Author's response to the reviewers' comments

Manuscript number TN-08-213
Title: Unfalsified Control of Linear Time-Invariant Systems
Author: Ivan Markovsky

August 25, 2008

We thank the reviewers and the associate editor for their relevant and useful comments. In this document, we quote in **bold face** statements from the reports. Our replies follow in ordinary print. First we summarize the main points from this review round.

1 Summary of response to the reviewers' comments and corrections/additions

Despite the severe criticism expressed by some of the reviewers, after studying their comments and the suggested references, we still believe that our results are *original, correct, and potentially important* for the systems and control community. Detailed answers to the reviewers' comments with rebuttal and explanation of the corrections/additions in the revised version of the paper are given in Sections 3–6 of this document. Here we summarize for the convenience of the editors the essential points from the reviewers' reports and state concisely our position and the main reasons for it. Please, refer to the corresponding sections below for more detailed justification of our position.

- *Novelty of the results in the paper*

  We disagree with the claim of reviewer 4 that question Q2 in our paper is already answered in references [A,B], because these references *do not take into account the LTI structure* of the plant (which is the aim of Q2).

- *"Complexity" of the results*

  We disagree with the claim of reviewer 4 that our solution is "more complicated" than the solution of [A,B], because the latter is a linear program and the former is a *computationally simpler* linear least-norm problem.

- *"Proper" problem definition and hypothesis*

  Reviewer 6 says that we lack "proper" problem definition and hypothesis. We do not know what he means by "proper" and why our Problem 1 and the hypothesis stated in Proposition 3 are not considered "proper".

- *Correctness of the results*

  Reviewer 6, gives an example that is claimed to be a counterexample for Proposition 3. This example, however, contradicts assumption 2 of Proposition 3 (persistency of excitation of the input), so it is not a counterexample.

- *Correctness of the simulation example*

  Reviewer 3, doubts the correctness of the simulation example, because unstable systems have unbounded responses. The simulation, however, is done on a *finite time interval* and *any* finite response of *any* LTI system is bounded. The MATLAB function that was used to generate the simulation results is made available to the reviewers and the editors in case they wish to reproduce the simulation results in the paper.

- *Importance/application of the results*

  Reviewers 3 and 6 have related questions about the application of our results to unfalsified switching adaptive control, i.e., what switching mechanism is used and how it is implemented in real-time. In the revised version, we have clarified that the scope of this *Technical Note* is limited to the answers of questions Q1 and Q2. We do not consider a complete unfalsified switching adaptive algorithm for LTI systems (an obvious next step).

Taking into account the points of the reviewers with which we agree, we have done the following main corrections/additions in the revised version of the paper:

- *Scope of the paper*

  We do agree with Reviewers 3 and 6 that the paper promises more than it actually delivers, because we talk about unfalsified control for LTI system but do not develop a complete unfalsified switching adaptive algorithm for such systems. We actually treat only the test for falsification of a potential controller, based on the fictitious reference signal, and develop tool that provide such test taking into account the LTI structure of the plant.

  In order to correctly emphasis the main contribution of our work, we have changed the title to "Closed-loop data data driven simulation", have reformulated questions Q1 and Q2, so that "unfalsified control" is replaced by "falsification test based on fictitious reference signal", and added the following sentence in the introduction:

  > In this technical note, we do not consider a complete unfalsified switching adaptive control method but focus instead on its core ingredient—the test of a potential controller's performance directly from data of the closed-loop system, operating with a (possibly) different controller.

  B.t.w., in the conclusion of the original version of the paper, we state that "The final goal [of our work] is applying the data-driven simulation algorithm in a (switching) adaptive control scheme.", a presentation of such an algorithm would probably go beyond the limitation of a 4 pages Technical Note in the IEEE-AC format.

- *References*

  We thank the reviewers for the suggested references. We have included 6 additional references in the revised version of the paper, which we believe are relevant for our work.

- *Theorem-proof style*

  In the original version of the paper, the proof of Proposition 3 was given *before* the statement of the result. We thought that this presentation style would make the paper more readable than the classical theorem-proof style. The questions of Reviewer 6, however, suggest that our choice was confusing for him, so in the revised version of the paper we reverted to the classical theorem-proof style.

- *Correction of typos*

  All indicated typos are corrected in the revised version of the paper. Thank you for pointing them to us.

2 Answer to the associate editor's report

Reviewer No. 3 commented that the citation should be more comprehensive and the example should be more strengthened.

The suggested references are included and the example is clarified in the revised version of the paper.

Reviewer No. 4 pointed out that authors should both explain the relation of their solution to that of Woodley

Quoting from the abstract of the paper of Woodley et al.

> "No assumptions about the plant are explicitly required ... In particular, no assumptions are made about the linearity of the system."

In contrast, as indicated already in the original title of our paper, we address the problem of taking into account the linear time-invariant structure of the plant. This is an important difference which makes the problem of Woodley et al. and our problem rather different. The solutions are also different and a comparison between them is bound to be unfair because our solution benefits from the extra knowledge of the LTI structure of the plant which is not used in the solution of Woodley et al.

and they should demonstrate that their apparently more complicated solution has advantages over that of Woodley

The solution of Woodley et al. (Section 4 of [A]) is a linear program. The solution of Kosut (Section 4 of [B]) is a convex feasibility problem in the case of an ARX model (and is a nonconvex feasibility problem in general). Our solution (see, Section V, Algorithm 1) requires solution of *one* overdetermined system of linear equations. In the sense of computational complexity (number of flops) as well as in the sense of theoretical complexity (duality theory) of the corresponding classes of problems (LP, Convex optimization, and linear system of equations, respectively) our solution is simpler than the ones presented in the references [A] and [B] of the reviewer's report.

Section "Details" in the reviewer's report clarifies that our solution is considered more complicated in the sense that it involves both Toeplitz *and Hankel* matrices, while the one of Woodley et al. involves positive definite Toeplitz matrices *only*. We do not understand why this fact (and not the computational or theoretical aspects of the solutions) is considered as the criterion for comparing the methods.

Please refer to page 6 of this document for further explanation on the points we disagree with the reviewer No. 4 comments.

Reviewer No. 6 raised a number of serious questions such as what hypotheses on the reference signal used in the experiment is needed

The hypotheses we make are stated in our main result—Proposition 3. According to assumption 2 of Proposition 3, the reference signal should be persistently exciting of a certain specified order. In addition, we state the hypotheses in the introduction

> "the procedure gives the exact answer assuming that the data is persistently exciting and the plant is controllable"

and discuss them in the second paragraph of the conclusion. We believe that the hypotheses in our work are sufficiently emphasized and their meaning is clearly explained. If the reviewer has specific advice on how to make the statement of the hypothesis in the paper clearer, we will be glad to consider it.

can it be shown that the worst-case (the H-infinity norm) can actually be found

The worst case is found in the example of Section IV. The purpose of this analysis on the first place is to answer question Q1. It turns out that the example suffices to answer Q1—namely, unfalsified control methods need to be specialized for the LTI case (when the plan is known to be LTI). Therefore, the worst-case analysis in the general case of an arbitrary LTI plant and controller is not needed for our purposes.

3 Answer to Reviewer 3

Concerning to Q1, now it is not novel at all to reveal that unfalsified control fails to detect instability of a switched-off destabilizing controller via a fictitious reference signal.

We agree that the inability of the fictitious reference signal to detect instability of a potential controller is well known and the suggested references are included in the paper. Note, however, that question Q1 is aiming for more than "detecting instability". We are interested in *quantifying* the conservatism of the fictitious reference test, which requires further analysis than the one in the suggested references. We have not done the analysis in its full generality but rather took a specific example. This is clearly stated in the abstract and in the introduction:

"The first question is answered in the special case of first order plant and static controller."

For this example, we were able to show that the fictitious reference signal can be arbitrarily conservative. To the best of our knowledge this result is novel and despite the fact that it is done for a specific example, it suffices to answer Q1—namely, unfalsified control methods need to be specialized for the LTI case when the plan is known to be LTI.

the definition of unfalsified control concept is generally incomplete. The author should reformulate the concept along with a cost detectability and a switching mechanism.

We do not consider a complete unfalsified control algorithm. The switching mechanism and the real-time implementation issues are indeed missing. The paper is focused on one aspect of the unfalsified control concept — the fictitious reference signal (which is an important ingredient in any cost detectability and switching mechanism). The reason for skipping the switching mechanism is that we prefer to keep this *Technical Note* short and focused on the theoretical questions Q1 and Q2. The revised version states this limitation in the introduction "In this technical note, we do not consider unfalsified switching adaptive control but focus instead on its core ingredient—the test of a potential controller's performance directly from data of the closed-loop system, operating with a (possibly) different controller." and in the end of Section III "In this note, we do not consider the switching strategy and its real-time implementation but focus on its core ingredient—the fictitious reference test—in the case of an LTI plant."

Regarding to [1], it was not clear how to choose an active controller in unfalsified controller set. This can be cited to the recent work proposed by Wang et al.

Thank you for the suggestion. The reference is included in the revised version of the paper, see the last paragraph of Section III.

In addition, notation of the fictitious reference signal and the true reference signal should make clear even though it is well-known if a controller is active, its associated fictitious reference signal can be computed exactly as the true reference signal.

Unfortunately we are unable to address this comment (on this review round) because we do not understand the meaning of the sentence above. If the reviewer can kindly reformulate the sentence, we hope to be able to understand its meaning and be able to address it.

e_d **should be defined as the fictitious error signal**

Done.

mistakes with square of gamma.

Thank you for pointing this out. These typos are corrected.

The conservatism measure c of its fictitious reference when c is switched-off destabilizing controller is not consistent with the measure that you have defined beforehand.

We do not define the conservatism measure c before page 7, line -5 in the original version (equation (8) in the revised version), so we do not see where is the inconsistency pointed out by the reviewer.

Finally, the problem of zero on Right-Half-Plane of fictitious reference signal for a destabilizing controller should be addressed.

Could the reviewer please clarify what problem a zero on Right-Half-Plane of fictitious reference signal can cause? From our analysis, we do not see how this can cause a problem.

it would be better to reformulate this example in a general case instead of a first-order system.

The example is actually a counterexample for the hypothesis that the fictitious reference test is non conservative. The example shows that the fictitious reference test may be conservative, so it is sufficient to answer question Q1 negatively. If the result of the example was that the test based on the fictitious reference signal is non conservative, then we would not be able to answer Q1 positively and in this case the general case would indeed be useful in *proving* that the test is non conservative in general. This, however, is not needed because already the simple example shows the conservatism of the test.

This approach is so-called 'the closed-loop data-driven simulation'. The technique was developed in [2].

Please note that the closed-loop data-driven simulation concept, its solution, and application to unfalsified control are contributions of *this paper*. The solution is inspired by our previous results and borrows ideas from the data-driven simulation of [2], however, we disagree with the statement that "The technique was developed in [2]."

The simulation is doubted whether or not it is correctly produced results. In simulation, the closed-loop system of an unknown plant and a controller beta = 0.1 which is destabilizing is excited by noise reference. Using this data to evaluate the performance of a controller alpha = 0.5 which is a destabilizing. However, the simulation results for closed-loop trajectories of y and u are bounded which is not true since a controller alpha is obviously destabilizing.

We do not follow the concern of the reviewer that "the simulation results for closed-loop trajectories of y and u are bounded". The simulated trajectory is finite and of course *any* finite response of *any* LTI system is bounded.

The reviewer is welcome to reproduce the simulation results reported in the paper. For this purpose, download and execute in MATLAB the function `http://users.ecs.soton.ac.uk/im/test.m`, which implements Algorithm 1 from the paper.

it is questionable how this approach works with a switching of controllers in real-time.

Switching strategies and real-time implementation are not considered in this Technical Note. These questions will be addressed elsewhere.

Finally, it is still not clear how this way will be useful for unfalsified control in order to detect instability of switched-off destabilizing controller in reasonable finite time.

The closed-loop data-driven simulation procedure is applicable in finite time, so we do not see any obstacles in using it to detect instability of a destabilizing controller in finite time. The only essential limitation is that the input signal should be persistently exciting (see assumption 2 of Proposition 3). Thoughts on how this limitation can be addressed/avoided are given in the conclusion.

The author should address a problem of solution using Moore-Penrose pseudoinverse of a matrix A.

We do not understand to what "problem of solution using Moore-Penrose pseudoinverse" the reviewer refers.

Using one example, this would lead to an inaccurate solution of impulse response g0.

Unfortunately this comment is also not clear. What is "one example" here? Note that in our method g_0 is not the impulse response of the system, but an auxiliary variable of the algorithm! Also note that Proposition 3 ensures that the solution is exact, so we do not understand what is meant by "inaccurate solution".

Typing Error: . . .

Thank you for pointing them to us. They are corrected in the revised version of the paper.

Additional references . . .

Thank you for pointing them to us. They are included in the revised version of the paper.

4 Answer to Reviewer 4

Summary: The paper examines to problem of how to test if data falsifies a controller under the assumption that the plant is LTI. Unfortunately, this problem has already been solved in references [A,B]

The solution in [A,B] do not take into account the assumption that the plant is LTI. (In fact, the solution in [A,B] explicitly avoids the LTI assumption!) We believe that the problem of taking into account the LTI assumption is important and to the best of our knowledge it has not been solved before. In particular, it is not solved in the suggested references [A,B]. Although the result in [A,B] is more general, when specialized to the LTI case, it is likely to be conservative because of ignoring the LTI assumption.

the solution of references [A,B] is apparently much simpler than the one offered by present paper.

Algorithmically our solution requires solving a *single* system of equations and is therefore computable *exactly* in a finite number of flops. The solution of [A], requires, solving an LP, and the one of [B], an LMI feasibility problem, in a special case, or nonconvex problem in general. Solving an LMI optimization problem involves iterative algorithms (so, in general, the problem is computed only approximately in any finite number of operations). Our solution is provably simpler in the computational sense of number of flops or execution time. In addition, it is conceptually simpler as it involves only simple linear algebra concepts and no advanced optimization theory and methods.

Details: Whereas the present paper gives only implicit conditions

As stated in Section V of the paper, our conditions are: 1) controllability, 2) persistency of excitation of order $T_\mathrm{r} + \mathbf{n}(\mathscr{B})$, where $\mathbf{n}(\mathscr{B})$ is the order of the system. In our opinion these conditions are as explicit as they could be.

requiring least square solution of a system of linear equations

In fact, in Algorithm 1 of the paper, we require a least *norm* solution of a system of linear equations. Note that the system is underdetermined and the assumptions of the paper ensure that it has infinitely many solutions.

involving both Toeplitz and Hankel matrices, the simpler solution of references [A,B] requires only that one test the positive definiteness of a matrix explicitly given in terms of Toeplitz matrices only.

Our conditions have system theoretic meaning while positive definiteness of Toeplitz matrices may be difficult to interpret in terms of the original problem. In this sense, we can argue that our conditions are more meaningful for the problem at hand. Apart from this, we do not follow the reasoning that existence of both Toeplitz and Hankel matrices in the solution implies higher complexity. A Hankel matrix is a flipped Toeplitz matrix, so why should the extra Hankel structure increase the complexity?

Additional References

Thank you for pointing them to us. They are included in the introduction of the revised version of the paper with the following explanation “Although tests for the controller performance are presented in the context of the direct unfalsified control [5,6], this work does not taking into account the LTI structure of the plant.”

5 Answer to Reviewer 5

In Example 4, should α be positive or negative?

The parameter α defines the controller, so the set of possible values of α defines the set of candidate controllers. In the example, we have $\alpha \in \mathbb{R}$, so that we allow positive as well as negative values. (This means that the set of candidate controllers contains stabilizing as well as destabilizing controllers.)

6 Answer to Reviewer 6

The method is not actually applied to the problem and no proof is given that shows that the non-conservative solution can actually be found.

If "the problem" is understood as a complete algorithm for unfalsified control of LTI systems, then we agree with the statement of the reviewer. Our goal in this *Technical Note*, however, is not to present a complete unfalsified switching adaptive algorithm. (Description of such an algorithm b.t.w. would probably require more than the 4 pages allowed for the Technical Note type publications in IEEE-AC.) Our goal instead is to pose and answer the questions Q1 and Q2. We believe that the questions are novel and important and we believe that the paper gives a definitive answer to them. Our hope is that our results will steer interest in the direction of developing unfalsified switching adaptive algorithm for LTI systems.

a counterexample is easily constructed that shows that the method will lead to an empty solution set for the specific example.

The example given below contradicts to our assumption that the reference signal should be persistently exciting. Therefore it is not a counterexample for the result in the paper.

A proper problem definition as well as necessary hypotheses are missing.

Problem definition is given in Problem 1 and necessary hypotheses are stated in its solution—Proposition 3. If the reviewer clarifies what is meant by "proper" (w.r.t. problem definition and hypotheses statement) we will better see why our statements are not proper and may be able to improve them.

A comparison and references to relevant publications in the field of unfalsified control as well as other data-driven controller tuning methods are lacking.

To the best of our knowledge the question of falsifying a controller assuming that the plant is LTI has not been addressed before, so we have nothing to compare with. (Comparison with other methods that do not assume LTI structure of the pant is bound to be unfair as such methods do not use the extra knowledge that we use.) We apologize for the any omissions in our references and will be happy to include relevant references suggested by the reviewer.

Conservatism is not clearly defined.

It is defined by the displayed equation on page 7, line -5, in the original version; equation (8) in the revised version.

The notion of conservatism for controller unfalsification seems closely related to convergence of the method.

We do not consider convergence of the control algorithm because we do not consider a switching adaptive algorithm. Still we have a relevant question of the conservatism of the unfalsification test (the fictitious reference).

hypotheses on the input signal are in general inevitable. A discussion of relevant results would greatly support the suggested method and show its additional value.

Our hypothesis is that the reference signal is persistently exciting of sufficiently high order. A relevant result in this respect is [WRMM05], which is cited and essentially used. If the reviewer has addition references in mind, we will be happy to know about them.

… For this problem it is well known that a step response cannot solve the problem, which is the setup used to show the conservatism in this paper.

The impulse response is one of the many possible parametrizations of an LTI system and is, in this sense, equivalent to any other—transfer function, state space, etc. We are not sure to which "well known result" the reviewer refers and we believe the problem is solvable in terms of the impulse response representation of the system just as well as in terms of the transfer function or (A,B,C,D) parameters. A simple argument to convince the reviewer that the impulse response is sufficient to solve the problem is that the impulse response contains all the necessary information about the system. In particular, one can derive the transfer function or the (A,B,C,D) parameters from the impulse response.

only the use of this method is suggested, the actual implementation in unfalsified control is missing.

As stated before and clarified in the revised version of the paper, we do not consider a complete unfalsified switching adaptive control algorithm. The switching mechanism and the real-time implementation issues are indeed missing in this *Technical Note*. The note is focused on one important aspect of the unfalsified control concept — the fictitious reference signal (which is an important ingredient in any cost detectability and switching mechanism). The reason for skipping the switching mechanism is that we prefer to keep this paper short and focused on the well defined theoretical questions Q1 and Q2.

results which seem to have been published before (Page 10: " we proved the following proposition" Where, which reference?).

The presentation in Section V contains the complete proof of the proposition. The reason for giving the proof before the statement of the result is that we thought that this will make the paper more readable. Since, it causes confusion, in the revised version, we followed the classical theorem-proof style.

Furthermore, a counter example is easily constructed when applying the method to the problem of section 4: For a chosen reference signal of r(t) = 0, . . .

$r = 0$ contradicts to assumption 2 of the proposition, so it is not a counter example for our result.

The claim that "closed-loop data-driven simulation" gives a complete and non-conservative answer to question Q2 is not supported by proofs nor by the examples.

Our claim is justified by the following explanation "Once computed, the behavior of the plant–controller closed-loop system can be tested against any desired performance criterion." (see the last paragraph of the introduction). We thought that it is rather obviously that once the behavior of the closed-loop system is available it is a matter of using standard analysis techniques to verify whether the performance specification is satisfied. Of course, the "analysis techniques" depend on the particular performance specification being used: some specifications may be difficult to verify while others may be trivial. A well known example of a difficult performance measure is μ. $\mathscr{H}_\infty$ and $\mathscr{H}_2$ are amenable to standard analysis tools. An example of a trivial one is the overshoot in the step response. If we have to check the overshoot, we simply need to find the step response of the closed-loop system (which is a direct application of the closed-loop data-driven simulation algorithm) and take the maximum of its absolute value, which is a straightforward computation in the discrete-time case.

A proper problem definition (with respect to for example initial conditions) is missing.

The problem definition is given in Problem 1. It is independent of the initial conditions, i.e., arbitrary initial conditions are allowed and no knowledge of them is required. (This does not mean that the initial conditions are ignored!)

the following questions need to be addressed in order to support the strong claim that the suggested method leads to a non-conservative answer to question Q2. What hypotheses on the reference signal used in the experiment are needed?

All hypothesis that are needed for our result are stated in Proposition 3.

Can it be shown that the worst-case (the H-infinity norm) can actually be found? How can you systematically find this worst case?

The worst case is found in the example of Section IV. The purpose of this analysis on the first place is to answer question Q1. It turns out that the example suffices to answer Q1—namely, unfalsified control methods need to be specialized for the LTI case (when the plan is known to be LTI). Therefore, the worst-case analysis in the general case of an arbitrary LTI plant and controller is not needed for our purposes.

Minor remarks: The paper is not clearly written. The introduction about unfalsified control is confusing.

We thought the paper is clearly written and will appreciate it very much if the reviewer kindly let us know what in particular is not clear or confusing. In this way we will be able to improve the presentation.

It is not clearly stated that results are only valid for noise-free data.

Quoting from the introduction:

> In the exact (noiseless) LTI case, the procedure gives the exact answer assuming that the data is persistently exciting and the plant is controllable.

Indeed, the results are derived under the assumption that the data is exact, however, this does not mean that they are applicable only for exact data. The closed-loop data-driven simulation method resembles the subspace identification methods, which are initially derived in the deterministic setting and are subsequently applied for approximate and stochastic system identification, see [VD96] and [MWVD06, Chapter 7]. Similarly our method for closed-loop data-driven simulation is derived in the exact setting, however, a small modification (replace solution of equations by least-squares approximate solution etc.) makes it applicable to noisy data.

Page 2: "unfalsified control ... exploiting the LTI structure could give". Which method for example (references)?

As far as we are aware this paper is the only reference on the unfalsified control for LTI systems.

What are the results you could expect for a method adapted to LTI systems as suggested with this phrase?

Quoting from the introduction:

> In the exact (noiseless) LTI case, the procedure gives the exact answer assuming that the data is persistently exciting and the plant is controllable.

We have added a note in the revised version of the paper that the procedure is also applicable with small modifications to the case of noisy data or data generated by a more complicated system, giving in this case only an approximate answer, see the last paragraph of the introduction.

References

[MWVD06] I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor. *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. Number 11 in Monographs on Mathematical Modeling and Computation. SIAM, March 2006.

[VD96] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer, Boston, 1996.

[WRMM05] J. C. Willems, P. Rapisarda, I. Markovsky, and B. De Moor. A note on persistency of excitation. *Control Lett.*, 54(4):325–329, 2005.