Article | Received 4 May 2026; Revised 25 May 2026; Accepted 25 May 2026; Published 22 June 2026
https://doi.org/10.55092/aic20260008

Quasi-closed-form MAP channel estimation with unknown noise covariance

Xin Meng

Huawei Technologies Co., Ltd., Shanghai, China; E-mail: xmeng@ieee.org

Highlights:

- Investigates the maximum a posteriori channel estimation problem under unknown noise covariance.
- Derives a quasi-closed-form solution that depends on a single scalar parameter obtained by solving a polynomial equation.
- Provides simplified approximations of the proposed solution for various special cases, with comprehensive analytical insights.

Abstract: In this paper, we investigate the maximum a posteriori channel estimation problem when the noise covariance is unknown. We develop a quasi-closed-form solution that requires only the determination of an unknown scalar parameter through solving a polynomial equation. Furthermore, we provide a comprehensive analysis of simplified approximations for this quasi-closed-form solution under several special cases.

Keywords: maximum a posteriori probability; channel estimation; co-channel interference

1. System model and problem formulation

We examine a fundamental training-based channel estimation problem with the following system model

$$\mathbf{y}_n = \mathbf{h}x_n + \mathbf{z}_n, \quad 0 \le n \le N-1, \tag{1}$$

where $\{\mathbf{y}_n\}_{n=0}^{N-1}$ represent the observed symbols, $\{x_n\}_{n=0}^{N-1}$ denote known pilot symbols, $\{\mathbf{z}_n\}_{n=0}^{N-1}$ are i.i.d. complex Gaussian noise vectors with $\mathbf{z}_n \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$, and $\mathbf{h}$ is an $M \times 1$ channel vector, respectively. Here both the channel vector $\mathbf{h}$ and the noise covariance $\boldsymbol{\Sigma}$ are unknown and need to be estimated.

Traditional Bayesian estimation typically assumes a Gaussian-inverse-Wishart joint prior for $\mathbf{h}$ and $\boldsymbol{\Sigma}$ [1,2], where $\boldsymbol{\Sigma}$ follows inverse-Wishart distribution and $\mathbf{h}$ follows Gaussian distribution conditioned on $\boldsymbol{\Sigma}$, respectively. While mathematically convenient for analysis, this assumption proves inadequate for practical problems, especially in wireless channel estimation with co-channel interference (CCI) [3,4], due to two critical mismatches:

(1) Parameter Dependence: The channel vector $\mathbf{h}$ (desired link) and interference-plus-noise

covariance $\mathbf{\Sigma}$ should be statistically independent, *i.e.*, $p(\mathbf{h}, \mathbf{\Sigma}) = p(\mathbf{h})p(\mathbf{\Sigma})$, which the Gaussian-inverse-Wishart prior violates.

(2) Distribution Mismatch: Practical CCI scenarios suggest $\mathbf{\Sigma} = \mathbf{G}\mathbf{G}^H + \sigma^2\mathbf{I}$, where the interference channel $\mathbf{G}$ can be assumed to be Gaussian. This implies $\mathbf{\Sigma}$ follows a non-central-Wishart distribution, contradicting the inverse-Wishart assumption.

Motivated by wireless channel estimation, we adopt the following realistic assumptions:

(1) Channel Prior: $\mathbf{h} \sim \mathcal{CN}(\bar{\mathbf{h}}, \mathbf{R})$ with known $\bar{\mathbf{h}}$ and $\mathbf{R}$, reflecting slow channel statistics variation in the desired link.

(2) Interference Uncertainty: The prior for $\mathbf{\Sigma}$ is treated as unknown, accounting for rapid interference pattern changes due to dynamic scheduling and resource allocation.

Within this framework, we will develop joint maximum a posteriori (MAP) estimation for both $\mathbf{h}$ and $\mathbf{\Sigma}$.

2. The solutions of joint MAP estimation

Based on Equation (1), the likelihood function can be expressed as

$$f(\mathbf{Y}|\mathbf{h}, \mathbf{\Sigma}) = |\pi\mathbf{\Sigma}|^{-N} \cdot \exp\left\{-\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\left(\mathbf{Y} - \mathbf{h}\mathbf{x}^H\right)\left(\mathbf{Y} - \mathbf{h}\mathbf{x}^H\right)^H\right)\right\} \quad (2)$$

where $\mathbf{Y} = [\mathbf{y}_0, \cdots, \mathbf{y}_{N-1}]$ and $\mathbf{x}^H = [x_0, \cdots, x_{N-1}]$. With the prior of $\mathbf{h}$, the joint posterior probability of $\mathbf{h}$ and $\mathbf{\Sigma}$ can be expressed as

$$\begin{aligned} f(\mathbf{h}, \mathbf{\Sigma}|\mathbf{Y}) &= \frac{f(\mathbf{h}, \mathbf{\Sigma}|\mathbf{Y}) f(\mathbf{h}) f(\mathbf{\Sigma})}{\iint f(\mathbf{h}, \mathbf{\Sigma}|\mathbf{Y}) f(\mathbf{h}) f(\mathbf{\Sigma})\, d\mathbf{h} d\mathbf{\Sigma}} \\ &\propto f(\mathbf{h}, \mathbf{\Sigma}|\mathbf{Y}) f(\mathbf{h}). \end{aligned} \quad (3)$$

Since the prior of the covariance matrix $\mathbf{\Sigma}$ is unknown, we assume an uninformative prior by setting its probability density function $f(\mathbf{\Sigma})$ to be constant. This leads to the $\propto$ in Equation (3).

Substituting Equation (2) and $f(\mathbf{h})$ into Equation (3), and then maximizing it, yields the MAP estimation of $\mathbf{h}$ and $\mathbf{\Sigma}$, which can be equivalent to the following minimization problem

$$\begin{aligned} \min_{\mathbf{h}, \mathbf{\Sigma}} \Big\{ & N \ln|\mathbf{\Sigma}| + \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\left(\mathbf{Y} - \mathbf{h}\mathbf{x}^H\right)\left(\mathbf{Y} - \mathbf{h}\mathbf{x}^H\right)^H\right) \\ & + \left(\mathbf{h} - \bar{\mathbf{h}}\right)^H \mathbf{R}^{-1}\left(\mathbf{h} - \bar{\mathbf{h}}\right)\Big\}. \end{aligned} \quad (4)$$

We will analyze the optimal solution of Equation (4) in subsequent subsections, presenting both a quasi-closed-form solution and an iterative solution approach.

2.1. *Proposed quasi-closed-form solution*

It can be observed that the optimal $\mathbf{\Sigma}$ has the form

$$\mathbf{\Sigma} = \frac{1}{N}\left(\mathbf{Y} - \mathbf{h}\mathbf{x}^H\right)\left(\mathbf{Y} - \mathbf{h}\mathbf{x}^H\right)^H. \quad (5)$$

Substituting Equation (5) into Equation (4) yields an equivalent problem

$$\min_{\mathbf{h}}\Big\{N\ln\left|\left(\mathbf{Y}-\mathbf{h}\mathbf{x}^H\right)\left(\mathbf{Y}-\mathbf{h}\mathbf{x}^H\right)^H\right| \\ +\left(\mathbf{h}-\bar{\mathbf{h}}\right)^H\mathbf{R}^{-1}\left(\mathbf{h}-\bar{\mathbf{h}}\right)\Big\}. \tag{6}$$

Once the optimal $\mathbf{h}$ is obtained, the optimal $\boldsymbol{\Sigma}$ can be obtained straightforwardly with Equation (5). Then we will mainly focus on the optimal $\mathbf{h}$ of Equation (6).

After performing the singular value decomposition of $\mathbf{x}^H$

$$\mathbf{x}^H=\left[\lambda_x,\mathbf{0}_{1\times(N-1)}\right]\mathbf{V}_x^H \tag{7}$$

where $\mathbf{V}_x$ is an $N\times N$ unitary matrix, we can obtain

$$\begin{aligned}\mathbf{Y}-\mathbf{h}\mathbf{x}^H&=\left(\mathbf{Y}\mathbf{V}_x-\mathbf{h}\left[\lambda_x,\mathbf{0}_{1\times(N-1)}\right]\right)\mathbf{V}_x^H\\&=\left[\mathbf{Y}\mathbf{V}_x\mathbf{e}_1-\lambda_x\mathbf{h},\mathbf{Y}\mathbf{V}_x\mathbf{E}_2\right]\mathbf{V}_x^H\end{aligned}$$

where $\mathbf{e}_1=[1,\mathbf{0}_{1\times(N-1)}]^T$ and $\mathbf{E}_2=[\mathbf{0}_{(N-1)\times1},\mathbf{I}_{N-1}]^T$. Therefore the minimization problem Equation (6) becomes

$$\min_{\mathbf{h}}\Big\{N\ln\Big|\mathbf{Y}\mathbf{V}_x\mathbf{E}_2\mathbf{E}_2^H\mathbf{V}_x^H\mathbf{Y}^H \\ +\left(\mathbf{Y}\mathbf{V}_x\mathbf{e}_1-\lambda_x\mathbf{h}\right)\left(\mathbf{Y}\mathbf{V}_x\mathbf{e}_1-\lambda_x\mathbf{h}\right)^H\Big| \\ +\left(\mathbf{h}-\bar{\mathbf{h}}\right)^H\mathbf{R}^{-1}\left(\mathbf{h}-\bar{\mathbf{h}}\right)\Big\}.$$

We let

$$\mathbf{Q}=\mathbf{Y}\mathbf{V}_x\mathbf{E}_2\mathbf{E}_2^H\mathbf{V}_x^H\mathbf{Y}^H \tag{8a}$$

$$\tilde{\mathbf{h}}=\lambda_x\mathbf{h}-\mathbf{Y}\mathbf{V}_x\mathbf{e}_1 \tag{8b}$$

$$\tilde{\bar{\mathbf{h}}}=\lambda_x\bar{\mathbf{h}}-\mathbf{Y}\mathbf{V}_x\mathbf{e}_1 \tag{8c}$$

$$\tilde{\mathbf{R}}=\lambda_x^2\mathbf{R} \tag{8d}$$

where $\tilde{\mathbf{h}}$ and $\tilde{\bar{\mathbf{h}}}$ denote the same linear transformation of $\mathbf{h}$ and $\bar{\mathbf{h}}$, respectively. Then utilizing the matrix determinant lemma

$$\left|\mathbf{A}+\mathbf{u}\mathbf{v}^H\right|=|\mathbf{A}|\left(1+\mathbf{v}^H\mathbf{A}^{-1}\mathbf{u}\right)$$

yields the equivalent problem

$$\min_{\tilde{\mathbf{h}}}\left\{N\ln\left(1+\tilde{\mathbf{h}}^H\mathbf{Q}^{-1}\tilde{\mathbf{h}}\right)+\left(\tilde{\mathbf{h}}-\tilde{\bar{\mathbf{h}}}\right)^H\tilde{\mathbf{R}}^{-1}\left(\tilde{\mathbf{h}}-\tilde{\bar{\mathbf{h}}}\right)\right\}. \tag{9}$$

We note that the analysis in Equation (9) assumes the invertibility of $\mathbf{Q}$. This condition holds if and only if $N\geq M+1$, as we rigorously prove in Appendix A.

Differential the objective function in Equation (9) with respect to $\tilde{\mathbf{h}}$ and letting the derivative equal to zero yields

$$\tilde{\mathbf{h}}=\nu\left(\mathbf{Q}^{-1}+\nu\tilde{\mathbf{R}}^{-1}\right)^{-1}\tilde{\mathbf{R}}^{-1}\tilde{\bar{\mathbf{h}}} \tag{10}$$

where

$$\nu = \frac{1}{N}\left(1 + \tilde{\mathbf{h}}^H \mathbf{Q}^{-1}\tilde{\mathbf{h}}\right). \tag{11}$$

With Equation (10) we can obtain that

$$\tilde{\mathbf{h}}^H \mathbf{Q}^{-1}\tilde{\mathbf{h}} = \nu^2 \mathbf{a}^H \mathbf{\Lambda}(\mathbf{\Lambda} + \nu\mathbf{I})^{-2}\mathbf{a} \tag{12}$$

where $\mathbf{\Lambda}$ and $\mathbf{a}$ are derived from the following eigenvalue decomposition and linear transformation

$$\tilde{\mathbf{R}}^{1/2}\mathbf{Q}^{-1}\tilde{\mathbf{R}}^{1/2} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \tag{13a}$$

$$\mathbf{a} = \mathbf{U}^H \tilde{\mathbf{R}}^{-1/2}\tilde{\tilde{\mathbf{h}}}. \tag{13b}$$

Since the right hand side of Equation (12) is the multiplication of a row vector, a diagonal matrix and a column vector, it can be expanded as

$$\tilde{\mathbf{h}}^H \mathbf{Q}^{-1}\tilde{\mathbf{h}} = \nu^2 \sum_{i=1}^{M} \frac{\lambda_i |a_i|^2}{(\nu + \lambda_i)^2} \tag{14}$$

where a_i and λ_i are the i-th element of $\mathbf{a}$ and the diagonal of $\mathbf{\Lambda}$ respectively. Substituting Equation (14) into Equation (11) and eliminating the denominator, we can obtain a $(2M+1)$-th degree polynomial equation of ν

$$(N\nu - 1)\prod_{i=1}^{M}(\nu + \lambda_i)^2 - \nu^2 \sum_{i=1}^{M} \lambda_i |a_i|^2 \prod_{j=1, j\neq i}^{M} (\nu + \lambda_j)^2 = 0. \tag{15}$$

As Equation (15) is a polynomial equation (unlike general nonlinear equations), efficient solution algorithms exist [5]. Among its $2M+1$ roots, the optimal ν must be positive and minimize Equation (9).

After obtaining the optimal ν, the partial MAP estimation $\hat{\mathbf{h}}_{\mathrm{PMAP}}$, *i.e.*, the optimum of Equation (4) can be calculated with Equations (8) and (10)

$$\hat{\mathbf{h}}_{\mathrm{PMAP}} = \left(\mathbf{x}^H\mathbf{x}\cdot(\nu\mathbf{Q})^{-1} + \mathbf{R}^{-1}\right)^{-1}(\nu\mathbf{Q})^{-1}\mathbf{Y}\mathbf{x} + \left(\mathbf{x}^H\mathbf{x}\cdot(\nu\mathbf{Q})^{-1} + \mathbf{R}^{-1}\right)^{-1}\mathbf{R}^{-1}\bar{\mathbf{h}}. \tag{16}$$

The name "partial MAP" means that the prior distribution of the noise covariance matrix is unknown, therefore only partial prior is used in MAP estimation.

2.2. Approximations for the quasi-closed-form solution in special cases

When the covariance matrix $\mathbf{\Sigma}$ is known a priori in Equation (4), the optimal solution for $\mathbf{h}$ reduces to the classical MAP estimation

$$\hat{\mathbf{h}}_{\mathrm{MAP}} = \left(\mathbf{x}^H\mathbf{x}\cdot\mathbf{\Sigma}^{-1} + \mathbf{R}^{-1}\right)^{-1}\mathbf{\Sigma}^{-1}\mathbf{Y}\mathbf{x} + \left(\mathbf{x}^H\mathbf{x}\cdot\mathbf{\Sigma}^{-1} + \mathbf{R}^{-1}\right)^{-1}\mathbf{R}^{-1}\bar{\mathbf{h}}. \tag{17}$$

Note that the sole difference between Equations (16) and (17) lies in the terms $\mathbf{\Sigma}$ and $\nu\mathbf{Q}$. This indicates that when $\mathbf{\Sigma}$ is unknown, $\nu\mathbf{Q}$ effectively serves as its substitute, while the remaining formulation maintains the structure of classical MAP estimation. Moreover, as demonstrated in Appendix A, $\frac{1}{N-1}\mathbf{Q}$ provides an unbiased estimate of $\mathbf{\Sigma}$. The parameter ν thus functions as a scaling factor to compensate for this substitution.

Since ν must simultaneously satisfy two conditions: being a positive root of the polynomial Equation (15) and minimizing Equation (9), deriving exact closed-form expressions for either ν or $\hat{\mathbf{h}}_{\text{PMAP}}$ may be intractable. Nevertheless, certain special cases admit simplified solutions. First, we define the signal-to-noise ratio (SNR) as

$$\text{SNR} = \frac{\mathbb{E}\left\{\text{tr}\left(\mathbf{h}\mathbf{x}^H\mathbf{x}\mathbf{h}^H\right)\right\}}{\mathbb{E}\left\{\text{tr}\left(\mathbf{z}_n\mathbf{z}_n^H\right)\right\}} = \frac{\lambda_x^2\text{tr}\left(\mathbf{R}\right)}{\text{tr}\left(\mathbf{\Sigma}\right)} = \frac{1}{\text{tr}\left(\mathbf{\Sigma}\right)}$$

where we have assumed that $\lambda_x^2 = 1$ and $\text{tr}\left(\mathbf{R}\right) = 1$ for simplicity. Then we consider three special cases: $N \to +\infty$, $\text{SNR} \to -\infty$ (in dB) and $\text{SNR} \to +\infty$, where the final two cases are obtained by setting $\text{tr}\left(\mathbf{\Sigma}\right) \to +\infty$ and $\text{tr}\left(\mathbf{\Sigma}\right) \to 0$, respectively.

2.2.1. $N \to +\infty$

It is noted that in the left hand side of Equation (13a),

$$\tilde{\mathbf{R}}^{1/2}\mathbf{Q}^{-1}\tilde{\mathbf{R}}^{1/2} = \frac{1}{N-1}\mathbf{R}^{1/2}\left(\frac{1}{N-1}\mathbf{Q}\right)^{-1}\mathbf{R}^{1/2}.$$

Since $\frac{1}{N-1}\mathbf{Q}$ is an unbias estimation of $\mathbf{\Sigma}$ (see Appendix A), when $N \to +\infty$, we have $\frac{1}{N-1}\mathbf{Q} \to \mathbf{\Sigma}$. Then in the right hand side of Equation (13a) each $\lambda_i \to 0$. Similarly, in the left hand side of Equation (13b),

$$\mathbf{a} = \mathbf{U}^H\mathbf{R}^{-1/2}\left(\bar{\mathbf{h}} - \mathbf{Y}\mathbf{V}_x\mathbf{e}_1\right).$$

When $N \to +\infty$, each $\frac{1}{N}|a_i|^2 \to 0$. Therefore Equation (15) can be approximated by

$$(N\nu - 1)\prod_{i=1}^{M}(\lambda_i + \nu)^2 = 0.$$

Since each λ_i is positive, the only positive root of this equation will be $\frac{1}{N}$. With $\nu \to \frac{1}{N}$ and $\frac{1}{N-1}\mathbf{Q} \to \mathbf{\Sigma}$, *i.e.*, when the training sample number is sufficiently large, we can obtain a highly accurate estimate of $\mathbf{\Sigma}$. Consequently, the partial MAP solution in Equation (16) approximates the standard MAP estimation given by Equation (17).

2.2.2. $\text{SNR} \to -\infty$

From Equation (11), we observe that ν remains bounded away from zero for fixed N. Consequently, in Equation (16), all elements of $(\nu\mathbf{Q})^{-1}$ diminish sufficiently as $\text{SNR} \to -\infty$. This leads to two key implications. Firstly, the partial MAP estimation (Equation (16)) converges to the prior mean $\bar{\mathbf{h}}$. Secondly, when the noise dominates (*i.e.*, at extremely low SNR), the optimal estimation naturally reverts to the prior mean $\bar{\mathbf{h}}$. This behavior aligns perfectly with classical MAP estimation theory, where the prior dominates the solution under high uncertainty conditions.

2.2.3. $\mathrm{SNR} \to +\infty$

As $\lambda_i \to +\infty$ for all i, dividing both sides of Equation (15) by $\prod_{i=1}^{M} \lambda_i^2$ yields the approximation $N\nu - 1 \to 0$, which implies $\nu \to \frac{1}{N}$. This result indicates that as $\mathrm{SNR} \to +\infty$, the partial MAP estimation of $\mathbf{\Sigma}$ converges to $\frac{1}{N}\mathbf{Q}$. Notably, Appendix A demonstrates that $\frac{1}{N}\mathbf{Q}$ coincides with the maximum likelihood (ML) estimation of $\mathbf{\Sigma}$. This agreement occurs because at high SNR, the data likelihood dominates the prior information about $\mathbf{h}$, driving the partial MAP solution to converge asymptotically to the ML estimation.

2.3. *Another iterative-form solution*

Equation (5) provides a direct solution for $\mathbf{\Sigma}$ when $\mathbf{h}$ is known, while Equation (17) yields a direct solution for $\mathbf{h}$ when $\mathbf{\Sigma}$ is known. This mutual dependence suggests an iterative estimation approach for jointly determining $\mathbf{h}$ and $\mathbf{\Sigma}$. This iterative approach corresponds to the expectation maximization (EM) solution [6]. While the convergence to the global optimum of iterative EM cannot be theoretically guaranteed, our simulation results demonstrate that the iterative EM solution and the propose quasi-closed-form solution yield identical performance.

2.4. *Complexity analysis*

Furthermore, we compare the computational complexity of our proposed quasi-closed-form solution and the iterative-form solution. For our proposed quasi-closed-form solution, the complexity consists of two parts. The complexity of root finding for the $(2M+1)$-th degree polynomial equation is $O(N_{\mathrm{iter}}(2M+1))$ [7]. The complexity of matrix inversion in Equation (16) is $O(M^3)$. Therefore the total complexity if $O(M^3) + O(N_{\mathrm{iter}}(2M+1))$. For the iterative EM solution, the total complexity is $O(N_{\mathrm{iter}}M^3)$ due to the matrix inversion in each iteration step. Usually the number of iterations N_{iter} should be greater than 5 to guarantee convergence. Therefore, our proposed quasi-closed-form solution has lower complexity.

3. Simulation results

Firstly, we consider a system configuration with dimension $M = 4$ and zero-mean prior $\bar{\mathbf{h}} = \mathbf{0}$. The covariance matrices for both the channel vector $\mathbf{h}$ and noise vectors $\mathbf{z}_n$ are modeled as Toeplitz matrices, with elements decaying exponentially according to their separation distance. Specifically, the (p,q)-th elements are given by $[\mathbf{R}]_{p,q} = {\rho_h}^{|p-q|}$ for the channel covariance and $[\mathbf{\Sigma}]_{p,q} = {\rho_\sigma}^{|p-q|}$ for the noise covariance, where ρ_h and ρ_σ represent the respective correlation coefficients. While $\mathbf{\Sigma}$ can be directly estimated using the closed-form solution in Equation (5), our simulation primarily examines the estimation performance of $\mathbf{h}$ through distinct approaches. The first approach is the conventional ML estimation. The second approach is the traditional MAP estimation which assumes perfect knowledge of $\mathbf{\Sigma}$. The third approach is our proposed quasi-closed-form partial MAP estimation that operates without knowledge of $\mathbf{\Sigma}$, by solving the polynomial Equation (15) using MATLAB's root-finding algorithm. The last approach is an iterative-form EM solution. The estimation performance is quantified using the mean quare serror (MSE) metric $\mathbb{E}\{||\hat{\mathbf{h}} - \mathbf{h}||_{\mathrm{F}}^2\}$, which measures the average squared Frobenius norm distance between the estimated channel vectors and the true channel vectors.

Figure 1 and Figure 2 demonstrate the performance comparison of these estimation approaches for the uncorrelated case ($\rho_h = \rho_\sigma = 0$) and the high correlated case ($\rho_h = -\rho_\sigma = 0.9$), respectively. The ML estimation, being solely dependent on SNR as theoretically expected, exhibits identical MSE performance in both scenarios. However, both the MAP and partial MAP estimations show significantly improved estimation accuracy in the correlated case compared to their performance in the uncorrelated scenario. This performance enhancement stems from the strong correlation structures present in both $\mathbf{h}$ and $\mathbf{z}_n$, which these approaches effectively exploit to achieve better estimation precision. Notably, the performance curves for the proposed quasi-closed-form solution and the iterative-form EM solution show complete overlap in Figure 1 and Figure 2. But our proposed quasi-closed-form solution has lower complexity as analyzed in Section II-D. The results further reveal that the performance gap between the MAP estimation with perfect $\mathbf{\Sigma}$ knowledge and our proposed partial MAP estimation narrows substantially in the correlated channel case, demonstrating that the proposed approach can effectively compensate for the lack of prior knowledge of $\mathbf{\Sigma}$ by leveraging the underlying correlation structure of the system parameters.

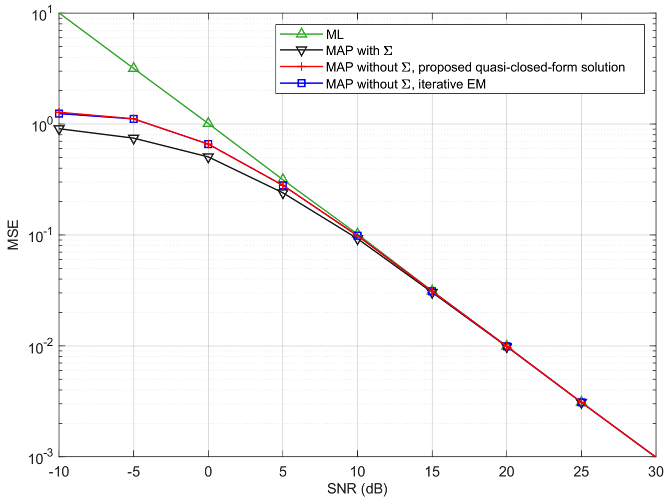

Figure 1. Channel estimation MSE *vs.* SNR under decaying exponentially channel covariance model with $M = 4$ and $\rho_h = \rho_\sigma = 0$.

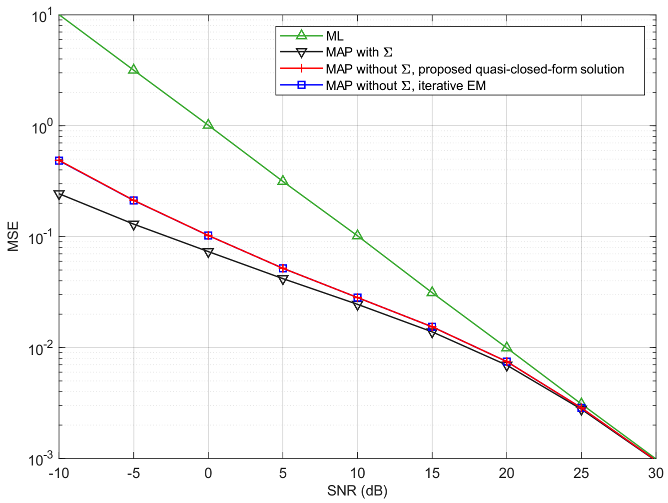

Figure 2. Channel estimation MSE *vs.* SNR under decaying exponentially channel covariance model with $M = 4$ and $\rho_h = -\rho_\sigma = 0.9$.

Figure 3 and Figure 4 illustrate the impact of training sequence length N on estimation performance at a fixed SNR of 0 dB. Under the constant training power constraint $\lambda_x^2 = 1$, both ML and MAP estimations exhibit invariant MSE performance across different values of N, consistent with theoretical expectations [8]. In contrast, the partial MAP estimation demonstrates progressively improved accuracy as N increases, eventually converging to the MAP performance. This empirical observation aligns perfectly with our theoretical analysis in Section 2.2 for the asymptotic case $N \to +\infty$. The results reveal an important trade-off between estimation accuracy and training duration: while maintaining fixed training power, enhanced estimation precision can be achieved through extended training sequences. This behavior stems from the partial MAP's ability to more effectively exploit the increasing sample size for covariance estimation, despite the constant power constraint.

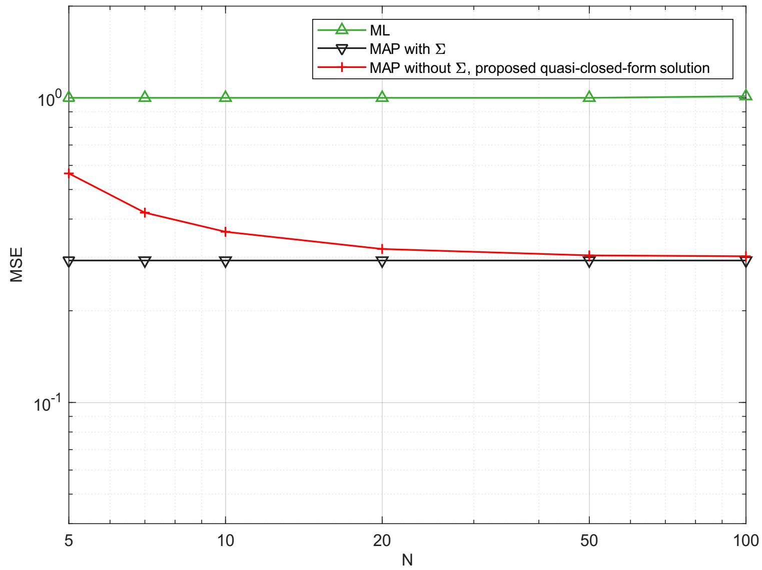

Figure 3. Channel estimation MSE *vs.* pilot length N under decaying exponentially channel covariance model with $M = 4$ and $\rho_h = \rho_\sigma = 0$.

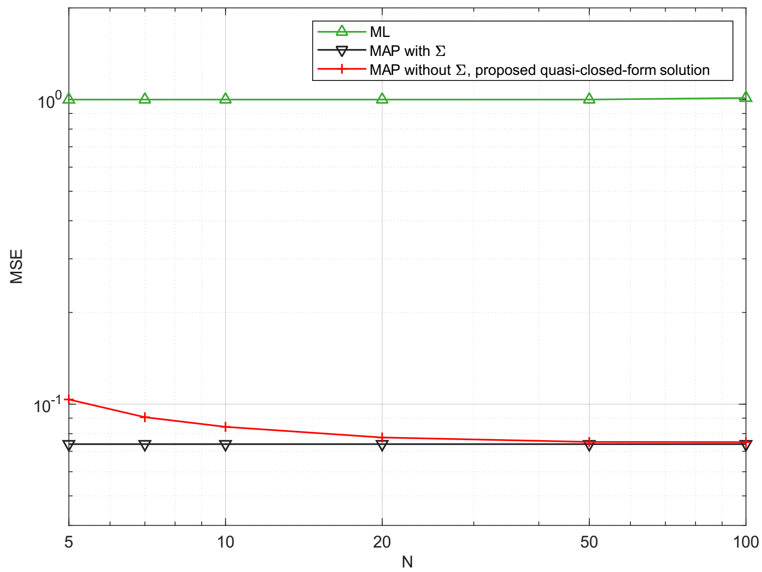

Figure 4. Channel estimation MSE *vs.* pilot length N under decaying exponentially channel covariance model with $M = 4$ and $\rho_h = -\rho_\sigma = 0.9$.

Figure 5 provides experimental validation of the asymptotic approximations developed in Section 2.2 for extreme SNR regimes. Under the uncorrelated scenario $\rho_h = \rho_\sigma = 0$, the partial MAP estimation demonstrates the predicted behaviors. As in the low-SNR regime $\mathrm{SNR} \to -\infty$, the estimation converges to the prior mean $\bar{\mathbf{h}}$, while in the high-SNR regime $\mathrm{SNR} \to +\infty$, it asymptotically approaches Equation (16) with $\nu = \frac{1}{N}$. The simulation results confirm excellent agreement between these theoretical approximations and the actual estimation performance in their respective SNR regions.

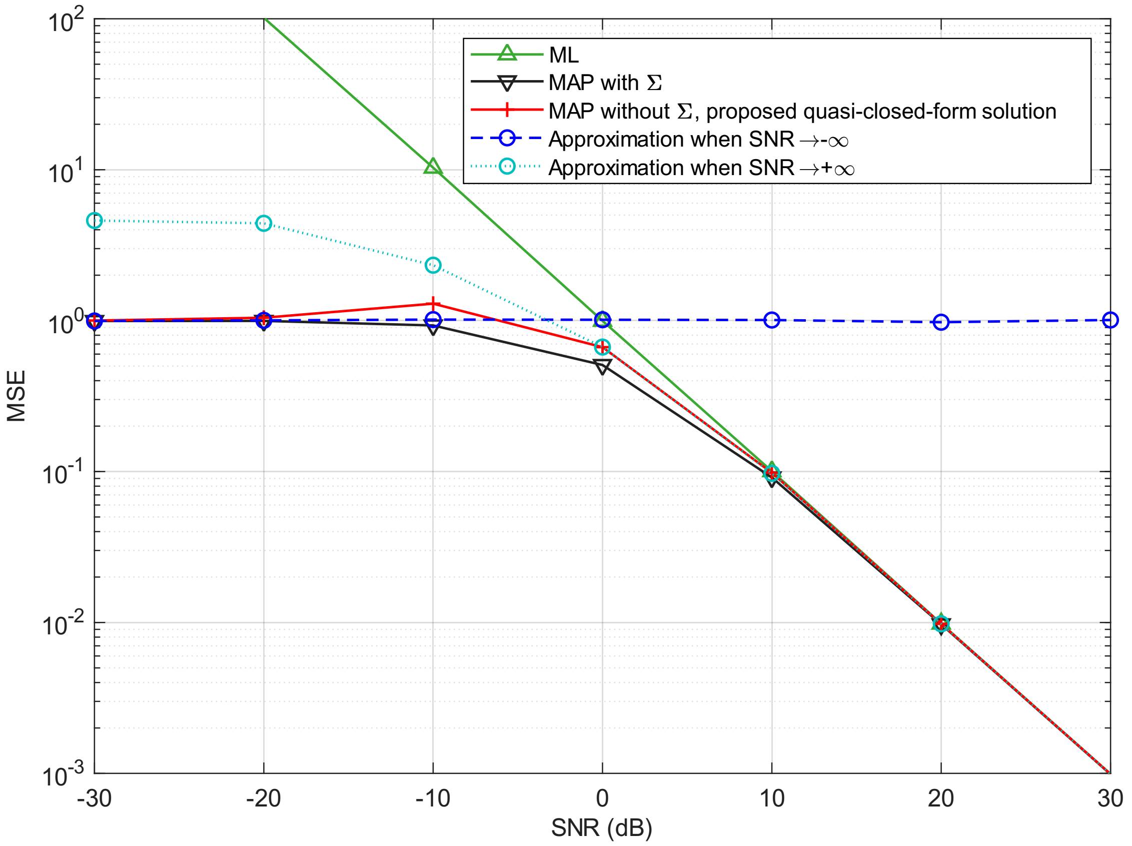

Figure 5. Channel estimation MSE approximation in extreme SNR under decaying exponentially channel covariance model with $M = 4$ and $\rho_h = \rho_\sigma = 0$.

Then, we evaluate the channel estimation performance under the 3GPP CDLC channel model. The receiver is equipped with a uniform linear array with $\lambda/2$ antenna space. The signal transmitter is at angle of arrive (AoA) 30 degree, and the interference transmitter is at AoA -30 degree. The angle spreads of both the signal channel and the interference channel are 20 degree. In Figue 6 and Figure 7, the number of receive antennas is set as $M = 4$ and $M = 8$, respectively. Besides the ML estimation, the MAP estimation with perfect knowledge of $\mathbf{\Sigma}$, the proposed quasi-closed-form partial MAP estimation without the knowledge of $\mathbf{\Sigma}$, and the iterative-form EM MAP estimation without the knowledge of $\mathbf{\Sigma}$, we also consider the MAP estimation with Gaussian-inverse-Wishart prior. It is shown that our proposed quasi-closed-form solution still has the best performance when without the knowledge of $\mathbf{\Sigma}$.

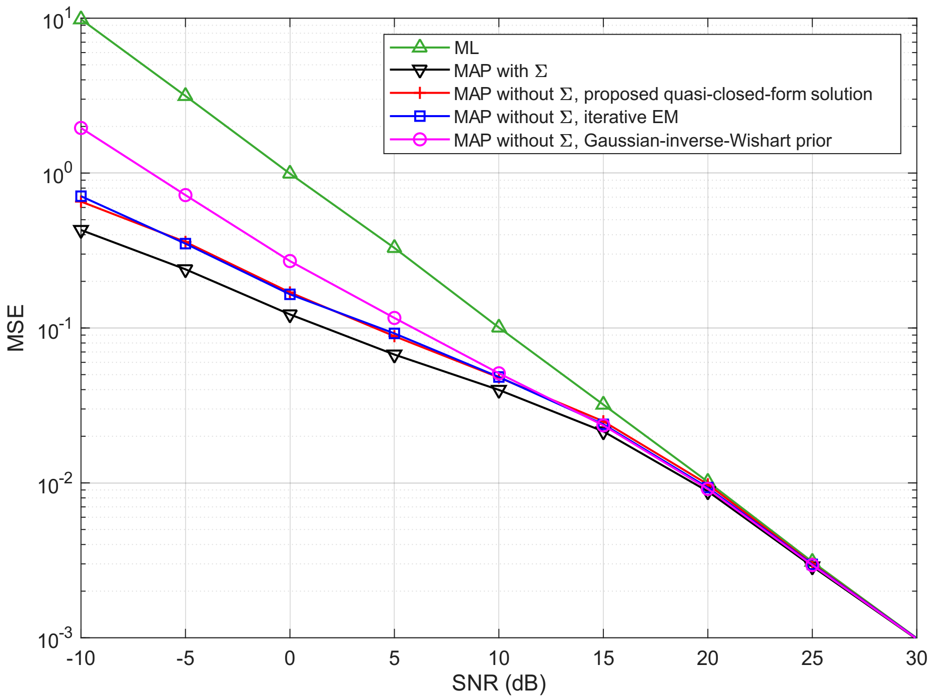

Figure 6. Channel estimation MSE *vs.* SNR under the 3GPP CDLC channel model for $M = 4$.

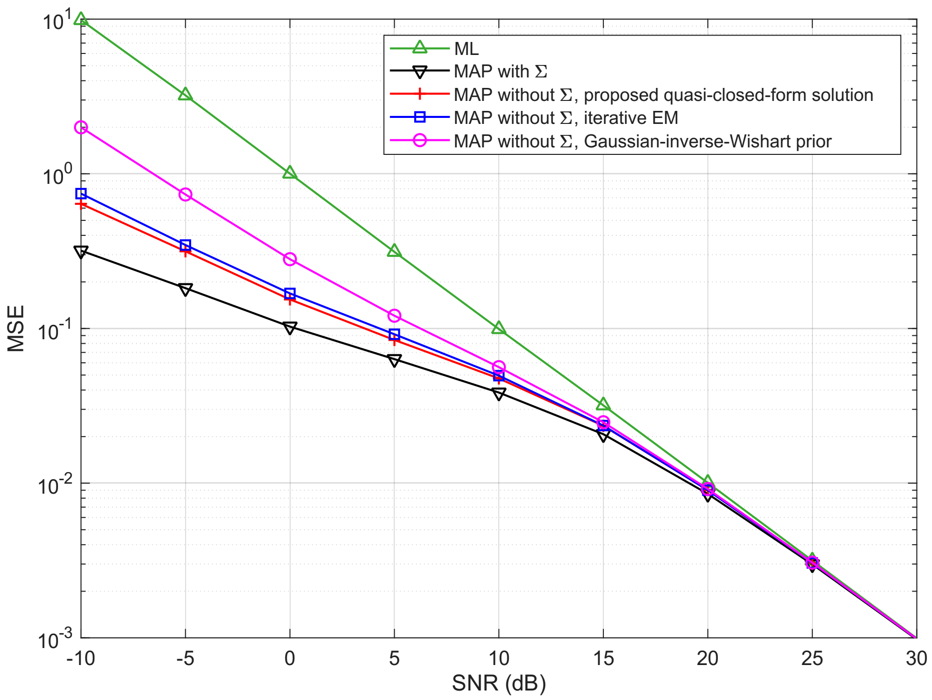

Figure 7. Channel estimation MSE *vs.* SNR under the 3GPP CDLC channel model for $M = 8$.

4. Conclusions

In our framework, we consider the scenario where only the prior distribution of $\mathbf{h}$ is known, while treating $\mathbf{\Sigma}$ as having an uninformative uniform prior. Under the joint MAP estimation criterion, we have derived a quasi-closed-form solution for the partial MAP estimations of both $\mathbf{h}$ and $\mathbf{\Sigma}$, where the key parameter ν

is obtained by solving the polynomial equation. Moreover, simplified closed-form expressions are derived for ν in several special cases where the general polynomial equation reduces to more tractable forms. This quasi-closed-form approach provides a balance between computational tractability and theoretical rigor, particularly when exact closed-form solutions are unavailable.

Appendix A. Some properties of Q

Property 1: With Equation (8a) we can obtain that

$$\begin{aligned}\mathbf{Q} &= \left(\mathbf{h}\mathbf{x}^H + \mathbf{Z}\right)\mathbf{V}_x\mathbf{E}_2\mathbf{E}_2^H\mathbf{V}_x^H\left(\mathbf{h}\mathbf{x}^H + \mathbf{Z}\right)^H \\ &= \left(\mathbf{h}\left[\lambda_x, \mathbf{0}_{1\times(N-1)}\right]\mathbf{E}_2 + \mathbf{Z}\mathbf{V}_x\mathbf{E}_2\right) \\ &\quad \cdot \left(\mathbf{h}\left[\lambda_x, \mathbf{0}_{1\times(N-1)}\right]\mathbf{E}_2 + \mathbf{Z}\mathbf{V}_x\mathbf{E}_2\right)^H \\ &= \mathbf{Z}\mathbf{V}_x\mathbf{E}_2\mathbf{E}_2^H\mathbf{V}_x^H\mathbf{Z}^H\end{aligned}$$

where $\mathbf{Z} = [\mathbf{z}_0, \cdots, \mathbf{z}_{N-1}]$. Since each column of $\mathbf{Z}$ is independent identically distributed (i.i.d.) and Gaussian, then $\mathbf{Z}\mathbf{V}_x$ follows the same distribution as $\mathbf{Z}$. Then $\mathbf{Z}\mathbf{V}_x\mathbf{E}_2$ is an $M \times (N-1)$ matrix and its every column is independent with each other. Therefore $\mathbf{Z}\mathbf{V}_x\mathbf{E}_2\mathbf{E}_2^H\mathbf{V}_x^H\mathbf{Z}^H$ will be positive definite with probability 1 as long as $N-1 \geq M$. Furthermore, since each column of $\mathbf{Z}\mathbf{V}_x\mathbf{E}_2$ is i.i.d. and follows the same distribution as the actual noise $\mathbf{z}_n$, hence $\frac{1}{N-1}\mathbf{Q}$ is an unbias estimation of $\boldsymbol{\Sigma}$.

Property 2: Considering the ML estimation of $\mathbf{h}$, *i.e.*, $\hat{\mathbf{h}}_{\mathrm{ML}} = \mathbf{Y}\mathbf{x}\left(\mathbf{x}^H\mathbf{x}\right)^{-1}$, then the ML estimation of $\boldsymbol{\Sigma}$ is showed to be

$$\begin{aligned}&\frac{1}{N}\left(\mathbf{Y} - \hat{\mathbf{h}}_{\mathrm{ML}}\mathbf{x}^H\right)\left(\mathbf{Y} - \hat{\mathbf{h}}_{\mathrm{ML}}\mathbf{x}^H\right)^H \\ &= \frac{1}{N}\left(\mathbf{Z} - \mathbf{Z}\mathbf{x}\left(\mathbf{x}^H\mathbf{x}\right)^{-1}\mathbf{x}^H\right)\left(\mathbf{Z} - \mathbf{Z}\mathbf{x}\left(\mathbf{x}^H\mathbf{x}\right)^{-1}\mathbf{x}^H\right)^H \\ &= \frac{1}{N}\mathbf{Z}\left(\mathbf{I} - \mathbf{V}_x\begin{bmatrix}1 \\ \mathbf{0}_{(N-1)\times 1}\end{bmatrix}\left[1, \mathbf{0}_{1\times(N-1)}\right]\mathbf{V}_x^H\right) \\ &\quad \cdot \left(\mathbf{I} - \mathbf{V}_x\begin{bmatrix}1 \\ \mathbf{0}_{(N-1)\times 1}\end{bmatrix}\left[1, \mathbf{0}_{1\times(N-1)}\right]\mathbf{V}_x^H\right)^H\mathbf{Z}^H \\ &= \frac{1}{N}\mathbf{Z}\mathbf{V}_x\mathbf{E}_2\mathbf{E}_2^H\mathbf{V}_x^H\mathbf{Z}^H = \frac{1}{N}\mathbf{Q}.\end{aligned}$$

Data availability statement

The data or datasets that support the findings of this study are available from the author upon reasonable request.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used generative AI tools only to improve language and readability. Specifically, the authors used DeepSeek for language polishing and readability enhancement in limited sections of the manuscript. The authors take full responsibility for the content of the manuscript.

Acknowledgments

We would like to thank the anonymous reviewers for their constructive comments and suggestions, which helped improve the quality of this paper.

References

[1] Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.
[2] Murphy KP. Conjugate Bayesian analysis of the Gaussian distribution. 2007. Available: http://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf (accessed on 3 October 2007).
[3] Dogandzic A, Mo W, Wang Z. Semi-blind SIMO flat-fading channel estimation in unknown spatially correlated noise using the EM algorithm. *IEEE Trans. Signal Process.* 2004, 52:1791–1797.
[4] Dogandzic A, Nehorai A. Generalized multivariate analysis of variance—a unified framework for signal processing in correlated noise. *IEEE Mag. Signal Process.* 2003, 2:39–54.
[5] Males F. A composite polynomial zero finding matrix algorithm. *Computers Math. Applic.* 1995, 30:37–47.
[6] Meng X, Rubin DB. Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika* 1993, 80(2):267–278.
[7] McNamee JM, Pan V. *Numerical Methods for Roots of Polynomials-Part II*. Oxford: Newnes, 2013.
[8] Kay SM. *Fundamentals of Statistical Signal Processing*. Englewood Cliffs: Prentice-Hall, 1993.