REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
19-01-2017	Final Report	29-May-2015 - 28-Feb-2016

4. TITLE AND SUBTITLE
Final Report: New Sliced Space-Filling Designs Construction and Theory

5a. CONTRACT NUMBER
W911NF-15-1-0156

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Qian, Zhiguang

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Wisconsin - Madison 1.00
RESEARCH & SPONSORED PROGRAMS
21 N. PARK STREET SUITE 6401
MADISON, WI 53715 -1218

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
66990-MA-II.7

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The purpose of this project is to improve sliced Latin hypercube designs which were previously developed by the PI and were successfully used by various DoD divisions including US Army.
The project has successfully developed some new classes of statistical designs for computer simulation, uncertainty quantification, sensitivity analysis and stochastic optimization. These include bi-directional sliced Latin hypercube designs and sliced Latin hypercube designs with controlled correlations. The project also derived new statistical

15. SUBJECT TERMS
Deisgn of experiments, computer experimetns, statistics, uncertainty quantification

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Zhiguang Qian
					19b. TELEPHONE NUMBER 608-262-2537

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Final Report: New Sliced Space-Filling Designs Construction and Theory

ABSTRACT

The purpose of this project is to improve sliced Latin hypercube designs which were previously developed by the PI and were successfully used by various DoD divisions including US Army.
The project has successfully developed some new classes of statistical designs for computer simulation, uncertainty quantification, sensitivity analysis and stochastic optimization. These include bi-directional sliced Latin hypercube designs and sliced Latin hypercube designs with controlled correlations. The project also derived new statistical theories to study sampling properties of the proposed designs and their use in integration and stochastic optimization. Specific applications of the designs include running simulation codes with both quantitative and qualitative factors, running simulation experiments in batches and ensemble evaluation of multiple computer models. In addition, some experimental design ideas have motivated the development of new computational algorithms for accelerating large-scale statistical computing.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received		Paper
10/06/2016	3	Shifeng Xiong, Bin Dai, Jared Huling, Peter Z. G. Qian. Orthogonalizing EM: A Design-Based Least Squares Algorithm, Technometrics, ( ): 285. doi:
10/09/2016	2	Qiang Zhou, Tian Jin, Peter Z. G. Qian, Shiyu Zhou. Bi-directional Sliced Latin Hypercube Designs, Statistica Sinica, ( ): 653. doi:
10/09/2016	9	Jiajie Chen, Cong Han Lim, Peter Z. G. Qian, Jeffrey T. Linderoth, Stephen J. Wright. Validating Sample Average Approximation Solutions with Negatively Dependent Batches, Computational Optimization and Applications, ( ): . doi:
10/09/2016	8	Xiao Nie, Jared Huling, Peter Z. G. Qian. Accelerating Large-scale Statistical Computation with the GOLEM Algorithm, Technometrics, ( ): . doi:
10/09/2016	10	Jiajie Chen, Peter Z. G. Qian. Controlling Correlations in Sliced Latin Hypercube Designs, Statistica Sinica, ( ): . doi:

TOTAL: **5**

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

The following papers were presented at meetings, but not published in conference proceedings.

1. Xiong, S., Dai, B., Huling, J. and Qian, P. Z. G. (2016), “Orthogonalizing EM: A Design-Based Least Squares Algorithm,” Technometrics, 58, 285-293.

2. Zhou, Q., Jin, T., Qian, P. Z. G. and Zhou, S. (2016), “Bi-Directional Sliced Latin Hypercube Designs,” Statistica Sinica, 26, 653-674.

3. Nie, X., Huling, J. and Qian, P. Z. G. (2016), “Accelerating Large-scale Statistical Computation with the GOLEM Algorithm,” Revised for Technometrics.

4. Chen, J., Lim, C. H., Qian, P. Z. G., Linderoth, J. T and Wright, S. J. (2016), “Validating Sample Average Approximation Solutions with Negatively Dependent Batches,” Submitted to Computational Optimization and Applications.

5. Chen, J. and Qian, P. Z. G. (2016), “Controlling Correlations in Sliced Latin Hypercube Designs,” Under Revision by Statistica Sinica.

These papers were presented in a number of invited talks at university and international conferences, including

1. Department of Statistics, Harvard University, 2015.
2. Department of Statistics, Virginia Tech, 2015.
3. Invited Speaker, 2016 Conference on Applied Statistics in Defense, October 24-28, Washington, DC.
4. Plenary Speaker, International Conference on Advances in Interdisciplinary Statistics and Combinatorics, Greensboro, NC, September 30 - October 2, 2016.
5. 2016 Joint Statistical Meetings, Chicago, IL, August 2-7, 2016.
6. The 23th ASA/IMS Spring Research Conference on Statistics in Industry and Technology, Chicago, IL, May 25-27, 2016.
7. Keynote speaker, American Family Analytics Forum, Madison, WI, Jan, 2016.
8. 59th Annual Fall Technical Conference, October 8-9, 2015, Houston, TX.
9. The 32th Quality and Productivity Research Conference, NCSU, NC, June, 2015.
10. The 22th ASA/IMS Spring Research Conference on Statistics in Industry and Technology in Cincinnati, Ohio, May, 2015.

Number of Presentations: 3.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received Paper

TOTAL:

Number of Manuscripts:

Books

Received Book

TOTAL:

Received Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Yan Chen	0.03	
Xiao Nie	0.17	
Youran Qi	0.11	
FTE Equivalent:	**0.31**	
Total Number:	**3**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Zhiguang "Peter" Qian	0.15	
FTE Equivalent:	**0.15**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME
Xiao Nie
Yan Chen

Total Number: **2**

Names of other research staff

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

See Attachment

Technology Transfer

Scientific progress and accomplishments

Statement of the problem studied

Qian (2012) introduced sliced Latin hypercube designs for computer experiments and uncertainty quantification with qualitative and quantitative factors. These design have been widely used in US Army and other defense agencies and companies. Example applications include a high visible project done by Douglas M. Ray (Mathematical Statistician Technical Lead) at QE&SA Reliability Management Branch of US Army RDECOM ARDEC. The problem studied by this project is to find refinements of the sliced Latin hypercube designs and investigate their statistical properties for computer experiments, uncertainty quantification, sensitivity analysis and stochastic optimization. Another related problem is to use experimental design ideas to create new computational algorithms for accelerating large-scale statistical computing.

List of Appendixes, Illustrations and Tables

The Appendix lists the following illustrations:

1. Illustration of Bi-directional sliced Latin hypercube design introduced by Zhou, Jin, Qian and Zhou (2016).
2. Illustration of the group coherence of the OEM algorithm introduced by Xiong, Dai, Huling, and Qian (2016).

Summary of the most important results

1. Xiong, S., Dai, B., Huling, J. and Qian, P. Z. G. (2016), "OEM: A Design-based Algorithm for Least Squares Problems," *Technometrics*, 58, 285-293.

Summary: We introduce an efficient iterative algorithm, intended for various least squares problems, based on a design of experiments perspective. The algorithm, called orthogonalizing EM (OEM), works for ordinary least squares (OLS) and can be easily extended to penalized least squares. The main idea of the procedure is to orthogonalize a design matrix by adding new rows and then solve the original problem by embedding the augmented design in a missing data framework. We establish several attractive theoretical properties concerning OEM. For the OLS with a singular regression matrix, an OEM sequence converges to the Moore-Penrose generalized inverse-based least squares estimator. For ordinary and penalized least squares with various penalties, it converges to a point having grouping coherence for fully aliased regression matrices. Convergence and the convergence rate of the algorithm are examined. Finally, we demonstrate that OEM is highly efficient for large-scale least squares and penalized least squares problems, and is considerably faster than competing methods when n is much larger than p.

2. Zhou, Q., Jin, T., Qian, P. Z. G. and Zhou, S. (2016), "Bi-Directional Sliced Latin Hypercube Designs," Statistica Sinica, 26, 653-674.

Summary: We propose a new type of design for computer experiments called bi-directional sliced Latin hypercube design (BSLHD). The proposed design is a special Latin hypercube design (LHD) that simultaneously accommodates two slicing structures. It consists of multiple LHDs of smaller sizes, which can be joined in alternative ways to form two sets of standard sliced LHDs. These new structures are useful for computer experiments with qualitative factors, experiments running in batch mode, and ensembles of multiple computer models. Some sampling properties of the designs in estimating function means are proved and illustrated through numerical examples.

3. Nie, X., Huling, J. and Qian, P. Z. G. (2016), "Accelerating Large-scale Statistical Computation with the GOEM Algorithm," Revised for *Technometics*.

Summary: Large-scale data analysis problems have become increasingly common across many disciplines. While large volume of data offers more statistical power, it also brings computational challenges. The orthogonalizing EM algorithm is an efficient method to deal with large-scale least squares problems from a design point of view. In this paper, we propose a reformulation and generalization of the orthogonalizing EM algorithm. Computational complexity and convergence guarantees are established. The reformulation of the orthogonalizing EM algorithm leads to a reduction in computational complexity for least squares problems and penalized least squares problems. The reformulation, named the GOEM (Generalized Orthogonalizing EM) algorithm, can incorporate a wide variety of convex and non-convex penalties, including the lasso, group lasso, and minimax concave penalty penalties. The GOEM algorithm is further extended to a wider class of models including generalized linear models and Cox's proportional hazards model. Synthetic and real data examples are included to illustrate its use and efficiency compared with standard techniques.

4. Chen, J., Lim, C. H., Qian, P. Z. G., Linderoth, J. T and Wright, S. J. (2016), "Validating Sample Average Approximation Solutions with Negatively Dependent Batches," Submitted to *Computational Optimization and Applications*.

Summary: Sample-average approximations (SAA) are a practical means of finding approximate solutions of stochastic programming problems involving an extremely large (or infinite) number of scenarios. SAA can also be used to find estimates of a lower bound on the optimal objective value of the true problem which, when coupled with an upper bound, provides confidence intervals for the true optimal objective value and valuable information about the quality of the approximate solutions. Specifically, the lower bound can be estimated by solving multiple SAA problems (each obtained using a particular sampling method) and averaging the obtained objective values. State-of-the-art methods for lower-bound estimation generate batches of scenarios for the SAA problems independently. In this paper, we describe sampling methods that produce negatively dependent batches, thus reducing the variance of the sample-averaged lower bound estimator and increasing its usefulness in defining a confidence interval for the optimal objective value. We provide conditions under which the new sampling methods can reduce the variance of the lower bound estimator, and present computational results to verify that our scheme can reduce the variance significantly, by comparison with the traditional Latin hypercube approach.

5. Chen, J. and Qian, P. Z. G. (2016), "Controlling Correlations in Sliced Latin Hypercube Designs," Under Revision by *Statistica Sinica*.

Summary: A sliced Latin hypercube design is a special Latin hypercube design that can be partitioned into smaller Latin hypercube designs. We propose an algorithm to construct sliced Latin hypercube designs with controlled column-wise correlations for each slice and the entire design. The proposed algorithm can significantly decrease the column-wise correlations in each slice as the number of slices increases even if the number of runs in each slice is fixed. The algorithm can also control the quadratic canonical correlations of the larger design. It has good convergence property and is flexible in sample size. The effectiveness of the algorithm is illustrated by several examples.

These results were presented in a number of invited talks at university and international conferences, including

1. Department of Statistics, Harvard University, 2015.
2. Department of Statistics, Virginia Tech, 2015.

3. Invited Speaker, 2016 Conference on Applied Statistics in Defense, October 24-28, Washington, DC.
4. Plenary Speaker, International Conference on Advances in Interdisciplinary Statistics and Combinatorics, Greensboro, NC, September 30 - October 2, 2016.
5. 2016 Joint Statistical Meetings, Chicago, IL, August 2-7, 2016.
6. The 23th ASA/IMS Spring Research Conference on Statistics in Industry and Technology, Chicago, IL, May 25-27, 2016.
7. Keynote speaker, American Family Analytics Forum, Madison, WI, Jan, 2016.
8. 59th Annual Fall Technical Conference, October 8-9, 2015, Houston, TX.
9. The 32th Quality and Productivity Research Conference, NCSU, NC, June, 2015.
10. The 22th ASA/IMS Spring Research Conference on Statistics in Industry and Technology in Cincinnati, Ohio, May, 2015.

Bibliography

1. Xiong, S., Dai, B., Huling, J. and **Qian, P. Z. G.** (2016), "OEM: A Design-based Algorithm for Least Squares Problems," *Technometrics*, 58, 285-293.

2. Zhou, Q., Jin, T., **Qian, P. Z. G.** and Zhou, S. (2016), "Bi-Directional Sliced Latin Hypercube Designs," *Statistica Sinica*, 26, 653-674.

3. Nie, X., Huling, J. and **Qian, P. Z. G.** (2016), "Accelerating Large-scale Statistical Computation with the GOEM Algorithm," Revised for *Technometics*.

4. Chen, J., Lim, C. H., **Qian, P. Z. G.**, Linderoth, J. T and Wright, S. J. (2016), "Validating Sample Average Approximation Solutions with Negatively Dependent Batches," Submitted to *Computational Optimization and Applications*.

5. Chen, J. and **Qian, P. Z. G.** (2016), "Controlling Correlations in Sliced Latin Hypercube Designs," Under Revision by *Statistica Sinica*.

6. **Qian, P. Z. G.** (2012), "Sliced Latin Hypercube Designs," *Journal of the American Statistical Association*, 107, 393-399.

Appendix

1. Illustration of Bi-directional sliced Latin hypercube design from our paper (Zhou, Jin, Qian and Zhou, 2016).

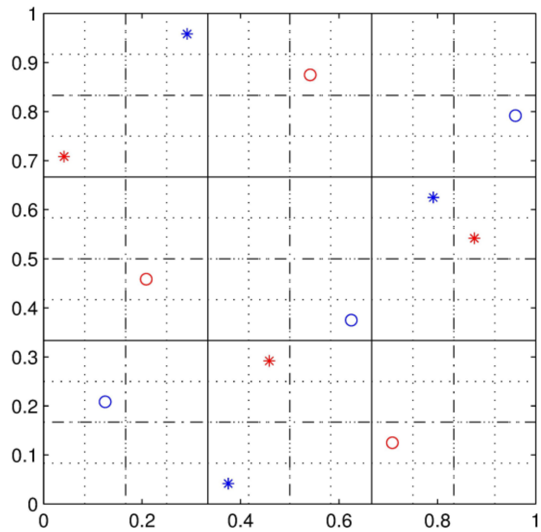
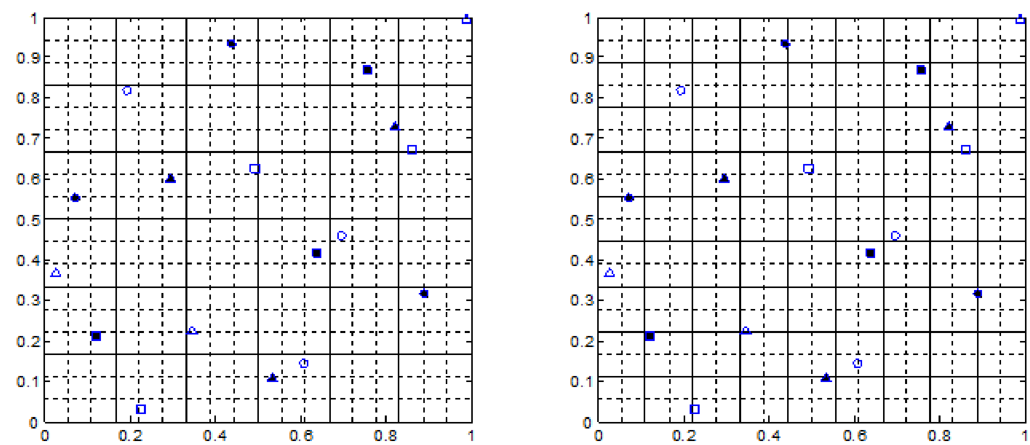


Figure 1: An example of DSLHD(3,2,2,2).



(a) three 6-run row-slices ($\mathbf{D}_{1\cdot}$, $\mathbf{D}_{2\cdot}$, $\mathbf{D}_{3\cdot}$)   (b) two 9-run column-slices ($\mathbf{D}_{\cdot 1}$, $\mathbf{D}_{\cdot 2}$)

Figure 2: An example of BSLHD (3,3,2,2).

Figures 1 and 2 present a bi-directional sliced Latin hypercube design. Such designs are useful for running expensive computer simulations and uncertainty quantification in batches or with mixed inputs and have more flexible slicing structure that sliced Latin hypercube designs (Qian, 2012) that previously invented by Prof. Qian's research group.

2. Illustration of the group coherence of the OEM algorithm introduced by Xiong, Dai, Huling, and Qian (2016).

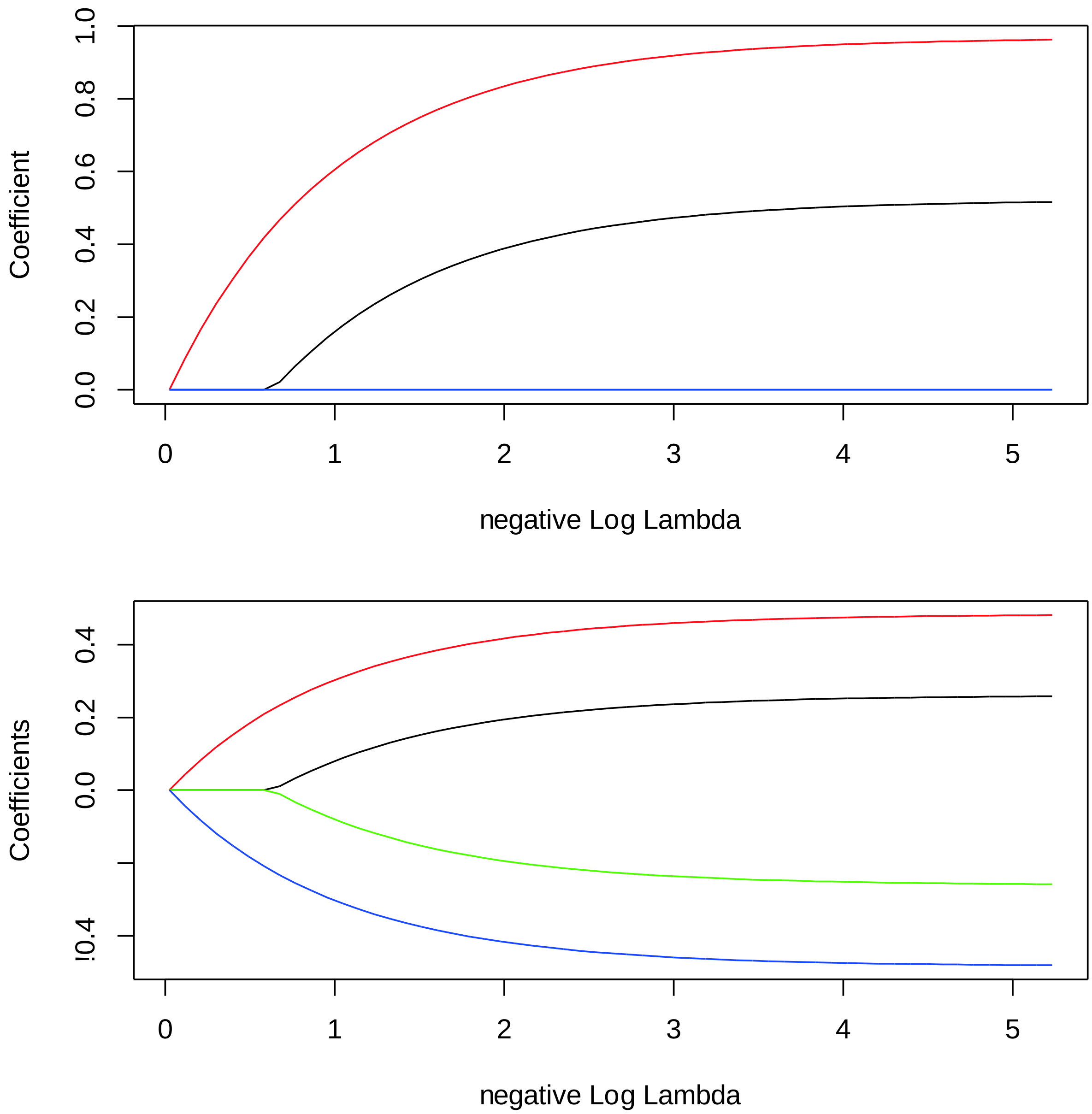


Figure 3: Solution paths of the lasso fitted by CD (the upper panel) and OEM (the lower panel) developed in Xiong, Dai, Huling, and Qian (2016).

Figure 3 displays the solution paths for the data using the lasso fitted by **R** packages **glmnet** and **oem** on the same set of tuning parameters lambda. The package **lars** gives the same solution path as **glmnet**. This figure reveals that OEM estimates the perfectly negative correlated pairs to have exactly the opposite signs but CD only has x1 and x2 in the model and fixes x3 and x4 to be zero for any lambda. This difference is due to the fact that in every iteration, both CD and LARS will find the predictor with the largest improvement on the target function and if more than one coordinates can give better results, only the one with the smallest index will enter the model. OEM considers all the predictors in every iteration equally, so the ones with same contribution to the target will receive equal steps. The grouping coherence property of OEM also holds for non-convex penalties such as SCAD, with the solution paths shown in Figure 3, where the same data are used as above for the lasso.