Robust Channel Estimation for RIS-Aided Communication Systems Under Unknown Mutual Coupling

Anas Talib, Yichuang Sun, and Oluyomi Simpson

***Abstract*—At sub-half-wavelength element spacing, mutual coupling degrades RIS channel estimation by several decibels when ignored, yet the passive surface cannot sense its own response and existing coupling-aware estimators require the scattering matrix a priori. This letter proposes a sparse channel estimator that remains accurate when the coupling is *entirely unknown*: translation invariance, reciprocity, and a single mirror plane confine the scattering matrix of a uniform planar array to a block-Toeplitz form with $(r+1)^2-1$ coefficients over a coupling radius r, reducing the joint problem to a few unknowns recoverable from a pilot budget far below the array size. On a 64-element RIS with 12 pilots, the estimator operates within 1.5 dB of a genie granted the exact scattering matrix, whereas a coupling-blind estimator saturates at an error floor more than 12 dB higher at high SNR. Under coupling coefficients reported for a fabricated prototype, which include interactions beyond the assumed model, the estimator retains a 4.6 dB advantage over the coupling-blind baseline.**

***Index Terms*—Channel estimation, compressed sensing, mutual coupling, orthogonal matching pursuit, reconfigurable intelligent surfaces.**

I. Introduction

RECONFIGURABLE intelligent surface (RIS) beamforming gains scale with the aperture, driving element pitches well below $\lambda/2$ toward the holographic regime [1]. At such spacing the current on one element is appreciably re-scattered by its neighbours, so the surface no longer behaves as a bank of independent reflectors [2]. Lacking radio-frequency chains, the RIS cannot observe its own response; the channel is inferred at the base station from reflected pilots [3], whose cost scales with the surface size unless structural priors are imposed [4], most commonly the angular sparsity of millimetre-wave links [5].

These sparsity-based methods rely on the diagonal reflection model, precisely what densification invalidates. Multiport network theory replaces the diagonal phase matrix by a non-diagonal response governed by an unknown scattering matrix [6], whose neglect inflates the estimation error by several decibels [7]. Coupling-aware estimators absorb the scattering matrix into the sensing operator but require it in advance, whether from full-wave simulation or chamber characterisation [8], [9] or from an analytical impedance model of the assumed element geometry [10]. Both routes fix the matrix offline, whereas the coupling drifts with manufacturing tolerances, ageing, temperature, and loading; classical compensation requires active receive chains [11] that the passive surface lacks.

(Corresponding author: Yichuang Sun.) The authors are with the School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, United Kingdom (e-mail: a.talib5@herts.ac.uk; y.sun@herts.ac.uk; o.simpson@herts.ac.uk).

This letter instead treats the scattering matrix as entirely unknown and recovers it jointly with the channel, exploiting its strong structure. The contributions are twofold.

- A structural reduction where translation invariance, reciprocity and a single mirror plane confine the scattering matrix of a holographic uniform planar array (UPA) to a block-Toeplitz form with $(r+1)^2-1$ coefficients over a coupling radius r.
- A coupling-robust sparse estimator (CRSE) that preserves channel estimation accuracy on an uncalibrated surface, treating the coupling as a nuisance parameter recovered from the same pilots as the channel, with no prior measurement and a pilot budget far below the array size.

II. Signal Model

A. Channel and Observation Model

In an uplink narrowband transmission, a single-antenna user equipment (UE) communicates with a base station (BS) exclusively through a holographic RIS, the direct link being obstructed. The BS is a UPA of $M = M_h \times M_v$ antennas at half-wavelength spacing; the RIS is a holographic UPA of $N = N_h \times N_v$ passive elements at sub-half-wavelength spacing $d_s < \lambda/2$. Both cascaded channels are block-fading and constant over the τ_p pilot slots.

The UE–RIS channel $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ follows a geometric L_k-path model reflecting the sparse multipath of millimetre-wave links [12],

$$\mathbf{h}_k = \sum_{\ell=1}^{L_k} \alpha_\ell \, \mathbf{a}_{\mathrm{RIS}}(\theta_\ell^{\mathrm{AoA}},\, \phi_\ell^{\mathrm{AoA}}), \tag{1}$$

where α_ℓ is the ℓ-th path gain, $(\theta_\ell^{\mathrm{AoA}}, \phi_\ell^{\mathrm{AoA}})$ the azimuth–elevation angle of arrival (AoA) at the RIS, and the UPA steering vector factors as

$$\mathbf{a}_{\mathrm{RIS}}(\theta,\phi) = \mathbf{a}(\cos\phi;\, N_v, \delta_s) \otimes \mathbf{a}(\sin\phi\cos\theta;\, N_h, \delta_s), \tag{2}$$

with $\mathbf{a}(u;\, n, \delta) = n^{-1/2}\big[1,\, e^{j2\pi\delta u},\, \ldots,\, e^{j2\pi(n-1)\delta u}\big]^T$, $\delta_s = d_s/\lambda < 1/2$, and $\mathbf{a}_{\mathrm{BS}}(\theta,\phi)$ built identically at the standard spacing $\delta = 1/2$. The RIS–BS channel has L_{br} paths,

$$\mathbf{H}_{br} = \sum_{i=1}^{L_{br}} \beta_i \, \mathbf{a}_{\mathrm{BS}}(\theta_i^{\mathrm{AoA}}, \phi_i^{\mathrm{AoA}})\, \mathbf{a}_{\mathrm{RIS}}^H(\theta_i^{\mathrm{AoD}}, \phi_i^{\mathrm{AoD}}) \in \mathbb{C}^{M\times N}, \tag{3}$$

with AoA at the BS and angles of departure (AoD) at the RIS. Both terminals are fixed infrastructure, so $\mathbf{H}_{br}$ varies orders of magnitude more slowly than $\mathbf{h}_k$. We therefore adopt the two-timescale framework [3], [13]. $\mathbf{H}_{br}$ is acquired in an infrequent long training phase and treated as known thereafter,

while $\mathbf{h}_k$ is re-estimated at every coherence block, which is the problem addressed here.

At sub-half-wavelength spacing the diagonal reflection model no longer holds, as the current induced on each element is re-scattered by its neighbours. In the S-parameter multiport framework [6], the effective RIS response at pilot slot t is

$$\mathbf{\Gamma}_t = \left(\mathbf{\Phi}_t^{-1} - \mathbf{S}\right)^{-1} \in \mathbb{C}^{N\times N}, \tag{4}$$

where $\mathbf{\Phi}_t = \mathrm{diag}(e^{j\varphi_{t,1}}, \ldots, e^{j\varphi_{t,N}})$ holds the intended phase shifts and $\mathbf{S}$ is the unknown $N \times N$ coupling matrix. The UE transmits a unit-power pilot while the RIS is driven through τ_p distinct configurations $\{\mathbf{\Phi}_t\}_{t=1}^{\tau_p}$; the BS observation at slot t is

$$\mathbf{y}_t = \mathbf{H}_{br}\,\mathbf{\Gamma}_t\,\mathbf{h}_k + \mathbf{n}_t \ \in \mathbb{C}^{M\times 1}, \tag{5}$$

with $\mathbf{n}_t \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_M)$ and phase shifts drawn i.i.d. as $\varphi_{t,n} \sim \mathcal{U}[0, 2\pi)$. Stacking all τ_p observations into $\mathbf{y} = [\mathbf{y}_1^T, \ldots, \mathbf{y}_{\tau_p}^T]^T \in \mathbb{C}^{M\tau_p\times 1}$,

$$\mathbf{y} = \mathbf{\Psi}\,\mathbf{h}_k + \mathbf{n}, \qquad \mathbf{\Psi} = \begin{bmatrix} \mathbf{H}_{br}\mathbf{\Gamma}_1 \\ \vdots \\ \mathbf{H}_{br}\mathbf{\Gamma}_{\tau_p} \end{bmatrix} \in \mathbb{C}^{M\tau_p\times N}. \tag{6}$$

Since $\mathrm{rank}(\mathbf{H}_{br}) \le \min(M, L_{br})$, the matrix $\mathbf{\Psi}$ has rank at most $\tau_p \min(M, L_{br})$, and least-squares recovery of $\mathbf{h}_k$ from (6) therefore requires $\tau_p \ge N/\min(M, L_{br})$ pilot slots. Moreover, $\mathbf{\Psi}$ is itself unknown because it depends on $\mathbf{S}$. Such a pilot budget is available during the long training phase, in which $\mathbf{S}$ and the cascade are jointly identifiable by least squares, but not within a coherence block, where $\tau_p \ll N$. The channel in (1), however, comprises only $L_k \ll N$ paths, and this redundancy becomes explicit in the angular domain.

With $\mathbf{F} = \mathbf{F}_{N_v} \otimes \mathbf{F}_{N_h} \in \mathbb{C}^{N\times N}$ the unitary two-dimensional DFT matrix and $\mathbf{g} = \mathbf{F}^H\mathbf{h}_k$ the beamspace channel, the geometric model (1) gives exact L_k-sparsity when path angles lie on the DFT grid,

$$\|\mathbf{g}\|_0 = L_k \ll N, \qquad \Omega \triangleq \mathrm{supp}(\mathbf{g}), \quad |\Omega| = L_k. \tag{7}$$

Substituting $\mathbf{h}_k = \mathbf{F}\mathbf{g}$ into (6) yields the beamspace observation model

$$\mathbf{y} = \mathbf{\Psi}_b\,\mathbf{g} + \mathbf{n}, \qquad \mathbf{\Psi}_b = \mathbf{\Psi}\mathbf{F} \in \mathbb{C}^{M\tau_p\times N}. \tag{8}$$

B. Coupling Matrix Parameterisation

Reciprocity of the passive surface enforces $\mathbf{S} = \mathbf{S}^T$, and since the controllable self-reflection of each element is already carried by $\mathbf{\Phi}_t$, the diagonal of $\mathbf{S}$ is zero. Setting $\mathbf{S} = \mathbf{0}$ in (4) recovers the ideal response $\mathbf{\Gamma}_t = \mathbf{\Phi}_t$. These constraints still leave $N(N-1)/2$ independent complex entries. Three modelling assumptions, each grounded in the physics of the surface, reduce this count to a number independent of N, namely the periodicity of the lattice, the mirror symmetry of the unit cell, and the finite range of the coupling.

First, index each element by its lattice coordinates $(m_h, m_v) \in \{0, \ldots, N_h - 1\} \times \{0, \ldots, N_v - 1\}$, with linear index $m = m_v N_h + m_h$. We model the aperture as a patch of an infinite periodic lattice, so every element sees the same surroundings and the coupling between m and n depends only on the displacement $(\Delta_h, \Delta_v) = (n_h - m_h,\, n_v - m_v)$ through a *coupling kernel* $f : \mathbb{Z}^2 \to \mathbb{C}$ set by the unit cell and the lattice pitch,

$$S_{mn} = f(\Delta_h, \Delta_v), \qquad f(0,0) = 0, \tag{9}$$

under which reciprocity reads $f(\Delta_h, \Delta_v) = f(-\Delta_h, -\Delta_v)$. Relation (9) is exact for interior elements; boundary elements see a truncated environment, and assigning them the interior values perturbs the corresponding entries of $\mathbf{S}$.

Second, we assume the unit cell has a single mirror plane, taken without loss of generality to contain the vertical array axis. Reflection in this plane reverses horizontal displacements only, so the kernel is even in its first argument,

$$f(\Delta_h, \Delta_v) = f(-\Delta_h, \Delta_v). \tag{10}$$

This symmetry is a property of the radiating element alone, imposing no equivalence between the two axes.

Third, we assume the coupling has finite range and truncate the kernel to the Chebyshev ball of radius r,

$$\mathcal{D}_r = \left\{(\Delta_h, \Delta_v) \neq (0,0) :\ \max(|\Delta_h|, |\Delta_v|) \le r\right\}, \tag{11}$$

with $f = 0$ outside $\mathcal{D}_r$. The integer r sets the model order, with $r = 1$ retaining the eight surrounding neighbours and $r = 2$ the first two rings.

The displacements in $\mathcal{D}_r$ sharing the pair $(|\Delta_h|, |\Delta_v|)$ form a displacement class, with indicator $\mathbf{B}_q \in \{0,1\}^{N\times N}$ whose (m, n) entry is one when the displacement from m to n lies in class q. The three properties then fix the structure of $\mathbf{S}$.

Lemma 1 (Structural reduction): Let $S_{mn} = f(n_h - m_h,\, n_v - m_v)$ as in (9), where the kernel $f : \mathbb{Z}^2 \to \mathbb{C}$ satisfies reciprocity $f(\Delta_h, \Delta_v) = f(-\Delta_h, -\Delta_v)$ and the mirror symmetry (10), and is zero outside the neighbourhood $\mathcal{D}_r$ of (11), with $r < \min(N_h, N_v)$. Then:

(i) f is constant on each of the $P = (r+1)^2 - 1$ displacement classes of $\mathcal{D}_r$, a count that depends on r alone;

(ii) $\mathbf{S} = \sum_{q=1}^{P} s_q \mathbf{B}_q$, where $s_q \in \mathbb{C}$ is the common value of f on class q and $\mathbf{B}_q$ is that class's indicator; the $\mathbf{B}_q$ are symmetric, have zero diagonal, and have pairwise disjoint supports;

(iii) $\mathbf{S}$ is block-Toeplitz with Toeplitz blocks (BTTB), comprising $N_v \times N_v$ blocks of size $N_h \times N_h$, with every block symmetric and the block sequence even in the block index.

Proof: Reciprocity acts on displacements as the point inversion $(\Delta_h, \Delta_v) \mapsto (-\Delta_h, -\Delta_v)$ and the mirror (10) as $(\Delta_h, \Delta_v) \mapsto (-\Delta_h, \Delta_v)$; their composition is the second mirror, which therefore need not be assumed, and the generated group is $\mathbb{Z}_2 \times \mathbb{Z}_2$ acting by independent sign changes. Hence f is even in each argument and constant on each orbit, and the orbits are exactly the displacement classes, each holding one representative in $\{0, \ldots, r\}^2 \setminus \{(0,0)\}$. Counting these gives (i), and (ii) follows because the orbits partition $\mathcal{D}_r$ exclude the origin and are closed under inversion. For (iii), the index map $m = m_v N_h + m_h$ places S_{mn} at position (m_h, n_h) of block (m_v, n_v), where it depends only on the differences $n_h - m_h$

and $n_v - m_v$; evenness of f in each argument gives the block symmetry and the even block sequence.

Because the group never exchanges the two coordinates, the classes $(1,0)$ and $(0,1)$ remain distinct, preserving the E-/H-plane anisotropy of planar arrays, and $r < \min(N_h, N_v)$ guarantees that every class is realised on the finite aperture. The hypotheses are consistent with fabricated hardware, where the 20×20 mmWave prototype of [9] is mirror-symmetric about a single axis only, yet its reported coupling reduces to eight parameters over the first two rings, which equals $P = (r+1)^2 - 1$ for $r = 2$.

At $r = 1$, Lemma 1 yields $P = 3$ classes, $(|\Delta_h|, |\Delta_v|) \in \{(1,0),(0,1),(1,1)\}$,

$$f(\pm 1, 0) = s_{01}, \quad f(0, \pm 1) = s_{10}, \quad f(\pm 1, \pm 1) = s_{11}, \tag{12}$$

where each subscript lists $|\Delta_v|$ then $|\Delta_h|$. The indicators admit the closed form

$$\begin{aligned} &\mathbf{B}_1 = \mathbf{I}_{N_v} \otimes \mathbf{N}_{N_h}, \qquad \mathbf{B}_2 = \mathbf{N}_{N_v} \otimes \mathbf{I}_{N_h}, \\ &\mathbf{B}_3 = \mathbf{N}_{N_v} \otimes \mathbf{N}_{N_h}, \end{aligned} \tag{13}$$

where $\mathbf{N}_n = \mathbf{T}_n + \mathbf{T}_n^T \in \{0,1\}^{n \times n}$ is the adjacency matrix of the length-n path graph and $\mathbf{T}_n$ is the $n \times n$ upper-shift matrix. Lemma 1(ii) therefore reduces the coupling matrix to

$$\mathbf{S} = s_{01}\mathbf{B}_1 + s_{10}\mathbf{B}_2 + s_{11}\mathbf{B}_3, \tag{14}$$

and we collect the unknowns in the coupling vector $\mathbf{s} = [s_1, \ldots, s_P]^T \in \mathbb{C}^P$; the development below is written for general P.

Remark 1 (Truncation bias): The radius r is a model order, not an assertion that coupling decays with distance. The coefficients reported for the prototype of [9] are not monotone in separation, their class at $2\sqrt{2}d_s$ exceeding both axial classes at $2d_s$ by at least $4.6\,\mathrm{dB}$. Whatever radius is adopted, the discarded classes occupy element pairs on which every retained indicator is zero, so no choice of $\mathbf{s}$ reproduces any part of them.

III. Coupling-Robust Sparse Channel Estimation

A. From Nonlinear Coupling to a Jointly Affine Model

The unknowns $\mathbf{g}$ and $\mathbf{s}$ enter the observation (5) asymmetrically. For a fixed scattering matrix the measurement is linear in the beamspace channel $\mathbf{g}$, whereas $\mathbf{s}$ acts through the matrix inverse in (4) and therefore enters nonlinearly. Since $\mathbf{\Phi}_t$ is unimodular, $\|\mathbf{S}\mathbf{\Phi}_t\|_2 = \|\mathbf{S}\|_2$, and for $\|\mathbf{S}\|_2 < 1$ the inverse in (4) admits the convergent Neumann series $\mathbf{\Gamma}_t = \mathbf{\Phi}_t \sum_{k \geq 0} (\mathbf{S}\mathbf{\Phi}_t)^k$. Truncating the series after its first-order term gives

$$\mathbf{\Gamma}_t \approx \mathbf{\Phi}_t + \mathbf{\Phi}_t\, \mathbf{S}\, \mathbf{\Phi}_t, \tag{15}$$

with a residual that collects the terms of order $\|\mathbf{S}\|_2^2$ and higher.

Substituting $\mathbf{S} = \sum_q s_q \mathbf{B}_q$ into the first-order term at slot t gives

$$\mathbf{H}_{br}\mathbf{\Phi}_t\mathbf{S}\mathbf{\Phi}_t\mathbf{h}_k = \sum_{q=1}^{P} s_q \underbrace{\mathbf{H}_{br}\mathbf{\Phi}_t\mathbf{B}_q\mathbf{\Phi}_t\mathbf{h}_k}_{[\mathbf{A}_t(\mathbf{h}_k)]_{:,q}} = \mathbf{A}_t(\mathbf{h}_k)\,\mathbf{s}, \tag{16}$$

Algorithm 1 Coupling-Robust Sparse Estimation (CRSE)

Require: $\mathbf{y},\ \mathbf{H}_{br},\ \{\mathbf{\Phi}_t\}_{t=1}^{\tau_p},\ \mathbf{F},\ L_k,\ r$
Ensure: channel $\hat{\mathbf{h}}_k$
1: **Blind matrix:** $\tilde{\mathbf{\Psi}}_b \leftarrow [\mathbf{H}_{br}\mathbf{\Phi}_1; \ldots; \mathbf{H}_{br}\mathbf{\Phi}_{\tau_p}]\,\mathbf{F}$
2: **Blind support:** $(\hat{\mathbf{g}}, \hat{\Omega}) \leftarrow \mathrm{OMP}(\mathbf{y}, \tilde{\mathbf{\Psi}}_b, L_k)$
3: **Coupling block:** $\hat{\mathbf{A}} \leftarrow \mathbf{A}(\mathbf{F}\hat{\mathbf{g}})$ by (16)
4: **Joint solve:** $[\hat{\mathbf{g}}_{\hat{\Omega}}; \hat{\mathbf{s}}] \leftarrow [\tilde{\mathbf{\Psi}}_b^{(\hat{\Omega})} \mid \hat{\mathbf{A}}]^{\dagger}\mathbf{y}$
5: **Correction:** $\hat{\mathbf{S}} \leftarrow \sum_{q=1}^{P} \hat{s}_q \mathbf{B}_q$; build $\hat{\mathbf{\Psi}}_b$ via exact (4)
6: **Refined support:** $(\hat{\mathbf{g}}, \hat{\Omega}') \leftarrow \mathrm{OMP}(\mathbf{y}, \hat{\mathbf{\Psi}}_b, L_k)$
7: **Channel solve:** $\hat{\mathbf{g}}_{\hat{\Omega}'} \leftarrow (\hat{\mathbf{\Psi}}_b^{(\hat{\Omega}')})^{\dagger}\mathbf{y},\ \ \hat{\mathbf{g}}_{\hat{\Omega}'^c} \leftarrow \mathbf{0}$
8: **return** $\hat{\mathbf{h}}_k = \mathbf{F}\hat{\mathbf{g}}$

where the q-th column of $\mathbf{A}_t(\mathbf{h}_k) \in \mathbb{C}^{M \times P}$ is the pilot response of the q-th displacement class. Inserting (15) and (16) into (5) and stacking the τ_p slots then turns the bilinear estimation problem into one *jointly affine* in $(\mathbf{g}, \mathbf{s})$,

$$\mathbf{y} \approx \tilde{\mathbf{\Psi}}_b\, \mathbf{g} + \mathbf{A}(\mathbf{F}\mathbf{g})\, \mathbf{s} + \mathbf{n}, \tag{17}$$

with $\mathbf{A}(\mathbf{h}_k) = [\mathbf{A}_1(\mathbf{h}_k), \ldots, \mathbf{A}_{\tau_p}(\mathbf{h}_k)] \in \mathbb{C}^{M\tau_p \times P}$ and coupling-blind sensing matrix

$$\tilde{\mathbf{\Psi}}_b = \begin{bmatrix} \mathbf{H}_{br}\mathbf{\Phi}_1 \\ \vdots \\ \mathbf{H}_{br}\mathbf{\Phi}_{\tau_p} \end{bmatrix} \mathbf{F} \in \mathbb{C}^{M\tau_p \times N}. \tag{18}$$

The first term of (17) propagates the channel through the ideal diagonal response; the second is the distortion that the diagonal model discards, now reduced to an additive perturbation with known structure and P unknown gains.

B. Reduced Joint Estimation and Identifiability

Although affine, model (17) is not yet estimable, as $\mathbf{g}$ has N coefficients and $\mathbf{A}(\mathbf{F}\mathbf{g})$ depends on the unknown channel. The sparsity prior (7) resolves the former by confining $\mathbf{g}$ to its support Ω, so that only the L_k columns $\tilde{\mathbf{\Psi}}_b^{(\Omega)}$ are retained, and freezing the argument of $\mathbf{A}(\cdot)$ at a provisional estimate $\hat{\mathbf{h}}_k$ resolves the latter by rendering the coupling block known. These two reductions yield the overdetermined linear model

$$\begin{aligned} &\mathbf{y} \approx \mathbf{J}\,\mathbf{x} + \mathbf{n}, \\ &\mathbf{J} = \left[\, \tilde{\mathbf{\Psi}}_b^{(\Omega)} \;\middle|\; \mathbf{A}(\hat{\mathbf{h}}_k) \,\right] \in \mathbb{C}^{M\tau_p \times (L_k + P)}, \\ &\mathbf{x} = \begin{bmatrix} \mathbf{g}_\Omega \\ \mathbf{s} \end{bmatrix} \in \mathbb{C}^{L_k + P}, \end{aligned} \tag{19}$$

with $M\tau_p$ equations for $L_k + P$ unknowns and the closed-form solution $\hat{\mathbf{x}} = \mathbf{J}^{\dagger}\mathbf{y}$, partitioned into $(\hat{\mathbf{g}}_\Omega, \hat{\mathbf{s}})$. By the rank argument of Sec. II, every M-row block of $\mathbf{J}$ lies in the range of $\mathbf{H}_{br}$, so $\operatorname{rank}\mathbf{J} \leq \min(M, L_{br})\,\tau_p$ and joint identifiability of $(\mathbf{g}_\Omega, \mathbf{s})$ from (19) requires

$$\min(M, L_{br})\,\tau_p \;\geq\; L_k + P. \tag{20}$$

This condition is met generically by i.i.d. random phase configurations, and the structure exploited above has reduced the unknown count from N to $L_k + P$, with each additional coupling class consuming one degree of freedom of the same pilot budget.

C. Two-Stage Coupling-Aware Estimation

The provisional estimate $\hat{\mathbf{h}}_k$ at which $\mathbf{A}$ is frozen remains to be specified, yet acquiring it appears to require a dictionary that is already coupling-compensated. Algorithm 1 resolves this circularity in two stages.

Acquisition (lines 1–4). Orthogonal matching pursuit (OMP) [14] on the coupling-blind dictionary $\tilde{\mathbf{\Psi}}_b$ returns a support $\hat{\Omega}$ and channel $\hat{\mathbf{g}}$; both are biased by the unmodelled coupling, yet suffice to build $\mathbf{A}(\mathbf{F}\hat{\mathbf{g}})$ and to restrict (19) to $\hat{\Omega}$, whose solution yields $\hat{\mathbf{s}}$.

Refinement (lines 5–7). The coupling matrix is reassembled as $\hat{\mathbf{S}} = \sum_{q=1}^{P} \hat{s}_q \mathbf{B}_q$ and the dictionary rebuilt from the exact response (4), so that a second OMP pass and a least-squares solve on the refined support $\hat{\Omega}'$ give the channel estimate. Because the rebuilt dictionary uses (4) and not the truncation (15), the linearisation error affects $\hat{\mathbf{s}}$ alone, not the operator from which the channel is recovered.

An erroneous $\hat{\Omega}$ degrades the estimate only mildly. Since $\mathbf{A}(\hat{\mathbf{h}}_k)$ is dominated by the strongest paths, a missed weak atom perturbs the column space of $\mathbf{J}$ little, while line 6 re-runs OMP unconstrained by $\hat{\Omega}$, leaving it recoverable once the coupling is compensated.

Remark 2 (Model order): Every step above holds for an arbitrary radius r, with only the width $P = (r+1)^2 - 1$ of $\mathbf{A}$ in (16) changing. By (20) the choice of r is a bias–variance trade-off against the pilot budget; the orders $r = 1$ ($P = 3$) and $r = 2$ ($P = 8$) are evaluated in Sec. IV.

Complexity. The estimator is non-iterative: two OMP passes and two least-squares solves, dominated by assembling the two sensing matrices at $\mathcal{O}(M\tau_p N^2)$. Rebuilding from (4) does not raise this order, since $(\mathbf{\Phi}_t^{-1} - \hat{\mathbf{S}})$ is banded BTTB and $\mathbf{H}_{br}\hat{\mathbf{\Gamma}}_t$ follows from a banded factorisation without forming any inverse.

IV. Simulation Results

Setup. We consider an 8×8 RIS ($N = 64$) and a 2×2 BS ($M = 4$), with $L_k = 3$ UE–RIS and $L_{br} = 5$ RIS–BS paths; the UE–RIS path angles lie on the DFT grid, while the RIS–BS angles are drawn uniformly over $[-\pi/2, \pi/2]$ and $[0, \pi]$, gains are i.i.d. $\mathcal{CN}(0, 1/L_k)$ and $\mathcal{CN}(0, 1/L_{br})$, and the phases $\varphi_{t,n}$ are i.i.d. $\mathcal{U}[0, 2\pi)$. The RIS pitch is that of the fabricated prototype of [9], $d_s = 0.35\lambda$, so the coefficients adopted in Fig. 2 apply at the element separations they were fitted for. The coupling follows (14) with nominal magnitudes $|s_{01}| = |s_{10}| = 0.10$, $|s_{11}| = 0.08$ and an independent phase per coefficient redrawn each trial, so $\mathbf{S}$ is unknown to the estimator; the resulting $\|\mathbf{S}\|_2 \leq 0.66$ lies within the convergence range of (15), and under this draw the truncation deviates from the exact response by a median 12.5% in relative spectral norm. With $\sigma^2 = \|\mathbf{\Psi}\mathbf{h}_k\|^2/(M\tau_p\,\mathrm{SNR})$, all results report the median NMSE $\|\hat{\mathbf{h}}_k - \mathbf{h}_k\|^2/\|\mathbf{h}_k\|^2$ over 300 trials. All schemes share the same OMP engine and the budget $\tau_p = 12$, differing only in the sensing matrix. *Genie* runs OMP on the $\mathbf{\Psi}_b$ of (8) built from the true $\mathbf{S}$ and serves as the performance bound, *Proposed* is CRSE of Sec. III-C, and *Blind* runs OMP on the coupling-blind $\tilde{\mathbf{\Psi}}_b$.

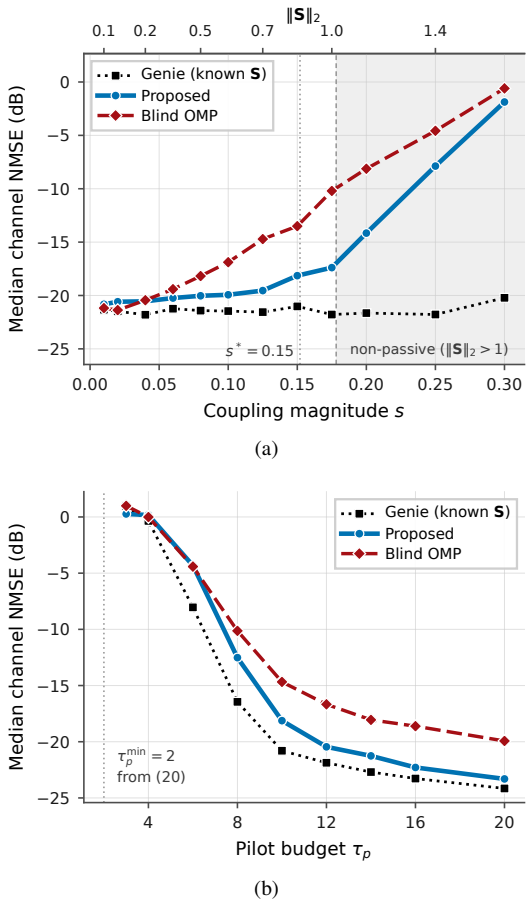


Fig. 1. Median NMSE at SNR $= 10$ dB. (a) Versus coupling magnitude ($|s_{01}| = |s_{10}| = s$, $|s_{11}| = 0.8s$) at $\tau_p = 12$. (b) Versus pilot budget τ_p at the nominal coupling.

Coupling strength (Fig. 1(a)). The Genie is flat at ≈ -21.5 dB across the sweep, so the degradation of the other schemes reflects missing coupling knowledge. The proposed estimator stays within 1.5 dB of this bound at the nominal $s = 0.10$, and its lead over Blind OMP grows to 6.0 dB at $s = 0.20$. The gap to the Genie first exceeds 3 dB at $s^* = 0.15$, while $\|\mathbf{S}\|_2$ reaches the passivity limit at $s = 0.18$; near-genie accuracy therefore extends over most of the admissible range.

Pilot overhead (Fig. 1(b)). Condition (20) is necessary but not sufficient. With $\min(M, L_{br}) = 4$ it admits $\tau_p \geq 2$, yet no estimator attains useful accuracy below $\tau_p = 6$. Between $\tau_p = 8$ and $\tau_p = 12$ the margin over Blind grows from 2.4 to 3.8 dB. Over the same interval the gap to the Genie narrows from 3.9 to 1.4 dB, so $\tau_p = 12$, used throughout, sits at the saturation point of the curve.

Prototype-derived kernels (Fig. 2). The preceding figures draw $\mathbf{S}$ from the estimator's own model (14), and a test whose ground truth obeys the assumed structure cannot expose the limits of that structure. Fig. 2 therefore assembles the ground truth from the eight coupling coefficients, magnitudes and phases alike, reported for the fabricated 20×20 prototype of [9]. These coefficients qualify as a reference because their fit is validated out of sample, on full-wave patterns and network-analyser measurements withheld from it; their absolute scale exceeds the passivity limit and is not transferred, each kernel

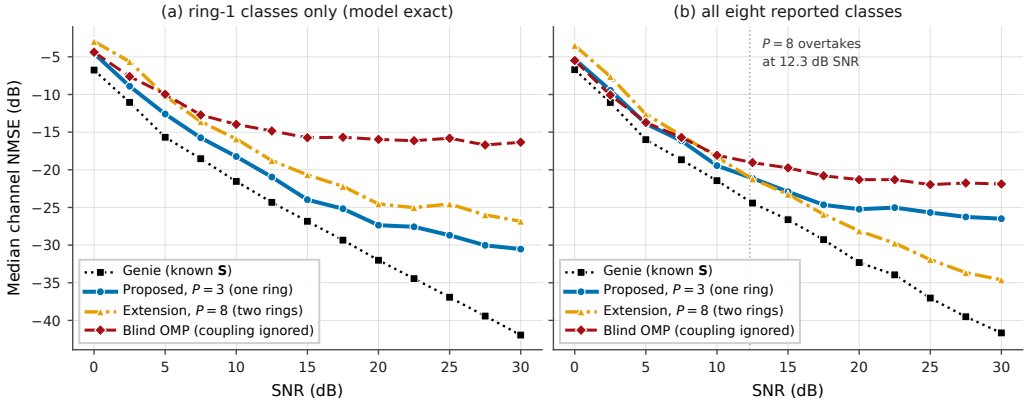


Fig. 2. Median NMSE vs. SNR at $\tau_p = 12$, prototype-derived kernel of [9]. (a) Ring-1 classes. (b) All eight classes.

being rescaled to the nominal $\|\mathbf{S}\|_2$. Fig. 2(a) retains the ring-1 classes alone, for which the model (12) is exact, whereas Fig. 2(b) uses all eight classes over two rings, extending beyond the support retained at $r = 1$. The estimator is applied unchanged to both, so the shape of the kernel is the only variable.

Under the matched kernel of Fig. 2(a) the schemes coincide in the noise-limited regime; beyond 12.5 dB SNR Blind OMP saturates on coupling bias while the proposed estimator improves monotonically. It trails the Genie by 3.3 dB at 10 dB SNR and leads Blind by 12.9 dB at 25 dB. Coupling mismatch is therefore the dominant high-SNR error source.

Under the two-ring kernel of Fig. 2(b) the margin over Blind widens from 1.4 dB at 10 dB SNR to 4.6 dB at 30 dB, where the truncation bias of Remark 1 sets a floor at -26.5 dB against a coupling-blind -21.9 dB. The gap to the Genie at 10 dB is 2.0 dB, whereas below it the two schemes lie within 0.7 dB of each other, Blind marginally ahead below 5 dB. The $r = 2$ extension ($P = 8$) covers the full support of the reported kernel, but adds five unknowns to (19) and therefore trades reduced bias for increased estimation variance under the budget of (20). It overtakes $P = 3$ only above 12.3 dB SNR at $\tau_p = 12$.

V. Conclusion

This letter treated channel estimation for a passive RIS whose mutual coupling is unknown. Rather than calibrating the scattering matrix offline, the proposed estimator reduces it to a few displacement-class coefficients and recovers them jointly with the sparse channel, so that the pilot budget grows with the number of propagation paths and coupling classes rather than with the size of the surface. The procedure is non-iterative and trails a genie-aided bound by 1.5 dB at 12 pilots on a 64-element surface. Applied unchanged to a kernel whose shape is that of a fabricated prototype and whose support exceeds the modelled radius, it keeps its margin over the coupling-blind baseline at the operating point, at the cost of a bias floor set by the discarded classes. Off-grid path angles and data-driven selection of the truncation radius remain open.

References

[1] Y. Liu *et al.*, "Densifying MIMO: Channel modeling, physical constraints, and performance evaluation for holographic communications," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 6, pp. 1504–1518, Jun. 2024.

[2] G. Gradoni and M. Di Renzo, "End-to-end mutual coupling aware communication model for reconfigurable intelligent surfaces: An electromagnetic-compliant approach based on mutual impedances," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 938–942, May 2021.

[3] C. Hu *et al.*, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.

[4] A. Talib, Y. Sun, and O. Simpson, "Pilot-efficient channel estimation for RIS-aided wireless communication systems: A systematization," *IEEE Open J. Commun. Soc.*, vol. 7, pp. 7957–7979, 2026.

[5] X. Bian *et al.*, "A sparse-based channel estimation scheme for XL-RIS-assisted wireless communications with low pilot overheads," *IEEE Trans. Commun.*, vol. 74, pp. 367–380, 2026.

[6] S. Shen, B. Clerckx, and R. D. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1229–1243, Feb. 2022.

[7] P. Zheng, X. Ma, and T. Y. Al-Naffouri, "On the impact of mutual coupling on RIS-assisted channel estimation," *IEEE Wireless Commun. Lett.*, vol. 13, no. 5, pp. 1275–1279, May 2024.

[8] P. Zheng *et al.*, "Mutual coupling-aware channel estimation and beamforming for RIS-assisted communications," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 4374–4390, 2026.

[9] P. Zheng *et al.*, "Mutual coupling in RIS-aided communication: Model training and experimental validation," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17174–17188, Nov. 2024.

[10] T. Qiu *et al.*, "Mutual coupling aware channel estimation for RIS-aided multi-user mmWave systems," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 15803–15817, 2026.

[11] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 273–284, Mar. 1991.

[12] A. Alkhateeb *et al.*, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[13] K. She and Y. Sun, "An optimal scheme of two-timescale channel estimation framework for RIS-assisted wireless communications," *Wireless Pers. Commun.*, vol. 142, no. 3–4, pp. 501–526, Jun. 2025.

[14] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.