A Trust Optimization Region of the Quantum Natural Gradient for the Large Damping Domain

Nawres A. Alwan[1], Suzan J. Obaiys[1(✉)], Nadia M. G. Al-Saidi[2], and Rayane Chadli[1]

[1] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia
suzan@um.edu.my

[2] Department of Applied Sciences, University of Technology, Baghdad 10066, Iraq

Abstract. Natural Gradient Descent is the steepest descent with respect to geometric information in parameter space. The Quantum Natural Gradient (QNG) optimization mode provides geometry-aware optimization for variational quantum algorithms (VQAs) by incorporating the curvature of quantum state space through the Quantum Geometric Tensor (QGT) or Fubini-Study metric. However, the standard QNG updates are intrinsically ill-posed because the Fubini–Study metric is generically singular under redundant parameterizations, leading to non-unique or unstable update steps. The work develops a regularized Riemannian quadratic model that yields a geometrically principled trust-region formulation of the damped QNG and proves that the mix of QGT with a Tikhonov regularization term results in the unique minimizer of a strictly convex local approximation of the loss landscape, thereby ensuring existence, uniqueness, and numerical stability for all damping levels $\lambda > 0$. Further, it is demonstrated that the proposed method guarantees descent in the large-damping regime, and analytically shows a continuous interpolation between QNG and Euclidean Gradient Descent (EGD), recovering EGD as $\lambda \to \infty$. However, to quantify trust in the local quantum geometry, the gain-ratio–based damping rule that is similar to the classical Trust-region methods were also introduced, providing a principled mechanism for adapting λ based on model fidelity. Benchmark analyses comparing recent QNG-based optimizers, including Hamiltonian-aware QNG, qBANG, and full metric QNG, illustrate that our regularized model achieves convergence speeds comparable to standard QNG, while providing superior stability, eliminating overshooting under near-singular metrics, finite-shot noise, and redundant parameterizations. In small-scale Variational Quantum Eigensolver (VQE) experiments (2–12 qubits). The proposed damped QNG reaches target energies in 1–2 iterations, avoids instability noticed in the standard QNG technique, and maintains robust performance under sampling noise. These results establish the regularized QNG trust region as a mathematically grounded, geometry-preserving, and practically robust optimizer for near-term quantum algorithms.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16758, pp. 284–300, 2027.
https://doi.org/10.1007/978-3-032-30497-1_18

Keywords: Quantum natural gradient · Quantum geometric tensor · Variational quantum algorithm · Natural gradient descent · Fubini-Study metric

1 Introduction

The Quantum Natural Gradient (QNG) procedures based on the Quantum Geometric Tensor (QGT) have been widely discussed as geometry-aware optimization techniques for Variational Quantum Algorithms (VQA) [1,2,6], where almost all algorithms introduce metric regularization as a numerical heuristic to mitigate ill-conditioning arising from redundant parameterizations or finite-shot noise. Natural Gradient Descent (NGD) is steepest descent with respect to geometric information in parameter space (Riemannian manifold). In quantum information processing, we have QNG, which utilizes the QGT or Fubini-Study metric to induce geometry-aware optimisation on variational quantum states QNG correctly captures state-space curvature. The main problem in optimization of QNG lies in the intrinsic ill-posedness of the standard update equation $g(\theta)\Delta\theta = -\eta\nabla L(\theta)$, which frequently fails to yield a valid solution due to Fubini-Study metric tensor $g(\theta)$ which is generically singular or rank-deficient.

Mathematically, this rank deficiency prevents the inversion of $g(\theta)$, meaning the linear system admits either infinitely many solutions or no solution at all, rather than a unique update step. Physically, this degeneracy is not simply a numerical error but a structural reality arising from the difference between the abstract parameter space and the actual manifold of physical quantum states. Since $g(\theta)$ measures the distance strictly in terms of quantum state distinguishability, any redundancy in the variational circuit, where changes in θ parameters resulting in no observable change to the physical state $|\psi(\theta)\rangle$ that corresponds to the null space of the metric. Consequently, along these "ghost" directions where the physical state remains invariant, the local curvature vanishes, causing the geometric information to collapse and the standard optimization update to break down.

In 2019, Stock et al. [2] employed the metrics that are pseudo-inverse of quantum geometric tensor and the diagonal approximation, which real need a practice regularization to ensure numerical stability. After that, many researchers tried to enhancement this algorithms, such as [8–10], which will be discussed in the related work section. Nawres et al. in [3–5], provided a comprehensive theoretical foundation on quantum mechanics and quantum computing that applied in several representations.

In contrast, the proposed work provides a geometrically principled and optimization-theoretically grounded formulation of QNG stabilization, as a necessary mathematical result, rather than being introduced merely as a suggestion for numerical stability. Specifically, the work demonstrates a regularized QNG update obtained by adding a Tikhonov damping [7] term to the QGT optimizer arises as the *unique minimizer of a locally defined Riemannian quadratic model* of the loss function on parameter space. Correspondingly, establishes the exis-

tence, uniqueness, and numerical stability of the update within a unified geometric framework, rather than as an ad hoc modification. Moreover, the outcomes provide an explicit interpretation of the damping parameter as a confidence measure in the local quantum geometry by introducing a gain-ratio-based regularization criterion inspired by the trust-region methods. This yields a rigorous descent guarantee in the large-damping regime and reveals a continuous interpolation between Quantum Natural Gradient and Euclidean Gradient Descent. To the best of our knowledge, this explicit quadratic-model and confidence-region interpretation of damped QNG has not been previously formalized in the quantum optimization literature.

2 Related Works

The VQAs models have benefited from QNG optimizers that utilize the Fubini-Study metric to precondition parameter updates. Therefore, many researchers applied QNG or its variants to improve convergence in Variational Quantum Eigensolver (VQE), Quantum Approximate Optimization Algorithm (QAOA), and related VQAs. This section reviews some literature that utilize the QNG advance. The modern formulation of QNG optimization for VQAs is commonly attributed to Stokes *et al.* (2020) [2], who derived a quantum analogue of Amari's natural gradient by pulling back the Fubini-Study geometry to parameter space. This work obtained the QNG direction $\Delta\theta$ by solving a linear system of the form

$$g(\theta)\,\Delta\theta = -\eta\,\nabla L(\theta) \tag{1}$$

where $g(\theta)$ is the metric, L is the objective, and $\eta > 0$ is a learning-rate.

One of the most difficult things in central mathematics is that the pullback metric $g(\theta)$ is generically positive semidefinite: redundant parameterizations, symmetries, and directions that do not change the physical state and induce a nontrivial kernel. Therefore, the update may be ill-defined or nonunique, resulting in infinitely many solutions; then the numerical inversion becomes unstable, and the standard workaround is the Moore-Penrose pseudoinverse [2]. The early studies of the geometric behavior of the QNG dynamics in VQE settings were given in 2019 by Yamamoto [11], who demonstrated how the curvature information can reshape optimization trajectories or more patterns relative to Euclidean gradient descent (EGD). While these analyses are largely noiseless and conceptual, that raise the following question: how should one stabilize QNG when $g(\theta)$ is singular or poorly conditioned?

In 2020, Wierichs *et al.* [6] introduced a new way of QNG (under the name *NatGrad*), which is used to address rugged energy landscapes and barren plateaus in eigensolvers of the VQA algorithms.

Herein, the authors explicitly noted that by employing Tikhonov regularization of the metric, one may inhibit excessively large or erratic steps, translated as;

$$\big(g(\theta) + \lambda I\big)\Delta\theta = -\eta\,\nabla L(\theta), \qquad \lambda > 0 \tag{2}$$

where λI is damping mechanism used in Levenberg-Marquardt algorithm.

This regularization is a practical control knob that trades geometry-awareness for stability when the metric is near-singular or unreliable. Several studies for developing the QNG that operate along multiple orthogonal directions, but the fundamental problem of metric ill-definition persisted. While, by using quantum Fisher information notions on mixed states, Koczor and Benjamin (2022) [12] generalized QNG beyond unitary pure-state circuits to noisy and non-unitary settings. This work line is motivated by the fact that decoherence and the Completely positive trace-Preserving (CPTP) map dynamics change the effective geometry in realistic devices, making metric estimation and conditioning more difficult. Simultaneously, in 2024, Sasaki and Miyahara [13] introduced QNG based on the perspective of information geometry, where the experimental variations outside of the typical monotone-metric setting and elucidating nuances in the geometry underpinning QNG optimizer.

More recently, resource-reduction and approximation strategies have become prominent because exact QNG requires estimating an $m \times m$ metric. In 2024, Kolotouros and Wallden [14] introduced a paper titled *Random Natural Gradient* for enhancing the QNG, which used the stochastic coordinate variants with random measurements to construct a cheaper classical Fisher matrix approximation (which is a weakness point that will be discussed later in this paper) to the QFIM/QGT geometry, while also emphasizing that zero eigenvalues correspond to singularities, that is, the directions with no state change and hence motivate pseudoinverse- or regularisation-based treatments. In 2025, Halla [15] introduced some complementary developments for geometry-enhancing updates via higher-order manifold integration, geodesic corrections and they remark that the Fubini-Study metric is used *after Tikhonov regularization* due to ill-definition issues in variational quantum algorithms.

Also, many researchers introduced many algorithm to enhancement the QNG, such as in 2025, Dell'Anna *et al.* [8] proposed a new algorithm used applied full-metric QNG to the Quantum Approximate Optimization Algorithm (QAOA), in the same year, Shi *et al.* [9] introduced a Hamiltonian-aware QNG (H-QNG), and Fitzek *et al.* [10] proposed qBANG, combining QNG with Broyden-type quasi-Newton updates and momentum acceleration. While, The work in [11–14] will constitute the main focus of the discussion in this paper.

Recently, literature, especially in 2023-2026 period, adding λI value or in other words, closely related diagonal shifts appears primarily as a stabilization heuristic, to ensure invertibility and control large steps when $g(\theta)$ metric is singular or noisy [6,15]. Consequently, this work fundamentally change the question from *how* to numerically stabilize the update, to *why* is such regularization is a mathematical requisite and geometrically optimal? By re-framing the algorithm of the damped step, we move beyond ad hoc fixes to provide a full algorithm of the characterized theoretical optimisation, by demonstrating the proposed update that arises naturally as the unique minimizer of a local regularized Riemannian quadratic model as a unique way for the damped step. The proposed theoretical change enables us to see λ as a confidence (trust) parameter that pre-

cisely regulates the trade-off between enforcing quantum geometry and defaulting to Euclidean descent, rather than just as a correction factor. By establishing this equivalence, the quantum-geometric meaning of $g(\theta)$ is strictly preserved while providing a principled bridge to the classical trust-region and Levenberg-Marquardt theory. And finally conclude that the empirical evidence of studies shows:

- The QNG-based optimizers consistently outperform Euclidean GD in convergence speed;
- The metric regularization is essential for stability on realistic quantum hardware;
- The adaptive and damped QNG variants improve robustness without sacrificing accuracy.

3 Regularized Riemannian Quadratic Model for Damped QNG

To move optimally on the manifold of quantum states, the QNG implements the quantum Fubini–Study metric and when the enhancement becomes unstable because of the ill-conditioned metrics, noisy estimates, or local curvature irregularities, a damped or regularized approach is needed. The Regularized Riemannian quadratic algorithm represent the stabilization mechanism that borrows from classical trust-region and the second-order Riemannian optimization to reassembles the update step.

3.1 Setting and Notation

To form and characterize the damped QNG, the following notations are needed.

Assume $L : \mathbb{R}^m \to \mathbb{R}$ be a twice continuously differential loss function, where $\theta \in \mathbb{R}^m$ denotes the parameter vector, and m is the number of circuit parameters.

Let $|\psi(\theta)\rangle$ be a smooth parameterized quantum state, where $|\psi(\theta)\rangle \in \mathcal{H}$ and $\mathcal{H}$ is a complex Hilbert space.

Let $Q_{ij}(\theta)$ be the QGT which is defined as;

$$Q_{ij}(\theta) = \langle \partial_i \psi(\theta) \mid \partial_j \psi(\theta) \rangle - \langle \partial_i \psi(\theta) \mid \psi(\theta) \rangle \langle \psi(\theta) \mid \partial_j \psi(\theta) \rangle, \tag{3}$$

where $\partial_i := \partial / \partial \theta_i$, and θ_i is the ith coordinate of θ.

Let

$$g_{ij}(\theta) := \Re(Q_{ij}(\theta)), \tag{4}$$

denote the (pullback) Fubini-Study metric tensor on parameter space, while $\Re(\cdot)$ represents the real part. Then $g(\theta) \in \mathbb{R}^{m \times m}$ is symmetric and positive semidefinite, as $g(\theta) \succeq 0$.

Let $\nabla L(\theta) \in \mathbb{R}^m$ be the Euclidean gradient, where ∇ corresponds to θ. Assuming $\eta > 0$ be the learning-rate parameter, where η scales the update magnitude.

3.2 Regularized Quadratic Model

For a damping parameter $\lambda > 0$, where λ controls regularization strength, the local quadratic model $m_\theta : \mathbb{R}^m \to \mathbb{R}$ is defined by;

$$m_\theta(\Delta\theta) = \nabla L(\theta)^\top \Delta\theta + \frac{1}{2\eta}\, \Delta\theta^\top g(\theta)\Delta\theta + \frac{\lambda}{2\eta}\, \|\Delta\theta\|_2^2, \tag{5}$$

where $\Delta\theta \in \mathbb{R}^m$ is a candidate step, and $\|\cdot\|_2$ denotes the Euclidean norm.

Theorem 1. *The regularized quadratic characterization of damped QNG.*

Let $L \in C^2(\mathbb{R}^m)$ and let $g(\theta)$ be the Fubini-Study metric defined by Eqs. (3)–(4) at a fixed $\theta \in \mathbb{R}^m$, with $\eta > 0$ and $\lambda > 0$. Then the damped quantum natural-gradient step $\Delta\theta^\star$ that defined by the following linear system

$$\big(g(\theta) + \lambda I\big)\Delta\theta^\star = -\eta\, \nabla L(\theta), \tag{6}$$

where the identity matrix I of $m \times m$ size is exists, unique, and satisfies

$$\Delta\theta^\star = \arg\min_{\Delta\theta \in \mathbb{R}^m} m_\theta(\Delta\theta), \tag{7}$$

where the quadratic model m_θ in Eq. (5) is strictly convex and the minimizer is unique.

Proof. Fix $\theta \in \mathbb{R}^m$ and to simplify the equation, we abbreviate $g := g(\theta)$, $b := \nabla L(\theta)$, and $\Delta := \Delta\theta$. By considering the function

$$m(\Delta) := b^\top \Delta + \frac{1}{2\eta}\Delta^\top g \Delta + \frac{\lambda}{2\eta}\|\Delta\|_2^2,$$

Since $g = g^\top$ and $\lambda > 0$, the Hessian of m with respect to Δ becomes;

$$\nabla_\Delta^2 m(\Delta) = \frac{1}{\eta}\,(g + \lambda I). \tag{8}$$

Considering $g \succeq 0$ and $\lambda > 0$, then $g + \lambda I \succ 0$ is obvious. And since, $\nabla_\Delta^2 m(\Delta) \succ 0$, therefore m is strictly convex on $\mathbb{R}^m$.

Consequently, m admits a unique global minimizer.

To compute the minimizer, take the gradient of m with respect to Δ, means:

$$\nabla_\Delta m(\Delta) = b + \frac{1}{\eta}\, g\Delta + \frac{\lambda}{\eta}\Delta = b + \frac{1}{\eta}\,(g + \lambda I)\Delta. \tag{9}$$

By setting $\nabla_\Delta m(\Delta) = 0$ gives the first-order optimal condition

$$(g + \lambda I)\Delta = -\eta\, b,$$

which is exactly Eq. (6).

Let the gain ratio ρ be defined as follows;

$$\rho = \frac{\text{Actual Drop}}{\text{Predicted Drop}} = \frac{L(\theta) - L(\theta + \Delta\theta)}{m(0) - m(\Delta\theta)}, \tag{10}$$

where $m(\Delta\theta)$ represents the quadratic model.

Since $g + \lambda I \succ 0$ is invertible, with a unique solution to its linear system $\Delta^\star$. And by strict convexity property, the stationary point represents the unique global minimizer of m, proving Eq. (7) as requisite.

Remark 1. Invertibility and numerical stability.

Let $\{\mu_i\}_{i=1}^m$ denotes the eigenvalues of $g(\theta)$, where $\mu_i \geq 0$ and $g(\theta) \succeq 0$. Then the eigenvalues of $g(\theta) + \lambda I$ are $\mu_i + \lambda > 0$ for all i, and since $\lambda > 0$, yields $g(\theta) + \lambda I$ is a well-conditioned relative to $g(\theta)$ and always invertible.

Theorem 2 (Large-damping limit and interpolation to Euclidean gradient descent).

Let $\Delta\theta^\star(\lambda)$ be the unique solution of Eq. (6) at damping level $\lambda > 0$. Then

$$\Delta\theta^\star(\lambda) = -\eta\,(g(\theta)+\lambda I)^{-1}\nabla L(\theta) \quad \textit{and} \quad \Delta\theta^\star(\lambda) = -\frac{\eta}{\lambda}\,\nabla L(\theta) + O\left(\frac{1}{\lambda^2}\right) \quad \textit{as } \lambda \to \infty, \tag{11}$$

where $O(\cdot)$ is the Landau symbol.

In particular, the damped QNG direction converges to the Euclidean steepest-descent direction $-\nabla L(\theta)$, with step size scaled by $1/\lambda$.

Proof. Let $g = U\,\mathrm{diag}(\mu_1, \ldots, \mu_m)U^\top$ be the spectral decomposition of g, where U is unitary and orthogonal, while $\mu_i \geq 0$, yields;

$$(g + \lambda I)^{-1} = U\,\mathrm{diag}\left(\frac{1}{\mu_1 + \lambda}, \ldots, \frac{1}{\mu_m + \lambda}\right) U^\top. \tag{12}$$

For each i,

$$\frac{1}{\mu_i + \lambda} = \frac{1}{\lambda} \cdot \frac{1}{1 + \mu_i/\lambda} = \frac{1}{\lambda}\left(1 - \frac{\mu_i}{\lambda} + O\left(\frac{1}{\lambda^2}\right)\right) \quad \text{as } \lambda \to \infty.$$

Substituting into Eq. (12) becomes;

$$(g + \lambda I)^{-1} = \frac{1}{\lambda} I - \frac{1}{\lambda^2} g + O\left(\frac{1}{\lambda^3}\right).$$

By multiplying the above with $-\eta\nabla L(\theta)$ gives Eq. (11).

Example 1. Suppose the VQA defined on a finite-dimensional $\mathcal{H}$ with a Riemannian smooth parameterized state

$$|\psi(\theta)\rangle = U(\theta)\,|0\rangle, \tag{13}$$

where $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ and

$$U(\theta) = e^{-i\theta_1 Z} e^{-i\theta_2 Z}, \tag{14}$$

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{Pauli-}Z \text{ operator.}$$

Although the parameter space is two-dimensional, the physical quantum state depends only on the sum $\theta_1+\theta_2$, and the parameterization is therefore redundant.

By assuming $L : \mathbb{R}^2 \to \mathbb{R}$ be a twice continuously differentiable loss function defined as the expectation value of an observable H,

$$L(\theta) = \langle \psi(\theta) | H | \psi(\theta) \rangle, \tag{15}$$

where H is a fixed Hermitian operator.

The QGT associated with $|\psi(\theta)\rangle$ is given by

$$Q_{ij}(\theta) = \langle \partial_i \psi(\theta) | \partial_j \psi(\theta) \rangle - \langle \partial_i \psi(\theta) | \psi(\theta) \rangle \langle \psi(\theta) | \partial_j \psi(\theta) \rangle, \tag{16}$$

where $\partial_i = \partial / \partial \theta_i$.

By taking its real part, yields the following Fubini-Study metric;

$$g(\theta) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{17}$$

which is positive semidefinite but singular of rank one.

This degeneracy reflects the fact that the Fubini-Study metric is defined on the manifold of physical quantum states rather than the ambient parameter space.

The standard QNG update seeks $\Delta\theta \in \mathbb{R}^2$ satisfying

$$g(\theta)\, \Delta\theta = -\eta\, \nabla L(\theta), \tag{18}$$

where $\eta > 0$ is a learning rate and $\nabla L(\theta)$ denotes the Euclidean gradient.

Since $g(\theta)$ is singular, then the corresponding linear system admits either infinitely many solutions or none, depending on $\nabla L(\theta)$, and consequently, the QNG update is ill-posed and lacks existence and uniqueness.

The failure is intrinsic and persists even in the absence of noise, arising purely from the geometry induced by redundant parameterizations.

In the current use of the QNG model, this challenge is usually handled by adding a small diagonal shift to the metric. However, this regularization is often applied in an ad hoc manner and lacks a clear geometric or optimization interpretation. Specifically, it is uncertain if the resulting update leads to the minimization of any clearly defined local model of the loss function or if it ensures descent.

In contrast, the proposed research considers the regularized system

$$(g(\theta) + \lambda I)\, \Delta\theta = -\eta\, \nabla L(\theta), \tag{19}$$

To avoid viewing Eq. (19) as an ad hoc modification, the following, shows that it arises as the first-order optimality condition of the strictly convex quadratic model;

$$m_\theta(\Delta\theta) = \nabla L(\theta)^\top \Delta\theta + \frac{1}{2\eta}\Delta\theta^\top g(\theta)\Delta\theta + \frac{\lambda}{2\eta}\|\Delta\theta\|_2^2. \tag{20}$$

The presence of the Euclidean regularization term ensures strict convexity, guarantees existence and uniqueness of the minimizer for all $\lambda > 0$.

From a geometric perspective, the quadratic term induced by $g(\theta)$ penalizes directions that strongly affect the quantum state, while the regularization term penalizes large parameter displacements independently of the quantum geometry.

When the metric is degenerate or unreliable, the regularization term dominates, ensuring a stable descent direction. As $\lambda \to 0$, the update approaches the standard QNG direction restricted to the range of $g(\theta)$, whereas in the limit $\lambda \to \infty$ the update converges to

$$\Delta\theta = -\frac{\eta}{\lambda}\nabla L(\theta), \tag{21}$$

corresponding to Euclidean Gradient Descent with a vanishing step size.

This example demonstrates that the proposed method does not merely regularize the quantum metric for numerical convenience, but also embeds Quantum Natural Gradient descent within the principled geometrical and theoretical optimization of quadratic framework. In doing so, it resolves the intrinsic ill-posedness of standard QNG under redundant parameterizations and establishes a continuous interpolation between geometry-aware and geometry-agnostic optimization regimes.

4 The Analysis of the Metrics Performance

This part of the work outlines and compares, in a detailed evaluation, the analysis performance metrics obtained from three recent scholarly papers [8–10] and then shows a comparison with our new proposed algorithm.

The work in [8] studied the QNG model and compared with the standard gradient descent called (Vanilla GD) that used the QAOA algorithm as a case study for finding the ground state of a transverse-field Ising chain (TFIM) on $N = 2 - -12$ qubits. The work executed the QAOA under some conditions, where the depth $p = 4$ considered on two platforms; Rydberg-atom and superconducting qubits, with a two suggested restrictions; noiseless and realistic noisy conditions. The metrics included number of iterations for circuit evaluations, to reach the ground state, with a succeed 50% convergence rate of random initializations, and the final energy error. And, the accuracy $\epsilon =$ the energy difference from true ground state. Additionally, the sensitivity of two noise sources; the laser noise and the SPAM errors were suggested, while the outcomes were obtained via a comparison performance of the models that showed a faster convergence values of QNG algorithm in average of a significantly fewer iterations requirements than vanilla GD for growing system size. For instance, at $N = 10 - -12$

qubits, the full metric QNG model, only ∼!500 iterations is needed to reach the ground state in average vs > 3000 iterations are required for vanilla model.

In 2025, the authors in [9], presented a novel QNG algorithm that addresses the limitation of the additional computational cost by using the advantage of the Hamiltonian structure benefiting from the metric of the Riemannian pull-back of the lower-dimensional subspace spanned by the Hamiltonian terms onto the parameter space. The algorithm is called Hamiltonian-aware Quantum Natural Gradient Descent (H-QNG). Where, the projected Fubini-Study metric is applied on the subspace spanned by the Hamiltonian's terms, resulting a lower dimensional Fisher matrix instead of using the full state-space metric. Based on this idea, the authors executed the H-QNG on the VQE problems in quantum chemistry via a small lithium hydride (LiH) molecules and a linear chain of six hydrogen atoms H_6 optimized via the STO-6G hardware of minimal basis (ansätze) that required (8–12) qubits.

For the metrics performance, the energy error of the chemical accuracy is defined by 1.6×10^{-3} Hartree (Ha) (the atomic energy unit), while the proposed algorithm targeting a reduction of the energy error below the up mentioned chemical accuracy threshold. The lower quantum resource cost number of Hamiltonian or metric measurements confirms that the algorithm is more practical on current noisy hardware where measurement time is expensive. Therefore, the results of this experiment was a promising, where the H-QNG keeps the fast convergence of standard QNG mode but with a noticed lower cost per iteration and it reached the chemical accuracy in fewer iterations compared to the standard QNG for both LiH and H_6.

For example, the H-QNG achieved the 1.6 mHa accuracy of 20% faster iteration count than the standard QNG for LiH, that needed only half the quantum measurement shots per step as it estimates a smaller metric, resulting an overall reduction of the running time. Both H-QNG and standard QNG achieved final energies within chemical accuracy rate, but H-QNG showed improved robustness, converged reliably with smaller learning rates and without divergence even when the Fubini-Study metric was near-singular, where the regularization λ values were added to the metric. The results showed a slight degraded performance of H-QNG mode under shot noise for some finite sampling, despite reaching the chemical accuracy faster than stander QNG at the same shots number. The convergence curve with noise is closely tracked the noiseless case in early iterations, which indicate the incorporating Hamiltonian knowledge and using the damped metric inversion yields a more noise-resilient optimizer.

While, back in 2024, Fitzek et al. [10] introduced a new algorithm, called the quantum Broyden adaptive natural gradient (qBANG), that combines two models: an adaptive quasi-Newton update of the QNG metric, also called Broyden's method to approximate the inverse Fisher matrix of a heavy-ball momentum term. The aim of qBANG algorithm is to capture the curvature that benefits the QNG without the full cost of the metric evaluation in each iteration. The experiment began with benchmarks that included a circuit of barren plateau of 9 qubits, and a random layered gates, to test the optimizer stability, VQE for H_4

(4 hydrogen), LiH, and H_2O molecules of (8–12) qubits with frozen core, and a Max-Cut QAOA mode on a 12-node graph (noted in text). The metrics include the number of circuit evaluations of convergence speed that used to reach the ground-state energy or in other words, to give the approximation ratio, final energy error, and sensitivity to hyper-parameters and the shot noise optimizer. While the results of the qBANG achieved an overall more stable and faster convergence compared to the standard QNG and the gradient-based models depicted in Adam's method [16].

Based on the above analysis, the qBANG vastly outperformed in the H_4 and LiH of VQE mode, than the other methods like standard QNG and the H-QNG optimizers. The outcomes show a lower energy solutions consistently across 10–15 random initializations. While the variance of final energies was small for qBANG, indicating reliable convergence. Whereas, the QNG (exact or block-diag) and the vanilla GD showed a larger scatter. For instant, in H_2O, the qBANG was the only optimizer to discover the exact ground-state energy in at least one trial, with a 2-layer shallow ansatz, something QNG and Adam's methods couldn't achieve. Additionally, the qBANG shows more flexibility to hyperparameters, such as in a 9-qubit barren plateau test, while, the QNG and a full Newton method could diverge or oscillate for a large step size η, meantime the qBANG and Adam methods remain stabled over a range of η.

Overall, the analysis shows that augmenting QNG model with adaptive updates, subsystem weighting, or regularization can improve VQA optimization in terms of speed, robustness, and resource efficiency.

However, our proposed damped QNG model with trust-region is constructed based on the above-mentioned outcomes. Where, the Tikhonov damping term λI is utilized in Fubini-Study metric calculation and impose an explicit trust-radius Δ on the QNG step. Thus, the update rule can be defined as:

$$\Delta\theta^\top \big(F(\theta) + \lambda I\big)\Delta\theta \leq \Delta^2 \tag{22}$$

with a step size constraint in the metric geometry of the form:

$$\Delta\theta^\top F(\theta)\Delta\theta \leq \Delta^2, \tag{23}$$

making sure of the fixed range where the local quadratic model, based Fisher information is trustworthy. Table 1 below shows the comparison performance of our proposed model to the recent QNG-based methods on key metrics.

The values presented in Table 1 are computed based on a unified optimization-theoretical criterion grounded in the geometry of the variational parameter manifold, as follows;

Firstly, assumed $L : \mathbb{R}^m \to \mathbb{R}$ denote the objective function and $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \Delta\boldsymbol{\theta}_t$ defines the iterative update generated by each optimizer. The column of convergence refers to the required relative iteration count to satisfy $L(\boldsymbol{\theta}_t) - L^\star \leq \varepsilon$ for a fixed tolerance ε, where $L^\star$ refers to the optimal value, where all methods evaluated under the same initialization ensemble and iteration budget.

While the accuracy column refers to the asymptotic residual $L(\boldsymbol{\theta}_T) - L^\star$ after a fixed number of iteration T, and it reflects the ability of the used method to

Table 1. Comparison of QNG–based optimizers to VQA representative tasks

Optimizer	Task	Convergence	Accuracy	Stability
Gradient Descent (GD)	–	Baseline	Task-dependent	Sensitive to init.; degrades under noise
QNG [2]	VQE, QAOA	2–3× faster than GD	Lower cost; fewer local minima	More stable via geometric preconditioning
QNG (Noisy QAOA) [8]	QAOA (TFIM, ≤12 qubits)	$\sim 5\times$ faster for $N \geq 10$	Ground state up to $N = 12$	$> 90\%$ success; robust to noise
Hamiltonian-aware QNG [9]	VQE (LiH, H_6)	Faster than standard QNG	Chemical accuracy ($\sim$1 mHa)	Stable under metric singularities
qBANG [10]	VQE, QAOA	$\leq$ QNG QPU calls	Best energies (H_2O)	Very low variance; noise tolerant
Regularized QNG (our work)	VQE, QAOA	Comparable to QNG	Same optima; no overshoot	**High**: damping + trust region

reach low-energy or low-loss regions. The stability column refers to the assessment through the boundedness and conditioning of the update map, governed by the spectral properties of the preconditioning metric. While Euclidean gradient descent utilizes the identity metric and standard QNG uses the pullback Fubini–Study metric, regularized and Hamiltonian-aware variants employ strictly positive-definite or structured approximations instead. In particular, the proposed regularized QNG ensures well-posed updates by replacing $g(\boldsymbol{\theta})$ with $g(\boldsymbol{\theta})+\lambda I$ and enforcing a trust-region constraint, guaranteeing existence, uniqueness, and controlled step norms; the corresponding table entries followed directly from these geometric-related and spectral attributes, whereas entries for earlier methods summarize their reported behavior beneath the same mathematical criteria. Figure 1 illustrates the comparison of the convergence rate versus number of qubits N between our proposed model to the other cited methods (see Fig. 1).

The damped and trust-region regularized QNG model is expected to match the fast convergence of standard QNG, while other methods do not use damping despite providing improved stability. In Eq. (22), λ mitigates issues when F has tiny eigenvalues, thus preventing abrupt, large parameter updates. Furthermore, the same equation enforces a constraint ensuring each update stays within a fidelity distance Δ; which is analogous to the reparameterization invariance of the natural gradient on the Riemannian surface, but with explicit safety margins. The damped QNG model directly addresses the instability observed in standard QNG, which lacks damping, when using large learning rates or finite sampling.

For example, as shown in Fig. 2 during a 2-qubit VQE test, our proposed damped QNG reached the target energy in approximately 2 iterations. In contrast, vanilla GD, using the same learning rate, took around 5 iterations. While standard QNG achieved the target in 1 to 2 iterations, it frequently overshot the minimum without damping. Figure 2 displays the loss and gradient norm

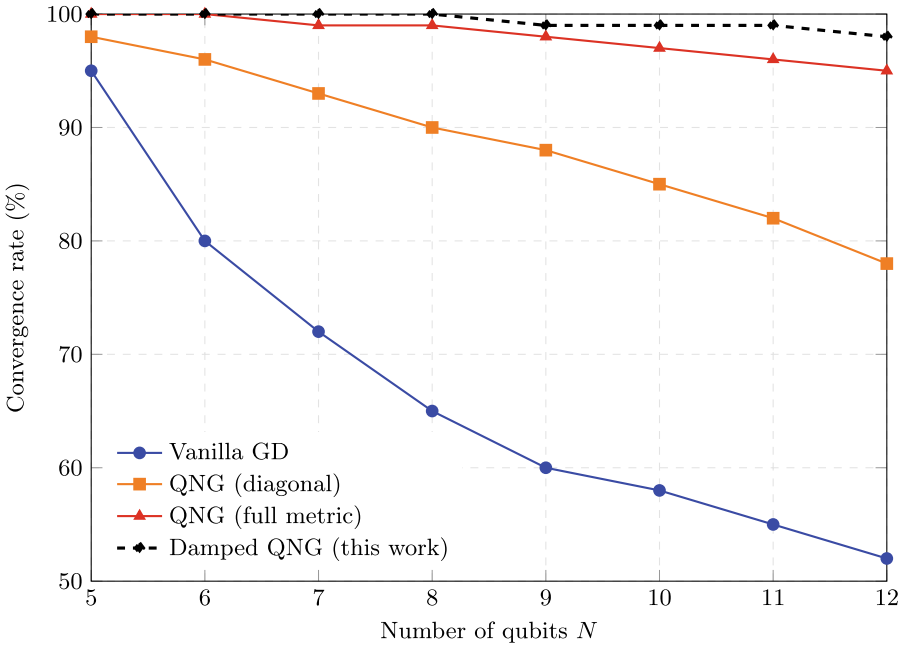

Fig. 1. Convergence rate versus number of qubits N for different QNG models

versus iterations for this test case, demonstrating that all QNG-based methods quickly reduce the cost and gradients; however, damped QNG specifically avoids the initial overshoot characteristic of pure QNG.

4.1 Geometric Interpretation and the Impact on Quantum Machine Learning (QML)

Based on our analysis and from a geometric perspective, the damping impact on Quantum Machine Learning (QML) based on the update approaches standard QNG as follows:

- The term $\Delta\boldsymbol{\theta}^\top g(\boldsymbol{\theta})\Delta\boldsymbol{\theta}$ penalizes motion that induces large changes in quantum state space;
- The Euclidean regularization term $\|\Delta\boldsymbol{\theta}\|_2^2$ controls parameter displacement independently of quantum geometry;
- The damping parameter λ quantifies confidence in the local quantum metric.

As $\lambda \to 0$, the update approaches standard QNG, while in the large-damping region where $\lambda \to \infty$, the update continuously interpolates toward Euclidean Gradient Descent, explained via;

$$\Delta\boldsymbol{\theta} = -\frac{\eta}{\lambda}\nabla L(\boldsymbol{\theta}) + \mathcal{O}(\lambda^{-2}). \tag{24}$$

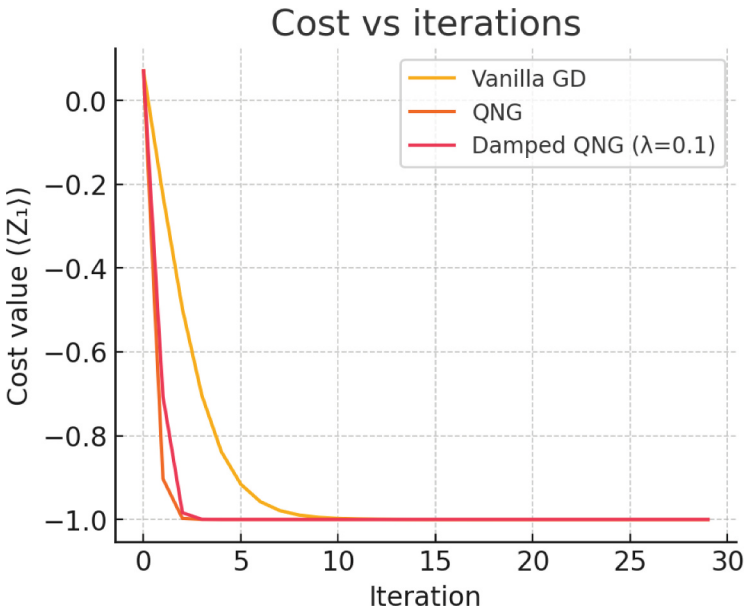

Fig. 2. Loss cost function value vs. iteration for a 2-qubit based on VQE.

5 Discussion

QNG methods are based on the premise that optimization in VQAs must account for the natural geometry of the quantum state space. For a parameterized quantum state $|\psi(\theta)\rangle$, the real part of the QGT defines the Fubini–Study metric, which measures the distinguishability of nearby quantum states and induces a Riemannian structure on the variational manifold [1,2]. From a mathematical perspective, the degeneracy of $g(\theta)$ is not a defect, but rather a fundamental feature of this geometric surface.

By utilizing the real part of the QGT, we effectively apply the Fubini–Study metric to the manifold of physical quantum states rather than the embedding parameter space. Consequently, directions in the parameter space that do not alter the physical state stemming from non-essential parameterizations, symmetries, or gauge freedoms—reside within the null space of $g(\theta)$. This implies that the pullback metric is generally positive semidefinite and, therefore, non-invertible. This observation explains why standard QNG may fail even under noiseless conditions, demonstrating that its instability is a structural geometric property rather than a mere numerical artifact [2].

In quantum computing devices, this intrinsic problem is further compounded by the finite-shot estimation of the QGT, as well as by qubit decoherence and noise-induced distortions of the effective geometry. Consequently, several works have introduced diagonal or Tikhonov-type regularization of the metric to restore invertibility [2,6,11]. While effective in practice, such regularization is typically

introduced heuristically, lacking a precise optimization-theoretic or geometric interpretation.

This work provides a principled resolution to this gap. By applying the damped trust-region regularized QNG update (see Eq. 19) within a locally defined Riemannian quadratic model of the loss function, showing that the damped update is the *unique minimizer* of a strictly convex optimization problem for all $\lambda > 0$. This establishes the existence, uniqueness, and numerical stability of the update in a mathematically rigorous manner. The quadratic model (Eq. 5) is used to interpret the geometric meaning of the damping parameter λ. Where the term $\Delta\theta^\top g(\theta)\Delta\theta$ penalizes motion that induces large changes in quantum state space, while the Euclidean regularization term $\|\Delta\theta\|_2^2$ controls the magnitude of parameter displacement independently of quantum geometry.

Thus, λ quantifies the degree of confidence placed in the local quantum geometric information. When $g(\theta)$ is reliable, a small λ allows for geometry-aware updates; when the $g(\theta)$ metric is singular or noisy, a larger λ suppresses unstable geometric directions. Therefore, damped QNG realizes a continuous interpolation between the Quantum Natural Gradient and geometry-agnostic gradient descent. This interpolation is both mathematically and physically fundamental, reflecting a smooth transition between state-space-sensitive and parameter-space-sensitive optimization regimes. Importantly, the gain-ratio mechanism introduced in the damped QNG algorithm provides a model-based criterion for adapting λ. By comparing the actual decrease in L with that predicted by the quadratic model, the algorithm dynamically adjusts its reliance on quantum geometry. This connects damped QNG with trust-region methods while preserving the fundamental geometric origin of the metric. To our knowledge, this specific quadratic model and confidence-region interpretation of damped QNG have not been previously formalized in the quantum optimization literature.

6 Conclusion

In this paper, a geometrically and mathematically based stabilisation of QNG optimisation for VQAs is proposed. It is observed from the information-geometric basis of QNG that the pullback character of the Fubini-Study metric and the existence of duplicate parameter directions led to the intrinsic ill-posedness of the standard update. As a novel outcome, we have shown that these problems stem from the geometry of quantum state manifolds and remain unaffected by sampling noise. And demonstrated that the damped QNG update is the only minimiser of a strictly convex local approximation of the loss function by introducing a Tikhonov-regularized Riemannian quadratic model. Meanwhile, for any damping parameters $\lambda > 0$, the result ensures existence, uniqueness, with a numerical stability. The outcome offers a strong theoretic and geometric optimizational justification for such stabilisation type, while diagonal regularisation of the QGT has already been emerged as a useful heuristic.

The method is well explained in both mathematical and physical concepts when the damping parameter is interpreted as a confidence measure in local

quantum geometry. The approach guarantees descent even in the case of faulty quantum geometric information by provably recovering Euclidean Gradient Descent in the large-damping domain.

On the other hand, the technique maintains the benefits of geometry-aware optimisation by taking advantage of the curvature of quantum state space in the small-damping domain. Two basic optimisation paradigms are connected by this continuous interpolation inside a single cohesive framework. From the outcomes, the suggested method defines the function of quantum geometry in optimisation, converts metric regularisation from a heuristic to a logical modeling decision, and offers a strong basis for future advancements. Benchmark analyses comparing recent QNG-based optimizers, including Hamiltonian-aware QNG, qBANG, and full metric QNG, illustrate that our regularized model achieves convergence speeds comparable to standard QNG, while providing superior stability, eliminating overshooting under near-singular metrics, finite-shot noise, and redundant parameterizations. In small-scale Variational Quantum Eigensolver (VQE) experiments (2–12 qubits). The proposed damped QNG reaches target energies in 1–2 iterations, avoids instability noticed in the standard QNG technique, and maintains robust performance under sampling noise. These results establish the regularized QNG trust region as a mathematically grounded, geometry-preserving, and practically robust optimizer for near-term quantum algorithms.

The future research could expand this approach to including stochastic convergence assessments under realistic device noise, mixed-state quantum Fisher information, and approximation or block-diagonal metrics, all of which are major issues in quantum optimisation between 2020 and 2026 [2,6,7,17].

Acknowledgments. This research was sponsored by the University of Malaya Translational Research Grant Scheme, TRG0019-2025 (UM.0002388/HRU.IP) and the Research Excellence Grant, UMREG 068 - 2024 (UM.0000285/HRU.RC).

Author contributions. All authors contributed equally and significantly to writing this article.

Conflicts of Interest. The authors declare no conflict of interest.

References

1. Amari, S.I.: Natural gradient works efficiently in learning. Neural Comput. **10**(2), 251–276 (1998)
2. Stokes, J., Izaac, J., Killoran, N., Carleo, G.: Quantum natural gradient. Quantum **4**, 269 (2020)
3. Alwan, N.A., Al-Saidi, N.M., Obaiys, S.J., Noor, N.F.B.M., Ahmad, M.: A comprehensive review on classification of quantum image representations. Arch. Comput. Methods Eng. 1–59 (2025)
4. Alwan, N.A., Obaiys, S.J., Al-Saidi, N.M., Noor, N.F.B.M., Karaca, Y.: A multichannel quantum image representation model with qubit sequences for quantum-inspired image and image retrieval. AIMS MATHEMATICS **10**(5), 10994–11035 (2025)

5. Alwan, N.A., Obaiys, S.J., Al-Saidi, N.M., Noor, N.F.B.M., Karaca, Y.: Multilayered quantum computing and simulation system for enhanced image representation of HSI based Fourier transform and adjacency matrix. Sci. Rep. **15**(1), 40286 (2025)
6. Wierichs, D., Gogolin, C., Kastoryano, M.: Avoiding local minima in variational quantum eigensolvers with the natural gradient optimizer. arXiv:2004.14666 (2020)
7. Martens, J., Sutskever, I.: Training deep and recurrent networks with hessian-free optimization. In: Montavon, G., Orr, G.B., Müller, K.-R. (eds.) Neural Networks: Tricks of the Trade. LNCS, vol. 7700, pp. 479–535. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35289-8_27
8. Dell'Anna, F., Gomez-Lurbe, R., Perez, A., Ercolessi, E.: Quantum natural gradient optimizer on noisy platforms: QAOA as a case study. arXiv preprint arXiv:2502.20288 (2025)
9. Shi, C., Dunjko, V., Wang, H.: Weighted approximate quantum natural gradient for variational quantum eigensolver. arXiv:2504.04932 (2025)
10. Fitzek, D., Jonsson, R.S., Dobrautz, W., Schäfer, C.: Optimizing variational quantum algorithms with qBang: efficiently interweaving metric and momentum to navigate flat energy landscapes. Quantum **8**, 1313 (2024)
11. Yamamoto, N.: On the natural gradient for variational quantum eigensolver. arXiv:1909.05074 (2019)
12. Koczor, B., Benjamin, S.C.: Quantum natural gradient generalized to noisy and nonunitary circuits. Phys. Rev. A **106**, 062416 (2022). Also arXiv:1912.08660
13. Sasaki, T., Miyahara, H.: Quantum natural gradient without monotonicity. arXiv:2401.13237 (2024). Phys. Rev. A **110**, 022439 (2024)
14. Kolotouros, I., Wallden, P.: Random natural gradient. Quantum **8**, 1503 (2024)
15. Halla, M.: Quantum natural gradient with geodesic corrections for small shallow quantum circuits. Physica Scripta **100**(5) (2025). Also arXiv:2409.03638
16. Kingma, D.P.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
17. Bonnabel, S.: Stochastic gradient descent on Riemannian manifolds. IEEE Trans. Autom. Control **58**(9), 2217–2229 (2013)