The authors have interesting and useful work! However, its presentation in this paper can be significantly improved with some relatively minor (but important) editing.

Here a quick/general overview, highlighting my main points (with more specific comments/suggestions saved for later):

The organization sometimes leaves out too much "intuition" until details are presented in appendices, which makes the paper more challenging to follow (and reduces the likely impact of the publication). Also, information on what content is new (compared with several past publications) should be a little more detailed, with a more specific "roadmap" of where this novel information appears within the paper. There should be a better description of the in-house simulation software used - and how it compares with "avatar" based modeling/tests from their previous work. More information about computers used (etc.) would be useful (to interpret benchmarking results), too.

*** Overall Review: ***

- Summary

This paper presents a framework for motion planning for a multi-limb robot to locomotion within complex environments (i.e., avoiding obstacles and achieving required contacts). Planning is broken into 3 basic problems (P1, P2, and P3):

P1: guide path for root/trunk of robot.
P2: discrete equilibrium poses, including limb contacts.
P3: interpolating between contact postures, for full trajectory.

* Note: Figure 1 provides an excellent "visual description" of P1 and P2! (Nicely done!)

This work focuses on P1 and P2, as decoupled problems, "while relying on [a?] state-of-the-art solution for P3" (in the authors' words), via use of a "reachability condition", to reduce dimensionality of the problem as a trade-off in speed vs accuracy, gaining improved algorithmic speed by solving only a decoupled and reduced-dimensionality approximation of the full problem.

An implementation of P3 is described/provided in Appendix C, to illustrate the full planning structure.

In essence, the method (for addressing P1 and P2) looks for paths in which a core "trunk-like" shape (e.g., cylinder or rectangular prism, etc.) is out of collision with the environment, while the reachable WS of the limbs is "in collision" (as a heuristic for likely achieving needed contacts with the environment). Adjusting (i.e., fictitiously super-sizing) the size of the "core" in turn affects how well this heuristic approach works, which is one issue the authors address through simulation.

More specifically, Ws is a scaled version of nominal trunk shape Wo,the polygonal shape of the trunk, used in P1 (guiding the root). S=1 is the "true" size, and S>1 is explored.

Full-body poses have only one contact change, between neighboring poses. Heuristics/criteria are described for ensuring "robustness" to slipping (via violation of an approximation of the friction cone), and for selecting "nice" limb configurations (e.g., away from singularities, or providing adequate end effector forces).

[C_contact: whole body configs in contact AND collision-free.
C_reach: an approximation, via simple geometries, which can be sampled efficiently.]

- Major Contributions

This paper collates previous pieces by the authors and others (for motion planning of legged systems across/within complex environments) into a more complete and whole presentation, to explore trade-offs in approximations used and general performance characteristics of the approach.

As for novel contributions, this did not seem to be addressed/explained until page 3 (in the last paragraph of Sec. I). Much of the methodology has already been presented in earlier work, and the particular aspects that set THIS paper apart seem (as claimed by the authors) to be:

(1) “ a criterion for robust static equilibrium” – but it isn’t described until page 12, in Appendix B, Section B… B is apparently based mostly on reference [29], except for its LP formulation in the present work. It is a pretty intuitive/typical check for quantifying a margin for friction cone–esque violations (using an approximation for the friction cone, with x and y direction forces considered independently). It would really help to describe this MUCH earlier in the paper (than page 13), if it is a significant novelty of your paper.

(2) Provides “complete pseudocode”. This is great! However,it’s not fully clear how much of this is novel vs just re-presenting the same algorithms previously publilshed. If only code in App. B and C are new, that should be clear. Conversely, if most of this presentation was NOT outlined in such detail before, that should also be clear (to the authors' credit). Either way, a brief clarification in the intro is recommended.

(3) The contribution to “provide a solution to P3” doesn’t happen until page 13 (Appendix C) – and as with contribution (1), it needs a better, intuitive summary (briefly) when first mentioned, as well as a “roadmap” for the reader as to where it will be described. It is also ambiguous in the paper whether to what extent (if any) this solution is novel, versus a presentation of work from [1]. "Provide a solution" seemed to imply novelty, in the context of this paragraph -- yet later, the authors (near the end of page 2) state they use a "state-of-the-art solution to P3", which implies the P3 solution is not a claim of novelty.

(4) The last contribution is validation using an in-house simulator, as opposed to more approximate (“unrealistic avatar”) models used in previous work. As with the previous contributions claimed, you should flesh out just a bit more detail, along with a roadmap to where a particular contribution appears in the paper, to prepare and guide the reader a bit more. More precisely, what are the differences between these two “simulation” approaches (previous and current)? You need a better description of what your in-house simulator does and does not do; a vague citation to another paper based on that simulator [38] is not sufficient.

- Technical Accuracy

I did not find any issues with technical accuracy, although presentation can be improved.

- Presentation

As mentioned (see "Major Contributions"), the authors need to better distinguish what the novel contributions are, and to highlight where these contributions will be presented in the paper. Also, many important details are in the appendices. The paper would be improved with more intuitive comments/summaries on these details, rather than just writing (for example), we use a bunch of heuristics, and they are described in Appendix X. For example, you can at least say that contact between feet and the ground use a heuristic (h_EFORT) for ensuring adequate forces are achieved, while hand/arm contacts are more successfully planned via a well-known heuristic (h_w) to avoid singularities... (Not necessarily that particular description, but however you would summarize your general approach, I mean.)

Also, where you describe "sensitivity" vs "specificity", it would be useful to give a more intuitive explanation, example, e.g., "meaning the heuristic predicts a successful full-body pose where the true model fails to find a solution"; or "when a solution exists, the heuristic also correctly predicts this". (Again, not necessarily THAT description, but something brief, in your words, for improved readability...)

There are other places where similar "brief intuitive descriptions" can help, but my review is getting a little long. (Check the attached "minor edit suggestions" for some of these...)

I fully realize the authors are probably up against page limits for this paper already. However, I suggest the paper will be improved if the presentation focuses more on highlighting (a) what is novel, and (b)
what take-away messages you have, about why the whole framework is designed the way it is [within the main text; not hidden in appendices].

- Adequate Citations?

Citations of relevant work were useful and sufficient.

- Multimedia

A nice video! More detail on what the simulator used would be useful to interpret the results, however.

*** Minor Suggestions, for consideration ***

-- SEE PDF ATTACHMENT...

Yikes. Sorry my review is already so long! You have done a nice body of work, and I'm only trying to help improve the long-term impact of the paper.... Anyway, I've attached specific potential edits as an attachment.

- Finally, I apologize for any typos (or confusion) of my own, within this review. (I've tried to be clear and helpful.)

On p.2 "to trade the combinatorial – [?]", should add a noun here. Maybe, "scaling issues"

p.4, Sec IV: "Given a start and a goal configurations" is awk. Maybe, "given start and goal configurations" (or else a start and a goal configuration).

p.5, on right, "less than epsilon" between end effector and world – but doesn't the orientation of the end effector matter tremendously here, too?

p.6 "a start and goal root configurations" – should drop last s

p.6 – this would be a great place to give more "intuition" for the heuristics. Hw seems to be used in essentially all foot contact, with h_efort for hand/arm contacts, and ALSO h_vel not seeming to be used at all.

p.6, Fig 7 – not really a "steep" (?) staircase… (especially compared with Table V data)

p.6 'three-fingered hand' – slightly misleading, since no rolling (as fully admitted in the text). I'd just cut this?

p.8– somehow, "Table II" doesn't appear until after "Table IV" – which is a frustrating choice for latex to make. Not sure what you can do to change that, but it would be preferable to have tables (and figures) appear in the correct order.

p.8 Sensitivity and Specificity discussion and data. Need to be more clear what these mean., e.g., via an intuitive example.

p.8 "requiring to exert important forces" is awkward. Perhaps change to, "requiring significant force production" or "in which significant forces must be exerted" or "requiring large magnitude contact forces" or maybe [to get your point across more clearly, as well]:
"in which achieving the necessary directions and magnitudes of contact forces intuitively requires some consideration of the Jacobians associated with the postures chosen in motion planning"

p.8. Need to define "discretization step" more precisely, I think(?)… Min distance? Max distance? Nominal distance? In time? In Distance? In all 6 DOF??

p.8 "analyzed their success rate" seems like "their" refers to "computation time". Just change to "analyzed success rate"

p.8 Near end, "thus allowing to consider" – also awkward. Just use "thus enabling"

p.8 Near end. "the more constrained… the less…" drop the "the's", e.g. "a more…" and then just "less" (w/out the "the")

Attachment uploaded Thu Aug 24 16:11:22 2017 PST

p.9 Not sure if Fig. 14 is worth including… It looks like its just Fig. 9 from the ISRR paper (?) [20], and there is still no "full solution" here – just an illustration of potential future work. (At a minimum, the figure/solution should probably be referenced as having already appeared in the previous ISSR paper?)

p.9 near end, Sec. VII – would replace "interest" with something else, maybe "purpose" – both in the section title and in the first sentence of that section.

p.9 'addresses highly constrained environment' – actually, not very well, correct? This was the "exception" case, for which results were rather poor, right? Also a case where "heuristics" are not so good for getting true (reliable) holds on grasp points, etc.?

p.12 The first part of App. B is referred to within this Appendix as "new minor contributions derived from previous work". These aren't so novel, perhaps, and what is more useful to the reader is a better intuition for when to use each within the MAIN BODY of the paper, where they are originally described/associated to particular simulations.

p.12 The "manipulability measure h_w" is said to be "also given by "Yoshikawa". Go ahead and provide that as a citation to an appropriate paper or other reference source, if possible, e.g., [32] (I believe?). […and if you found out about this older work through some newer publication, it's best to mention both.]

p.12 pyramid friction cone – it's not clear this is "conservative"… e.g., a square "within" a circle vs "outside" a circle… this sounds like the pyramid is outside, which means sometimes things would actually slip – which is not "conservative"…

p.13 after Eq. (15), "rather than solving directly (19)…." – do you mean "(15)"?

p.13, perhaps be more clear (very briefly) in mapping (scalar) solution b0 to solution of v (R^6) in (16).

p.13 – github link seems to be dead? (The link on p. 15 is OK… should they be the same?)

p.13 – after equation (17), suppress indentation.

p.14 – not a good description of the simulation software used… Ref [38] focuses on robustness of simulations, but presumably, the testing in the present work is deterministic – just using the same simulator? Not much description is given of how/whether contacts/slipping are modeled… why isn't a 3rd-part software used, or at least software with better documentation? Better info is needed to understand what the "simulations" truly represent.

Attachment uploaded Thu Aug 24 16:11:22 2017 PST

The paper concerns the problem of planning legged locomotion tasks, assuming multi-contacts (i.e. not only feet can touch the environment) and arbitrary environment type, including complex obstacles (stairs, rubble). The work focuses on efficient solutions to two sub-problems, namely planning a guide path for the root of the robot, and planning a discrete sequence of equilibrium configurations of the robot along the guide path. For the interpolation of motion between the consecutive configurations this work relies upon already known methods, including the one described by the same authors in [1]. The main main scientific contribution is the reachability condition, which defines a geometric approximation of the contact manifold that can be efficiently sampled and explored. The use of this concept enables to split the motion planning problem into two sub-problems that are much easier to solve from the algorithmic point of view. The technological contribution are the two planning algorithm: the first one is a RRT-based fast guide path planner, which samples from an approximation of the contact reachable space, and the second one is a method that extends a contact reachable path into a sequence of whole-body statically stable configurations. It is remarkable, that the proposed algorithms are provided as open source software, which is a standard in robotic vision, SLAM, etc., but still quite rare in planning and control. The contributions are well explained in the introduction, making a good use of the presentation of related work.

The significance of the problem being considered in this research, and the practical implications of the proposed solutions are also well explained in the introduction. A good point is that the authors motivate their solution from one side by a clear analysis of the relevant publications, but from the other side, they draw a practical motivation for the solution from the recent DARPA Robotics Challenge. Hence, their solution, which is not complete from the algorithmic point of view, is arguably still the best known approximation of the planning system we would like to have for our legged robots.

although the paper is clearly written and well organized, I am a bit skeptical about the readability of the figures. Some of them seem to bee too small (e.g. Fig. 3, Fig. 4), some are of quite poor quality (why the robot silhouettes in Fig. 4 are blurred ?), and the purpose and meaning of Fig. 6 is rather unclear. In general, the simulation results presented qualitatively in the paper are a bit hard to interpret without the accompanying video, particularly if someone looks at a b/w print of the paper (as it may be looking in T-RO).

The references are relevant and complete if it comes to the literature directly related to humanoid motion planning and planning of multi-contact sequences. However, I suggest that a few papers related to more general motion and contact planning for multi-legged robots can be added. Particularly, the problem of ``foothold selection'' for 4 or 6-legged robots travelling over uneven terrains is a problem related to the one tackled in the paper. It would be nice to select a few representative works (e.g. from the DARPA Learning Locomotion project or
papers related to the ETHZ StarlETH and ANYmal robots) and explain the differences between those approaches and the problem solved in the paper.

The paper is technically sound, but in my opinion a T-RO regular paper should be more self-containing. In several points the authors heavily rely on previous publications, e.g. when presenting the RRT-based planner (Section IV-C). This part should be expanded, describing in more detail what exactly has been changed with respect to the reference paper. I also found the description of the environment model being used,
and the way it is used in planning quite superficial. It is said, that ``an octree data structure'' is used to keep the sampled the end-effector positions, and it is ``intersect'' with the environment. The idea is OK, I understand that the octree is a sample-based representation of the effector's workspace, and using some environment model it is possible to decide if the given configuration can provide a contact point or not. but how is the environment described ? Howe

efficient is this intersection operation ? Does this method require the environment to be described by some specific objects (e.g. polygons) ? What about the (unavoidable) uncertainty if a world model acquired via on-line sensing is considered ? Perhaps some of these issues (e.g. uncertainty) are not in the scope of the paper, but I can imagine that a robotic researcher who reads the paper and is about to use the open source software it advertises is very interested in the answers.

I appreciate the thorough presentation of results in the paper, however it seems that this presentation is somewhat biased towards simulation results. Simulation are perfectly OK for planning algorithms, but if you decided to present the experimental results (HRP-2), then at least some illustrations should be included in the paper. now they are only in the accompanying video. also a comparison between the results of the same task(s) executed in dynamic simulation and on the real robot should convince the reader that the planning method's good performance does not depend too much on the idealized simulation environment.

This paper presents a general planner to produce whole-body motions in multi-contact scenarios. The planner is divided into several stages: First, a trajectory is chosen for the root body of the robot using a sampling-based planner, with a reachability heuristic used to guide the search towards configurations which are likely to admit feasible quasistatic contacts. Second, each root configuration is extended into a whole-body configuration by searching for the configuration of each limb independently, guided by additional heuristics. Finally, the configurations are interpolated by solving an optimal control problem. Results are presented on several simulated multiped robots and a physical HRP2 humanoids.

This paper is clear and thorough in its description of the methods and results, and the level of detail is good. My primary concern, however, is that the elements described in this paper have been described in previous published work, which makes the particular contributions of this work unclear. To be specific, the reachability-guided planner appears identical to the planner in [20] (as stated by the authors in the end of Sect. I.B.), and the optimal control formulation is from [1] (again as stated in I.B.). The authors highlight a particular contribution of this work to be the efficient robust balance criterion, but I had trouble understanding this particular contribution. Does this refer to the approximation of equilibrium feasibility by contact reachability? In that case, how does this differ from just applying the planner in [20] directly? Or does this refer to the robustness parameter b_0 in Appendix B Eq. (15)?
In addition, sections IV and V are very similar to the description of the planner in [20]. That's not surprising, of course, but there is no indication in those sections that the work being presented is a rephrasing of an earlier publication.

All of this is not to say that I don't believe this work has publication value. The authors have clearly presented their approach and thoroughly documented their algorithms, heuristics, and results (which are impressive and certainly worth sharing). I do, however, believe that the fact that the technical approaches are rephrasings of prior work should be made more clear (or that sections which can be replaced by citations of [20] and [1] be removed).

If the authors choose to follow this advice, then I have a few other suggestions and requests for clarification:

* I think I understand what the authors mean by solving at "interactive" rates, but this is worth defining. In my mind, it means something like "before the human operator gets bored", or roughly > 0.5 Hz. But since the authors make the claim that this is the "first interactive implementation of a contact planner", then that statement should be made quantitative and verifiable

* In Section I.A. the authors refer to Bretl's work on multi-contact planning as the inspiration for the P1,P2,P3 decomposition of the problem. But it is not, as the authors state, always necessary to decompose motion planning into choosing a series of quasi-static configurations. There are a variety of feasible motions with no static equilibrium configurations, so it is worth mentioning that this decomposition only solves a subset of motion planning problem.

* Also in Sect. I.A, the authors state that "optimization-based approaches only converge locally". This is not necessarily true, as convex or mixed-integer convex optimizations can (sometimes) be solved to a verifiable global optimum. I would suggest removing this sentence entirely.

* I am not sure why the C_reach hypothesis is described as the "strong" hypothesis in, e.g., Sect. IV. Is there a corresponding weak hypothesis?

* In Sect. IV.B, the authors state that the ideal shape B* "has no

explicit definition". For a particular configuration q_0, B* is simply the surface of the robot, right? So I assume that this claim refers to there being no explicit definition of B* which is the volume around B*(q_0) for every q_0? I think I agree with the authors, but I found this claim hard to follow and would appreciate some clarification

* In Sect. V.A., the authors mention that the ability to plan a contact break and creation with the same limb allows the creation of dynamic motions. This is quite an important ability for an otherwise quasi-static planner, but this ability isn't described again until the appendix. It would be worth explaining why this is true in the paper itself.

* In Sect. V.B, the authors mention inserting an additional state between contact configurations if two or more contacts cannot be maintained. Shouldn't this step appear somewhere in Alg. 1?

* In Appendix A.A, it is mentioned that step 2 in the workspace computation is conservative, with reference to figure 17-right. But Fig. 17-right is the simplification of the convex hull, corresponding to step 3, and it's actually hard to tell from the figure that it is conservative. Not knowing anything about Blender's decimate tool, I assume it is operating by removing vertices, which would indeed be conservative, but some clarification would be useful here.

* Finally, I would suggest providing explicit git SHAs or tags for all of the github links. I've personally made the mistake of linking to a github project in a paper, only to have that project change and eventually delete the specific code I was citing.