AO-MM:

The objective function of $\mathsf{OP4}$ is given as

$$\frac{\boldsymbol{\varphi}}{\beta_e^2+\rho_e^2\|\mathbf{\Theta}\mathbf{H}\mathbf{f}\|^2}=R\left[\frac{\|\left(\rho_b\widehat{\mathbf{g}}_b\mathbf{\Theta}\mathbf{H}+\beta_b\mathbf{h}_b\right)\mathbf{f}\|^2}{\|\sqrt{1-\rho_b^2}\widetilde{\mathbf{g}}_b\mathbf{\Theta}\mathbf{H}\mathbf{f}\|^2+\frac{\sigma_b^2}{P_A}}+\mathcal{B}_1\right]. \tag{1}$$

where, $R=\frac{\alpha}{\beta_e^2+\rho_e^2\|\mathbf{\Theta}\mathbf{H}\mathbf{f}\|^2}$. Now, let us perform optimization over $\mathbf{\Theta}$ by fixing $\mathbf{f}$. With (1), $\mathsf{OP4}$ can be transformed as

$$\begin{aligned}\mathsf{P1}\colon &\max_{\mathbf{\Theta}}\ \frac{\|\left(\rho_b\widehat{\mathbf{g}}_b\mathbf{\Theta}\mathbf{H}+\beta_b\mathbf{h}_b\right)\mathbf{f}\|^2}{\|\sqrt{1-\rho_b^2}\widetilde{\mathbf{g}}_b\mathbf{\Theta}\mathbf{H}\mathbf{f}\|^2+\frac{\sigma_b^2}{P_A}},\\ &\text{s.t. } 12.\end{aligned} \tag{2}$$

Using (37) and (38), we have

$$\mathsf{P2}\colon \max_{\boldsymbol{\upsilon},|\upsilon_m|=1}\ \frac{\boldsymbol{\upsilon}^H\mathbf{y}_1\boldsymbol{\upsilon}+|\alpha_A|^2+2\rho_b\Re\{\mathbf{e}_1^H\boldsymbol{\upsilon}^H\alpha_A^*\}}{\boldsymbol{\upsilon}^H\mathbf{y}_2\boldsymbol{\upsilon}+\frac{\sigma_b^2}{P_A}}, \tag{3}$$

where $\mathbf{y}_1=\rho_b^2\mathbf{e}_1^H\mathbf{e}_1$, $\mathbf{y}_2=(1-\rho_b^2)\mathbf{e}_2^H\mathbf{e}_2$, and $\alpha_A=\beta_b\mathbf{h}_b\mathbf{f}$. Rewrite (3), we have

$$\mathsf{P3}\colon \min_{\boldsymbol{\upsilon},|\upsilon_m|=1}\ \frac{\boldsymbol{\upsilon}^H\mathbf{y}_2\boldsymbol{\upsilon}+\frac{\sigma_b^2}{P_A}}{\boldsymbol{\upsilon}^H\mathbf{y}_1\boldsymbol{\upsilon}+|\alpha_A|^2+2\rho_b\Re\{\mathbf{e}_1^H\boldsymbol{\upsilon}^H\alpha_A^*\}}. \tag{4}$$

This problem belongs to fractional programming. Following [33], we consider the corresponding parametric program:

$$\mathsf{P4}\colon \min_{\boldsymbol{\upsilon},|\upsilon_m|=1}\ \boldsymbol{\upsilon}^H\boldsymbol{\theta}\boldsymbol{\upsilon}+\frac{\sigma_b^2}{P_A}-\mu|\alpha_A|^2-2\mu\rho_b\Re\{\mathbf{e}_1^H\boldsymbol{\upsilon}^H\alpha_A^*\}, \tag{5}$$

where $\mu>0$ is an introduced parameter and $\boldsymbol{\theta}=\mathbf{y}_2-\mu\rho_b^2\mathbf{y}_1$. Here, we adopt MM algorithm to solve $\mathsf{P4}$ due to its closed-form solution at each iteration [9],[11],[18].[30]. At iteration t, for any feasible point $\boldsymbol{\upsilon}(t)$,

$$\begin{aligned}&\boldsymbol{\upsilon}^H\boldsymbol{\theta}\boldsymbol{\upsilon}+\frac{\sigma_b^2}{P_A}-\mu|\alpha_A|^2-2\Re\{\mu\rho_b\mathbf{e}_1\boldsymbol{\upsilon}^H\alpha_A^*\}\\ &\le\lambda_{\max}\left(\boldsymbol{\theta}\right)\|\boldsymbol{\upsilon}\|^2-\Re\{\boldsymbol{\upsilon}^H\boldsymbol{\varphi}\}+q_1,\end{aligned} \tag{6}$$

where $\boldsymbol{\varphi}=(\lambda_{\max}(\boldsymbol{\theta})\mathbf{I}-\boldsymbol{\theta})\boldsymbol{\upsilon}(t)+\mu\rho_b\mathbf{e}_1\alpha_A^*$, $q_1=(\lambda_{\max}(\boldsymbol{\theta})\mathbf{I}-\boldsymbol{\theta})\boldsymbol{\upsilon}+\frac{\sigma_b^2}{P_A}-\mu|\alpha_A|^2$. The simplified optimization problem becomes

$$\begin{aligned}\mathsf{P5}\colon &\max_{\boldsymbol{\upsilon}}\ \Re\{\boldsymbol{\upsilon}^H\boldsymbol{\varphi}\},\\ &\text{s.t. } |\upsilon_m|=1,\forall\quad m=1,\ldots,M.\end{aligned} \tag{7}$$

The optimal solution of $\mathsf{P5}$ at iteration t is

$$\boldsymbol{\upsilon}(t+1)=\left[e^{\jmath\arg(\boldsymbol{\varphi}_1)},\ldots,e^{\jmath\arg(\boldsymbol{\varphi}_M)}\right]. \tag{8}$$

Algorithm A MM-Based Algorithm for Phase-shift Matrix

Input: $\mathbf{H}, \widehat{\mathbf{g}}_b, \widetilde{\mathbf{g}}_b, \mathbf{h}_b, \mathbf{f}, \rho_b, \beta_b, \sigma_b^2, P_A, \epsilon, T_{\max}$
Initialize: Random unit-modulus vector $\boldsymbol{\upsilon}^{(0)}$, set $\mu_{\min}, \mu_{\max}, \mu \leftarrow (\mu_{\min} + \mu_{\max})/2$
repeat
  Compute $\boldsymbol{\theta} = \mathbf{y}_2 - \mu \mathbf{y}_1$
  $t \leftarrow 0$
  repeat
    Update $\boldsymbol{\varphi} = (\lambda_{\max}\mathbf{I} - \boldsymbol{\theta})\boldsymbol{\upsilon}^{(t)} + \mu\rho_b \mathbf{e}_1 \alpha_A^*$, where

$$\lambda_{\max} = \lambda_{\max}(\boldsymbol{\theta});$$

    Update each element: $\upsilon_m^{(t+1)} = e^{j \arg(\varphi_m)}, \quad \forall m = 1, \ldots, M$
    $t \leftarrow t + 1$
  until $\|\boldsymbol{\upsilon}^{(t)} - \boldsymbol{\upsilon}^{(t-1)}\| < \epsilon$ or $t \geq T_{\max}$
  Compute numerator and denominator:

$$N(\mu) = \boldsymbol{\upsilon}^H \mathbf{y}_1 \boldsymbol{\upsilon} + |\alpha_A|^2 + 2\rho_b \Re\{\mathbf{e}_1^H \boldsymbol{\upsilon}^H \alpha_A^*\}$$
$$D(\mu) = \boldsymbol{\upsilon}^H \mathbf{y}_2 \boldsymbol{\upsilon} + \frac{\sigma_b^2}{P_A}$$

  if $N(\mu) - \mu D(\mu) \geq 0$ **then**
    $\mu_{\min} \leftarrow \mu$
  else
    $\mu_{\max} \leftarrow \mu$
  end if
  Update $\mu \leftarrow (\mu_{\min} + \mu_{\max})/2$
until $|\mu_{\max} - \mu_{\min}| < \epsilon_\mu$ or max iterations reached
Output: Optimal $\boldsymbol{\upsilon}$

According to [11], the optimal solution $\boldsymbol{\upsilon}$ can be obtained via an Algorithm A with a computational complexity of $\mathcal{O}\left(M^3 + T_{\max} M^2\right)$, where $T_{\max}$ denotes the number of iterations required for convergence [11, 34].

Algorithm B Alternating Algorithm

Require: $P_A, \rho_b, \rho_e, R_s, \sigma_b^2, \sigma_e^2, \mathbf{H}, \mathbf{h}_b, \mathbf{h}_e, N_A, M, \delta$
Ensure: $\boldsymbol{\Theta}^*, w^*$
Initialize $\mathbf{f}_o$ as a random unitary vector of size $N_A \times 1$.
Set iteration counter $i = 1$ and maximum iteration limit $i_{\max}$.
repeat
  Obtain $\mathbf{f}_i^*$ via generalized Rayleigh quotient.
  Obtain $\boldsymbol{\upsilon}^*$ using MM algorithm
  Compute $P_{\text{out}}(i)$.
  Update iteration counter: $i \leftarrow i + 1$.
until $(i < i_{\max})$ and $|P_{\text{out}}(i) - P_{\text{out}}(i-1)| \leq \delta$
Set $\boldsymbol{\upsilon}^* \leftarrow \boldsymbol{\upsilon}_i^*$, $\mathbf{w}^* \leftarrow \mathbf{w}_i^*$.
Compute $\boldsymbol{\Theta}^* \leftarrow \text{diag}(\boldsymbol{\upsilon}^*)$.
return $\boldsymbol{\Theta}^*, \mathbf{w}^*$.

Algorithm B shows alternating algorithm for phase-shifter matrix and beamforming vector. The MM method is used to find optimal phase-shift matrix and generalized Rayleigh to find optimal beamforming vector.