APPENDIX

A. Proof of Theorem 1

To prove theorem 1, we use results from [43] to show the asymptotic convergence of SAA expressions to the ergodic quantities. We start by showing that some necessary technical assumptions are fulfilled in our problem. We focus on scenarios for which the problem is feasible, i.e., without loss of generality we focus on the scenario for which the C-RAN is able to allocate resources such that the QoS requirements of all users can be satisfied, i.e., the feasible set is compact and not empty. Let us assume that the noise variance is non-zero, i.e., $\sigma > 0$ and the transmit power is finite, i.e., $P^{\text{Tr}}(\mathbf{w}) < \infty \quad \forall n \in \mathcal{N}$. Note that such assumptions are quite natural as the transmit power in practical systems is physically limited to a certain value. Then the SINR expressions in (9) and (10) are finite, given that the channel realizations of each user are bounded. That is, we have the following: $\gamma_k^p(\mathbf{w}) < \infty \quad \forall k \in \mathcal{K}$, $\gamma_{k,i}^c(\mathbf{w}) < \infty \quad \forall i \in \mathcal{M}_k, \forall k \in \mathcal{K}$, where the dependency of SINR expressions on the beamforming vectors is made explicit. Then, the limit of average spectral efficiency when the sample size tends to infinity exists. From ergodicity assumption of the channel distribution, and the law of large numbers the following holds [43, Theorem 7.48]

$$\sup_{\mathbf{w}\in\mathcal{W}} \left| \frac{1}{M} \log_2(1+\gamma_k^p(\mathbf{w})) - \mathbb{E}_{\mathbf{h}} \left\{ \log_2(1+\gamma_k^p(\mathbf{w})) \right\} \right| \to 0, \quad \text{as } M \to \infty \tag{1}$$

$$\sup_{\mathbf{w}\in\mathcal{W}} \left| \frac{1}{M} \log_2(1+\gamma_{k,i}^c(\mathbf{w})) - \mathbb{E}_{\mathbf{h}} \left\{ \log_2(1+\gamma_{k,i}^c(\mathbf{w})) \right\} \right| \to 0, \quad \text{as } M \to \infty \tag{2}$$

where $\mathcal{W}$ is the feasible set for the beamforming vectors, determined with $P^{\text{Tr}}(\mathbf{w})$ and given by the objective function of problem (P_2) and it is convex and compact given the aforementioned assumptions are full filled. That is (1) and (2) indicate that the SAA estimate of the rates converges to the ergodic rate uniformly on the compact set $\mathcal{W}$ with probability one as the sample size goes to infinity. By [43, Theorem 5.3], we conclude that the set of optimal solutions of problem $\text{P}_3(\text{M})$ converges uniformly to the the set of optimal solutions of problem P_2 with probability one as $M \to \infty$, which completes the proof.

B. Proof of Theorem 2

The proof follows similar steps as used in [48, Theorem 2]. First, we note that the problem $\text{P}_3(\text{M})$ has a convex objective function, but the feasible set is non convex as the functions in constraints (48c) and (48d) are non convex. In each iteration r of Algorithm 1, we solve the following convex optimization problem P_4. Let us define the objective function of problem P_4

as $Q(\mathbf{w})$. Moreover, we define $\mathbf{u}^r = \Upsilon(\mathbf{w}^{r-1})$ and $\boldsymbol{\rho}^r = \Phi(\mathbf{w}^{r-1})$, where the mappings $\Upsilon(\cdot)$ and $\Phi(\cdot)$ are given in (50) and (51), respectively. Note that all the auxiliary coefficients in (52)-(57) are already defined in terms of $\mathbf{u}^r$ and $\boldsymbol{\rho}^r$. We note that the sequence $\{Q(\mathbf{w}^r)\}_{r=0}^{\infty}$ is monotonically decreasing after each iteration and converges. This is because it is a convex function in the variables, let $\bar{Q}$ denotes the limit of this function. Due to the compactness of the convex feasible set defined by the constraints (16c), (16d), the iterates $\{\mathbf{w}^r\}_{r=0}^{\infty}$ must have a cluster point, denoted as $\bar{\mathbf{w}}$. That is, it exists a subsequence $\{\mathbf{w}^{r_1}\}_{r_1=J}^{\infty}$ for some $J > 0$ which converges to $\bar{\mathbf{w}}$. Since the functions $\Upsilon(\cdot)$ and $\Phi(\cdot)$ are continuous we have the following result

$$\lim_{r_1 \to \infty} \{\mathbf{w}^{r_1}, \mathbf{u}^{r_1}, \boldsymbol{\rho}^{r_1}\} = \{\bar{\mathbf{w}}, \Upsilon(\bar{\mathbf{w}}), \Phi(\bar{\mathbf{w}})\} \triangleq \{\bar{\mathbf{w}}, \bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}\}. \tag{3}$$

That is, $\{\bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}\}$ is optimal when $\mathbf{w} = \bar{\mathbf{w}}$. Now, we need to prove that the beamforming vectors $\bar{\mathbf{w}}$ are optimal when $\{\mathbf{u}, \boldsymbol{\rho}\} = \{\bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}\}$. To see this, we note that the $\{\mathbf{w}^{r_1+1}\}$ is given as the optimal solution of problem $\mathrm{P}_4^{\mathrm{r}_1}$. Hence, from the optimality of $\{\mathbf{w}^{r_1+1}\}$ and the monotonicity of the objective function we conclude the following

$$Q(\mathbf{w}^{r_1+1}) \leq Q(\mathbf{w}^{r_1+1}) \leq Q(\mathbf{w}), \quad \forall \mathbf{w}. \tag{4}$$

By taking the limits of both sides of equation (4), we get

$$\bar{Q} = Q(\bar{\mathbf{w}}, \bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}) \geq Q(\mathbf{w}, \bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}), \quad \forall \mathbf{w}. \tag{5}$$

Thus, $\bar{\mathbf{w}}$ must be the optimal solution to the problem $\mathrm{P}_3(\mathrm{M})$ when $\{\mathbf{u}, \boldsymbol{\rho}\} = \{\bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}\}$, and we already have shown that $\{\bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}\}$ is the optimal solution to problem $\mathrm{P}_3(\mathrm{M})$ when $\mathbf{w} = \bar{\mathbf{w}}$. Based on these observations we can easily show that $\{\bar{\mathbf{w}}, \bar{\mathbf{u}}, \bar{\boldsymbol{\rho}}\}$ is a KKT solution to the optimization problem $\mathrm{P}_3(\mathrm{M})$ by checking the KKT conditions. To this end, we have shown that any cluster point of the iterates generated by Algorithm 1 converges to a KKT solution of the optimization problem $\mathrm{P}_3(\mathrm{M})$. Although the KKT points are not necessarily unique, the distance between any sequence $\{\mathbf{w}^r, \mathbf{u}^r, \boldsymbol{\rho}^r\}_{r=0}^{\infty}$ and the KKT solution set goes to zero in the limit. This completes the proof.