}(z), & \frac{2c_\ell}{3} \leq c_k \leq \frac{3c_\ell}{4} \\ f_{6j}(z), & \frac{3c_\ell}{4} \leq c_k \leq c_\ell, \end{cases} \qquad z = ||\boldsymbol{r} - \boldsymbol{s}||, \quad j = 1, 2, ..., 12. \tag{D.5}$$

The six different cases and twelve corresponding subcases can be determined by considering figures such as Fig. D.1. This figure shows the $(z, r)-$plane (with z and r defined as in Eq. 8.14), where the solid black lines mark the regions that determine the bounds of Eq. (8.14) and the piecewise portion of $h_\ell(z; a_\ell, c_\ell)$. The intersections of $r = c_k/2$ and $r = c_k$ in grey (which for Fig. D.1 are shown for $0 < c_k \leq c_\ell/4$) with the solid black lines divide the z-axis into twelve different intervals which define twelve different sets of Eq. (8.14) integrals to evaluate. These are the $j = 1, 2, ..., 12$ cases in $f_{pj}(z)$. The twelve intersections shift along the $z-$axis as c_k changes with respect to c_ℓ, defining the six functions $f_{pj}(z)$, $p = 1, 2, ..., 6$.

All functions $f_{pj}(z)$ for $p = 1, 2, ..., 6$ and $j = 1, 2, ..., 12$ in Eq. (D.5) are of the form

$$f_{pj}(z) = n_k n_\ell \left(b_{pj}^{(5)} z^5 + b_{pj}^{(4)} z^4 + b_{pj}^{(3)} z^3 + b_{pj}^{(2)} z^2 + b_{pj}^{(1)} z + b_{pj}^{(0)} + b_{pj}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}, \tag{D.6}$$

where

$$n_k = (44a_k^2 + 6a_k + 2)^{-1/2}, \tag{D.7}$$

and similarly for n_ℓ. The coefficients $b_{pj}^{(i)}$ for $p = 1, 2, ..., 6$, $j = 1, 2, ..., 12$, $i = -1, 0, 1, ..., 5$ are given in Tables D.1–D.19 at the end of this appendix. When $c_k = c_\ell$, GenGC reduces to Eq. (33) of G06, see Appendix C of G06 for the explicit formulae.

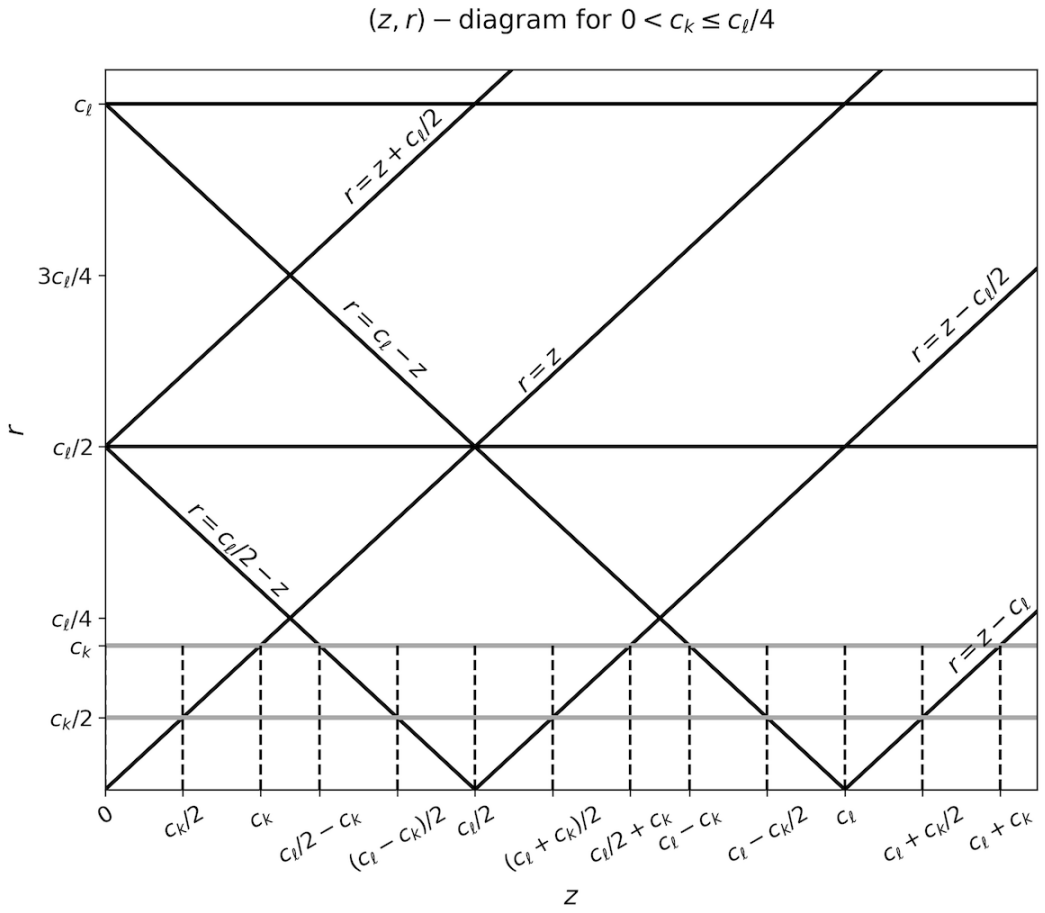

Figure D.1: Diagram in the $(z, r)-$plane used to evaluate the integrals for the function $f_{1j}(z)$ $(0 < c_k \leq c_\ell/4)$. Solid black lines denote boundaries in z and r that define either the bounds in Eq. (8.14) or different piecewise definitions of h_ℓ. Grey horizontal lines mark $r = c_k/2$ and $r = c_k$ for the case that $0 < c_k \leq c_\ell/4$. The intersections of the solid black lines and $r = c_k/2$ and the $r = c_k$ grey lines (indicated by the black dashed lines) mark the different sub-intervals in z that define $f_{1j}(z)$, where the evaluation of Eq. (8.14) changes.

We begin by computing the normalization constant $\sqrt{B^0_{kk}(0)B^0_{\ell\ell}(0)}$ in Eq. (8.17). To do so, we first compute $B^0_{kk}(0)$, which is equivalent to $B^0_{\ell\ell}(0)$ for $k = \ell$. From Eq. (8.15), we evaluate the following integral,

$$B^0_{kk}(0) = 4\pi \int_0^{c_k} r^2 h_k^2(r)\, dr = 4\pi \left[\int_0^{c_k/2} r^2 h_{k,1}^2(r)\, dr + \int_{c_k/2}^{c_k} r^2 h_{k,2}^2(r)\, dr \right] = \frac{\pi c_k^3}{120}, \tag{D.8}$$

where for simplicity we have written $h_{k,1}(z) = \big(2(a_k - 1)z/c_k + 1\big)n_k$ for $0 \le z \le c_k/2$ and $h_{k,2}(z) = 2a_k n_k(1 - z/c_k)$ for $c_k/2 \le z \le c_k$; see Eq. (8.18). Therefore, the normalization constant becomes

$$\sqrt{B^0_{kk}(0)B^0_{\ell}(0)} = \frac{\pi\sqrt{c_k^3 c_\ell^3}}{120}. \tag{D.9}$$

For illustration, we now show how to compute $f_{11}(z)$. The $j = 1$ interval in z, which determines $f_{11}(z)$, is $0 \le z \le c_k/2$. From Fig. D.1, we see three distinct regions which determine the bounds on the outer-most integral: $r = z \le c_k/2$, $r = c_k/2$ and $r = c_k$. All three regions satisfy $r < c_\ell/2 - z$, implying that $r + z < c_\ell/2$, so $h_\ell(s; a_\ell, c_\ell) = h_{\ell,1}(s)$ for the inter-most integral in all cases. The lower bound $|r - z|$ is split into three cases: $r - z < 0$, $0 \le r - z \le c_k/2$, and $c_k/2 \le r - z \le c_k$. Thus, Eq. (8.14) reduces to the following:

$$B^0_{k\ell}(z) = \frac{2\pi}{z}\left[\int_0^z r h_{k,1}(r) \int_{z-r}^{r+z} s h_{\ell,1}(s)\, ds\, dr + \int_z^{c_k/2} r h_{k,1}(r) \int_{r-z}^{r+z} s h_{\ell,1}(s)\, ds\, dr \right.$$
$$\left. + \int_{c_k/2}^{c_k} r h_{k,2}(r) \int_{r-z}^{r+z} s h_{\ell,1}(s)\, ds\, dr \right], \tag{D.10}$$

where again for simplicity we have used $h_{k,1}(z)$ and $h_{k,2}(z)$ as defined above. The coefficients for $f_{11}(z)$ are computed by evaluating Eq. (D.10) and normalizing by Eq. (D.9). These are given explicitly in the first row of Table D.1.

All other functions $f_{pj}(z)$ are computed in a similar manner, although the integrals become more complicated, as do the resulting coefficients. Tables D.1–D.19 summarize the coefficients in each case. Again, in the case that $c_k = c_\ell$, GenGC reduces to Eq. (33) of G06, see Appendix C of G06 for the formulae in this case. For these computations, we evaluated the integrals symbolically using Mathematica (Wolfram Research, Inc., 2022).

Since an explicit, analytical formula for GenGC is defined by Tables D.1–D.19, it can be implemented efficiently in a discrete setting. Once the coefficients in Tables D.1–D.19 are written, two sets of conditionals are necessary to then determine which coefficients to evaluate. Since GenGC is a scalar covariance function, only scalar operators are necessary for its evaluation. Therefore, GenGC can be evaluated efficiently in practice, typically as a linear operator acting on a vector (e.g., Gilpin, 2023).

Table D.1: Coefficients for the function
$f_{1j}(z) = n_k n_\ell \left(b_{1j}^{(5)} z^5 + b_{1j}^{(4)} z^4 + b_{1j}^{(3)} z^3 + b_{1j}^{(2)} z^2 + b_{1j}^{(1)} z + b_{1j}^{(0)} + b_{1j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $0 < c_k \leq c_\ell/4$.

Interval	$\boldsymbol{f_{1j}}$	**Coefficients**
$0 \leq z \leq \frac{c_k}{2}$	f_{11}	$b_{11}^{(5)} = \frac{32}{3c_k c_\ell}(-1 + a_k + a_\ell - a_k a_\ell)$, $b_{11}^{(4)} = \frac{16}{c_\ell}(1 - a_\ell)$, $b_{11}^{(3)} = 0$, $b_{11}^{(2)} = \frac{40c_k^2}{3c_\ell}(-1 - 6a_k + a_\ell + 6a_k a_\ell)$, $b_{11}^{(1)} = 0$, $b_{11}^{(0)} = 5c_k^3(1 + 14a_k) + \frac{3c_k^4}{c_\ell}(-1 - 30a_k + a_\ell + 30a_k a_\ell)$, $b_{1j}^{(-1)} = 0$
$\frac{c_k}{2} \leq z \leq c_k$	f_{12}	$b_{12}^{(5)} = \frac{32}{3c_k c_\ell}(a_k a_\ell - a_k)$, $b_{12}^{(4)} = \frac{32}{c_\ell}(a_k - a_k a_\ell)$, $b_{12}^{(3)} = 0$, $b_{12}^{(2)} = \frac{320c_k^2}{3c_\ell}(a_k a_\ell - a_k)$, $b_{12}^{(1)} = \frac{10c_k^2}{c_\ell}$, $b_{12}^{(0)} = (-1 + 2a_k + a_\ell - 2a_k a_\ell) + 5c_k^3(1 + 14a_k) + \frac{96c_k^4}{c_\ell}(a_k a_\ell - a_k)$, $b_{12}^{(-1)} = \frac{c_k^5}{3c_\ell}(-1 + 2a_k + a_\ell - 2a_k a_\ell)$
$c_k \leq z \leq \frac{c_\ell}{2} - c_k$	f_{13}	$b_{13}^{(5)} = 0$, $b_{13}^{(4)} = 0$, $b_{13}^{(3)} = 0$, $b_{13}^{(2)} = 0$, $b_{13}^{(1)} = \frac{10c_k^3}{c_\ell}(-1 - 14a_k + a_\ell + 14a_k a_\ell)$, $b_{13}^{(0)} = 5c_k^3(1 + 14a_k)$, $b_{13}^{(-1)} = \frac{c_k^5}{3c_\ell}(-1 - 62a_k + a_\ell + 62a_k a_\ell)$
$\frac{c_\ell}{2} - c_k \leq z \leq \frac{c_\ell - c_k}{2}$	f_{14}	$b_{14}^{(5)} = \frac{16}{3c_k c_\ell}(2a_k a_\ell - a_k)$, $b_{14}^{(4)} = \frac{8}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(2a_k a_\ell - a_k)$, $b_{14}^{(3)} = 20(a_k - 2a_k a_\ell)$, $b_{14}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k - 2a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_k a_\ell - a_k)$, $b_{14}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{10c_k^3}{c_\ell}(-1 - 6a_k + a_\ell - 2a_k a_\ell) + 10c_\ell^2(2a_k a_\ell - a_k)$ $+\frac{5c_\ell^3}{c_k}(a_k - 2a_k a_\ell)$, $b_{14}^{(0)} = 5c_k^3(1 + 6a_k + 16a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k - 2a_k a_\ell) + 5c_\ell^3(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(2a_k a_\ell - a_k)$, $b_{14}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k)\frac{c_k^5}{3c_\ell}(-1 - 30a_k + a_\ell - 2a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(a_k - 2a_k a_\ell)$
$\frac{c_\ell - c_k}{2} \leq z \leq \frac{c_\ell}{2}$	f_{15}	$b_{15}^{(5)} = \frac{16}{3c_k c_\ell}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{15}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}(2a_\ell - 1)$, $b_{15}^{(3)} = 10(1 - 2a_\ell)$, $b_{15}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{15}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k) + 5c_\ell^2(2a_\ell - 1)$ $+\frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{15}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{15}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_k^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$

Table D.2: Coefficients for the function
$f_{1j}(z) = n_k n_\ell \left(b_{1j}^{(5)} z^5 + b_{1j}^{(4)} z^4 + b_{1j}^{(3)} z^3 + b_{1j}^{(2)} z^2 + b_{1j}^{(1)} z + b_{1j}^{(0)} + b_{1j}^{(-1)} z^{-1}\right) (c_k c_\ell)^{-3/2}$ for $0 < c_k \leq c_\ell/4$.

Interval	$\boldsymbol{f_{1j}}$	**Coefficients**
$\frac{c_\ell}{2} \leq z \leq \frac{c_\ell + c_k}{2}$	f_{16}	$b_{16}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{16}^{(4)} = \frac{8}{c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell) + \frac{8}{c_\ell}(2a_\ell - 1)$, $b_{16}^{(3)} = 10(1 - 2a_\ell)$, $b_{16}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 12a_k a_\ell)$ $+\frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{16}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k) + 5c_\ell^2(2a_\ell - 1)$ $+\frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{16}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{16}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k) +$ $\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$
$\frac{c_\ell + c_k}{2} \leq z \leq \frac{c_\ell}{2} + c_k$	f_{17}	$b_{17}^{(5)} = \frac{16}{3c_k c_\ell}(a_k - 2a_k a_\ell)$, $b_{17}^{(4)} = \frac{8}{c_k}(2a_k a_\ell - a_k) + \frac{16}{c_\ell}(2a_k a_\ell - a_k)$, $b_{17}^{(3)} = 20(a_k - 2a_k a_\ell)$, $b_{17}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k - 2a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k - 2a_k a_\ell)$, $b_{17}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{10c_k^3}{c_\ell}(2a_k a_\ell - a_\ell - 8a_k) + 10c_\ell^2(2a_k a_\ell - a_k)$ $+\frac{5c_\ell^3}{c_k}(2a_k a_\ell - a_k)$, $b_{17}^{(0)} = 10c_k^3(4a_k + a_\ell + 6a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k - 2a_k a_\ell) + 5c_\ell^3(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(a_k - 2a_k a_\ell)$, $b_{17}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{c_k^5}{3c_\ell}(-32a_k - a_\ell + 2a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(2a_k a_\ell - a_k)$
$\frac{c_\ell}{2} + c_k \leq z \leq c_\ell - c_k$	f_{18}	$b_{18}^{(5)} = 0$, $b_{18}^{(4)} = 0$, $b_{1j}^{(3)} = 0$, $b_{18}^{(2)} = 0$, $b_{18}^{(1)} = \frac{10c_k^3}{c_\ell}(-a_\ell - 14a_k a_\ell)$, $b_{18}^{(0)} = 10c_k^3(a_\ell + 14a_k a_\ell)$, $b_{18}^{(-1)} = \frac{c_k^5}{3}(-a_\ell - 62a_k a_\ell)$
$c_\ell - c_k \leq z \leq c_\ell - \frac{c_k}{2}$	f_{19}	$b_{19}^{(5)} = \frac{-16}{3c_k c_\ell}(a_k a_\ell)$, $b_{19}^{(4)} = \frac{16}{c_k}(a_k a_\ell) + \frac{-16}{c_\ell}(a_k a_\ell)$, $b_{19}^{(3)} = 40a_k a_\ell$, $b_{19}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k a_\ell) + \frac{-160c_\ell^2}{3c_k}(a_k a_\ell)$, $b_{19}^{(1)} = -80c_k^2(a_k a_\ell) + \frac{10c_k^3}{c_\ell}(-a_\ell - 6a_k a_\ell) + 80c_\ell^2(-a_k a_\ell) + \frac{80c_\ell^3}{c_k}(a_k a_\ell)$, $b_{19}^{(0)} = 10c_k^3(a_\ell + 6a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k a_\ell) + 80c_\ell^3(a_k a_\ell) + \frac{48c_\ell^4}{c_k}(-a_k a_\ell)$, $b_{19}^{(-1)} = -24c_k^4(a_k a_\ell) + \frac{c_k^5}{3c_\ell}(-a_\ell - 30a_k a_\ell) + \frac{80c_k^2 c_\ell^2}{3}(a_k a_\ell) + 24c_\ell^4(-a_k a_\ell)$ $+\frac{32c_\ell^5}{3c_k}(a_k a_\ell)$

Table D.3: Coefficients for the function
$f_{1j}(z) = n_k n_\ell \left(b_{1j}^{(5)} z^5 + b_{1j}^{(4)} z^4 + b_{1j}^{(3)} z^3 + b_{1j}^{(2)} z^2 + b_{1j}^{(1)} z + b_{1j}^{(0)} + b_{1j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $0 < c_k \leq c_\ell/4$.

Interval	$\boldsymbol{f_{1j}}$	**Coefficients**
$c_\ell - \frac{c_k}{2} \leq z \leq c_\ell$	f_{110}	$b_{1j}^{(5)} = \frac{16}{3c_k c_\ell}(a_k a_\ell - a_\ell)$, $b_{1j}^{(4)} = \frac{16}{c_k}(a_\ell - a_k a_\ell) + \frac{8}{c_\ell}(-a_\ell)$, $b_{110}^{(3)} = 20a_\ell$, $b_{110}^{(2)} = \frac{20c_k^2}{3c_\ell}(a_\ell + 6a_k a_\ell) + \frac{160c_\ell^2}{3c_k}(a_k a_\ell - a_\ell)$, $b_{110}^{(1)} = 10c_k^2(-a_\ell - 6a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-a_\ell - 14a_k a_\ell) + 40c_\ell^2(-a_\ell)$ $+\frac{80c_\ell^3}{c_k}(a_\ell - a_k a_\ell)$, $b_{110}^{(0)} = 5c_k^3(14a_k a_k + a_\ell) + \frac{3c_k^4}{2c_\ell}(a_\ell + 30a_k a_\ell) + 40c_\ell^3(a_\ell) + \frac{48c_\ell^4}{c_k}(a_k a_\ell - a_\ell)$, $b_{110}^{(-1)} = \frac{3c_k^4}{4}(-a_\ell - 30a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-a_\ell - 62a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_\ell + 6a_k a_\ell)$ $+12c_\ell^4(-a_\ell) + \frac{32c_\ell^5}{3c_k}(a_\ell - a_k a_\ell)$
$c_\ell \leq z \leq c_\ell + \frac{c_k}{2}$	f_{111}	$b_{111}^{(5)} = \frac{16}{3c_k c_\ell}(a_\ell - a_k a_\ell)$, $b_{111}^{(4)} = \frac{16}{c_k}(a_k a_\ell - a_\ell) + \frac{8}{c_\ell}(-a_\ell)$, $b_{111}^{(3)} = 20a_\ell$, $b_{111}^{(2)} = \frac{20c_k^2}{3c_\ell}(a_\ell + 6a_k a_\ell) + \frac{160c_\ell^2}{3c_k}(a_\ell - a_k a_\ell)$, $b_{111}^{(1)} = 10c_k^2(-a_\ell - 6a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-a_\ell - 14a_k a_\ell) + 40c_\ell^2(-a_\ell)$ $+\frac{80c_\ell^3}{c_k}(-a_\ell + a_k a_\ell)$, $b_{111}^{(0)} = 5c_k^3(a_\ell + 14a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(a_\ell + 30a_k a_\ell) + 40c_\ell^3(a_\ell) + \frac{48c_\ell^4}{c_k}(a_\ell - a_k a_\ell)$, $b_{111}^{(-1)} = \frac{3c_k^4}{4}(-a_\ell - 30a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-a_\ell - 62a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_\ell + 6a_k a_\ell)$ $+12c_\ell^4(-a_\ell) + \frac{32c_\ell^5}{3c_k}(a_k a_\ell - a_\ell)$
$c_\ell + \frac{c_k}{2} \leq z \leq c_\ell + c_k$	f_{112}	$b_{112}^{(5)} = \frac{16}{3c_k c_\ell} a_k a_\ell$, $b_{112}^{(4)} = \left(\frac{-16}{c_k} + \frac{-16}{c_\ell}\right) a_k a_\ell$, $b_{112}^{(3)} = 40a_k a_\ell$, $b_{112}^{(2)} = \left(\frac{160c_k^2}{3c_\ell} + \frac{160c_\ell^2}{3c_k}\right) a_k a_\ell$, $b_{112}^{(1)} = -80\left(c_k^2 + \frac{c_k^3}{c_\ell} + c_\ell^2 + \frac{c_\ell^3}{c_k}\right) a_k a_\ell$, $b_{112}^{(0)} = (80c_k^3 + \frac{48c_k^4}{c_\ell} + 80c_\ell^3 + \frac{48c_\ell^4}{c_k}) a_k a_\ell$, $b_{112}^{(-1)} = \left(-24c_k^4 - \frac{32c_k^5}{3c_\ell} + \frac{80c_k^2 c_\ell^2}{3} - 24c_\ell^4 - \frac{32c_\ell^5}{3c_k}\right) a_k a_\ell$

Table D.4: Coefficients for the function
$f_{2j}(z) = n_k n_\ell \left(b_{2j}^{(5)} z^5 + b_{2j}^{(4)} z^4 + b_{2j}^{(3)} z^3 + b_{2j}^{(2)} z^2 + b_{2j}^{(1)} z + b_{2j}^{(0)} + b_{2j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/4 \le c_k \le c_\ell/3$.

Interval	f_{2j}	**Coefficients**
$0 \le z \le \frac{c_k}{2}$	f_{21}	$b_{21}^{(5)} = \frac{32}{3c_k c_\ell}(a_k + a_\ell - a_k a_\ell - 1), \quad b_{21}^{(4)} = \frac{16}{c_\ell}(1 - a_\ell), \quad b_{21}^{(3)} = 0,$ $b_{21}^{(2)} = \frac{40c_k^2}{3c_\ell}(-1 - 6a_k + a_\ell + 6a_k a_\ell), \quad b_{21}^{(1)} = 0,$ $b_{21}^{(0)} = 5c_k^3(1 + 14a_k) + \frac{3c_k^4}{c_\ell}(-1 - 30a_k + a_\ell + 30a_k a_\ell),$ $b_{21}^{(-1)} = 0,$
$c_k/2 \le z \le \frac{c_\ell}{2} - c_k$	f_{22}	$b_{22}^{(5)} = \frac{32}{3c_k c_\ell}(a_k a_\ell - a_k), \quad b_{22}^{(4)} = \frac{32}{c_\ell}(a_k - a_k a_\ell), \quad b_{22}^{(3)} = 0,$ $b_{22}^{(2)} = \frac{320c_k^2}{3c_\ell}(a_k a_\ell - a_k), \quad b_{22}^{(1)} = \frac{10c_k^3}{c_\ell}(-1 + 2a_k + a_\ell - 2a_k a_\ell),$ $b_{22}^{(0)} = 5c_k^3(1+14a_k) + \frac{96c_k^4}{c_\ell}(a_k a_\ell - a_k), \quad b_{22}^{(-1)} = \frac{c_k^5}{3c_\ell}(-1+2a_k+a_\ell-2a_k a_\ell)$
$\frac{c_\ell}{2} - c_k \le z \le c_k$	f_{23}	$b_{23}^{(5)} = \frac{16}{3c_k c_\ell}(4a_k a_\ell - 3a_k), \quad b_{23}^{(4)} = \frac{8}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}a_k,$ $b_{23}^{(3)} = 20(a_k - 2a_k a_\ell), \quad b_{23}^{(2)} = \frac{160c_k^2}{3c_\ell}(-a_k) + \frac{20c_\ell^2}{3c_k}(2a_k a_\ell - a_k),$ $b_{23}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{10c_k^3}{c_\ell}(-1 + 10a_k + a_\ell - 18a_k a_\ell)$ $+10c_\ell^2(2a_k a_\ell - a_k) + \frac{5c_\ell^3}{c_k}(a_k - 2a_k a_\ell),$ $b_{23}^{(0)} = 5c_k^3(1 + 6a_k + 16a_k a_\ell) + \frac{48c_k^4}{c_\ell}(-a_k) + 5c_\ell^3(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(2a_k a_\ell - a_k),$ $b_{23}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{c_k^5}{3c_\ell}(-1 + 34a_k + a_\ell - 66a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(a_k - 2a_k a_\ell)$
$c_k \le z \le \frac{c_\ell - c_k}{2}$	f_{24}	$b_{24}^{(5)} = \frac{16}{3c_k c_\ell}(2a_k a_\ell - a_k), \quad b_{24}^{(4)} = \frac{8}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(2a_k a_\ell - a_k),$ $b_{24}^{(3)} = 20(a_k - 2a_k a_\ell), \quad b_{24}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k - 2a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_k a_\ell - a_k),$ $b_{24}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{10c_k^3}{c_\ell}(-1 - 6a_k + a_\ell - 2a_k a_\ell) + 10c_\ell^2(2a_k a_\ell - a_k)$ $+\frac{5c_\ell^3}{c_k}(a_k - 2a_k a_\ell),$ $b_{24}^{(0)} = 5c_k^3(1 + 6a_k + 16a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k - 2a_k a_\ell) + 5c_\ell^3(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(2a_k a_\ell - a_k),$ $b_{24}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{c_k^5}{3c_\ell}(-1 - 30a_k + a_\ell - 2a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(a_k - 2a_k a_\ell)$
$\frac{c_\ell - c_k}{2} \le z \le \frac{c_\ell}{2}$	f_{25}	$b_{25}^{(5)} = \frac{16}{3c_k c_\ell}(-1 + a_k + 2a_\ell - 2a_k a_\ell),$ $b_{25}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}(2a_\ell - 1), \quad b_{25}^{(3)} = 10(1 - 2a_\ell),$ $b_{25}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell),$ $b_{25}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k) + 5c_\ell^2(2a_\ell - 1)$ $+\frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell),$ $b_{25}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell),$ $b_{25}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$

Table D.5: Coefficients for the function
$f_{2j}(z) = n_k n_\ell \left(b_{2j}^{(5)} z^5 + b_{2j}^{(4)} z^4 + b_{2j}^{(3)} z^3 + b_{2j}^{(2)} z^2 + b_{2j}^{(1)} z + b_{2j}^{(0)} + b_{2j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/4 \le c_k \le c_\ell/3$.

Interval	f_{2j}	**Coefficients**
$\frac{c_\ell}{2} \le z \le \frac{c_\ell + c_k}{2}$	f_{26}	$b_{26}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{26}^{(4)} = \frac{8}{c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell) + \frac{8}{c_\ell}(2a_\ell - 1), \quad b_{26}^{(3)} = 10(1 - 2a_\ell)$, $b_{26}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{26}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k) + 5c_\ell^2(2a_\ell - 1)$ $+\frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{26}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{26}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$
$\frac{c_\ell + c_k}{2} \le z \le c_\ell - c_k$	f_{27}	$b_{27}^{(5)} = \frac{16}{3c_k c_\ell}(a_k - 2a_k a_\ell), \quad b_{27}^{(4)} = \frac{8}{c_k}(2a_k a_\ell - a_k) + \frac{16}{c_\ell}(2a_k a_\ell - a_k)$, $b_{27}^{(3)} = 20(a_k - 2a_k a_\ell), b_{27}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k - 2a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k - 2a_k a_\ell)$, $b_{27}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{10c_k^3}{c_\ell}(-8a_k - a_\ell + 2a_k a_\ell) + 10c_\ell^2(2a_k a_\ell - a_k)$ $+\frac{5c_\ell^3}{c_k}(2a_k a_\ell - a_k)$, $b_{27}^{(0)} = 10c_k^3(4a_k + a_\ell + 6a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k - 2a_k a_\ell) + 5c_\ell^3(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(a_k - 2a_k a_\ell)$, $b_{27}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{c_k^5}{3c_\ell}(2a_k a_\ell - a_\ell - 32a_k) + \frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(2a_k a_\ell - a_k)$
$c_\ell - c_k \le z \le c_k + \frac{c_\ell}{2}$	f_{28}	$b_{28}^{(5)} = \frac{16}{3c_k c_\ell}(a_k - 3a_k a_\ell), \quad b_{28}^{(4)} = \frac{8}{c_k}(4a_k a_\ell - a_k) + \frac{16}{c_\ell}(a_k a_\ell - a_k)$, $b_{28}^{(3)} = 20a_k, \quad b_{28}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k - a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k - 10a_k a_\ell)$, $b_{28}^{(1)} = 40c_k^2(-a_k) + \frac{10c_k^3}{c_\ell}(10a_k a_\ell - a_\ell - 8a_k) + 10c_\ell^2(-a_k - 6a_k a_\ell)$ $+\frac{5c_\ell^3}{c_k}(18a_k a_\ell - a_k)$, $b_{28}^{(0)} = 10c_k^3(4a_k + a_\ell - 2a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k - a_k a_\ell) + 5c_\ell^3(a_k + 14a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(a_k - 34a_k a_\ell)$, $b_{28}^{(-1)} = 12c_k^4(-a_k) + \frac{c_k^5}{3c_\ell}(34a_k a_\ell - a_\ell - 32a_k) + \frac{10c_k^2 c_\ell^2}{3}(a_k + 6a_k a_\ell)$ $+\frac{3c_\ell^4}{4}(-a_k - 30a_k a_\ell) + \frac{c_\ell^5}{6c_k}(-a_k + 66a_k a_\ell)$

Table D.6: Coefficients for the function
$f_{2j}(z) = n_k n_\ell \left(b_{2j}^{(5)} z^5 + b_{2j}^{(4)} z^4 + b_{2j}^{(3)} z^3 + b_{2j}^{(2)} z^2 + b_{2j}^{(1)} z + b_{2j}^{(0)} + b_{2j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/4 \le c_k \le c_\ell/3$.

Interval	$\boldsymbol{f_{2j}}$	**Coefficients**
$c_k + \frac{c_\ell}{2} \le z \le c_\ell - \frac{c_k}{2}$	f_{29}	$b_{29}^{(5)} = \frac{-16}{3c_k c_\ell} a_k a_\ell, \quad b_{29}^{(4)} = \frac{16}{c_k} a_k a_\ell - \frac{16}{c_\ell} a_k a_\ell, \quad b_{29}^{(3)} = 40 a_k a_\ell,$ $b_{29}^{(2)} = \frac{160 c_k^2}{3c_\ell}(a_k a_\ell) - \frac{160 c_\ell^2}{3c_k}(a_k a_\ell),$ $b_{29}^{(1)} = -80c_k^2(a_k) + \frac{10c_k^3}{c_\ell}(-a_\ell - 6a_k a_\ell) - 80c_\ell^2(a_k a_\ell) + \frac{80c_\ell^3}{c_k}(a_k a_\ell),$ $b_{29}^{(0)} = 10c_k^3(a_\ell + 6a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k a_\ell) + 80c_\ell^3(a_k a_\ell) - \frac{48c_\ell^4}{c_k}(a_k a_\ell),$ $b_{29}^{(-1)} = -24c_k^4(a_k a_\ell) + \frac{c_k^5}{3c_\ell}(-a_\ell - 30a_k a_\ell) + \frac{80c_k^2 c_\ell^2}{3}(a_k a_\ell) - 24c_\ell^4(a_k a_\ell)$ $+\frac{32c_\ell^5}{3c_k}(a_k a_\ell)$
$c_\ell - \frac{c_k}{2} \le z \le c_\ell$	f_{210}	Same as f_{110}, see Table D.3.
$c_\ell \le z \le c_\ell + \frac{c_k}{2}$	f_{211}	Same as f_{111}, see Table D.3.
$c_\ell + \frac{c_k}{2} \le z \le c_\ell + c_k$	f_{212}	Same as f_{112}, see Table D.3.

Table D.7: Coefficients for the function
$f_{3j}(z) = n_k n_\ell \left(b_{3j}^{(5)} z^5 + b_{3j}^{(4)} z^4 + b_{3j}^{(3)} z^3 + b_{3j}^{(2)} z^2 + b_{3j}^{(1)} z + b_{3j}^{(0)} + b_{3j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/3 \le c_k \le c_\ell/2$.

Interval	$\boldsymbol{f_{3j}}$	**Coefficients**
$0 \le z \le \frac{c_\ell}{2} - c_k$	f_{31}	$b_{31}^{(5)} = \frac{32}{3c_k c_\ell}(-1 - a_k a_\ell + a_k + a_\ell)$, $\quad b_{31}^{(4)} = \frac{16}{c_\ell}(1 - a_\ell)$, $\quad b_{31}^{(3)} = 0$, $b_{31}^{(2)} = \frac{40c_k^2}{3c_\ell}(a_\ell - 1 + 6a_k a_\ell - 6a_k)$, $\quad b_{31}^{(1)} = 0$, $b_{31}^{(0)} = 5c_k^3(1 + 14a_k) + \frac{3c_k^4}{c_\ell}(-1 - 30a_k + a_\ell + 30a_k a_\ell)$, $\quad b_{31}^{(-1)} = 0$
$\frac{c_\ell}{2} - c_k \le z \le \frac{c_k}{2}$	f_{32}	$b_{32}^{(5)} = \frac{16}{3c_k c_\ell}(a_k + 2a_\ell - 2)$, $b_{32}^{(4)} = \frac{8}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(1 - a_k - a_\ell + 2a_k a_\ell)$, $\quad b_{32}^{(3)} = 20(a_k - 2a_k a_\ell)$, $b_{32}^{(2)} = \frac{40c_k^2}{3c_\ell}(a_\ell - 2a_k - 1 - 2a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_k a_\ell - a_k)$, $b_{32}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{80c_k^3}{c_\ell}(a_k - 2a_k a_\ell) + 10c_\ell^2(2a_k a_\ell - a_k)$ $+\frac{5c_\ell^3}{c_k}(a_k - 2a_k a_\ell)$, $b_{32}^{(0)} = 5c_k^3(1 + 6a_k + 16a_k a_\ell) + \frac{3c_k^4}{c_\ell}(-1 - 14a_k + a_\ell - 2a_k a_\ell)$ $+5c_\ell^3(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(2a_k a_\ell - a_k)$, $b_{32}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{32c_k^5}{3c_\ell}(a_k - 2a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(a_k - 2a_k a_\ell)$
$\frac{c_k}{2} \le z \le \frac{c_\ell - c_k}{2}$	f_{33}	$b_{33}^{(5)} = \frac{16}{3c_k c_\ell}(-3a_k + 4a_k a_\ell)$, $\quad b_{33}^{(4)} = \frac{8}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(a_k)$, $b_{33}^{(3)} = 20(1 - 2a_\ell)$, $\quad b_{33}^{(2)} = \frac{160c_k^2}{3c_\ell}(-a_k) + \frac{20c_\ell^2}{3c_k}(2a_k a_\ell - a_k)$, $b_{33}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{10c_k^3}{c_\ell}(-1 + 10a_k + a_\ell - 18a_k a_\ell)$ $+10c_\ell^2(2a_k a_\ell - a_k) + \frac{5c_\ell^3}{c_k}(a_k - 2a_k a_\ell)$, $b_{33}^{(0)} = 5c_k^3(1 + 6a_k + 16a_k a_\ell) + \frac{48c_k^4}{c_\ell}(-a_k)$ $+5c_\ell^3(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(2a_k a_\ell - a_k)$, $b_{33}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{c_k^5}{3c_\ell}(-1 + 34a_k + a_\ell - 66a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(a_k - 2a_k a_\ell)$
$\frac{c_\ell - c_k}{2} \le z \le c_k$	f_{34}	$b_{34}^{(5)} = \frac{16}{3c_k c_\ell}(2a_\ell - a_k - 1)$, $b_{34}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 4a_k a_\ell)$, $b_{34}^{(3)} = 10(1 - 2a_\ell)$, $b_{34}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 4a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{34}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(18a_k - 1 - 32a_k a_\ell)$ $+5c_\ell^2(2a_\ell - 1) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{34}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 4a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{34}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 128a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$

Table D.8: Coefficients for the function
$f_{3j}(z) = n_k n_\ell \left(b_{3j}^{(5)} z^5 + b_{3j}^{(4)} z^4 + b_{3j}^{(3)} z^3 + b_{3j}^{(2)} z^2 + b_{3j}^{(1)} z + b_{3j}^{(0)} + b_{3j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/3 \le c_k \le c_\ell/2$.

Interval	$\boldsymbol{f_{3j}}$	**Coefficients**
$c_k \le z \le \frac{c_\ell}{2}$	f_{35}	$b_{35}^{(5)} = \frac{16}{3c_k c_\ell}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{35}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}(2a_\ell - 1), \quad b_{35}^{(3)} = 10(1 - 2a_\ell)$, $b_{35}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{35}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k) + 5c_\ell^2(2a_\ell - 1)$ $+\frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{35}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{35}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$,
$\frac{c_\ell}{2} \le z \le c_\ell - c_k$	f_{36}	$b_{36}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{36}^{(4)} = \frac{8}{c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell) + \frac{8}{c_\ell}(2a_\ell - 1), \quad b_{36}^{(3)} = 10(1 - 2a_\ell)$, $b_{36}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{36}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k) + 5c_\ell^2(2a_\ell - 1)$ $+\frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{36}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{36}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$
$c_\ell - c_k \le z \le \frac{c_\ell + c_k}{2}$	f_{37}	$b_{37}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 2a_\ell + a_k a_\ell)$, $b_{37}^{(4)} = \frac{8}{c_k}(-1 + a_k + 2a_\ell) + \frac{8}{c_\ell}(-1 + 2a_\ell - 2a_k a_\ell)$, $b_{37}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{37}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 4a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell - 6a_k a_\ell)$, $b_{37}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k + 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell + 14a_k a_\ell)$, $b_{37}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 28a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell - 30a_k a_\ell)$, $b_{37}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k + 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell + 62a_k a_\ell)$

Table D.9: Coefficients for the function
$f_{3j}(z) = n_k n_\ell \left(b_{3j}^{(5)} z^5 + b_{3j}^{(4)} z^4 + b_{3j}^{(3)} z^3 + b_{3j}^{(2)} z^2 + b_{3j}^{(1)} z + b_{3j}^{(0)} + b_{3j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/3 \le c_k \le c_\ell/2$.

Interval	$\boldsymbol{f_{3j}}$	**Coefficients**
$\frac{c_\ell + c_k}{2} \le z \le c_\ell - \frac{c_k}{2}$	f_{38}	$b_{38}^{(5)} = \frac{16}{3c_k c_\ell}(a_k - 3a_k a_\ell)$, $\quad b_{38}^{(4)} = \frac{8}{c_k}(4a_k a_\ell - a_k) + \frac{16}{c_\ell}(a_k a_\ell - a_k)$, $b_{38}^{(3)} = 20a_k$, $\quad b_{38}^{(2)} = \frac{160c_k^2}{3c_\ell}(a_k - a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k - 10a_k a_\ell)$, $b_{38}^{(1)} = -40c_k^2(a_k) + \frac{10c_k^3}{c_\ell}(-8a_k - a_\ell + 10a_k a_\ell) + 10c_\ell^2(-a_k - 6a_k a_\ell)$ $+\frac{5c_\ell^3}{c_k}(18a_k a_\ell - a_k)$, $b_{38}^{(0)} = 10c_k^3(a_\ell + 4a_k - 2a_k a_\ell) + \frac{48c_k^4}{c_\ell}(a_k - a_k a_\ell) + 5c_\ell^3(a_k + 14a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(a_k - 34a_k a_\ell)$, $b_{38}^{(-1)} = -12c_k^4(a_k) + \frac{c_k^5}{3c_\ell}(-32a_k - a_\ell + 34a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_k + 6a_k a_\ell)$ $+\frac{3c_\ell^4}{4}(-a_k - 30a_k a_\ell) + \frac{c_\ell^5}{6c_k}(-a_k + 66a_k a_\ell)$
$c_\ell - \frac{c_k}{2} \le z \le c_k + \frac{c_\ell}{2}$	f_{39}	$b_{39}^{(5)} = \frac{16}{3c_k c_\ell}(a_k - a_\ell - a_k a_\ell)$, $\quad b_{39}^{(4)} = \frac{8}{c_k}(2a_\ell - a_k) + \frac{8}{c_\ell}(4a_k a_\ell - 2a_k - a_\ell)$, $b_{39}^{(3)} = 20(a_k + a_\ell - 2a_k a_\ell)$, $b_{39}^{(2)} = \frac{20c_k^2}{3c_\ell}(8a_k + a_\ell - 10a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k - 8a_\ell + 6a_k a_\ell)$, $b_{39}^{(1)} = 10c_k^2(-4a_k - a_\ell + 2a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-16a_k - a_\ell + 18a_k a_\ell)$ $+10c_\ell^2(-a_k - 4a_\ell + 2a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-a_k + 16a_\ell - 14a_k a_\ell$, $b_{39}^{(0)} = 5c_k^3(8a_k + a_\ell - 2a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(32a_k + a_\ell - 34a_k a_\ell)$ $+5c_\ell^3(a_k + 8a_\ell - 2a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(a_k - 32a_\ell + 30a_k a_\ell)$, $b_{39}^{(-1)} = \frac{3c_k^4}{4}(-16a_k - a_\ell + 2a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-64a_k - a_\ell + 66a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k + a_\ell + 4a_k a_\ell) + \frac{3c_\ell^4}{4}(-a_k - 16a_\ell + 2a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-a_k + 64a_\ell - 62a_k a_\ell)$
$c_k + \frac{c_\ell}{2} \le z \le c_\ell$	f_{310}	$b_{310}^{(5)} = \frac{16}{3c_k c_\ell}(a_k a_\ell - a_\ell)$, $\quad b_{310}^{(4)} = \frac{16}{c_k}(a_\ell - a_k a_\ell) + \frac{8}{c_\ell}(-a_\ell)$, $b_{310}^{(3)} = 20a_\ell$, $\quad b_{310}^{(2)} = \frac{20c_k^2}{3c_\ell}(a_\ell + 6a_k a_\ell) + \frac{160c_\ell^2}{3c_k}(a_k a_\ell - a_\ell)$, $b_{310}^{(1)} = 10c_k^2(-a_\ell - 6a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-a_\ell - 14a_k a_\ell) + 40c_\ell^2(-a_\ell)$ $+\frac{80c_\ell^3}{c_k}(a_\ell - a_k a_\ell)$, $b_{310}^{(0)} = 5c_k^3(a_\ell + 14a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(a_\ell + 30a_k a_\ell) + 40c_\ell^3(a_\ell) + \frac{48c_\ell^4}{c_k}(a_k a_\ell - a_\ell)$, $b_{310}^{(-1)} = \frac{3c_k^4}{4}(-a_\ell - 30a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-a_\ell - 62a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_\ell + 6a_k a_\ell) + 12c_\ell^4(-a_\ell) + \frac{32c_\ell^5}{3c_k}(a_\ell - a_k a_\ell)$
$c_\ell \le z \le c_\ell + \frac{c_k}{2}$	f_{311}	Same as f_{111}, see Table D.3.
$c_\ell + \frac{c_k}{2} \le z \le c_\ell + c_k$	f_{312}	Same as f_{112}, see Table D.3.

Table D.10: Coefficients for the function
$f_{4j}(z) = n_k n_\ell \left(b_{4j}^{(5)} z^5 + b_{4j}^{(4)} z^4 + b_{4j}^{(3)} z^3 + b_{4j}^{(2)} z^2 + b_{4j}^{(1)} z + b_{4j}^{(0)} + b_{4j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/2 \le c_k \le 2c_\ell/3$.

Interval	$\boldsymbol{f_{4j}}$	**Coefficients**
$0 \le z \le c_k - \frac{c_\ell}{2}$	f_{41}	$b_{41}^{(5)} = \frac{32}{3c_k c_\ell}(-1 + a_k + a_\ell - a_k a_\ell)$, $b_{41}^{(4)} = \frac{16}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(1 - 2a_k - a_\ell + 4a_k a_\ell), \quad b_{41}^{(3)} = 0$, $b_{41}^{(2)} = \frac{40c_k^2}{3c_\ell}(-1 + 2a_k + a_\ell - 10a_k a_\ell) + \frac{40c_\ell^2}{3c_k}(2a_k a_\ell - a_k), \quad b_{41}^{(1)} = 0$, $b_{41}^{(0)} = 5c_k^3(1 - 2a_k + 32a_k a_\ell) + \frac{3c_k^4}{c_\ell}(-1 + 2a_k + a_\ell - 34a_k a_\ell)$ $+10c_\ell^3(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{c_k}(2a_k a_\ell - a_k), \quad b_{41}^{(-1)} = 0$
$c_k - \frac{c_\ell}{2} \le z \le \frac{c_\ell - c_k}{2}$	f_{42}	$b_{42}^{(5)} = \frac{16}{3c_k c_\ell}(-2 + a_k + 2a_\ell)$, $b_{42}^{(4)} = \frac{8}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(1 - a_k - a_\ell + 2a_k a_\ell), \quad b_{42}^{(3)} = 20(a_k - 2a_k a_\ell)$, $b_{42}^{(2)} = \frac{40c_k^2}{3c_\ell}(-1 - 2a_k + a_\ell - 2a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_k a_\ell - a_k)$, $b_{42}^{(1)} = 40c_k^2(2a_k a_\ell - a_k) + \frac{80c_k^3}{c_\ell}(a_k - 2a_k a_\ell) + 10c_\ell^2(2a_k a_\ell - a_k)$ $+\frac{5c_\ell^3}{c_k}(a_k - 2a_k a_\ell)$, $b_{42}^{(0)} = 5c_k^3(1 + 6a_k + 16a_k a_\ell) + \frac{3c_k^4}{c_\ell}(-1 - 14a_k + a_\ell - 2a_k a_\ell)$ $+5c_\ell^3(a_k - 2a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(2a_k a_\ell - a_k)$, $b_{42}^{(-1)} = 12c_k^4(2a_k a_\ell - a_k) + \frac{32c_k^5}{3c_\ell}(a_k - 2a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{4}(2a_k a_\ell - a_k) + \frac{c_\ell^5}{6c_k}(a_k - 2a_k a_\ell)$
$\frac{c_\ell - c_k}{2} \le z \le \frac{c_k}{2}$	f_{43}	$b_{43}^{(5)} = \frac{16}{3c_k c_\ell}(-3 + 3a_k + 4a_\ell - 4a_k a_\ell)$, $b_{43}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}, \quad b_{43}^{(3)} = 10(1 - 2a_\ell)$, $b_{43}^{(2)} = \frac{20c_k^2}{3c_\ell}(-1 - 6a_k) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{43}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1 + 14a_k - 2a_\ell - 28a_k a_\ell)$ $+5c_\ell^2(2a_\ell - 1) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{43}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(-1 - 30a_k) + \frac{5c_\ell^3}{2}(1 - 2a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{43}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1 + 62a_k - 2a_\ell - 124a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$
$\frac{c_k}{2} \le z \le c_\ell - c_k$	f_{44}	$b_{44}^{(5)} = \frac{16}{3c_k c_\ell}(-1 - a_k + 2a_\ell)$, $b_{44}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 4a_k a_\ell)$, $b_{44}^{(3)} = 10(1 - 2a_\ell)$, $b_{44}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 4a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{44}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell + 12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 32a_k a_\ell)$ $+5c_\ell^2(2a_\ell - 1) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$, $b_{44}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 4a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 2a_k a_\ell)$, $b_{44}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell + 60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 128a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell - 12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 2a_k a_\ell)$

Table D.11: Coefficients for the function
$f_{4j}(z) = n_k n_\ell \left(b_{4j}^{(5)} z^5 + b_{4j}^{(4)} z^4 + b_{4j}^{(3)} z^3 + b_{4j}^{(2)} z^2 + b_{4j}^{(1)} z + b_{4j}^{(0)} + b_{4j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/2 \le c_k \le 2c_\ell/3$.

Interval	f_{4j}	**Coefficients**
$c_\ell - c_k \le z \le \frac{c_\ell}{2}$	f_{45}	$b_{45}^{(5)} = \frac{16}{3c_k c_\ell}(-1 - a_k + 2a_\ell - a_k a_\ell)$, $b_{45}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 4a_k a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 6a_k a_\ell)$, $b_{45}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{45}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 10a_k a_\ell)$, $b_{45}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 18a_k a_\ell)$, $b_{45}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 36a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 34a_k a_\ell)$, $b_{45}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 66a_k a_\ell)$
$\frac{c_\ell}{2} \le z \le c_k$	f_{46}	$b_{46}^{(5)} = \frac{16}{3c_k c_\ell}(1 - 3a_k - 2a_\ell + 3a_k a_\ell)$, $b_{46}^{(4)} = \frac{8}{c_k}(-1 + a_k + 2a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 6a_k a_\ell)$, $b_{46}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{46}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell - 6a_k a_\ell)$, $b_{46}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell + 14a_k a_\ell)$, $b_{46}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 36a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell - 30a_k a_\ell)$, $b_{46}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell + 62a_k a_\ell)$
$c_k \le z \le c_\ell - \frac{c_k}{2}$	f_{47}	$b_{47}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 2a_\ell + a_k a_\ell)$, $b_{47}^{(4)} = \frac{8}{c_k}(-1+a_k+2a_\ell)+\frac{8}{c_\ell}(-1+2a_\ell-2a_k a_\ell)$, $b_{47}^{(3)} = 10(1-2a_\ell+4a_k a_\ell)$, $b_{47}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - 2a_\ell - 4a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell - 6a_k a_\ell)$, $b_{47}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k + 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell + 14a_k a_\ell)$, $b_{47}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - 2a_\ell - 28a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell - 30a_k a_\ell)$, $b_{47}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k + 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell + 62a_k a_\ell)$

Table D.12: Coefficients for the function
$f_{4j}(z) = n_k n_\ell \left(b_{4j}^{(5)} z^5 + b_{4j}^{(4)} z^4 + b_{4j}^{(3)} z^3 + b_{4j}^{(2)} z^2 + b_{4j}^{(1)} z + b_{4j}^{(0)} + b_{4j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/2 \le c_k \le 2c_\ell/3$.

Interval	$\boldsymbol{f_{4j}}$	**Coefficients**
$c_\ell - \frac{c_k}{2} \le z \le \frac{c_\ell + c_k}{2}$	f_{48}	$b_{48}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 3a_\ell + 3a_k a_\ell)$, $b_{48}^{(4)} = \frac{8}{c_k}(-1 + a_k + 4a_\ell - 4a_k a_\ell) + \frac{8}{c_\ell}(a_\ell - 1), \quad b_{48}^{(3)} = 10$, $b_{48}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - a_\ell - 6a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 10a_\ell + 10a_k a_\ell)$, $b_{48}^{(1)} = 5c_k^2(-1 - 6a_k) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k + a_\ell + 14a_k a_\ell) + 5c_\ell^2(-1 - 6a_\ell)$ $+\frac{5c_\ell^3}{c_k}(-1 + a_k + 18a_\ell - 18a_k a_\ell)$, $b_{48}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - a_\ell - 30a_k a_\ell) + \frac{5c_\ell^3}{2}(1 + 14a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(1 - a_k - 34a_\ell + 34a_k a_\ell)$, $b_{48}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k + a_\ell + 62a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k + 6a_\ell + 36a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 - 30a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 66a_\ell - 66a_k a_\ell)$
$\frac{c_\ell + c_k}{2} \le z \le c_\ell$	f_{49}	$b_{49}^{(5)} = \frac{16}{3c_k c_\ell}(a_k - a_\ell - a_k a_\ell)$, $b_{49}^{(4)} = \frac{8}{c_k}(2a_\ell - a_k) + \frac{8}{c_\ell}(4a_k a_\ell - a_\ell - 2a_k), \quad b_{49}^{(3)} = 20(a_k + a_\ell - 2a_k a_\ell)$, $b_{49}^{(2)} = \frac{20c_k^2}{3c_\ell}(8a_k + a_\ell - 10a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k - 8a_\ell + 6a_k a_\ell)$, $b_{49}^{(1)} = 10c_k^2(-4a_k - a_\ell + 2a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-16a_k - a_\ell + 18a_k a_\ell)$ $+10c_\ell^2(-a_k - 4a_\ell + 2a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-a_k + 16a_\ell - 14a_k a_\ell)$, $b_{49}^{(0)} = 5c_k^3(8a_k + a_\ell - 2a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(32a_k + a_\ell - 34a_k a_\ell) + 5c_\ell^3(a_k + 8a_\ell - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(a_k - 32a_\ell + 30a_k a_\ell)$, $b_{49}^{(-1)} = \frac{3c_k^4}{4}(-16a_k - a_\ell + 2a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-64a_k - a_\ell + 66a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k + a_\ell + 4a_k a_\ell) + \frac{3c_\ell^4}{4}(-a_k - 16a_\ell + 2a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(64a_\ell - a_k - 62a_k a_\ell)$
$c_\ell \le z \le c_k + \frac{c_\ell}{2}$	f_{410}	$b_{410}^{(5)} = \frac{16}{3c_k c_\ell}(a_k + a_\ell - 3a_k a_\ell)$, $b_{410}^{(4)} = \frac{8}{c_k}(-a_k - 2a_\ell + 4a_k a_\ell) + \frac{8}{c_\ell}(-2a_k - a_\ell + 4a_k a_\ell)$, $b_{410}^{(3)} = 20(a_k + a_\ell - 2a_k a_\ell)$, $b_{410}^{(2)} = \frac{20c_k^2}{3c_\ell}(8a_k + a_\ell - 10a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(a_k + 8a_\ell - 10a_k a_\ell)$, $b_{410}^{(1)} = 10c_k^2(-4a_k - a_\ell + 2a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-16a_k - a_\ell + 18a_k a_\ell)$ $+10c_\ell^2(-a_k - 4a_\ell + 2a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-a_k - 16a_\ell + 18a_k a_\ell)$, $b_{410}^{(0)} = 5c_k^3(8a_k + a_\ell - 2a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(32a_k + a_\ell - 34a_k a_\ell) + 5c_\ell^3(a_k + 8a_\ell - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(a_k + 32a_\ell - 34a_k a_\ell)$, $b_{410}^{(-1)} = \frac{3c_k^4}{4}(2a_k a_\ell - a_\ell - 16a_k) + \frac{c_k^5}{6c_\ell}(-64a_k - a_\ell + 66a_k a_\ell)$ $+\frac{10c_k^2 c_\ell^2}{3}(a_k + a_\ell + 4a_k a_\ell) + \frac{3c_\ell^4}{4}(2a_k a_\ell - 16a_\ell - a_k)$ $+\frac{c_\ell^5}{6c_k}(66a_k a_\ell - a_k - 64a_\ell)$

Table D.13: Coefficients for the function
$f_{4j}(z) = n_k n_\ell \left(b_{4j}^{(5)} z^5 + b_{4j}^{(4)} z^4 + b_{4j}^{(3)} z^3 + b_{4j}^{(2)} z^2 + b_{4j}^{(1)} z + b_{4j}^{(0)} + b_{4j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $c_\ell/2 \le c_k \le 2c_\ell/3$.

Interval	$\boldsymbol{f_{4j}}$	**Coefficients**
$c_k + \frac{c_\ell}{2} \le z \le c_\ell + \frac{c_k}{2}$	f_{411}	$b_{411}^{(5)} = \frac{16}{3c_k c_\ell}(a_\ell - a_k a_\ell)$, $\quad b_{411}^{(4)} = \frac{16}{c_k}(a_k a_\ell - a_\ell) + \frac{8}{c_\ell}(-a_\ell)$, $b_{411}^{(3)} = 20a_\ell$, $\quad b_{411}^{(2)} = \frac{20c_k^2}{3c_\ell}(a_\ell + 6a_k a_\ell) + \frac{160c_\ell^2}{3c_k}(a_\ell - a_k a_\ell)$, $b_{411}^{(1)} = 10c_k^2(-a_\ell - 6a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-a_\ell - 14a_k a_\ell) + 40c_\ell^2(-a_\ell) + \frac{80c_\ell^3}{c_k}(a_k a_\ell - a_\ell)$, $b_{411}^{(0)} = 5c_k^3(a_\ell + 14a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(a_\ell + 30a_k a_\ell) + 40c_\ell^3(a_\ell) + \frac{48c_\ell^4}{c_k}(a_\ell - a_k a_\ell)$, $b_{411}^{(-1)} = \frac{3c_k^4}{4}(-a_\ell - 30a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-a_\ell - 62a_k a_\ell) + \frac{10c_k^2 c_\ell^2}{3}(a_\ell + 6a_k a_\ell)$ $+12c_\ell^4(-a_\ell) + \frac{32c_\ell^5}{3c_k}(a_k a_\ell - a_\ell)$
$c_\ell + \frac{c_k}{2} \le z \le c_\ell + c_k$	f_{412}	$b_{412}^{(5)} = \frac{16}{3c_k c_\ell} a_k a_\ell$, $\quad b_{412}^{(4)} = \left(-\frac{16}{c_k} - \frac{16}{c_\ell}\right) a_k a_\ell$, $\quad b_{412}^{(3)} = 40 a_k a_\ell$, $b_{412}^{(2)} = \left(\frac{160c_k^2}{3c_\ell} + \frac{160c_\ell^2}{3c_k}\right) a_k a_\ell, b_{412}^{(1)} = 80\left(-c_k^2 - \frac{c_k^3}{c_\ell} - c_\ell^2 - \frac{c_\ell^3}{c_k}\right) a_k a_\ell$, $b_{412}^{(0)} = \left(80c_k^3 + 48\frac{c_k^4}{c_\ell} + 80c_\ell^3 + 48\frac{c_\ell^4}{c_k}\right) a_k a_\ell$, $b_{412}^{(-1)} = \left(-24c_k^4 - \frac{32c_k^5}{3c_\ell} + \frac{80c_k^2 c_\ell^2}{3} - 24c_\ell^4 - \frac{32c_\ell^5}{3c_k})\right) a_k a_\ell$

Table D.14: Coefficients for the function $f_{5j}(z) = n_k n_\ell \left(b_{5j}^{(5)} z^5 + b_{5j}^{(4)} z^4 + b_{5j}^{(3)} z^3 + b_{5j}^{(2)} z^2 + b_{5j}^{(1)} z + b_{5j}^{(0)} + b_{5j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $2c_\ell/3 \leq c_k \leq 3c_\ell/4$.

Interval	f_{5j}	**Coefficients**
$0 \leq z \leq \frac{c_\ell - c_k}{2}$	f_{51}	$b_{51}^{(5)} = \frac{32}{3c_k c_\ell}(-1 + a_k + a_\ell - a_k a_\ell)$, $b_{51}^{(4)} = \frac{16}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(1 - 2a_k - a_\ell + 4a_k a_\ell)$, $b_{51}^{(3)} = 0$, $b_{51}^{(2)} = \frac{40c_k^2}{3c_\ell}(-1+2a_k+a_\ell-10a_k a_\ell) + \frac{40c_\ell^2}{3c_k}(2a_k a_\ell - a_k)$, $b_{51}^{(1)} = 0$, $b_{51}^{(0)} = 5c_k^3(1-2a_k+32a_k a_\ell) + \frac{3c_k^4}{c_\ell}(a_\ell+2a_k-1-34a_k a_\ell) + 10c_\ell^3(a_k-2a_k a_\ell)$ $+\frac{3c_\ell^4}{c_k}(2a_k a_\ell - a_k)$, $b_{51}^{(-1)} = 0$
$\frac{c_\ell - c_k}{2} \leq z \leq c_k - \frac{c_\ell}{2}$	f_{52}	$b_{52}^{(5)} = \frac{16}{3c_k c_\ell}(-3 + 4a_k + 4a_\ell - 6a_k a_\ell)$, $b_{52}^{(4)} = \frac{8}{c_k}(1-2a_\ell) + \frac{8}{c_\ell}(1-2a_k+4a_k a_\ell)$, $b_{52}^{(3)} = 10(1-2a_k-2a_\ell+4a_k a_\ell)$, $b_{52}^{(2)} = 20\frac{c_k^2}{3c_\ell}(-1+2a_k-16a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_\ell - 1)$, $b_{52}^{(1)} = 5c_k^2(-1+2a_k+2a_\ell-4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1-2a_k-2a_\ell+4a_k a_\ell)$ $+5c_\ell^2(-1+2a_k+2a_\ell-4a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1-2a_k-2a_\ell+4a_k a_\ell)$, $b_{52}^{(0)} = \frac{5c_k^3}{2}(1-2a_k+2a_\ell+60a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(2a_k - 1 - 64a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1+2a_k-2a_\ell-4a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(2a_\ell - 1)$, $b_{52}^{(-1)} = \frac{3c_k^4}{8}(-1+2a_k+2a_\ell-4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1-2a_k-2a_\ell+4a_k a_\ell)$ $+\frac{5c_k^2c_\ell^2}{12}(1-2a_k-2a_\ell+4a_k a_\ell) + \frac{3c_\ell^4}{8}(-1+2a_k+2a_\ell-4a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1-2a_k-2a_\ell+4a_k a_\ell)$
$c_k - \frac{c_\ell}{2} \leq z \leq c_\ell - c_k$	f_{53}	$b_{53}^{(5)} = \frac{16}{3c_k c_\ell}(-3+3a_k+4a_\ell-4a_k a_\ell)$, $b_{53}^{(4)} = \frac{8}{c_k}(1-a_k-2a_\ell+2a_k a_\ell) + \frac{8}{c_\ell}$, $b_{53}^{(3)} = 10(1-2a_\ell)$, $b_{53}^{(2)} = \frac{20c_k^2}{3c_\ell}(-1-6a_k) + \frac{20c_\ell^2}{3c_k}(-1+a_k+2a_\ell-2a_k a_\ell)$, $b_{53}^{(1)} = 5c_k^2(-1-6a_k+2a_\ell+12a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1+14a_k-2a_\ell-28a_k a_\ell)$ $+5c_\ell^2(2a_\ell - 1) + \frac{5c_\ell^3}{c_k}(1-a_k-2a_\ell+2a_k a_\ell)$, $b_{53}^{(0)} = \frac{5c_k^3}{2}(1+14a_k+2a_\ell+28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(-1-30a_k) + \frac{5c_\ell^3}{2}(1-2a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(-1+a_k+2a_\ell-2a_k a_\ell)$, $b_{53}^{(-1)} = \frac{3c_k^4}{8}(-1-30a_k+2a_\ell+60a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1+62a_k-2a_\ell-124a_k a_\ell)$ $+\frac{5c_k^2c_\ell^2}{12}(1+6a_k-2a_\ell-12a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell-1) + \frac{c_k^5}{6c_k}(1-a_k-2a_\ell+2a_k a_\ell)$
$c_\ell - c_k \leq z \leq \frac{c_k}{2}$	f_{54}	$b_{54}^{(5)} = \frac{16}{3c_k c_\ell}(-3+3a_k+4a_\ell-5a_k a_\ell)$, $b_{54}^{(4)} = \frac{8}{c_k}(1-a_k-2a_\ell+4a_k a_\ell) + \frac{8}{c_\ell}(1-2a_k a_\ell)$, $b_{54}^{(3)} = 10(1-2a_\ell+4a_k a_\ell)$, $b_{54}^{(2)} = \frac{20c_k^2}{3c_\ell}(-1-6a_k+8a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1+a_k+2a_\ell-10a_k a_\ell)$, $b_{54}^{(1)} = 5c_k^2(-1-6a_k+2a_\ell-4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1+14a_k-2a_\ell-12a_k a_\ell)$ $+5c_\ell^2(-1+2a_\ell-16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1-a_k-2a_\ell+18a_k a_\ell)$, $b_{54}^{(0)} = \frac{5c_k^3}{2}(1+14a_k+2a_\ell-4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(-1-30a_k+32a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1-2a_\ell+32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(-1+a_k+2a_\ell-34a_k a_\ell)$, $b_{54}^{(-1)} = \frac{3c_k^4}{8}(-1-30a_k+2a_\ell-4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1+62a_k-2a_\ell-60a_k a_\ell)$ $+\frac{5c_k^2c_\ell^2}{12}(1+6a_k-2a_\ell+52a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell-1-64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1-a_k-2a_\ell+66a_k a_\ell)$

Table D.15: Coefficients for the function
$f_{5j}(z) = n_k n_\ell \left(b_{5j}^{(5)} z^5 + b_{5j}^{(4)} z^4 + b_{5j}^{(3)} z^3 + b_{5j}^{(2)} z^2 + b_{5j}^{(1)} z + b_{5j}^{(0)} + b_{5j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $2c_\ell/3 \leq c_k \leq 3c_\ell/4$.

Interval	f_{5j}	**Coefficients**
$\frac{c_k}{2} \leq z \leq \frac{c_\ell}{2}$	f_{55}	$b_{55}^{(5)} = \frac{16}{3c_k c_\ell}(-1 - a_k + 2a_\ell - a_k a_\ell)$, $b_{55}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 4a_k a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 6a_k a_\ell)$, $b_{55}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{55}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 10a_k a_\ell)$, $b_{55}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 18a_k a_\ell)$, $b_{55}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 36a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 34a_k a_\ell)$, $b_{55}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 66a_k a_\ell)$
$\frac{c_\ell}{2} \leq z \leq c_\ell - \frac{c_k}{2}$	f_{56}	$b_{56}^{(5)} = \frac{16}{3c_k c_\ell}(1 - 3a_k - 2a_\ell + 3a_k a_\ell)$, $b_{56}^{(4)} = \frac{8}{c_k}(a_k - 1 + 2a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 6a_k a_\ell)$, $b_{56}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{56}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell - 6a_k a_\ell)$, $b_{56}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 16a_k a_\ell)$ $+5c_\ell^2(2a_\ell - 1 - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(a_k - 1 + 2a_\ell + 14a_k a_\ell)$, $b_{56}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 36a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell - 30a_k a_\ell)$, $b_{56}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(2a_\ell - 1 - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(a_k - 1 + 2a_\ell + 62a_k a_\ell)$
$c_\ell - \frac{c_k}{2} \leq z \leq c_k$	f_{57}	$b_{57}^{(5)} = \frac{16}{3c_k c_\ell}(1 - 3a_k - 3a_\ell + 5a_k a_\ell)$, $b_{57}^{(4)} = \frac{8}{c_k}(a_k - 1 + 4a_\ell - 4a_k a_\ell) + \frac{8}{c_\ell}(4a_k - 1 + a_\ell - 4a_k a_\ell)$, $\quad b_{57}^{(3)} = 10$, $b_{57}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - a_\ell + 10a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 10a_\ell + 10a_k a_\ell)$, $b_{57}^{(1)} = 5c_k^2(-1 - 6a_k) + \frac{5c_k^3}{c_\ell}(18a_k - 1 + a_\ell - 18a_k a_\ell) + 5c_\ell^2(-1 - 6a_\ell)$ $+\frac{5c_\ell^3}{c_k}(a_k - 1 + 18a_\ell - 18a_k a_\ell)$, $b_{57}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - a_\ell + 34a_k a_\ell) + \frac{5c_\ell^3}{2}(1 + 14a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(1 - a_k - 34a_\ell + 34a_k a_\ell)$, $b_{57}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k + a_\ell - 66a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k + 6a_\ell + 36a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 - 30a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 66a_\ell - 66a_k a_\ell)$

Table D.16: Coefficients for the function
$f_{5j}(z) = n_k n_\ell \left(b_{5j}^{(5)} z^5 + b_{5j}^{(4)} z^4 + b_{5j}^{(3)} z^3 + b_{5j}^{(2)} z^2 + b_{5j}^{(1)} z + b_{5j}^{(0)} + b_{5j}^{(-1)} z^{-1}\right) (c_k c_\ell)^{-3/2}$ for $2c_\ell/3 \leq c_k \leq 3c_\ell/4$.

Interval	$\boldsymbol{f_{5j}}$	**Coefficients**
$c_k \leq z \leq \frac{c_\ell + c_k}{2}$	f_{58}	$b_{58}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 3a_\ell + 3a_k a_\ell)$, $b_{58}^{(4)} = \frac{8}{c_k}(a_k - 1 + 4a_\ell - 4a_k a_\ell) + \frac{8}{c_\ell}(a_\ell - 1), \quad b_{58}^{(3)} = 10$, $b_{58}^{(2)} = (\frac{20c_k^2}{3c_\ell}(1 + 6a_k - a_\ell - 6a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 10a_\ell + 10a_k a_\ell)$, $b_{58}^{(1)} = 5c_k^2(-1 - 6a_k) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k + a_\ell + 14a_k a_\ell) + 5c_\ell^2(-1 - 6a_\ell)$ $+\frac{5c_\ell^3}{c_k}(-1 + a_k + 18a_\ell - 18a_k a_\ell)$, $b_{58}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - a_\ell - 30a_k a_\ell) + \frac{5c_\ell^3}{2}(1 + 14a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(1 - a_k - 34a_\ell + 34a_k a_\ell)$, $b_{58}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k + a_\ell + 62a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k + 6a_\ell + 36a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 - 30a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 66a_\ell - 66a_k a_\ell)$
$\frac{c_\ell + c_k}{2} \leq z \leq c_\ell$	f_{59}	Same as f_{49}, see Table D.12.
$c_\ell \leq z \leq c_k + \frac{c_\ell}{2}$	f_{510}	Same as f_{410}, see Table D.12.
$c_k + \frac{c_\ell}{2} \leq z \leq c_\ell + \frac{c_k}{2}$	f_{511}	Same as f_{411}, see Table D.13.
$c_\ell + \frac{c_k}{2} \leq z \leq c_\ell + c_k$	f_{512}	Same as f_{412}, see Table D.13.

Table D.17: Coefficients for the function $f_{6j}(z) = n_k n_\ell \left(b_{6j}^{(5)} z^5 + b_{6j}^{(4)} z^4 + b_{6j}^{(3)} z^3 + b_{6j}^{(2)} z^2 + b_{6j}^{(1)} z + b_{6j}^{(0)} + b_{6j}^{(-1)} z^{-1}\right) (c_k c_\ell)^{-3/2}$ for $3c_\ell/4 \le c_k \le c_\ell$. For $c_k = c_\ell$, the function f_{63} is equal to f_1 of Appendix C in G06.

Interval	f_{6j}	**Coefficients**
$0 \le z \le \frac{c_\ell - c_k}{2}$	f_{61}	$b_{61}^{(5)} = \frac{32}{3c_k c_\ell}(-1 + a_k + a_\ell - a_k a_\ell)$, $b_{61}^{(4)} = \frac{16}{c_k}(a_k - 2a_k a_\ell) + \frac{16}{c_\ell}(1 - 2a_k - a_\ell + 4a_k a_\ell)$, $b_{61}^{(3)} = 0$, $b_{61}^{(2)} = \frac{40c_k^2}{3c_\ell}(-1+2a_k+a_\ell-10a_k a_\ell) + \frac{40c_\ell^2}{3c_k}(2a_k a_\ell - a_k)$, $b_{61}^{(1)} = 0$, $b_{61}^{(0)} = 5c_k^3(1-2a_k+32a_k a_\ell) + \frac{3c_k^4}{c_\ell}(-1+2a_k+a_\ell-34a_k a_\ell) + 10c_\ell^3(a_k - 2a_k a_\ell)$ $+\frac{3c_\ell^4}{c_k}(2a_k a_\ell - a_k)$, $b_{61}^{(-1)} = 0$
$\frac{c_\ell - c_k}{2} \le z \le c_\ell - c_k$	f_{62}	$b_{62}^{(5)} = \frac{16}{3c_k c_\ell}(-3 + 4a_k + 4a_\ell - 6a_k a_\ell)$, $b_{62}^{(4)} = \frac{8}{c_k}(1 - 2a_\ell) + \frac{8}{c_\ell}(1 - 2a_k + 4a_k a_\ell)$, $b_{62}^{(3)} = 10(1 - 2a_\ell - 2a_k + 4a_k a_\ell)$, $b_{62}^{(2)} = \frac{20c_k^2}{3c_\ell}(-1 + 2a_k - 16a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_\ell - 1)$, $b_{62}^{(1)} = 5c_k^2(-1 + 2a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1 - 2a_k - 2a_\ell + 4a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1 - 2a_k - 2a_\ell + 4a_k a_\ell)$, $b_{62}^{(0)} = \frac{5c_k^3}{2}(1 - 2a_k + 2a_\ell + 60a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(-1 + 2a_k - 64a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 + 2a_k - 2a_\ell - 4a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(2a_\ell - 1)$, $b_{62}^{(-1)} = \frac{3c_k^4}{8}(-1 + 2a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1 - 2a_k - 2a_\ell + 4a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 - 2a_k - 2a_\ell + 4a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_k + 2a_\ell - 4a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1 - 2a_k - 2a_\ell + 4a_k a_\ell)$
$c_\ell - c_k \le z \le c_k - \frac{c_\ell}{2}$	f_{63}	$b_{63}^{(5)} = \frac{16}{3c_k c_\ell}(-3 + 4a_k + 4a_\ell - 7a_k a_\ell)$, $b_{63}^{(4)} = \frac{8}{c_k}(1 - 2a_\ell + 2a_k a_\ell) + \frac{8}{c_\ell}(1 - 2a_k + 2a_k a_\ell)$, $b_{63}^{(3)} = 10(1 - 2a_\ell - 2a_k + 8a_k a_\ell)$, $b_{63}^{(2)} = \frac{20c_k^2}{3c_\ell}(2a_k - 1 - 8a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(2a_\ell - 1 - 8a_k a_\ell)$, $b_{63}^{(1)} = 5c_k^2(-1 + 2a_k + 2a_\ell - 20a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1 - 2a_k - 2a_\ell + 20a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_k + 2a_\ell - 20a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1 - 2a_k - 2a_\ell + 20a_k a_\ell)$, $b_{63}^{(0)} = \frac{5c_k^3}{2}(1 - 2a_k + 2a_\ell + 28a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(-1 + 2a_k - 32a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 + 2a_k - 2a_\ell + 28a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + 2a_\ell - 32a_k a_\ell)$, $b_{63}^{(-1)} = \frac{3c_k^4}{8}(-1 + 2a_k + 2a_\ell - 68a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1 - 2a_k - 2a_\ell + 68a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 - 2a_k - 2a_\ell + 68a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_k + 2a_\ell - 68a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1 - 2a_k - 2a_\ell + 68a_k a_\ell)$
$c_k - \frac{c_\ell}{2} \le z \le \frac{c_k}{2}$	f_{64}	$b_{64}^{(5)} = \frac{16}{3c_k c_\ell}(-3 + 3a_k + 4a_\ell - 5a_k a_\ell)$, $b_{64}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 4a_k a_\ell) + \frac{8}{c_\ell}(1 - 2a_k a_\ell)$, $b_{64}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{64}^{(2)} = \frac{20c_k^2}{3c_\ell}(-1 - 6a_k + 8a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 10a_k a_\ell)$, $b_{64}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(1 + 14a_k - 2a_\ell - 12a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 18a_k a_\ell)$, $b_{64}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(-1 - 30a_k + 32a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 34a_k a_\ell)$, $b_{64}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(1 + 62a_k - 2a_\ell - 60a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 66a_k a_\ell)$

Table D.18: Coefficients for the function $f_{6j}(z) = n_k n_\ell \left(b_{6j}^{(5)} z^5 + b_{6j}^{(4)} z^4 + b_{6j}^{(3)} z^3 + b_{6j}^{(2)} z^2 + b_{6j}^{(1)} z + b_{6j}^{(0)} + b_{6j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $3c_\ell/4 \le c_k \le c_\ell$. For $c_k = c_\ell$, the function f_{67} is equal to f_2 of Appendix C in G06.

Interval	f_{6j}	**Coefficients**
$\frac{c_k}{2} \le z \le \frac{c_\ell}{2}$	f_{65}	$b_{65}^{(5)} = \frac{16}{3c_k c_\ell}(-1 - a_k + 2a_\ell - a_k a_\ell)$, $b_{65}^{(4)} = \frac{8}{c_k}(1 - a_k - 2a_\ell + 4a_k a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 6a_k a_\ell)$, $b_{65}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{65}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(-1 + a_k + 2a_\ell - 10a_k a_\ell)$, $b_{65}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(1 - a_k - 2a_\ell + 18a_k a_\ell)$, $b_{65}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 36a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(-1 + a_k + 2a_\ell - 34a_k a_\ell)$, $b_{65}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(1 - a_k - 2a_\ell + 66a_k a_\ell)$
$\frac{c_\ell}{2} \le z \le c_\ell - \frac{c_k}{2}$	f_{66}	$b_{66}^{(5)} = \frac{16}{3c_k c_\ell}(1 - 3a_k - 2a_\ell + 3a_k a_\ell)$, $b_{66}^{(4)} = \frac{8}{c_k}(-1 + a_k + 2a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + 2a_\ell - 6a_k a_\ell)$, $b_{66}^{(3)} = 10(1 - 2a_\ell + 4a_k a_\ell)$, $b_{66}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - 2a_\ell + 12a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 2a_\ell - 6a_k a_\ell)$, $b_{66}^{(1)} = 5c_k^2(-1 - 6a_k + 2a_\ell - 4a_k a_\ell) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k - 16a_k a_\ell)$ $+5c_\ell^2(-1 + 2a_\ell - 16a_k a_\ell) + \frac{5c_\ell^3}{c_k}(-1 + a_k + 2a_\ell + 14a_k a_\ell)$, $b_{66}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k + 2a_\ell - 4a_k a_\ell) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - 2a_\ell + 36a_k a_\ell)$ $+\frac{5c_\ell^3}{2}(1 - 2a_\ell + 32a_k a_\ell) + \frac{3c_\ell^4}{2c_k}(1 - a_k - 2a_\ell - 30a_k a_\ell)$, $b_{66}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k + 2a_\ell - 4a_k a_\ell) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k - 64a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k - 2a_\ell + 52a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 + 2a_\ell - 64a_k a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 2a_\ell + 62a_k a_\ell)$
$c_\ell - \frac{c_k}{2} \le z \le c_k$	f_{67}	$b_{67}^{(5)} = \frac{16}{3c_k c_\ell}(1 - 3a_k - 3a_\ell + 5a_k a_\ell)$, $b_{67}^{(4)} = \frac{8}{c_k}(-1 + a_k + 4a_\ell - 4a_k a_\ell) + \frac{8}{c_\ell}(-1 + 4a_k + a_\ell - 4a_k a_\ell)$, $b_{67}^{(3)} = 10$, $b_{67}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 - 10a_k - a_\ell + 10a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 10a_\ell + 10a_k a_\ell)$, $b_{67}^{(1)} = 5c_k^2(-1 - 6a_k) + \frac{5c_k^3}{c_\ell}(-1 + 18a_k + a_\ell - 18a_k a_\ell) + 5c_\ell^2(-1 - 6a_\ell)$ $+\frac{5c_\ell^3}{c_k}(-1 + a_k + 18a_\ell - 18a_k a_\ell)$, $b_{67}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k) + \frac{3c_k^4}{2c_\ell}(1 - 34a_k - a_\ell + 34a_k a_\ell) + \frac{5c_\ell^3}{2}(1 + 14a_\ell)$ $+\frac{3c_\ell^4}{2c_k}(1 - a_k - 34a_\ell + 34a_k a_\ell)$, $b_{67}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k) + \frac{c_k^5}{6c_\ell}(-1 + 66a_k + a_\ell - 66a_k a_\ell)$ $+\frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k + 6a_\ell + 36a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 - 30a_\ell)$ $+\frac{c_\ell^5}{6c_k}(-1 + a_k + 66a_\ell - 66a_k a_\ell)$

Table D.19: Coefficients for the function $f_{6j}(z) = n_k n_\ell \left(b_{6j}^{(5)} z^5 + b_{6j}^{(4)} z^4 + b_{6j}^{(3)} z^3 + b_{6j}^{(2)} z^2 + b_{6j}^{(1)} z + b_{6j}^{(0)} + b_{6j}^{(-1)} z^{-1} \right) (c_k c_\ell)^{-3/2}$ for $3c_\ell/4 \leq c_k \leq c_\ell$. For $c_k = c_\ell$, the functions f_{610} and f_{612} are equal to f_3, and f_4 of Appendix C in G06, respectively.

Interval	$\boldsymbol{f_{6j}}$	**Coefficients**
$c_k \leq z \leq \frac{c_\ell + c_k}{2}$	f_{68}	$b_{68}^{(5)} = \frac{16}{3c_k c_\ell}(1 - a_k - 3a_\ell + 3a_k a_\ell)$, $b_{68}^{(4)} = \frac{8}{c_k}(-1 + a_k + 4a_\ell - 4a_k a_\ell) + \frac{8}{c_\ell}(a_\ell - 1), \quad b_{68}^{(3)} = 10$, $b_{68}^{(2)} = \frac{20c_k^2}{3c_\ell}(1 + 6a_k - a_\ell - 6a_k a_\ell) + \frac{20c_\ell^2}{3c_k}(1 - a_k - 10a_\ell + 10a_k a_\ell)$, $b_{68}^{(1)} = 5c_k^2(-1 - 6a_k) + \frac{5c_k^3}{c_\ell}(-1 - 14a_k + a_\ell + 14a_k a_\ell) + 5c_\ell^2(-1 - 6a_\ell)$ $+ \frac{5c_\ell^3}{c_k}(-1 + a_k + 18a_\ell - 18a_k a_\ell)$, $b_{68}^{(0)} = \frac{5c_k^3}{2}(1 + 14a_k) + \frac{3c_k^4}{2c_\ell}(1 + 30a_k - a_\ell - 30a_k a_\ell) + \frac{5c_\ell^3}{2}(1 + 14a_\ell)$ $+ \frac{3c_\ell^4}{2c_k}(1 - a_k - 34a_\ell + 34a_k a_\ell)$, $b_{68}^{(-1)} = \frac{3c_k^4}{8}(-1 - 30a_k) + \frac{c_k^5}{6c_\ell}(-1 - 62a_k + a_\ell + 62a_k a_\ell)$ $+ \frac{5c_k^2 c_\ell^2}{12}(1 + 6a_k + 6a_\ell + 36a_k a_\ell) + \frac{3c_\ell^4}{8}(-1 - 30a_\ell)$ $+ \frac{c_\ell^5}{6c_k}(-1 + a_k + 66a_\ell - 66a_k a_\ell)$
$\frac{c_\ell + c_k}{2} \leq z \leq c_\ell$	f_{69}	Same as f_{49}, see Table D.12.
$c_\ell \leq z \leq c_k + \frac{c_\ell}{2}$	f_{610}	Same as f_{410}, see Table D.12.
$c_k + \frac{c_\ell}{2} \leq z \leq c_\ell + \frac{c_k}{2}$	f_{611}	Same as f_{411}, see Table D.13.
$c_\ell + \frac{c_k}{2} \leq z \leq c_\ell + c_k$	f_{612}	Same as f_{412}, see Table D.13.