

Quantum-ready formulation for uncompromised data assimilation with high-rank covariances

Takemasa Miyoshi **RIKEN Center for Computational Science, Kobe, Japan**and RIKEN Center for Interdisciplinary Theoretical and Mathematical Sciences, Kobe, Japan*

(Received 11 May 2026; accepted 11 August 2026; published 8 September 2026)

Data assimilation (DA) optimally fuses observational data with mathematical models, serving as a core computational framework for forecasting complex dynamical systems like the weather. Because manipulating large-scale high-rank error covariances is classically intractable, conventional DA is forced into severe structural compromises such as artificial localization or static parametrizations that destroy genuine cross-variable correlations. Future fault-tolerant quantum computing (FTQC) could offer transformative advantages to overcome these classical computational limits. While quantum linear system solvers theoretically provide exponential speedups for matrix inversion, the practical mapping of large-scale, uncompromised DA onto quantum architectures remains largely unexplored. Our framework builds upon two emerging technologies: FTQC hardware and efficient block-encoded oracles of the error covariance and observation operators, which we envision being supplied by future artificial intelligence surrogate models. Under this premise, we identify a potential bottleneck in the quantum implementation. We show that a direct mapping of the DA equations leads to a quadratic increase in the subnormalization factor (α) associated with block encoding when typical nonlocal observations are utilized. To mitigate this algorithmic overhead, we propose a quantum-ready, combined state-observation space formulation. By reconfiguring the DA problem into an expanded saddle-point block-matrix system, our approach avoids the explicit multiplication of nonlocal observation operators. This maintains a linear scaling of α and confines the conditioning of the quantum solver to that of the observation error covariance, decoupled from the severely ill-conditioned prior. Numerical verification using the chaotic Lorenz-96 model demonstrates that our formulation achieves perfect fidelity with the uncompromised DA solution. Although reading out the full analysis state reduces the exponential speedup to a polynomial one, the framework retains a substantial advantage over the classical matrix inversion. This work bridges computational geosciences and quantum information science, providing the essential algorithmic foundation for large-scale DA to become a compelling application for FTQC.

DOI: [10.1103/x91t-v327](https://doi.org/10.1103/x91t-v327)

I. INTRODUCTION

Since the foundational formulation of variational objective analysis [1], data assimilation (DA) has developed primarily within numerical weather prediction, driving a decades-long continuous revolution in forecasting capabilities [2], to optimally fuse mathematical models with noisy observations, from historically sparse *in situ* networks to increasingly dense satellite measurements [3]. Today, it serves as a core computational engine for forecasting a wide range of complex dynamical systems, catalyzing the development of large-scale digital twins across broad scientific disciplines [4]. To optimally extract observational information and accurately track the true system trajectory, modern DA fundamentally relies on the error covariance matrix of the prior model state estimate. Ideally, to capture the intricate cross-variable correlations that dictate the system's multivariate dynamics, this matrix must

be both dynamically evolving and untruncated. However, for state spaces reaching millions or billions of variables ($N \sim 10^8 - 10^9$, where N represents the state-space dimension), explicit algebraic manipulation poses insurmountable barriers for classical high-performance computing (HPC), demanding an $\mathcal{O}(N^2)$ memory footprint and cubic $\mathcal{O}(N^3)$ matrix inversion scaling.

To maintain tractability under these HPC barriers, conventional DA is forced into severe structural compromises. Variational DA, a mainstream approach, typically approximates the covariance using static or heavily parametrized representations, prioritizing computational feasibility at the expense of critical dynamical evolution and by imposing structural constraints on the prior error covariance [3]. Ensemble-based DA, such as the ensemble Kalman filter (EnKF) [5,6], is the other mainstream approach that utilizes low-rank approximations. While this approach successfully exploits the inherently low-dimensional error structure of local dynamics [7], the inevitable sampling noise requires artificial localization. Although localization mathematically suppresses this noise and increases the rank, it fundamentally destroys genuine cross-variable correlations. The cost of this compromise was explicitly demonstrated when an unprecedented (10 240)-member EnKF successfully removed localization, yielding profound improvements in state

*Contact author: takemasa.miyoshi@riken.jp

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

estimation and forecast accuracy by fully capturing these long-range multivariate dependencies [8–10]. Thus, classical DA remains trapped in an inescapable computational dilemma: To maintain tractability, the uncompromised covariance structure must be either approximated by parametrized static models or artificially localized, preventing an uncompromised state estimation.

Recent advancements in artificial intelligence (AI) offer a potential pathway to break this classical paradigm. Deep learning surrogates can now inexpensively evolve massive data-driven ensembles without conventional sampling errors, directly generating uncompromised, dynamically evolving covariances [11,12]. Yet, the $\mathcal{O}(N^2)$ memory footprint and $\mathcal{O}(N^3)$ matrix inversion bottlenecks remain. While recent exploratory studies have introduced the concept of quantum DA by utilizing quantum annealing to heuristically minimize the variational cost function [13], such analog optimization approaches do not mathematically resolve the explicit cubic scaling barrier. Here, quantum linear system solvers (QLSS) [14] offer a theoretical breakthrough by exponentially reducing the $\mathcal{O}(N^3)$ scaling to a logarithmic dependence $\mathcal{O}(\kappa \text{polylog} N)$, where κ is the condition number of the matrix. We propose that large-scale DA represents a highly compelling, real-world computational challenge for QLSS, providing a critical use case where overcoming the $\mathcal{O}(N^3)$ bottleneck translates to profound societal impact. Executing the deep quantum circuits required for QLSS relies on the future realization of fully error-corrected, fault-tolerant quantum computers (FTQC). As the hardware community aggressively pursues this error-free quantum computing era, it is imperative to establish the theoretical algorithms that will run on them.

However, exploiting this quantum speedup requires overcoming two fundamental barriers. While recent advances in quantum algorithm engineering have successfully leveraged QLSS to solve specific classical differential equations [15,16], the fundamental mathematical structure of these problems differs from DA. Closer to our setting, recent studies formulated the classical Kalman filter on block-encoding arithmetic: a quantum Kalman filter composing the required matrix additions, multiplications, and inversions as block-encoded operations [17], subsequently extended to the full linear-quadratic-Gaussian control pipeline via quantum singular value transformation [18]. These pioneering works demonstrate that the sequential Kalman filter equations are expressible in quantum linear algebra in proof-of-concept settings. However, block-encoded products multiply their subnormalization factors and block-encoded inversions import the operand’s condition number, so that such direct compositions inherit the bottlenecks that dominate the large-scale DA regime: severely ill-conditioned, dynamically evolving priors and dense nonlocal observation operators.

Formulating large-scale DA for quantum execution introduces its own unique bottlenecks. First is the input problem, or state preparation bottleneck: Classically loading the uncompromised covariance matrix demands $\mathcal{O}(N^2)$ time, destroying the exponential quantum advantage. To solve this, we adopt an oracle-based approach, a standard paradigm in quantum algorithm design. We postulate a future where advanced AI surrogates directly output the block-encoded quantum oracle of the uncompromised covariance, without classically

instantiating the matrix elements. By establishing this AI-oracle premise alongside the anticipation of FTQC hardware, we isolate the true algorithmic barriers. The uncompromised covariance usually exhibits extreme eigenvalue disparities reflecting complex dynamics, creating severe ill-conditioning. Furthermore, when assimilating modern remote-sensing observations with spatially broad footprints, their nonlocal observation operators destroy the mathematical sparsity required to bound the quantum subnormalization factor (α). Because the gate complexity of QLSS scales proportionally with α [19], this loss of sparsity triggers a quadratic increase in quantum circuit depth, potentially limiting the exponential quantum advantage. Consequently, directly applying naive QLSS to these ill-conditioned and dense classical formulations results in computational failure.

Therefore, realizing uncompromised DA requires finding its quantum-ready formulation. To facilitate this interdisciplinary endeavor, this paper first provides a concise introduction to the essence of the DA problem for nonexperts in the broader computational science and quantum computing communities (Sec. II). Building upon this foundation, we present a quantum-ready mathematical framework designed to overcome these intertwined algorithmic and conditioning barriers (Sec. III). We systematically introduce three distinct pathways for quantum execution: mapping the classical state-space and observation-space DA formulations into linear systems to expose their bottlenecks and proposing a quantum-ready combined state-observation space formulation. This combined approach constructs an expanded block-matrix system that preserves sparsity and avoids explicitly forming the ill-conditioned, dense intermediate matrices required by classical formulations. Through proof-of-concept numerical experiments with the chaotic Lorenz-96 model [20,21], we demonstrate that this framework achieves perfect fidelity with the uncompromised DA solution, drastically outperforming classically compromised methods under broad observational footprints. By proving that quantum algorithms can execute the exact algebraic fusion previously deemed impossible, this work unlocks the full potential of high-rank error covariances for optimal state estimation in large-scale, complex dynamical systems.

II. THE DATA ASSIMILATION PROBLEM

To establish the mathematical foundation for our quantum-ready framework, we first define the exact DA problem, explain why explicit computation is classically intractable, and review how conventional HPC has historically approximated it.

A. The exact problem

DA solves the state-estimation problem of a dynamical system by optimally fusing two independent sources of information: a prior model forecast and observations. Following the standard DA convention, we refer to this prior forecast as the *background* state throughout; readers from the dynamical-systems community may read “background” interchangeably as the model forecast or the prior. Let $\mathbf{x} \in \mathbb{R}^N$ be the state vector. We denote the background state as $\mathbf{x}^b \in \mathbb{R}^N$. Assuming

the background errors are unbiased and Gaussian, the prior probability density function (PDF) of the state given the background is

$$\Pr(\mathbf{x}|\mathbf{x}^b) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b)\right), \quad (1)$$

where $\mathbf{B} \in \mathbb{R}^{N \times N}$ is the background error covariance matrix. In our exact formulation, we assume $\mathbf{B}$ is untruncated, capturing all genuine cross-variable correlations.

Next, let $\mathbf{y}^o \in \mathbb{R}^p$ be the observation vector, where p is the number of observations. The relationship between the state and observations is described by a generally nonlinear observation operator $\mathcal{H}$, such that $\mathbf{y}^o = \mathcal{H}(\mathbf{x}^t) + \boldsymbol{\epsilon}$, where superscript t represents truth. Assuming the observation errors $\boldsymbol{\epsilon}$ are unbiased, Gaussian, and independent of the background errors, with the observation error covariance matrix $\mathbf{R} \in \mathbb{R}^{p \times p}$, the likelihood of obtaining the specific observations $\mathbf{y}^o$ given a hypothetical state $\mathbf{x}$ is

$$\Pr(\mathbf{y}^o|\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{y}^o - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathcal{H}(\mathbf{x}))\right). \quad (2)$$

Because the background forecast and the new observations represent independent sources of information, the posterior PDF, representing the optimal probabilistic estimate of the state given both sources, is given by Bayes' theorem as the product of the likelihood and the prior:

$$\Pr(\mathbf{x}|\mathbf{x}^b, \mathbf{y}^o) \propto \Pr(\mathbf{y}^o|\mathbf{x})\Pr(\mathbf{x}|\mathbf{x}^b). \quad (3)$$

To find the optimal state estimate, known as the analysis state $\mathbf{x}^a$, we seek the mode of this joint posterior PDF (maximum *a posteriori* or simply MAP estimate). Maximizing $\Pr(\mathbf{x}|\mathbf{x}^b, \mathbf{y}^o)$ is mathematically equivalent to minimizing its negative natural logarithm. This derives the fundamental variational cost function $J(\mathbf{x})$ [1,3], separated into the background term J_b and the observation term J_o :

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y}^o - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathcal{H}(\mathbf{x})). \quad (4)$$

To solve this minimization problem efficiently and avoid direct nonlinear optimization, $J(\mathbf{x})$ is typically linearized around the background state $\mathbf{x}^b$. Defining the state increment $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$, the innovation vector $\mathbf{d} = \mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)$, and the linearized observation operator $\mathbf{H} = \frac{\partial \mathcal{H}}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{x}^b} \in \mathbb{R}^{p \times N}$, the incremental cost function becomes quadratic:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T \mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{d} - \mathbf{H}\delta\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\delta\mathbf{x}). \quad (5)$$

The exact analytical minimum, yielding the optimal analysis increment $\delta\mathbf{x}^a$, is found by setting the gradient $\nabla J(\delta\mathbf{x})$ to zero. This leads to the fundamental formulation of exact DA:

$$(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})\delta\mathbf{x} = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{d}. \quad (6)$$

By applying the Sherman-Morrison-Woodbury matrix identity, this state-space formulation can be transformed into an equivalent formulation in the observation space [22]. This transformation yields the state increment as

$$\delta\mathbf{x} = \mathbf{B}\mathbf{H}^T \mathbf{z}, \quad (7)$$

where the vector $\mathbf{z} \in \mathbb{R}^p$ is the solution to the following observation-space linear system:

$$(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})\mathbf{z} = \mathbf{d}. \quad (8)$$

By substituting $\mathbf{z}$ from Eq. (8) into Eq. (7), the explicit analysis update can be compactly written as $\delta\mathbf{x} = \mathbf{K}\mathbf{d}$, where $\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$ is defined as the Kalman gain. This matrix serves as the optimal weight, dictating exactly how much the prior state should be corrected by balancing the background and observation error covariances.

When the observation operator $\mathcal{H}$ is mildly nonlinear, variational systems may iterate this linearization: The quadratic problem is solved, the state is updated, and $\mathbf{H}$ and $\mathbf{d}$ are reevaluated about the new state, forming so-called outer loops that realize an incremental optimization [23]. Each outer iteration is algebraically identical to the linear system above, so all the formulations developed in this paper apply iterationwise to mildly nonlinear problems. In the quantum setting discussed below, each outer iteration would additionally require reading out the analysis increment to relinearize the operators, multiplying the total cost by the number of outer iterations, which is typically only a few.

Both Eqs. (6) and (8) represent the exact algebraic fusion of the model state and observations. The observation-space formulation [Eq. (8)] is efficient when observations are sparse, i.e., $p \ll N$. However, in the modern era of big data, driven by, e.g., massive satellite constellations and ubiquitous sensor networks, the observation dimension p is generally exploding, potentially rivaling N in some applications. Consequently, solving either system exposes fundamental classical impossibilities. Specifically, realizing exact DA in large-scale applications via conventional HPC is precluded because the system inevitably encounters at least one of the following five classical barriers:

- (1) *Memory footprint*: Explicitly storing $\mathbf{B}$ demands an $\mathcal{O}(N^2)$ memory capacity.
- (2) *State-space computational cost*: Computing the inverse $\mathbf{B}^{-1}$ in Eq. (6) requires $\mathcal{O}(N^3)$ operations.
- (3) *Observation-space computational cost*: While reducing the dimension to p avoids the explicit inversion of $\mathbf{B}$, classically constructing the operator $(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})$ in Eq. (8) requires the explicit matrix multiplication $\mathbf{H}\mathbf{B}\mathbf{H}^T$ with $\mathcal{O}(p^2N)$ operations.
- (4) *Prior ill-conditioning*: $\mathbf{B}$ tends to present extreme eigenvalue disparities reflecting complex dynamics. This severe ill-conditioning destabilizes numerical solvers.
- (5) *Observation complexity*: In advanced DA accommodating complex sensors, the observation error covariance $\mathbf{R}$ itself may become an ill-conditioned matrix, imposing further $\mathcal{O}(p^3)$ inversion bottlenecks with overall numerical instability.

Crucially, these bottlenecks are insurmountable regardless of future classical HPC advancements. As computational power increases, we usually increase the complexity of dynamical models by, e.g., increasing the spatial resolution and including more precise physical representations, continuously driving up N . Therefore, explicitly holding and manipulating the full matrices remains an eternally moving target.

B. Classical DA approximations

Faced with the fundamental impossibility of solving the exact problem, the historical trajectory of classical DA research has been dedicated to developing practical approximations that make the computation feasible for operational HPC.

In variational DA, this computational bottleneck is resolved by approximating $\mathbf{B}$ with a static and heavily parametrized representation $\mathbf{B}_c$. To avoid storing the massive $N \times N$ matrix and computing the explicit inversion $\mathbf{B}_c^{-1}$, a control variable transformation is typically introduced. By defining $\mathbf{B}_c = \mathbf{U}\mathbf{U}^T$ and expressing the state increment as $\delta\mathbf{x} = \mathbf{U}\mathbf{v}$, where $\mathbf{v}$ is a preconditioned control variable, the incremental cost function [Eq. (5)] is transformed into

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{d} - \mathbf{H}\mathbf{U}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\mathbf{U}\mathbf{v}). \quad (9)$$

This mathematical transformation successfully eliminates the $\mathcal{O}(N^3)$ bottleneck by reducing the background term to a simple inner product. Moreover, to simultaneously evade the $\mathcal{O}(N^2)$ memory bottleneck, the full matrix $\mathbf{B}_c$ is never explicitly constructed or stored. Instead, variational DA research has been profoundly focused on the intricate design of $\mathbf{U}$. By engineering $\mathbf{U}$ not as an explicit matrix, but as a sequence of matrix-free, physically based operational steps such as spatial smoothers and dynamical balance constraints, practitioners can implicitly apply the covariance structure. While this ingenious operator design makes large-scale DA computationally feasible, it fundamentally forces $\mathbf{B}_c$ to remain a structurally simplified, heavily parametrized approximation, sacrificing the untruncated, flow-dependent multivariate correlations of the exact $\mathbf{B}$.

In addition, $\mathbf{R}^{-1}$ remains. To ensure that the $\mathcal{O}(p^3)$ inversion of $\mathbf{R}$ does not become a new intractable barrier, classical DA has historically relied on another approximation: treating $\mathbf{R}$ as a purely diagonal matrix, implying that all observation errors are mutually independent. Although recent advanced operational systems have begun incorporating block-diagonal structures to account for specific cross-correlations such as interchannel error correlations in satellite radiances, the massive global $\mathbf{R}$ matrix remains overwhelmingly diagonal in practice.

Alternatively, ensemble-based DA methods such as the EnKF attempt to capture the dynamical flow dependence by estimating the covariance from a finite ensemble of m model forecasts. To maintain computational feasibility, m is typically chosen to be much smaller than the system dimension ($m \ll N$); for instance, in operational numerical weather prediction, $m \sim 100$. The ensemble-based sample covariance is defined as

$$\mathbf{P}_e = \frac{1}{m-1}\mathbf{X}\mathbf{X}^T, \quad (10)$$

where $\mathbf{X} \in \mathbb{R}^{N \times m}$ is the ensemble perturbation matrix. Denoting the ensemble of model forecasts by $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ and their mean by $\bar{\mathbf{x}} = \frac{1}{m}\sum_{i=1}^m \mathbf{x}_i$, the i th column of $\mathbf{X}$ is the ensemble perturbation $\mathbf{x}_i - \bar{\mathbf{x}}$, i.e., $\mathbf{X} = (\mathbf{x}_1 - \bar{\mathbf{x}}, \dots, \mathbf{x}_m - \bar{\mathbf{x}})$. In addition, we define the observation-space ensemble perturbations $\mathbf{Y} = \mathbf{H}\mathbf{X} \in \mathbb{R}^{p \times m}$. Substituting Eq. (10) into the observation-space formulations [Eqs. (7) and (8)] yields

$$\left(\frac{1}{m-1}\mathbf{Y}\mathbf{Y}^T + \mathbf{R} \right) \mathbf{z} = \mathbf{d}, \quad (11)$$

$$\delta\mathbf{x} = \frac{1}{m-1}\mathbf{X}\mathbf{Y}^T\mathbf{z}. \quad (12)$$

This mathematical substitution reveals two major classical computational savings. First, the $\mathcal{O}(p^2N)$ dense matrix multiplication $\mathbf{H}\mathbf{B}\mathbf{H}^T$ is replaced by $\mathbf{Y}\mathbf{Y}^T$, reducing the cost to

$\mathcal{O}(p^2m)$. Because $m \ll N$, the dimensional dependence on N is eliminated from the matrix construction. Second, by applying the Sherman-Morrison-Woodbury identity to Eq. (11), the required explicit inversion is fundamentally shifted from a $p \times p$ space to a computationally trivial $m \times m$ matrix in the ensemble subspace [24,25].

By utilizing this formulation, the approach drastically compresses the intractable $\mathcal{O}(N^2)$ memory footprint into a manageable $N \times m$ matrix and entirely bypasses the cubic inversion bottlenecks. However, because operational ensemble sizes are severely limited by massive computational costs ($m \ll N$), the sample covariance $\mathbf{P}_e$ is fundamentally rank deficient with a maximum rank of $m - 1$. Furthermore, it suffers from severe sampling noise, which manifests as spurious correlations.

To mathematically suppress this sampling noise and increase the rank, ensemble methods necessitate artificial localization. Conceptually, this is defined via a Schur product with a tapering matrix $\boldsymbol{\rho} \in \mathbb{R}^{N \times N}$:

$$\mathbf{B} \approx \boldsymbol{\rho} \circ \mathbf{P}_e. \quad (13)$$

However, explicitly constructing and computing this $N \times N$ Schur product would immediately reintroduce the intractable $\mathcal{O}(N^2)$ memory wall. To circumvent this contradiction, practical strategies have been developed to avoid forming the massive full matrix, broadly categorized into two main approaches. The first applies the tapering function during the computation of the Kalman gain, avoiding dense $N \times N$ operations by exploiting sparse matrix algorithms [26]. The second is known as observation localization or domain localization, which performs the analysis independently for each local state variable (or local patch). This is achieved by effectively tapering $\mathbf{R}^{-1}$ based on physical distance, e.g., the local ensemble transform Kalman filter (LETKF) [24].

These localized computations successfully circumvent the memory wall, reduce sampling noise, and mathematically force the effectively truncated local matrices to become high rank and invertible, allowing the classical DA system to function despite the extremely small ensemble size. However, they destroy genuine, physically meaningful correlations.

These classical structural compromises, whether parametrized static models or artificial localization, inadvertently served a profound secondary mathematical purpose: They heavily regularized the system. The untruncated covariance matrix often reflects complex dynamics with extreme eigenvalue disparities, meaning the $\mathbf{B}$ matrix tends to be severely ill-conditioned. By artificially truncating the correlations via localization or heavily parametrizing them via control variables, classical methods implicitly mitigate this severe ill-conditioning. Consequently, the moment we abandon these classical approximations to execute DA with an uncompromised $\mathbf{B}$, this underlying mathematical pathology is exposed, creating a critical numerical barrier for new computational paradigms such as quantum algorithms.

III. QUANTUM-READY FORMULATION

As established in Sec. II, the goal of executing uncompromised DA has historically been blocked by classical computational walls. In our envisioned quantum-AI hybrid

framework, we assume that future AI can directly generate quantum circuits that explicitly represent the uncompromised, dynamically evolving background error covariance $\mathbf{B}$. Because the massive $N \times N$ matrix is embedded directly as quantum operators, this AI-driven circuit generation fundamentally eradicates the intractable $\mathcal{O}(N^2)$ memory footprint (Barrier 1).

With the memory barrier resolved by AI, the bottleneck shifts to the algebraic fusion of this uncompromised $\mathbf{B}$ with massive observation networks. Here, QLSS provides the essential theoretical breakthrough to overcome the classical cubic computational walls (Barrier 2). Modern optimal QLSS algorithms rely on a block-encoding technique [19,27], which embeds a nonunitary matrix $\mathbf{A}$ into a larger unitary quantum operator. To satisfy the strict unitarity requirement of quantum circuits, $\mathbf{A}$ must be scaled down by a subnormalization factor [19] α such that $\alpha \geq \|\mathbf{A}\|_2$, where $\|\mathbf{A}\|_2$ is the spectral norm (i.e., the maximum singular value) of $\mathbf{A}$. For a given linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$, these block-encoding-based QLSS algorithms can find the solution state with a computational complexity scaling proportionally to $\mathcal{O}(\alpha\kappa \text{polylog}N)$, where κ is the condition number of $\mathbf{A}$. By executing the matrix inversion exponentially faster, QLSS reduces the $\mathcal{O}(N^3)$ scaling to a logarithmic dependence on the system dimension.

However, extracting this exponential quantum advantage requires careful mathematical formulation. Since the uncompromised DA problems are already expressed as linear systems [Eqs. (6) and (8)], mapping them to a quantum linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ appears mathematically straightforward. Yet, executing these conventional formulations on quantum hardware for realistic, large-scale DA reveals critical algorithmic bottlenecks.

If we map the state-space formulation [Eq. (6)] directly to a quantum circuit, we hit the mathematical wall of ill-conditioning. Constructing the quantum oracle for $\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ requires the explicit quantum inversion of the uncompromised covariances. This severely amplifies the circuit depth proportionally to the extreme condition numbers $\kappa(\mathbf{B})$ and $\kappa(\mathbf{R})$ (Barriers 4 and 5).

Alternatively, the observation-space formulation [Eq. (8)] avoids the explicit inversion of $\mathbf{B}$. Indeed, for highly simplified models or when $\mathbf{R}$ is well conditioned, the addition of $\mathbf{R}$ acts as a Tikhonov regularization, mitigating the ill-conditioning of $\mathbf{B}$ and allowing QLSS to solve the system efficiently. However, in some applications including the operational numerical weather prediction, the observation operator $\mathbf{H}$ is rarely a simple sparse matrix; it is typically highly complex, especially for remote-sensing data (e.g., radiative transfer models) that dominate modern numerical weather prediction.

This complexity exposes two critical algorithmic bottlenecks in Eq. (8). First, QLSS only yields the intermediate quantum state $\mathbf{z}$. Extracting the analysis increment requires an additional operation of Eq. (7). In quantum mechanics, applying such a nonunitary operator requires a probabilistic postselection step; if the measurement fails, the entire QLSS must be rerun from the beginning, severely degrading overall algorithmic efficiency. Second, constructing the explicitly fused oracle $\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}$ via the linear combination of unitaries (LCU) [19,27] exponentially inflates the

subnormalization factor α . In the block-encoding framework, explicitly multiplying encoded operators inherently requires the multiplication of their respective subnormalization factors. Consequently, the block encoding of $\mathbf{H}\mathbf{B}\mathbf{H}^T$ triggers a quadratic explosion with respect to the observation operator, scaling as $\mathcal{O}(\alpha_B \alpha_H^2)$. This quadratic explosion in the subnormalization factor destroys the efficiency of QLSS for large-scale, complex observation scenarios.

To overcome the major classical barriers and simultaneously resolve the quantum barrier for complex observations, we must reformulate the exact DA equations into a quantum-ready formulation. Here, we propose a quantum-ready combined state-observation space formulation. Augmented saddle-point systems of this type are well established in classical numerical linear algebra and optimization [28]; they are a standard approach to avoid the squaring of condition numbers incurred when a least-squares problem is reduced to its normal equations, the single linear system obtained by setting the gradient of the quadratic cost to zero [29]; Eq. (6) is the DA analog of the normal equations. Our contribution is not the augmentation itself but its identification as the natural quantum-ready structure for DA. As shown below, in the block-encoding framework the same single algebraic rearrangement simultaneously controls the multiplicative composition of subnormalization factors, confines the conditioning to that of $\mathbf{R}$, and eliminates the probabilistic postselection step. These three obstacles are distinct in the quantum setting. In classical operational DA, such expanded systems have historically been avoided due to their massively expanded $(N + p)$ dimensionality. However, in the quantum regime, representing this expanded $(N + p)$ -dimensional space requires at most one additional data qubit (i.e., $\lceil \log_2(N + p) \rceil \leq \lceil \log_2(N) \rceil + 1$ for $p \leq N$). This makes the additional overhead computationally negligible. Rather than projecting the problem solely into the state space or the observation space, we retain both $\delta\mathbf{x}$ and $\mathbf{z}$ simultaneously. By algebraically uncoupling the exact DA equations, we can construct an expanded block-matrix linear system that naturally circumvents both matrix densification and extreme ill-conditioning:

$$\begin{pmatrix} \mathbf{I} & -\mathbf{B}\mathbf{H}^T \\ \mathbf{H} & \mathbf{R} \end{pmatrix} \begin{pmatrix} \delta\mathbf{x} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{d} \end{pmatrix}. \quad (14)$$

The practical implementation of this framework relies on the assumption that the entire block matrix can be efficiently generated as a single quantum oracle. Just as advanced AI provides the oracle for $\mathbf{B}$, we assume AI-based surrogate models can provide the block-encoded oracle for complex $\mathbf{H}$. Furthermore, by the foundational principles of quantum mechanics, once the unitary oracle for $\mathbf{H}$ is constructed, its transpose $\mathbf{H}^T$ is obtained virtually for free by simply applying the inverse quantum circuit. Consequently, the entire expanded block matrix can be systematically constructed using standard quantum matrix arithmetic and block-encoding techniques, such as LCU.

By structuring the quantum oracle in this way, the combined formulation eliminates explicit matrix inversions, mathematically immunizing the quantum solver against the severe ill-conditioning of $\mathbf{B}$. The inverse of our expanded

block matrix is analytically expressed as

$$\begin{pmatrix} \mathbf{I} & -\mathbf{B}\mathbf{H}^T \\ \mathbf{H} & \mathbf{R} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I} - \mathbf{B}\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H} & \mathbf{B}\mathbf{H}^T\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{H} & \mathbf{S}^{-1} \end{pmatrix}, \quad (15)$$

where $\mathbf{S} = \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T$ is the Schur complement of the upper-left block. Because $\mathbf{B}$ is a positive semidefinite covariance matrix, $\mathbf{H}\mathbf{B}\mathbf{H}^T$ is also positive semidefinite, which guarantees $\mathbf{S} \geq \mathbf{R}$ in the sense of the Loewner partial order. This matrix inequality ensures that the minimum eigenvalue of $\mathbf{S}$ is bounded by the minimum eigenvalue of $\mathbf{R}$, implying $\|\mathbf{S}^{-1}\|_2 \leq \|\mathbf{R}^{-1}\|_2$. Since the analytical form of Eq. (15) exclusively involves forward multiplications of $\mathbf{B}$ and the bounded operator $\mathbf{S}^{-1}$, the overall norm of the system remains decoupled from the near-zero eigenvalues of $\mathbf{B}$. Consequently, the extreme eigenvalue disparities of $\mathbf{B}$ do not govern the combined matrix's spectrum. Instead, the condition number is primarily bounded by the condition number of the observation error covariance $\mathbf{R}$. A mathematical upper bound formalizing this statement is derived in Appendix B and verified numerically in Sec. IV D: The condition number of the combined system is controlled by $\|\mathbf{R}^{-1}\|_2$ and the operator norms α_B , α_H , and α_R , with no dependence on $\kappa(\mathbf{B})$. In traditional DA regimes where observations are sparse and assumed mutually independent, this property perfectly isolates the quantum solver. However, as highlighted in Barrier 5, $\mathbf{R}$ may remain severely ill-conditioned. While our formulation does not eliminate this intrinsic issue, it reduces the fundamental challenge to managing the static conditioning of $\mathbf{R}$ [30], rather than confronting the intractable, dynamically evolving $\mathbf{B}$.

In addition, by isolating the unmultiplied operators into separate subblocks, the combined formulation completely avoids the explicit computation of $\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}$. This isolation inherently bounds the total subnormalization factor to a linear dependence on the subblocks, i.e., $\mathcal{O}(\alpha_B\alpha_H + \alpha_R)$, preventing the quadratic explosion that renders the observation-space formulation intractable. Furthermore, executing QLSS on this system yields the combined quantum state directly. This allows for the one-shot preparation of the quantum state $\delta\mathbf{x}$, avoiding the probabilistic postselection penalties that would arise from explicitly applying the nonunitary operator $\mathbf{B}\mathbf{H}^T\mathbf{z}$ to intermediate quantum states.

Consequently, this combined formulation provides the first universally scalable pathway to algebraically fuse AI-generated untruncated error covariances with massive, complex observation networks, achieving uncompromised DA without artificial localization or structural compromises.

IV. PROOF-OF-CONCEPT EXPERIMENTS

To empirically validate the theoretical advantages of the proposed quantum-ready framework in Sec. III, we perform proof-of-concept experiments using the chaotic Lorenz-96 model [20,21] widely used in various DA studies. First, Sec. IV A details the experimental setup, adapting the general DA problem in Sec. II to this specific dynamical system. Next, Sec. IV B focuses on the structural complexity of the observation operator $\mathbf{H}$, illustrating how the subnormalization factor α severely inflates in the conventional formulation and how the proposed formulation avoids it. Finally, Sec. IV C

evaluates the power of uncompromised DA compared to the classical EnKF with localization under two contrasting regimes: a sparse observation network where α is small mimicking *in situ* observations and a highly dense network where α is prohibitively large mimicking satellite radiance observations.

A. Experimental setup and the Lorenz-96 model

The Lorenz-96 model [20,21] is a low-dimensional surrogate for atmospheric dynamics along a latitude circle, widely used in the DA community to test novel algorithms. It represents an unspecified scalar meteorological quantity x_i at N equidistant grid points along a one-dimensional periodic latitude circle. The system's evolution is governed by a set of ordinary differential equations:

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F \quad (i = 1, \dots, N), \quad (16)$$

where the cyclic boundary conditions are defined as $x_{-1} = x_{N-1}$, $x_0 = x_N$, and $x_{N+1} = x_1$. The three terms on the right-hand side conceptually simulate nonlinear advection, linear dissipation, and external forcing, respectively. We adopt the canonical configuration with $N = 40$ and a forcing parameter of $F = 8.0$, which drives the system into a strongly chaotic regime. The equations are integrated using a fourth-order Runge-Kutta scheme with a nondimensional time step of $\Delta t = 0.01$.

To connect this dynamical system to the DA problem in Sec. II, we define the model state vector as $\mathbf{x} \in \mathbb{R}^{40}$. We conduct an identical twin experiment where a “nature run,” representing the true state trajectory of the system, is generated by integrating the model from a spun-up initial condition. Following the measurement equation in Sec. II, synthetic observation vectors $\mathbf{y}^o \in \mathbb{R}^p$ are periodically extracted from this true state every $5\Delta t$, i.e., a nondimensional time interval of 0.05 that represents 6 h in analogy of synoptic weather whose error doubles in approximately 2.5 days. These observations are generated by applying an observation operator $\mathbf{H}$ and injecting random Gaussian noise with the observation error covariance $\mathbf{R}$.

The goal of these experiments is to perform the optimal state estimation using the true, untruncated background error covariance $\mathbf{B}$. In the subsequent sections, we systematically manipulate the spatial footprint of $\mathbf{H}$ to expose its profound impact on both the theoretical quantum computational cost of the linear systems derived in Sec. III and the practical fidelity of the state estimation.

B. Avoidance of the quadratic subnormalization explosion

As established in Sec. III, the efficiency of QLSS based on block-encoding techniques is bounded by the subnormalization factor α of the encoded matrix $\mathbf{A}$. Guided by the mathematical requirement $\alpha \geq \|\mathbf{A}\|_2$ [14], we employ the spectral norm $\|\mathbf{A}\|_2$ as a metric to evaluate how the structural densification of the observation operator $\mathbf{H}$ inflates the quantum computational cost.

We construct an irregularly spaced sparse observation network with $p = 15$ within the Lorenz-96 domain. Specifically, the (i, j) entry of the observation operator $\mathbf{H}$ is given by a

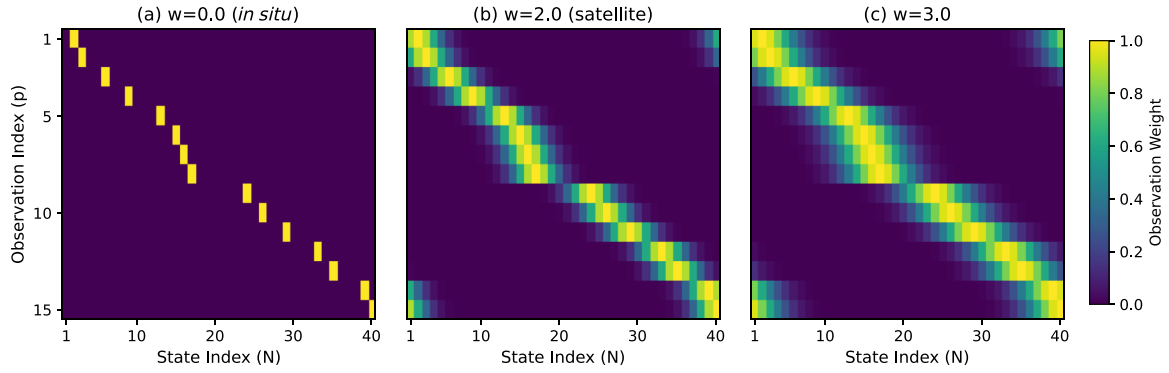


FIG. 1. Heatmaps illustrating the nonzero elements of the 15×40 observation operator $\mathbf{H}$ under varying footprint widths w within the Lorenz-96 domain. (a) $w = 0$, representing idealized *in situ* point measurements where $\mathbf{H}$ reduces to Kronecker deltas. (b) $w = 2.0$ and (c) $w = 3.0$, simulating overlapping remote-sensing footprints. The specific configurations shown in panels (a) and (b) serve as the baseline physical structures for the empirical DA cycle experiments in Sec. IV C.

Gaussian function:

$$H_{i,j} = \exp\left(-\frac{d(c_i, j)^2}{2w^2}\right), \quad (17)$$

where c_i is the state grid index corresponding to the center of the i th observation and $d(i, j)$ is the grid distance between i th and j th grid indexes considering the periodic boundary. Figure 1 shows three examples including $w = 0$ with which $H_{i,j}$ reduces to the Kronecker delta $\delta_{c_i,j}$ that mimics *in situ* sensors [Fig. 1(a)] and $w > 0$ that mimics sensors sensitive to multiple locations like satellite radiances except nonlinearity [Figs. 1(b) and 1(c)]. Equation (17) fixes the peak amplitude of each row at unity as shown in Fig. 1; a row-normalized variant, in which each row is divided by its sum to represent a weighted spatial average, is introduced in Sec. IV C.

We assume a diagonal observation error covariance matrix $\mathbf{R} = 0.25\mathbf{I}$, matching the DA experiments in Sec. IV C; the effect of spatially correlated, ill-conditioned $\mathbf{R}$ is examined in Sec. IV D. The background error covariance matrix $\mathbf{B}$ is assumed to represent untruncated long-range cross-correlations by a Gaussian correlation structure:

$$B_{i,j} = a \exp\left(-\frac{d(i, j)^2}{2l^2}\right), \quad (18)$$

where the spatial length scale is chosen to be $l = 2.0$. To consider physical consistency in asymptotic DA cycles where the background and observation errors are balanced, the background error variance is set to $a = 0.25$, matching the observation error magnitude.

Figure 2 illustrates the scaling of the subnormalization proxy $\|\mathbf{A}\|_2$ against the observation footprint width w . The result exposes a severe algorithmic bottleneck of the conventional observation-space formulation (red solid line). As the footprint widens and observations begin to spatially overlap ($w > w_c \sim 0.6$), the untruncated matrix multiplications inherent in $\mathbf{H}\mathbf{B}\mathbf{H}^T$ accumulate the cross-correlations, triggering a severe quadratic explosion in the quantum computational cost following the $\mathcal{O}(\alpha_B \alpha_H^2)$ scaling described in Sec. III. Therefore, directly mapping the classical observation-space formulation to quantum circuits becomes computationally prohibitive for modern DA systems that rely on complex, overlapping remote-sensing observations. In contrast, the pro-

posed combined formulation [Eq. (14)] avoids this quadratic explosion, maintaining a linear scaling $\mathcal{O}(\alpha_B \alpha_H + \alpha_R)$ as described in Sec. III (blue dashed line). It is insightful to observe the crossover in the highly localized regime ($0 \leq w < w_c$), where the combined formulation exhibits a slightly higher $\|\mathbf{A}\|_2$ than the observation-space formulation. This reflects the fixed base overhead required by the combined formulation to explicitly hold the full, untruncated state-space covariance structure $\mathbf{B}$. In this regime, the observation-space formulation compresses the system into a small observation subspace, yielding a smaller spectral norm. However, this compression breaks down as the observation operator becomes denser. Investing in the initial base overhead of the combined formulation is a prerequisite for gaining computational benefits with complex observations.

One caveat concerns the absolute scale of $\|\mathbf{A}\|_2$. The DA analysis is invariant, also known as a gauge freedom, under a diagonal rescaling of the observation equation, $(\mathbf{H}, \mathbf{R}, \mathbf{d}) \rightarrow (\mathbf{D}\mathbf{H}, \mathbf{D}\mathbf{R}\mathbf{D}, \mathbf{D}\mathbf{d})$ for any diagonal $\mathbf{D}$ with nonzero diagonal entries, which amounts to a change of the physical units of each observation; the row normalization employed in the dense regime of Sec. IV C is one such choice. The subnormalization factor, in contrast, is not invariant: Choosing $\mathbf{D}$ to normalize the rows of $\mathbf{H}$ suppresses the growth of $\|\mathbf{A}\|_2$ in Fig. 2. This, however, is not a remedy but a reallocation of the difficulty. The same rescaling inflates $\|\mathbf{R}^{-1}\|_2$ and thereby the condition number of the rescaled system by the corresponding factor; in our homogeneous network, where $\mathbf{D}$ reduces to a global scalar, we verified numerically that the product $\alpha\kappa$ governing the QLSS complexity is not improved but rather increased (Table I). The unit-independent content of the problem is carried by the error-normalized observation operator $\mathbf{R}^{-1/2}\mathbf{H}$, whose norm grows as the observations have more overlaps. Moreover, under any choice of units, constructing $\mathbf{H}\mathbf{B}\mathbf{H}^T$ composes the subnormalization factors as $\mathcal{O}(\alpha_B \alpha_H^2)$, whereas the combined formulation requires only $\mathcal{O}(\alpha_B \alpha_H)$, and the postselection penalty of the observation-space route persists in every unit system. Figure 2 is drawn in the physically natural units in which $\mathbf{H}$ carries the instrument response with unit peak amplitude and $\mathbf{R}$ the instrument error variance; for AI-generated oracles, the choice of units is fixed at oracle-construction time, since composing a rescaling

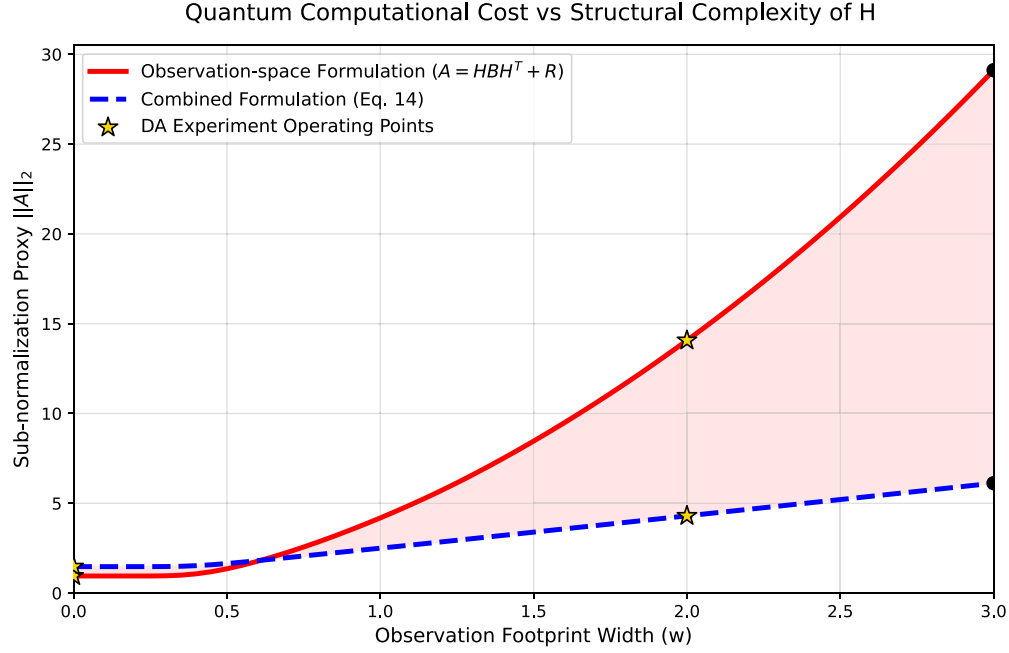


FIG. 2. Subnormalization proxy $\|A\|_2$ as a function of the observation footprint width w . Stars denote the specific operating points [$w = 0.0$ and 2.0 , corresponding to Figs. 1(a) and 1(b)] evaluated in the subsequent DA experiments. Closed circles denote $w = 3.0$, corresponding to Fig. 1(c).

onto an existing block encoding multiplies and increases its subnormalization factor. Optimizing this choice is a design consideration for the oracle-generating surrogate models. For heterogeneous observing networks, where the natural units of different instruments differ by orders of magnitude, unit rescaling of H resurfaces as ill-conditioning of DRD ; this is the regime addressed in the following subsection.

To seamlessly bridge this theoretical cost analysis with practical state-estimation performance, we highlighted specific operating points ($w = 0.0$ and 2.0) with stars in Fig. 2. These configurations correspond to the small α and large α regimes, respectively, which are utilized in the subsequent empirical DA cycle experiments in Sec. IV C.

C. Perfect fidelity of the combined formulation in continuous DA cycles

Following the experimental setup detailed in Sec. IV A, we perform continuous DA cycles on the Lorenz-96 model. The primary objective is to numerically verify that the proposed

combined formulation perfectly replicates the analytical solution of the uncompromised DA. Since their mathematical equivalence was established in Sec. III, these cycle experiments serve as an empirical sanity check of the quantum-ready framework's fidelity. Furthermore, by contrasting this uncompromised solution against classical baselines, a static three-dimensional (3D)-Var and an EnKF with localization, we illustrate the specific observational conditions under which executing this uncompromised DA provides a substantial advantage.

We perform the DA experiments under two distinct observational networks. One is a sparse *in situ* regime ($w = 0.0$) with an irregularly spaced network of $p = 15$ point observations corresponding to H shown in Fig. 1(a). The other is a dense satellite regime [$w = 2.0$, Fig. 1(b)]. Since data are sensitive to multiple locations with blurred sampling, we expand the number to be $p = 40$ with overlapping observations covering all state variables. To maintain the physical magnitude of the observational updates, the rows of H are normalized to sum to 1.0, representing a weighted spatial average.

We compare four DA methods as summarized in Table II. Throughout the experiments in this subsection, the observation error covariance is fixed to a diagonal matrix $R = 0.25I$; spatially correlated R is examined in Sec. IV D. For the classical baseline, both the static 3D-Var and the EnKF compute the analysis by solving the observation-space update [Eqs. (7) and (8)]. The 3D-Var utilizes a simple, static diagonal background error covariance ($B = 0.25I$). In contrast, the EnKF explicitly constructs B from the ensemble with $m = 10$ and artificially restricts it using a physical distance-based localization matrix ρ [Eqs. (10) and (13)]. The elements of ρ are defined by a Gaussian function $\rho_{i,j} = \exp(-\frac{d(i,j)^2}{2L^2})$, with the length scale L manually optimized along with another tuning parameter known as multiplicative covariance inflation (see

TABLE I. Subnormalization proxy α , condition number κ , and their product $\alpha\kappa$ for the combined formulation of Eq. (14) in the sparse observation network, with the observation operator of Eq. (17) as shown in Fig. 1 and with its row-normalized counterpart.

w	Equation (17)			Row normalized		
	α	κ	$\alpha\kappa$	α	κ	$\alpha\kappa$
0.5	1.65	8.65	14.2	1.46	11.3	16.5
1.0	2.50	10.9	27.1	1.44	36.1	51.9
2.0	4.30	17.2	74.0	1.40	141	197
3.0	6.12	24.5	150	1.38	312	431

TABLE II. Summary of the four DA experiments.

DA methods	DA equations	B	Plot style
3D-Var	Equations (7) and (8)	$0.25I$	Blue dashed line
Classical EnKF	Equations (7) and (8)	Equations (10) and (13) with $m = 10$, $\rho_{i,j} = \exp(-d(i,j)^2/(2L^2))$, $L = 3.0$, and inflation factor of 1.05 (sparse) or 1.15 (dense) (cf. Appendix A)	Green solid line
Uncompromised DA	Equations (7) and (8)	Equation (10) with $m = 400$ and inflation factor of 1.02	Red solid line
Quantum combined	Equation (14)	P_e from uncompromised DA	Black circles

Appendix A for details). While operational large-scale EnKFs typically bypass the explicit $N \times N$ construction of Eq. (13), we explicitly compute the full Schur product here since the system dimension $N = 40$ is computationally trivial. The uncompromised DA employs a large ensemble of $m = 400$ with a mild multiplicative inflation factor of 1.02 and explicitly constructs the untruncated ensemble covariance [Eq. (10)]. Our quantum-ready combined formulation [Eq. (14)] directly ingests this untruncated covariance matrix. The ensemble members in these experiments are updated with the stochastic perturbed-observation method [26].

To establish the nature run or the true state trajectory, the system is integrated from random initial conditions for a sufficiently long spin-up period (1000 integration steps). To ensure a fair evaluation, the DA experiments must be initialized completely independently of the nature run. Therefore, the initial ensemble states for both the EnKF and uncompromised DA are generated from completely independent random states and integrated for a 1000-step spin-up period. This allows the initial ensembles to represent climatological background error structures on the model attractor. The static 3D-Var is then initialized using one of the members from the spun-up EnKF ensemble. The continuous DA cycles are performed over a one-year period, i.e., 1460 cycles with a DA interval of $\Delta t = 0.05$.

Figure 3(a) illustrates the DA accuracy in the sparse *in situ* regime. Most importantly, the states estimated by our combined formulation (black circles) perfectly align with the uncompromised DA (red line), confirming the absolute algebraic fidelity of the quantum-ready framework. A critical distinction between the methods is immediately apparent in their initial spin-up dynamics from the climatological cold start. The classical EnKF (green line) requires a prolonged transient period to fully capture the true trajectory and stabilize. This behavior is likely due to the rigorously tuned optimal localization and inflation parameters, which push the boundaries of filter stability to achieve the best possible asymptotic accuracy at the expense of rapid initial convergence. In contrast, the uncompromised DA exhibits rapid convergence that demonstrates the informational efficiency of utilizing the genuine, long-range dynamic correlations, allowing sparse observations to effectively constrain the entire state space. As the cycles progress, the artificial distance-based localization applied to the EnKF is physically well justified by the sparse network, allowing it to eventually robustly track the true state and outperform the static 3D-Var (blue dashed line). However, uncompromised DA consistently yields superior accuracy throughout.

The physical mechanism behind this superior accuracy is supported by the estimated background error covariance structures, captured as a snapshot at the final step of the one-year experiment [Figs. 3(b)–3(d)]. As shown in Fig. 3(c), the classical EnKF with localization exhibits a physically reasonable tendency where the error variance naturally increases in regions lacking direct observations as indicated by the gaps between the green observation markers. However, the uncompromised DA [Fig. 3(d)] successfully retains all off-diagonal components without artificial truncation. The preservation and exploitation of these genuine long-range dynamical correlations directly contribute to its optimal DA accuracy. This empirical advantage of utilizing uncompromised covariances aligns perfectly with previous findings that demonstrated drastic accuracy improvements by removing covariance localization in the SPEEDY atmospheric general circulation model with 10 240 ensembles [10].

The structural limitations of classical DA are clearly exposed in the dense satellite regime (Fig. 4). First, the performance of the static 3D-Var degrades noticeably compared to the sparse *in situ* regime, demonstrating the inherent difficulty of utilizing highly overlapping, dense observations without flow-dependent correlations. For the classical EnKF, in the sparse regime, the localization scale $L = 3.0$ was effectively tuned to suppress the spurious long-range correlations inherent in the small ensemble. However, in the dense regime, the $p = 40$ observation footprints broadly overlap and structurally conflict with the localization. This creates an unavoidable structural dilemma: Maintaining the relatively tight localization artificially fragments the cohesive, nonlocal observational information, whereas expanding L to accommodate the broad footprints exposes the filter to severe sampling errors. Constrained by this fundamental trade-off, the classical EnKF with localization suffers a severe degradation in its asymptotic tracking capability. While its initial spin-up phase is not unreasonable, it ultimately fails to converge to the highly accurate state achieved in the sparse regime. Nevertheless, thanks to its flow-dependent variance, the EnKF still bounds the errors at a level much lower than the static 3D-Var [Fig. 4(a)]. However, its accuracy remains significantly degraded compared to the uncompromised DA trajectory. The estimated covariance matrix [Fig. 4(c)] becomes structurally distorted; the artificial spatial tapering forces an unnatural fragmentation of the cohesive cross-variable correlations. In contrast, uncompromised DA digests the densely overlapping satellite information. Without localization approximations, it captures the intricate cross-variable correlations [Fig. 4(d)] and maintains a low and stable RMSE. Once again, the

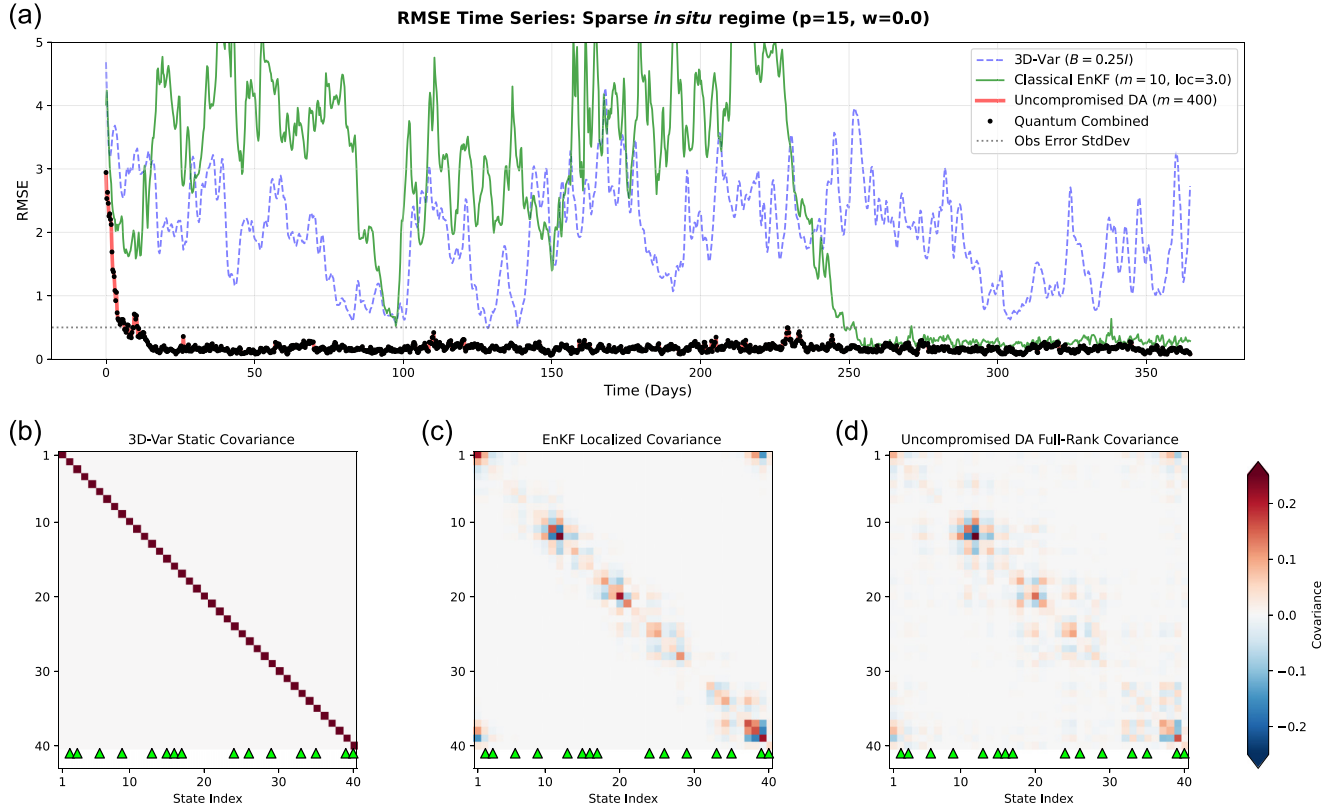


FIG. 3. (a) Root Mean Square Error (RMSE) time series over a one-year DA period (1460 cycles) in the sparse *in situ* regime ($p = 15$, $w = 0.0$) for four DA methods as shown in legend. The proposed quantum combined formulation (black circles) perfectly reproduces the uncompromised DA solution (red solid line). The gray dotted line indicates the observation error standard deviation. Snapshots of the background error covariance matrices at the final assimilation step for (b) 3D-Var, (c) classical EnKF, and (d) uncompromised DA. Green triangles indicate the spatial locations of the 15 point observations.

proposed combined formulation perfectly reproduces this uncompromised DA.

D. Robustness under ill-conditioned observation error covariances

The experiments above employed a diagonal, well-conditioned $\mathbf{R}$, leaving the conditioning claim of Sec. III, the resolution of Barrier 5, tested only in a benign regime. Here we construct ill-conditioned observation error covariances and numerically verify that the condition number of the combined system is governed by the static conditioning of $\mathbf{R}$, decoupled from the severe ill-conditioning of the untruncated $\mathbf{B}$, as formalized by the analytical bound of Appendix B.

Spatially correlated observation errors, characteristic of dense remote-sensing data, are introduced through a Gaussian correlation structure:

$$R_{i,j} = \sigma_R^2 \exp\left(-\frac{d(c_i, c_j)^2}{2l_R^2}\right), \quad (19)$$

with $\sigma_R^2 = 0.25$, where l_R is the observation error correlation length scale and c_i is the state grid index of the i th observation center as in Eq. (17). Sweeping l_R drives $\kappa(\mathbf{R})$ from unity, corresponding to mutually independent errors, to $\sim 10^{13}$. The sweep is restricted to $l_R \leq 2.5$, within which $\mathbf{R}$ remains positive definite: On the discrete periodic grid, the minimum

eigenvalue of the Gaussian correlation matrix becomes negative near $l^* = \sqrt{N/(2\pi)} \approx 2.52$ for $N = 40$, beyond which $\mathbf{R}$ ceases to be a valid covariance matrix. This intrinsic fragility of correlation structures on discrete grids is a known motivation for covariance reconditioning in DA [30,31], to which we return below. This is a static matrix analysis in the same spirit as Fig. 2: With the untruncated $\mathbf{B}$ of Eq. (18) and the two observational networks of Fig. 1, including the row normalization of the dense regime, we compute for each l_R the condition numbers of the coefficient matrices of the three formulations, $\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ of Eq. (6), $\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R}$ of Eq. (8), and the combined block matrix of Eq. (14).

Figures 5(a) and 5(b) present the condition numbers of the three formulations. The state-space formulation (purple) stays around $\kappa(\mathbf{B}) \approx 1.9 \times 10^8$ regardless of l_R , confirming that it is dominated by the ill-conditioned prior. In contrast, the combined formulation (blue dashed) tracks $\kappa(\mathbf{R})$ (gray dotted) throughout: For a diagonal $\mathbf{R}$, it achieves $\kappa \approx 8$, and as the observation error correlations strengthen it follows $\kappa(\mathbf{R})$ up to $\sim 1.6 \times 10^{13}$ in the dense regime, always within the analytical bound of Appendix B (thin black dashed), which holds uniformly within about one order of magnitude. The observation-space formulation (red solid) exhibits a comparable condition number to the combined formulation; the conditioning advantage of the combined formulation is specifically over the state-space formulation,

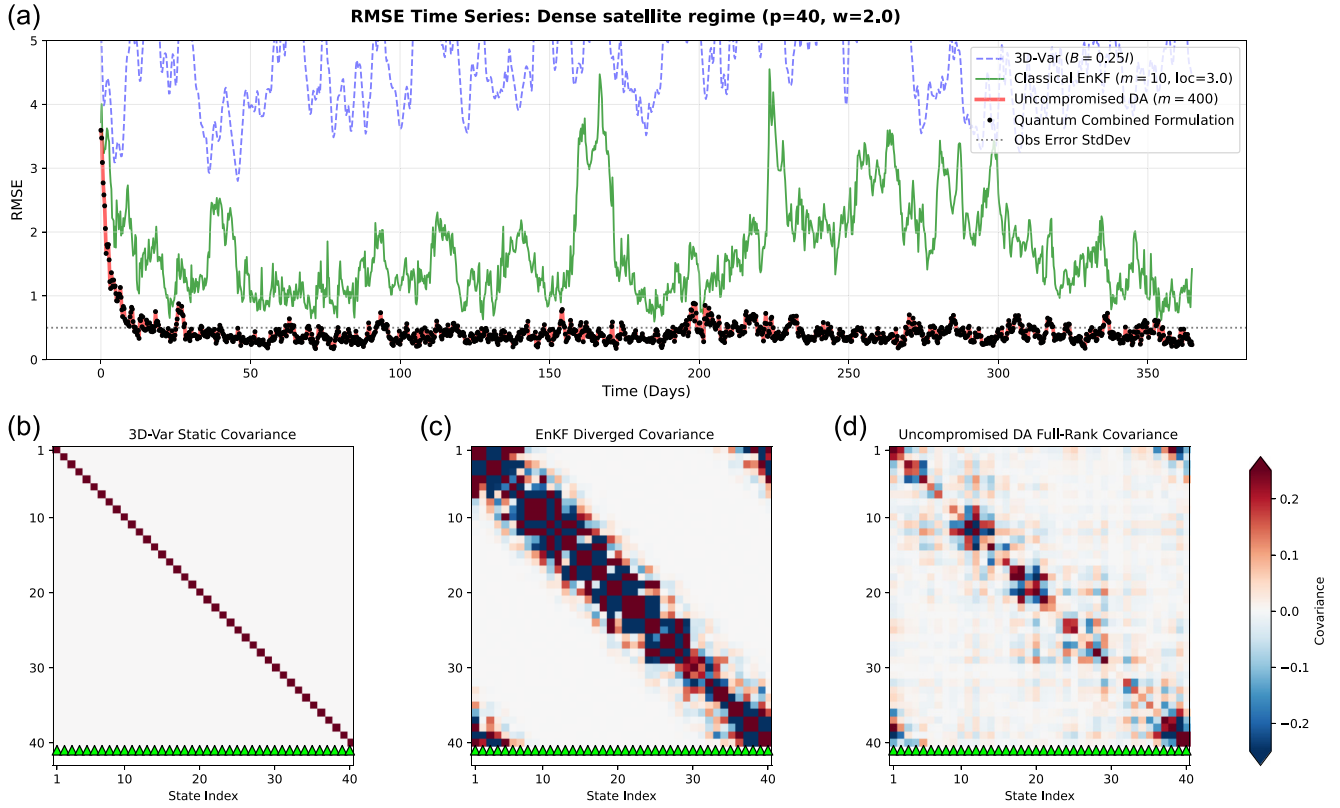


FIG. 4. Similar to Fig. 3, but for the dense satellite regime. Unlike Fig. 3, gray dotted line indicates the observation error standard deviation purely as a reference baseline. Because the broad observation footprints involve spatial averaging, the observation error does not directly translate into the state-estimation error in the model space due to the nature of the inverse problem.

while its advantages over the observation-space formulation are the linear subnormalization scaling and the avoidance of postselection established in Secs. III and IV B.

The residual dependence on $\kappa(\mathbf{R})$ is not an intrinsic limitation. Figure 5(c) demonstrates diagonal reconditioning [30,31], $\mathbf{R} \rightarrow \mathbf{R} + \delta \mathbf{I}$, at the worst valid operating point

(dense regime, $l_R = 2.5$). Since $\lambda_{\min}(\mathbf{R} + \delta \mathbf{I}) \geq \delta$, the reconditioning caps $\|(\mathbf{R} + \delta \mathbf{I})^{-1}\|_2 \leq 1/\delta$, rendering the bound of Appendix B computable *a priori*. The condition number of the combined system (blue dashed, left axis) collapses in proportion to $1/\delta$, from 1.6×10^{13} to 8×10^2 at $\delta = 10^{-2} \sigma_R^2$. To quantify the accuracy cost, we repeat the con-

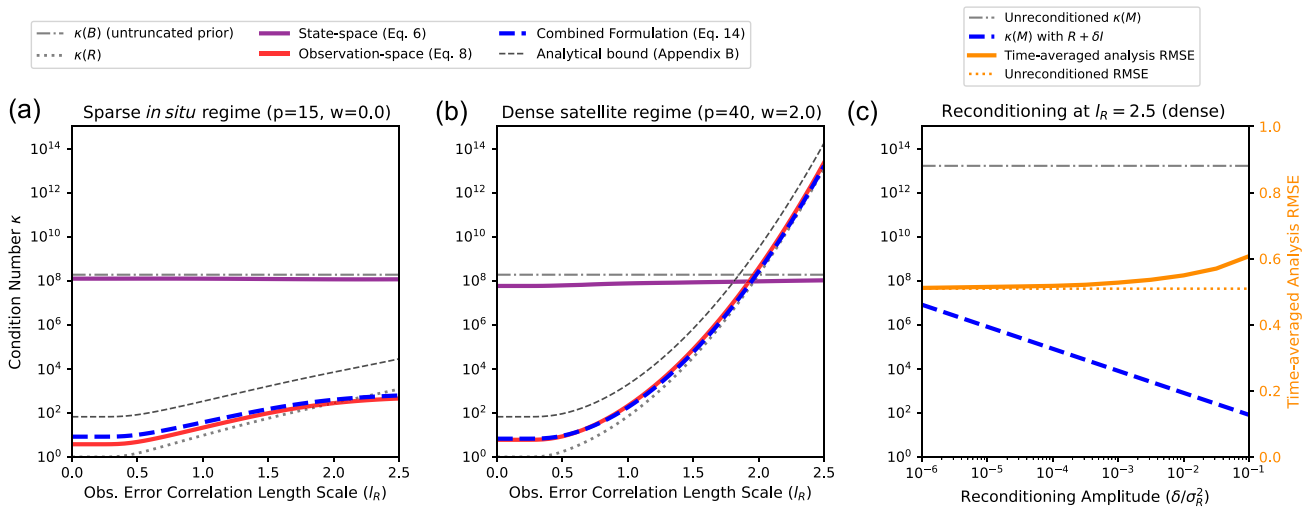


FIG. 5. Condition numbers of the three formulations as functions of the observation error correlation length scale l_R for (a) the sparse *in situ* regime ($p = 15$, $w = 0.0$) and (b) the dense satellite regime ($p = 40$, $w = 2.0$), with line styles as in the legend. (c) Diagonal reconditioning $\mathbf{R} \rightarrow \mathbf{R} + \delta \mathbf{I}$ in the dense regime at the worst valid operating point $l_R = 2.5$; the condition number of the combined system is shown on the left axis and the time-averaged analysis RMSE of the corresponding DA cycles on the right axis.

tinuous DA cycles of Sec. IV C in the dense regime with the correlated $\mathbf{R}$ at $l_R = 2.5$, applying $\mathbf{R} + \delta\mathbf{I}$ in the analysis while the observation errors are drawn from the correlated $\mathbf{R}$ of Eq. (19). The time-averaged analysis RMSE (orange, right axis) increases by less than 2% for $\delta \leq 10^{-4}\sigma_R^2$ and by about 8% at $\delta = 10^{-2}\sigma_R^2$, consistent with the small accuracy impact of reconditioning reported in realistic settings [30]. In the quantum implementation, the oracle for $\mathbf{R} + \delta\mathbf{I}$ adds only a trivial identity term via LCU at negligible cost. The parameter δ provides an explicit, user-controlled trade-off between a quantifiably small analysis perturbation and an arbitrarily large reduction in quantum circuit depth, completing the reduction of the conditioning challenge from the intractable, dynamically evolving $\mathbf{B}$ to an established static practice on $\mathbf{R}$.

V. CONCLUSION AND DISCUSSION

Classical DA methods, encompassing both ensemble and variational approaches, are fundamentally constrained by computational bottlenecks. Because the uncompromised DA with an untruncated, dynamically evolving error covariance matrix demands an $\mathcal{O}(N^2)$ memory footprint and $\mathcal{O}(N^3)$ matrix inversion scaling, decades of DA research have been forced into a continuous struggle to squeeze marginal accuracy improvements from fundamentally approximated systems. Historically, the theoretical and applied evolution of large-scale DA has been dictated by the necessity to evade this massive matrix bottleneck, shackling the pursuit of uncompromised state estimation.

To overcome this limitation, we introduced the quantum-ready combined formulation. While mathematically a straightforward rearrangement of DA equations into an augmented system, this structural shift is the crucial bridge connecting DA to quantum computation. We demonstrated that standard observation-space equations trigger a catastrophic quadratic explosion in the quantum subnormalization factor under dense nonlocal observation operators. The combined formulation not only avoids this barrier but also mathematically circumvents the extreme ill-conditioning that plagues the state-space formulation. Empirically, it perfectly reproduced the uncompromised DA solution with enhanced predictability.

A fundamental conclusion of this study is that the synergy between QLSS and AI-driven oracles offers a theoretical pathway to this uncompromised DA. While the explicit generation of high-rank error covariances remains classically prohibitive, it aligns with the trajectory of quantum machine learning. Emerging technologies, such as parametrized quantum circuits [32] and quantum generative adversarial networks [33], offer realistic pathways to implicitly learn and directly generate these massive, flow-dependent error structures as quantum states. Our combined formulation provides a mathematical architecture for these future AI surrogate models to optimize toward. While efficient quantum state preparation and the construction of these oracles remain significant technical bottlenecks in quantum computing, our framework successfully isolates this challenge by resolving the downstream mathematical barriers of matrix inversion and nonlocal observation. Consequently, developing scalable, AI-driven state preparation techniques emerges as the next grand challenge for the interdisciplinary community, i.e., the critical step required

to actualize this framework and flawlessly fuse AI-generated priors with massive observation networks.

Independently of how the oracles are realized, the structural analysis of Sec. III delineates the algebra that any quantum implementation of DA must follow. In parallel, an explicit construction of the block encoding of Eq. (14) and its end-to-end execution on quantum simulators at reduced dimensions is a natural step toward hardware validation; the present $N = 40$ system, requiring only seven data qubits, makes such a demonstration a concrete target for follow-up work.

It is crucial to acknowledge the readout bottleneck inherent in current quantum computing paradigms. Extracting the full state vector $\delta\mathbf{x}$ from the generated quantum state requires quantum state tomography, which incurs an $\mathcal{O}(N)$ measurement cost. However, replacing the classical $\mathcal{O}(N^3)$ matrix inversion bottleneck with an $\mathcal{O}(N)$ tomography cost still guarantees a massive polynomial speedup of $\mathcal{O}(N^2)$. For state-of-the-art DA applications ranging from global Earth system models to large-scale digital twins that routinely assimilate massive observations into systems exceeding $N \sim \mathcal{O}(10^8)$ degrees of freedom, overcoming this cubic scaling barrier represents a transformative computational breakthrough.

Ultimately, this study serves a dual purpose: to mathematically prepare the broad DA community for the FTQC era and to introduce DA to the quantum computational science community as a promising application. By structurally surviving the densification of observations and enabling the exploitation of true high-rank covariances, this interdisciplinary framework removes the historical approximations constraining classical DA. It transforms uncompromised DA from a theoretical ideal into a concrete algorithmic target, paving the way toward redefining the ultimate limits of predictability for chaotic and complex systems.

ACKNOWLEDGMENTS

The author gratefully acknowledges Erik Lötstedt, Tetso Hatsuda, and Satoshi Iso for stimulating discussions at RIKEN, which inspired the foundational ideas of this work. This work was supported by JST SATREPS (JPMJSA2109), JST CREST (JPMJCR24Q3), JSPS KAKENHI (JP24H00021), Japan Aerospace Exploration Agency (EORA4), the COE research grant in computational science from Hyogo Prefecture and Kobe City through Foundation for Computational Science, the UK Advanced Research + Invention Agency (ARIA) under Project No. FPCW-PR01-P007, and RIKEN TRIP initiative (RIKEN Prediction Science).

DATA AVAILABILITY

The Python code that reproduces all figures and experiments in this paper is openly available on GitHub and permanently archived on Zenodo [34].

APPENDIX A: PARAMETER OPTIMIZATION OF CLASSICAL EnKF

For fair comparison of the DA accuracy in Sec. IV C, we performed a comprehensive two-dimensional grid search to

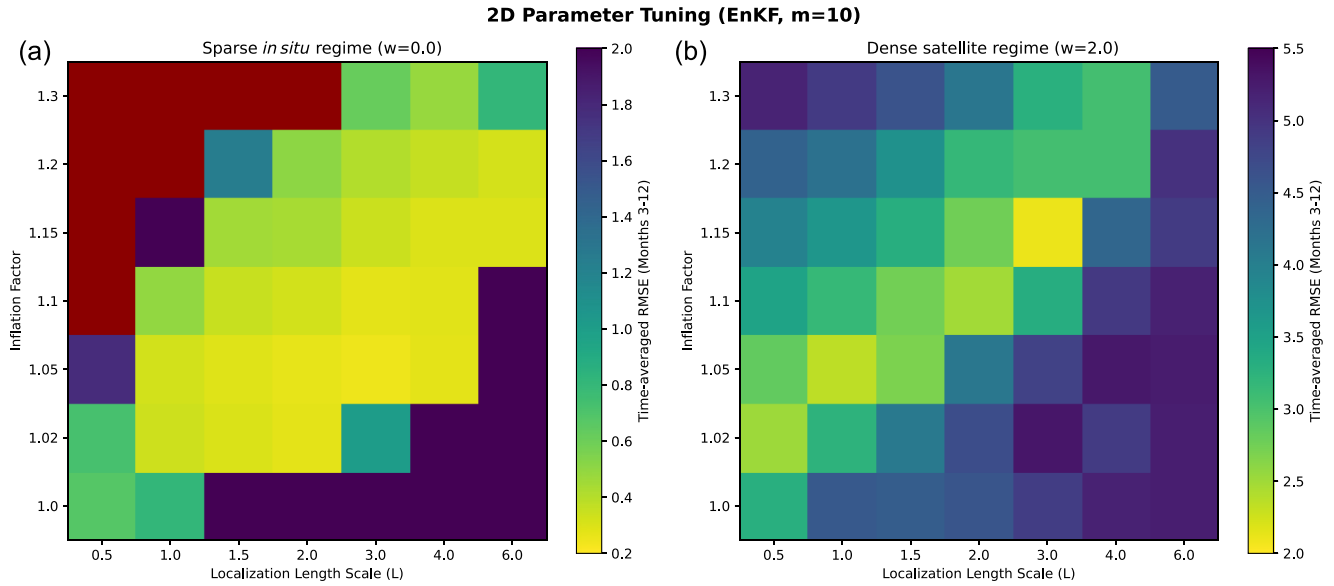


FIG. 6. Heatmaps of the time-mean analysis RMSE of the classical EnKF with $m = 10$ as a function of the localization length scale L (horizontal axis) and the multiplicative covariance inflation factor (vertical axis) for (a) sparse *in situ* regime and (b) dense satellite regime. Color scales are adjusted per panel, with dark red indicating unstable filter divergence.

optimize the classical EnKF. For small ensemble sizes, stability and accuracy of DA heavily depend on two empirical parameters: the localization length scale L and the multiplicative covariance inflation factor.

While spatial localization effectively tapers spurious long-range correlations, finite ensemble sizes inherently underestimate the background error variance. To compensate for this variance loss and prevent filter divergence, multiplicative inflation is applied to the prior ensemble perturbations prior to the analysis step, such that the ensemble perturbations are multiplied by a factor larger than 1.

To identify the optimal pair of these two parameters, we evaluated a parameter grid ($L \in \{0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 6.0\}$ and inflation factor $\in \{1.00, 1.02, 1.05, 1.10, 1.15, 1.20, 1.30\}$) under both the sparse and dense observation regimes. For each parameter combination, the DA cycles were run for a total of one year (1460 cycles). To reduce transient artifacts from the initial cold start, the time-averaged RMSE was calculated over the final ten months.

Figure 6 displays the resulting RMSE heatmaps. For the sparse *in situ* regime, a clear basin of optimal performance emerges, identifying $L = 3.0$ and an inflation factor of 1.05 as the optimal parameters [Fig. 6(a)]. For the dense satellite regime, $L = 3.0$ and an inflation factor of 1.15 yield the most accurate DA analysis [Fig. 6(b)]. These optimized parameters were utilized for the classical baseline in Figs. 3 and 4.

APPENDIX B: CONDITION-NUMBER BOUND FOR THE COMBINED FORMULATION

We derive an upper bound on the condition number of the combined block matrix in Eq. (14),

$$\mathbf{M} = \begin{pmatrix} \mathbf{I} & -\mathbf{B}\mathbf{H}^T \\ \mathbf{H} & \mathbf{R} \end{pmatrix}, \quad (\text{B1})$$

demonstrating that $\kappa(\mathbf{M}) = \|\mathbf{M}\|_2 \|\mathbf{M}^{-1}\|_2$ is controlled by the operator norms $\alpha_B = \|\mathbf{B}\|_2$, $\alpha_H = \|\mathbf{H}\|_2$, $\alpha_R = \|\mathbf{R}\|_2$, and $\|\mathbf{R}^{-1}\|_2$, with no dependence on the condition number of $\mathbf{B}$. Let $\mathbf{B}$ be symmetric positive semidefinite and $\mathbf{R}$ symmetric positive definite. We show that

$$\kappa(\mathbf{M}) \leq (1 + \alpha_H + \alpha_B \alpha_H + \alpha_R) \times [1 + (1 + \alpha_H)(1 + \alpha_B \alpha_H) \|\mathbf{R}^{-1}\|_2]. \quad (\text{B2})$$

For any block matrix, the spectral norm is bounded by the sum of the spectral norms of its blocks, since each block can be isolated by a triangle inequality over matrices that retain a single block and vanish elsewhere. Applying this to $\mathbf{M}$,

$$\begin{aligned} \|\mathbf{M}\|_2 &\leq \|\mathbf{I}\|_2 + \|\mathbf{B}\mathbf{H}^T\|_2 + \|\mathbf{H}\|_2 + \|\mathbf{R}\|_2 \\ &\leq 1 + \alpha_B \alpha_H + \alpha_H + \alpha_R. \end{aligned} \quad (\text{B3})$$

For the inverse, the analytical block form of Eq. (15) applies with the Schur complement $\mathbf{S} = \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T$. Because $\mathbf{H}\mathbf{B}\mathbf{H}^T$ is positive semidefinite, $\lambda_{\min}(\mathbf{S}) \geq \lambda_{\min}(\mathbf{R}) > 0$; hence, $\|\mathbf{S}^{-1}\|_2 \leq \|\mathbf{R}^{-1}\|_2$. Bounding each block of Eq. (15) by submultiplicativity,

$$\begin{aligned} \|\mathbf{I} - \mathbf{B}\mathbf{H}^T \mathbf{S}^{-1} \mathbf{H}\|_2 &\leq 1 + \alpha_B \alpha_H^2 \|\mathbf{R}^{-1}\|_2, \\ \|\mathbf{B}\mathbf{H}^T \mathbf{S}^{-1}\|_2 &\leq \alpha_B \alpha_H \|\mathbf{R}^{-1}\|_2, \\ \|\mathbf{S}^{-1} \mathbf{H}\|_2 &\leq \alpha_H \|\mathbf{R}^{-1}\|_2, \\ \|\mathbf{S}^{-1}\|_2 &\leq \|\mathbf{R}^{-1}\|_2, \end{aligned}$$

and summing,

$$\begin{aligned} \|\mathbf{M}^{-1}\|_2 &\leq 1 + (\alpha_B \alpha_H^2 + \alpha_B \alpha_H + \alpha_H + 1) \|\mathbf{R}^{-1}\|_2 \\ &= 1 + (1 + \alpha_H)(1 + \alpha_B \alpha_H) \|\mathbf{R}^{-1}\|_2. \end{aligned} \quad (\text{B4})$$

Multiplying Eqs. (B3) and (B4) yields Eq. (B2).

Three remarks are in order. First, the bound involves $\mathbf{B}$ only through its norm α_B , a physically bounded error variance

scale; the near-zero eigenvalues of $\mathbf{B}$, which drive $\kappa(\mathbf{B})$ to extreme values, do not enter. Second, since $\kappa(\mathbf{R}) = \alpha_R \|\mathbf{R}^{-1}\|_2$, the bound scales linearly with $\kappa(\mathbf{R})$ at fixed operator norms, formalizing the statement of Sec. III that the conditioning of the combined system is confined to that of the observation error covariance. Third, under diagonal reconditioning

$\mathbf{R} \rightarrow \mathbf{R} + \delta \mathbf{I}$, the substitution $\|\mathbf{R}^{-1}\|_2 \rightarrow \|(\mathbf{R} + \delta \mathbf{I})^{-1}\|_2 \leq 1/\delta$ makes the bound computable *a priori* from the user-chosen δ and the oracle norms alone. Section IV D verifies numerically that the bound holds uniformly, within about one order of magnitude of the measured condition numbers, across thirteen orders of magnitude of $\kappa(\mathbf{R})$.

-
- [1] Y. Sasaki, Some basic formalisms in numerical variational analysis, *Mon. Weather Rev.* **98**, 875 (1970).
 - [2] P. Bauer, A. Thorpe, and G. Brunet, The quiet revolution of numerical weather prediction, *Nature (London)* **525**, 47 (2015).
 - [3] A. C. Lorenc, Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.* **112**, 1177 (1986).
 - [4] P. Bauer *et al.*, The digital revolution of Earth-system science, *Nat. Comput. Sci.* **1**, 104 (2021).
 - [5] G. Evensen, Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.* **99**, 10143 (1994).
 - [6] P. L. Houtekamer and F. Zhang, Review of the ensemble Kalman filter for atmospheric data assimilation, *Mon. Weather Rev.* **144**, 4489 (2016).
 - [7] D. J. Patil *et al.*, Local low dimensionality of atmospheric dynamics, *Phys. Rev. Lett.* **86**, 5878 (2001).
 - [8] T. Miyoshi, K. Kondo, and T. Imamura, The 10 240-member ensemble Kalman filtering with an intermediate AGCM, *Geophys. Res. Lett.* **41**, 5264 (2014).
 - [9] T. Miyoshi, K. Kondo, and K. Terasaki, Big ensemble data assimilation in numerical weather prediction, *Computer* **48**, 15 (2015).
 - [10] K. Kondo and T. Miyoshi, Impact of removing covariance localization in an ensemble Kalman filter: Experiments with 10 240 members using an intermediate AGCM, *Mon. Weather Rev.* **144**, 4849 (2016).
 - [11] A. Chattopadhyay, E. Nabizadeh, E. Bach, and P. Hassanzadeh, Deep learning-enhanced ensemble-based data assimilation for high-dimensional nonlinear dynamical systems, *J. Comput. Phys.* **477**, 111918 (2023).
 - [12] M. A. Sacco, J. J. Ruiz, M. Pulido, and P. Tandeo, Evaluation of machine learning techniques for forecast uncertainty quantification, *Q. J. R. Meteorol. Soc.* **148**, 3470 (2022).
 - [13] S. Kotsuki, F. Kawasaki, and M. Ohashi, Quantum data assimilation: A new approach to solving data assimilation on quantum annealers, *Nonlinear Processes Geophys.* **31**, 237 (2024).
 - [14] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum algorithm for linear systems of equations, *Phys. Rev. Lett.* **103**, 150502 (2009).
 - [15] J.-P. Liu, H. Ø. Kolden, H. K. Krovi, N. F. Loureiro, K. Trivisa, and A. M. Childs, Efficient quantum algorithm for dissipative nonlinear differential equations, *Proc. Natl. Acad. Sci. USA* **118**, e2026805118 (2021).
 - [16] J.-P. Liu, D. An, D. Fang, J. Wang, G. H. Low, and S. Jordan, Efficient quantum algorithm for nonlinear reaction-diffusion equations and energy estimation, *Commun. Math. Phys.* **404**, 963 (2023).
 - [17] H. Shi, G. Zhang, and M. Zhang, A quantum algorithm for the Kalman filter using block encoding, *arXiv:2404.04554*.
 - [18] N. Binandeh Dehaghani, R. Wisniewski, and A. P. Aguiar, Quantum solution framework for finite-horizon LQG control via block encodings and QSVT, in *Proceedings of 2025 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Piscataway, NJ, 2025), p. 282.
 - [19] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (Association for Computing Machinery, New York, NY, 2019), p. 193.
 - [20] E. N. Lorenz, Predictability: A problem partly solved, *Proc. Seminar on Predictability 1* (ECMWF, Reading, UK, 1996).
 - [21] E. N. Lorenz and K. A. Emanuel, Optimal sites for supplementary weather observations: Simulation with a small model, *J. Atmos. Sci.* **55**, 399 (1998).
 - [22] P. Courtier, Dual formulation of four-dimensional variational assimilation, *Q. J. R. Meteorol. Soc.* **123**, 2449 (1997).
 - [23] P. Courtier, J.-N. Thépaut, and A. Hollingsworth, A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.* **120**, 1367 (1994).
 - [24] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D* **230**, 112 (2007).
 - [25] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, Ensemble square root filters, *Mon. Weather Rev.* **131**, 1485 (2003).
 - [26] P. L. Houtekamer and H. L. Mitchell, A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Weather Rev.* **129**, 123 (2001).
 - [27] A. M. Childs, R. Kothari, and R. D. Somma, Quantum algorithm for systems of linear equations with exponentially improved dependence on precision, *SIAM J. Comput.* **46**, 1920 (2017).
 - [28] M. Benzi, G. H. Golub, and J. Liesen, Numerical solution of saddle point problems, *Acta Numer.* **14**, 1 (2005).
 - [29] Å. Björck, *Numerical Methods for Least Squares Problems* (SIAM, Philadelphia, 1996).
 - [30] K. Terasaki and T. Miyoshi, Including the horizontal observation error correlation in the ensemble Kalman filter: Idealized experiments with NICAM-LETKF, *Mon. Weather Rev.* **152**, 277 (2024).
 - [31] P. P. Weston, W. Bell, and J. R. Eyre, Accounting for correlated error in the assimilation of high-resolution sounder data, *Q. J. R. Meteorol. Soc.* **140**, 2420 (2014).
 - [32] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, Parameterized quantum circuits as machine learning models, *Quantum Sci. Technol.* **4**, 043001 (2019).

- [33] C. Zoufal, A. Lucchi, and S. Woerner, Quantum generative adversarial networks for learning and loading random distributions, *npj Quantum Inf.* **5**, 103 (2019).
- [34] T. Miyoshi, takemasa-miyoshi/quantum-da-lorenz96: v1.0.0: Code for PRR publication, Zenodo, 2026, <https://doi.org/10.5281/zenodo.21972189>; development version available at <https://github.com/takemasa-miyoshi/quantum-da-lorenz96>.