In this work, the authors propose a robust optimization scheme to solve the EE maximization problem with channel uncertainty in a D2D communication-supplied distributed antenna system. To slove the non-linear optimization problem which is involved with an intractable probability constraint, Bernstein method and **D.C.** approximation-based CCCP algorithm are adopted. In the proposed scheme, the two tractable approximate constraints under two forms of channel uncertainty are reformulated to guarantee a well-functioning communication system. The results show the impact of channel parameter uncertainty on system performances.

I consider that this paper addresses an attractive research topic, but the contributions of this work are quite limited and important discussion is missing.

In conclusion, I believe that there are major flaws that need to be addressed. The weak aspects of the paper are summarized as follows:

1. The abstract is confusing and should be rewritten. Most noteworthy, the description of the simulation results in this abstract looks feeble.

2. The introduction isn’t well organized and more in-depth research for the related work should be done. Here are some typical examples I've listed.

1) It’s very abrupt that " […] There have been many contributions in the literature investigating the problem of maximizing EE in DAS […] ", since the discussions ahead of this sentence focus on the spectral efficiency instead of EE.

2) The grounds of argument that " […] Given the fact that channel estimation errors always exist in practice, it is imperative to examine the impact of channel uncertainty on systems performance and design a systems with good robustness […]" is insufficient, more advantages of imperfect CSI should be discussed here. Besides, in my opinion, the impact of channel uncertainty on system performances is not a particularly important topic since the ultimate goal is to demonstrate the effectiveness of the proposed algorithm.

3) It’s not true that " […] However, the deterministic channel error model assumes needs to feed back the instantaneous CSI and require large feedback overhead[…] " since the deterministic channel error model is based on an estimated CSI and is restricted by the predefined error boundaries.

4) The sentence that " […] However, to the best of our knowledge, there has been little work studying the EE optimization […] "is inaccurate since there are many articles that investigate EE maximization. Besides, the involved EE optimization can be regarded as a contribution point only if a novel optimization algorithm which aims to solve the fractional programming problem is proposed.

3. An introduction which exhaustively overviews the characteristics and limitations of DAS with D2D communications should be included in Sec. I.

4. Contributions of this paper should be systematically summarized, and some notations (e.g., Exponential distribution-*E,*complex Gaussian distribution- $\mathcal{CN}$) should be included in the end of Sec. I.

5. There are some flaws in Sec. II-Subsec. B. First, the sentence " […]The CSI between the CUE and RAUs is assumed perfectly known[…] "contradicts the Eq. (1) because Eq. (1) contains the random variable $g\_n$. Hence, the classification criteria based on the perfect or imperfect CSI situation is inappropriate. In addition, the ref [18] in this paper is misquoted since ref [18] has nothing with user’s mobility and the delayed feedback.

6. The system model containing DAS is complicated and has its characteristic. However, the corresponding mathematical optimization problems have little to do with the characteristics of DAS.

7. In this work, *l*2-norm approximation is first obtained by Bernstein method, and then *l*1-norm approximation is deduced by a norm inequality. However, the refs [15] and [18] in this paper are focus on *l*1- approximation and *l*∞- approximation which are both involved with a low-complexity separable structure. Based on these, I’m very confused why this paper selects *l*1- and *l*2- approximation to have a comparison, instead of *l*1- approximation and *l*∞- approximation.

8. To solve the similar non-convex problem with D.C. structure, three optimization algorithms have been proposed in the past (e.g., D. T. Ngo, S. Khakurel and T. Le-Ngoc, "Joint Subchannel Assignment and Power Allocation for OFDMA Femtocell Networks," in *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 342-355, January 2014.). The involved three algorithms are based on arithmetic-geometric mean (AGM) approximation、successive convex approximation (SCA) and difference-of-two-concave-functions (**D.C.**) approximation respectively. And the **D.C.** approximation-based algorithm has the slowest convergence rate among these involved algorithms. Hence, I have a question that why did you use the algorithm with the slowest convergence rate, are there any other advantages of CCCP algorithm?

9. Since there are two identical legends which correspond to different lines in a graph, legend entries in Figs. 2 and 8 are not really clear and should be edited. Besides，the label "outage probability" in Fig. 3 and 4 should be "Real outage percentage (%)".

10. These simulation results are insufficient to support a high-quality paper. The paper would have gained much more value if the performance of the proposed optimization scheme were validated through a comparison with the performance of state-of-the-art methods.

Due to the above reasons, I recommend that this paper should be rejected.

Also, there are many vague statements and grammatical mistakes. I have listed some of them.

- page 1, column 1, line 14/15: overlaid -> underlaid

- page 1, column 1, line 21/22: I could not understand what the term ' it ' in "In order to solve it" is.

- page 1, column 1, line 37/38: It is not better to put "due to environmental concerns" in the middle of a sentence.

- page 1, column 2, line 22: improves->improve

- page 1, column 2, line 24: There have been-> There are

- page 1, column 2, line 37: systems performance-> system performances; a systems-> a system

- page 1, column 2, line 38: exist-> exists

- page 1, column 2, line 40: the receiver assumes-> the receiver is assumed

- page 1, column 2, line 47: "which controls the interference of the secondary users (SUs) to the primary users (PUs) below a tolerable threshold " should be revised.

- page 1, column 2, line 49/50: "the deterministic channel error model assumes needs to feed back the instantaneous CSI and require large feedback overhead " should be revised.

- page 2, column 1, line 43: the downlink-> a downlink

- page 2, column 1, line 45: RAU-> (RAU)

- page 2, column 1, line 56: channel links-> communication links; the model-> the system model

- page 2, column 2, line 51: c->$c$

- page 3, column 1, line 48:P\_{i,i}-> P\_i

- page 3, column 1, line 53:unit-> unit 1; the sum SE-> the SE

- page 6, column 1, line49: '*h*1' should be illustrated in advance.

A space should be added between all 'Fig. ' and number.( e.g., Fig.1-> Fig. 1)

In terms of readability and correctness, the paper really needs to undergo a thorough revision.