Optimization of a computationally expensive simulator with quantitative and qualitative inputs

Jhouben Janyk Cuesta Ramirez

▶ **To cite this version:**

Jhouben Janyk Cuesta Ramirez. Optimization of a computationally expensive simulator with quantitative and qualitative inputs. Other. Université de Lyon, 2022. English. NNT : 2022LYSEM010 . tel-03845141

HAL Id: tel-03845141
https://theses.hal.science/tel-03845141

Submitted on 9 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre NNT : 2022LYSEM010

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
L'Ecole des Mines de Saint-Etienne

Ecole Doctorale N° 488
Sciences, Ingénierie, Santé

Spécialité de doctorat : Mathématiques Appliquées
Discipline : Science des Données

Soutenue publiquement le 14/06/2022, par :

Jhouben Janyk Cuesta Ramirez

Optimisation de codes numériques coûteux en présence de variables quantitatives et qualitatives

Devant le jury composé de :

Pascal Lafon	*Professeur, Univ. Technologie de Troyes*	*Président*
Mathieu Balesdent	*Ingénieur de recherche HDR, ONERA*	*Rapporteur*
Joseph Morlier	*Professeur, ISAE Toulouse*	*Examinateur*
Delphine Sinoquet	*Ingénieure de recherche, IFPEN Rueil Malmaison*	*Examinatrice*
Rodolphe Le Riche	*DR CNRS, Mines St-Etienne et LIMOS*	*Directeur*
Olivier Roustant	*Professeur, INSA Toulouse*	*Co-Directeur*
Guillaume Perrin	*Chercheur, Univ. Gustave Eiffel Paris-Est*	*Encadrant*
Cédric Durantin	*Ingénieur de recherche, CEA DAM Bruyères-le-Châtel*	*Encadrant*
Thomas Perrillat-Bottonet	*Ingénieur de recherche, CEA LETI Grenoble*	*Invité*
Alain Glière	*ex-Ingénieur de recherche, CEA LETI Grenoble*	*Invité*

Spécialités doctorales	Responsables :	Spécialités doctorales	Responsables
SCIENCES ET GENIE DES MATERIAUX	K. Wolski Directeur de recherche	MATHEMATIQUES APPLIQUEES	M. Batton-Hubert
MECANIQUE ET INGENIERIE	S. Drapier, professeur	INFORMATIQUE	O. Boissier, Professeur
GENIE DES PROCEDES	F. Gruy, Maître de recherche	SCIENCES DES IMAGES ET DES FORMES	JC. Pinoli, Professeur
SCIENCES DE LA TERRE	B. Guy, Directeur de recherche	GENIE INDUSTRIEL	N. Absi, Maitre de recherche
SCIENCES ET GENIE DE L'ENVIRONNEMENT	V.Laforest, Directeur de recherche	MICROELECTRONIQUE	Ph. Lalevée, Professeur

EMSE : Enseignants-chercheurs et chercheurs autorisés à diriger des thèses de doctorat (titulaires d'un doctorat d'État ou d'une HDR)

ABSI	Nabil	MR	Génie industriel	CMP
AUGUSTO	Vincent	MR	Génie industriel	CIS
AVRIL	Stéphane	PR	Mécanique et ingénierie	CIS
BADEL	Pierre	PR	Mécanique et ingénierie	CIS
BALBO	Flavien	PR	Informatique	FAYOL
BASSEREAU	Jean-François	PR	Sciences et génie des matériaux	SMS
BATTON-HUBERT	Mireille	PR	Mathématiques appliquées	FAYOL
BEIGBEDER	Michel	MA	Informatique	FAYOL
BILAL	Blayac	DR	Sciences et génie de l'environnement	SPIN
BLAYAC	Sylvain	PR	Microélectronique	CMP
BOISSIER	Olivier	PR	Informatique	FAYOL
BONNEFOY	Olivier	PR	Génie des Procédés	SPIN
BORBELY	Andras	DR	Sciences et génie des matériaux	SMS
BOUCHER	Xavier	PR	Génie Industriel	FAYOL
BRUCHON	Julien	PR	Mécanique et ingénierie	SMS
CAMEIRAO	Ana	PR	Génie des Procédés	SPIN
CHRISTIEN	Frédéric	PR	Science et génie des matériaux	SMS
DAUZERE-PERES	Stéphane	PR	Génie Industriel	CMP
DEBAYLE	Johan	MR	Sciences des Images et des Formes	SPIN
DEGEORGE	Jean-Michel	MA	Génie industriel	Fayol
DELAFOSSE	David	PR	Sciences et génie des matériaux	SMS
DELORME	Xavier	PR	Génie industriel	FAYOL
DESRAYAUD	Christophe	PR	Mécanique et ingénierie	SMS
DJENIZIAN	Thierry	PR	Science et génie des matériaux	CMP
BERGER-DOUCE	Sandrine	PR	Sciences de gestion	FAYOL
DRAPIER	Sylvain	PR	Mécanique et ingénierie	SMS
DUTERTRE	Jean-Max	PR	Microélectronique	CMP
EL MRABET	Nadia	MA	Microélectronique	CMP
FAUCHEU	Jenny	MA	Sciences et génie des matériaux	SMS
FAVERGEON	Loïc	MR	Génie des Procédés	SPIN
FEILLET	Dominique	PR	Génie Industriel	CMP
FOREST	Valérie	PR	Génie des Procédés	CIS
FRACZKIEWICZ	Anna	DR	Sciences et génie des matériaux	SMS
GAVET	Yann	MA	Sciences des Images et des Formes	SPIN
GERINGER	Jean	MA	Sciences et génie des matériaux	CIS
GONDRAN	Natacha	MA	Sciences et génie de l'environnement	FAYOL
GONZALEZ FELIU	Jesus	MA	Sciences économiques	FAYOL
GRAILLOT	Didier	DR	Sciences et génie de l'environnement	SPIN
GRIMAUD	Frederic	EC	Génie mathématiques et industriel	FAYOL
GROSSEAU	Philippe	DR	Génie des Procédés	SPIN
GRUY	Frédéric	PR	Génie des Procédés	SPIN
HAN	Woo-Suck	MR	Mécanique et ingénierie	SMS
HERRI	Jean Michel	PR	Génie des Procédés	SPIN
ISMAILOVA	Esma	MC	Microélectronique	CMP
KERMOUCHE	Guillaume	PR	Mécanique et Ingénierie	SMS
KLOCKER	Helmut	DR	Sciences et génie des matériaux	SMS
LAFOREST	Valérie	DR	Sciences et génie de l'environnement	FAYOL
LERICHE	Rodolphe	DR	Mécanique et ingénierie	FAYOL
LIOTIER	Pierre-Jacques	MA	Mécanique et ingénierie	SMS
MEDINI	Khaled	EC	Sciences et génie de l'environnement	FAYOL
MOLIMARD	Jérôme	PR	Mécanique et ingénierie	CIS
MOULIN	Nicolas	MA	Mécanique et ingénierie	SMS
MOUTTE	Jacques	MR	Génie des Procédés	SPIN
NAVARRO	Laurent	MR	Mécanique et ingénierie	CIS
NEUBERT	Gilles	PR	Génie industriel	FAYOL
NIKOLOVSKI	Jean-Pierre	Ingénieur de recherche	Mécanique et ingénierie	CMP
O CONNOR	Rodney Philip	PR	Microélectronique	CMP
PICARD	Gauthier	PR	Informatique	FAYOL
PINOLI	Jean Charles	PR	Sciences des Images et des Formes	SPIN
POURCHEZ	Jérémy	DR	Génie des Procédés	CIS
ROUSSY	Agnès	MA	Microélectronique	CMP
SANAUR	Sébastien	MA	Microélectronique	CMP
SERRIS	Eric	IRD	Génie des Procédés	FAYOL
STOLARZ	Jacques	CR	Sciences et génie des matériaux	SMS
VALDIVIESO	François	PR	Sciences et génie des matériaux	SMS
VIRICELLE	Jean Paul	DR	Génie des Procédés	SPIN
WOLSKI	Krzystof	DR	Sciences et génie des matériaux	SMS
XIE	Xiaolan	PR	Génie industriel	CIS
YUGMA	Gallian	MR	Génie industriel	CMP

Mise à jour : 04/02/2021

Supervisor
Rodolphe Le Riche, DR CNRS, LIMOS et École des Mines Saint-Étienne, France

Co-supervisor
Olivier Roustant, Professeur, INSA Toulouse
Guillaume Perrin, Chercheur, Univ. Gustave Eiffel Paris-Est
Cédric Durantin, ingénieur de recherche, CEA DAM Bruyères-le-Châtel

Jury
Pascal Lafon, Professeur, Univ. Technologie de Troyes
Delphine Sinoquet, ingénieure de recherche, IFPEN Rueil Malmaison
Joseph Morlier, professeur, ISAE Toulouse
Mathieu Balesdent, ingénieur de recherche HDR, ONERA

Guest
Thomas Perrillat-Bottonet, Ingénieur de recherche, CEA-LETI
Alain Glière, ex-Ingénieur de recherche, CEA-LETI

Opponent
Jhouben Custa-Ramirez

Contact information

Department of Mathematics and Industrial Engineering
Institut Fayol, Mines Saint-Étienne
158 cours Fauriel, F-42023 Saint-Étienne cedex 2, France

Email address: webmaster@emse.fr
URL: https://www.mines-stetienne.fr/
Telephone: +33 (0)4 77 42 01 23
Fax: +33 (0)4 77 42 00 00

Copyright © 2022 Jhouben Janyk Cuesta Ramirez

Acknowledgements

The more you experience life, the harder it gets to build a clear picture of all the decisions that made you who you currently are. All the small interactions, the doubts, the fears, the determinations, the dreams and the most important part, the people around you. Even though it is my intention in this paragraph to preserve the names of those who contributed to my choices even without noticing, such a memory exercise would require an expensive simulator with millions of calls. So, instead of names I want to provide you, a cheaper, faster but accurate enough list of contributions. I want to thank everyone if possible, from the woman who became the reflection of my soul for being there even in the most difficult moments to people whose appearance in my life was brief. I also want to thank the amazing friends I made, they surely still remember me smiling with them. Also I want to thank all the people who provide me with guidance, academically and personally; yes even a small guidance of "take this bus to get there" classify as a guidance. In short, I thank to my fiancée "Muricata", to family (mine and hers), friends, colleagues, directors, neighbors, and all the people whom I can still have a memory of.

Grenoble - France, June 16 of 2022
Jhouben Janyk Cuesta Ramirez

Optimisation de codes numériques coûteux en présence de variables quantitatives et qualitatives

Jhouben Janyk Cuesta Ramirez

Génie Mathématique et Industriel
Institut Fayol, Mines Saint-Étienne
158 cours Fauriel, F-42023 Saint-Étienne cedex 2, France
jhouben.cuesta@emse.fr

Dans cette thèse, les problèmes d'optimisation mixtes coûteux sont abordés par le biais de processus gaussiens où les variables discrètes sont relaxées en variables latentes continues. L'espace continu est plus facilement exploité par les techniques classiques d'optimisation bayésienne que ne le serait un espace mixte. Les variables discrètes sont récupérées soit après l'optimisation continue, soit simultanément avec une contrainte supplémentaire de compatibilité continue-discrète qui est traitée avec des Lagrangiens augmentés. Plusieurs implémentations possibles de ces optimiseurs mixtes bayésiens sont comparées. En particulier, la reformulation du problème avec des variables latentes continues est mise en concurrence avec des recherches travaillant directement dans l'espace mixte. Parmi les algorithmes impliquant des variables latentes et un Lagrangien augmenté, une attention particulière est portée aux multiplicateurs de Lagrange pour lesquels des techniques d'estimation locale et globale sont étudiées. Les comparaisons sont basées sur l'optimisation répétée de trois fonctions analytiques et sur une application mécanique concernant la conception d'une poutre. Une étude supplémentaire dans le domaine de l'auto-calibrage est faite, dont une des perspectives est l'application de la stratégie d'optimisation mixte proposée. Cette application concerne la quantification des radionucléides, qui définit une fonction inverse spécifique nécessitant l'étude de ses multiples propriétés. Nous réalisons cette étude dans un scénario continu. Une proposition de différentes stratégies déterministes et bayésiennes a été faite en vue d'une définition ultérieure dans un contexte de variables mixtes.

Plus précisément, le chapitre 2 examine le cadre de l'optimisation bayésienne pour les fonctions continues coûteuses. Ensuite, les principaux ingrédients de l'optimisation bayésienne, la technique de remplissage d'espace, le métamodèle du processus gaussien et les critères d'acquisition de l'amélioration attendue sont présentés.

Dans le chapitre 3, nous passons en revue les approches déterministes, bayésiennes et stochastiques les plus courantes pour traiter les problèmes inverses mal posés. De plus, nous présentons la mise en place de techniques telles que les moindres carrés régularisés, le maximum a posteriori et le Monte Carlo par chaîne de Markov.

Le chapitre 4 présente la principale contribution de cette thèse, l'algorithme LV-EGO,

une méthodologie capable d'effectuer une optimisation bayésienne de fonctions mixtes coûteuses. Cette méthode propose une relaxation de l'espace mixte vers un espace continu par le biais de la cartographie des variables latentes, tout en préservant le lien de cette relaxation comme une contrainte discrète lors de l'optimisation des critères d'acquisition. Cela nous a conduit à proposer une variante plus robuste basée sur les Lagrangiens augmentés et de nombreuses autres variantes pour traiter cette contrainte. Toutes les méthodes proposées ont été comparées aux stratégies de l'état de l'art sans exigence de cartographie et/ou de méta-modèle. Toutes les stratégies ont été comparées entre différentes fonctions déterministes et une application en mécanique.

Dans le chapitre 5, nous considérons une famille de problèmes inverses. Le Chapitre se concentre sur l'étude de différents scénarios pour une forme spécifique de problème inverse qui n'a pas été étudiée dans l'état de l'art et qui apparaît couramment dans le domaine de la spectrométrie gamma. Les scénarios proposés impliquent une fonction déterministe et un simulateur non linéaire de type boîte noire. Bien qu'ils soient définis pour des entrées continues uniquement, ces scénarios représentent un cadre à partir duquel la méthodologie LV-EGO pourrait être appliquée. Enfin, dans le chapitre 6, nous discutons de plusieurs lignes de recherche futures, à la fois théoriques et concernant les applications mixtes réelles au CEA.

Contents

List of Figures **xi**

List of Tables **xv**

1 Engineering problems with mixed variables **1**
1.1 Structure of the manuscript 4
1.2 Scientific contributions 4

2 Global Continuous Optimization of Expensive Functions **7**
2.1 Expensive Functions and Global Optimum 7
2.2 Bayesian Optimization 9
2.2.1 Space-filling techniques 9
2.2.2 Gaussian Process Metamodel 10
2.2.3 Acquisition Criterion 11

3 Solution of continuous ill-posed inverse problems **13**
3.1 Deterministic Approach 13
3.2 Bayesian Approach 14
3.2.1 Stochastic Sampling 15

4 Optimization of an expensive simulator with mixed variables **19**
4.1 Latent Variable EGO 22
4.1.1 The vanilla LV-EGO algorithm 23
4.1.2 LV-EGO algorithms with Augmented Lagrangian 25
4.2 Description of the numerical experiments 30
4.2.1 Test cases 32
4.2.2 Experiments setup and metrics 34
4.3 Results and discussion 34
4.3.1 Analytical test functions 35
4.3.2 Beam bending application 38
4.3.3 Summarized results 41
4.4 Complementary heuristics 43
4.4.1 Static exponential penalty 44

4.4.2 Adaptive exponential penalty 46
4.4.3 Updated results 48
4.4.4 Re-visiting the Beam Bending Problem 49
4.5 Conclusions 50

5 Inversion of a bi-linear black-box function 53
5.1 Scenario 1: $f(x) = X(x)\beta$ 54
5.1.1 Least-Squares approach 54
5.1.2 Gaussian approach 56
5.1.3 MCMC-based Bayesian approach 57
5.1.4 Computer Experiments 58
5.2 Scenario 2: $M(x)$, a non-linear black-box function 67
5.2.1 Experiments MCMC-based Bayesian approach 68
5.3 Conclusions 69

6 Conclusions and perspective 75

Appendices 79

A Mixed Optimization 81
A.1 Complements on the augmented Lagrangians 81
A.1.1 Random Forest Regression 85
A.2 Supplementary results Including Heuristics 87

B Mixed Inversion 91
B.1 Deterministic expression for m 91
B.2 Including m penalty term 96
B.3 General expression for any $f(X) = X\beta$ 97
B.4 Complete Results By Configuration 102

References 109

List of Figures

1.1 Diagram of a gamma spectrometry setup (Guillot [2015]). 4

2.1 Diagram of population-based optimization. 8
2.2 Diagram of Surrogate-based Optimization. 9

3.1 Sample path representing a good mixture. 17
3.2 Sample path representing bad mixture. 17
3.3 Sample path representing a good autocorrelation. 17
3.4 Sample path representing bad autocorrelation. 17

4.1 Two of the test functions with 1 discrete variable. 32
4.2 Comparison of all 9 algorithms on the Branin function. $y^\star = 2.79118$. . . 36
4.3 Comparison of the 9 algorithms on the Goldstein function. $y^\star = 3$. 37
4.4 Comparison of the 9 algorithms on the Hartmann function (for which $y^\star = -3.32237$). 38
4.5 Comparison of all 9 algorithms on the beam design test case ($y^\star = 1.28738$). 39
4.6 Shapes of the considered beam profiles. The scale differs from one picture to another, as the areas are supposed to be the same for each cross-section. From Roustant et al. [2020]. 40
4.7 Representation of the correlation between the latent variables at various iterations t. The correlations correspond to the categorical kernel of Equation (4.5). The levels were grouped according to $\tilde{I}$: $\{1, 4, 7, 10\}$, $\{2, 5, 8, 11\}$, $\{3, 6, 9, 12\}$. 40
4.8 Comparison of the 9 algorithms tested with results averaged over all test cases. 42
4.9 Comparison of LV-EGO with and without (NR-LV-EGO) a repeated estimation of the latent variables. Results for the beam design application. 42
4.10 Comparison of the 11 algorithms tested with results averaged over all test cases. By including LV-EGO-a and -s, the relative targets differ with respect to Figure 4.8 which explains changes in total performance. 48
4.11 Comparison of all 11 algorithms on the beam design test case ($y^\star = 1.28738$). By including LV-EGO-a and -s, the relative targets differ with respect to Figure 4.5 which explains changes in total performance. 49

4.12 Representation of the correlation between the latent variables at various iterations t. The correlations correspond to the categorical kernel of Equation (4.5). The levels were grouped according to $\tilde{I}$: $\{1, 4, 7, 10\}$, $\{2, 5, 8, 11\}$, $\{3, 6, 9, 12\}$. 50

5.1 Results for the strategy 1 $n_c + 1$ MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 61
5.2 Trace plots for the strategy 1 $n_c + 1$ MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). 62
5.3 Autocorrelation plots for the strategy 1 $n_c + 1$ MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). 63
5.4 Results for the strategy 2 MC within MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 65
5.5 Trace plots for the strategy 2 MC within MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). 66
5.6 Autocorrelation plots for the strategy 2 MC within MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). 67
5.7 Results for the strategy 3 point-wise MC, for configurations C1 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$. 71
5.8 Obtained densities for the point-wise MC strategy. Red dashed line represents the MAP estimate. Golden dashed line represents $m^\star$ 72
5.9 Obtained densities for the strategies MCMC (left), MC within MCMC (middle) and point-wise MC (right) for configurations C1 (top), C2 (middle) and CT (bottom). Red dashed line represents the MAP estimate. Golden dashed line represents $m^\star$ 72
5.10 Obtained trace (left) and autocorrelation (right) plots for the MCMC strategy. 73
5.11 Obtained trace (left) and autocorrelation (right) plots for the MC within MCMC strategy. 74

6.1 (a) Conventional camera system based on a color filter array (Bayer mosaic), an image sensor with the IR cut-off filter (IRCF). (b) Spectral sensitivity of the camera system. Source: Park and Kang [2016]. 76

A.1 Sketch of Rockafellar's augmented Lagrangian for $\rho \approx 0$ in blue and $\rho > 0$ in red. x^1 is infeasible, x^2 feasible (and $g(x^2) < -\lambda/\rho$) and $x^\star$ is an optimum with $g(x^\star) = 0$. The black highlighted curves are the approximation to the dual function, $\widehat{D}(\lambda)$ for $\mathbf{X} = \{x^1, x^2, x^\star\}$, for $\rho \approx 0$ and $\rho > 0$. There is no saddle point and a duality gap with the blue set of curves in that $x^\star \notin \arg\min_x L_A(x; \lambda^\star, \rho \approx 0)$ and $\widehat{D}(\lambda^\star) = min_x L_A(x; \lambda^\star, \rho \approx 0) < L_A(x^\star; \lambda^\star, \rho \approx 0)$, i.e., minimizing the augmented Lagrangian does not lead to the result of the problem. However, by increasing ρ, it is visible that the y-intercept of the infeasible points increase so that one always reaches a state where $x^\star = \arg\min_x L_A(x; \lambda^\star, \rho)$ as in the red set of curves. A similar illustration can be done with the augmented Lagrangian with equality constraint: $f(x) + \rho/2h^2(x)$ is the y-intercept and $h(x)$ is the slope of the augmented Lagrangian associated to x. The main difference is that all points contribute linearly in terms of λ to $L_A(x; \lambda, \rho)$. 84

A.2 Comparison of all 11 algorithms on the Branin function. $y^\star = 2.79118$. . 87

A.3 Comparison of the 11 algorithms on the Golstein function. $y^\star = 3$. 88

A.4 Comparison of the 11 algorithms on the Hartmann function (for which $y^\star = -3.32237$). 89

B.1 Graphical representation of the function $J(m)$, for real values of m. This function reach its minimum value on the model estimator $\hat{m}$ 92

B.2 Estimations of the m for different configurations. Top Left:$\gamma = 0.05$, $m = 0.3$, $\beta = 0.1$,$\hat{m} = 0.30824$. Top Right: $\gamma = 0.05$, $m = 0.3$, $\beta = 3$, $\hat{m} = 0.2991$. Bottom Left: $\gamma = 0.5$, $m = 0.3$, $\beta = -0.1$, $\hat{m} = 0.2045$. Bottom Right: $\gamma = 0.5$, $m = 0.3$, $\beta = 4$, $\hat{m} = 0.3071$. 94

B.3 Results for the strategy 1 $n_c + 1$ MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 102

B.4 Results for the strategy 1 $n_c + 1$ MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 103

B.5 Results for the strategy 1 $n_c + 1$ MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 104

B.6 Results for the strategy 2 MC within MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 105

B.7 Results for the strategy 2 MC within MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 106
B.8 Results for the strategy 2 MC within MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 107
B.9 Results for the strategy 3 point-wise MC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$ 108

List of Tables

4.1 Numerical complexities of the algorithms compared at each iteration (for a given t). 25
4.2 Summary of the 9 algorithms tested: name, space over which it is defined (mixed versus continuous with latent variables), metamodel used, acquisition criterion, optimizer of the acquisition criterion. 30
4.3 Dimensions and DoE size of the test cases. 35
4.4 $P^{(t)}$ expected variations 44

5.1 Summarized results for the different configurations under 1000 samples from prior 59
5.2 Summarized results Strategy 1 at 90 confidence interval 64
5.3 Summarized results Strategy 2 at 90 confidence interval 64
5.4 Summarized results Strategy 3 at 90 confidence interval 68

Chapter 1

Engineering problems with mixed variables

Contents

1.1 Structure of the manuscript 4

1.2 Scientific contributions 4

A key task in engineering design is to find an optimal configuration from a very large set of alternatives. When the performance of the candidate solutions is measured through a realistic simulation, the numerical cost of the procedure becomes a bottleneck. The optimization of computationally expensive simulators is a topic widely studied in the literature Thi et al. [2019]. In this field of study, we focus on Bayesian optimization (BO), which is particularly suitable for solving such problems Frazier [2018]. Bayesian optimization is a sequential design strategy that requires a data-driven mathematical model or metamodel that provides predictions along with their uncertainty Bartz-Beielstein et al. [2019]. The metamodel replaces some of the calls to the expensive simulation and is a key ingredient to the optimization of costly functions. An acquisition criterion Wilson et al. [2018] aggregates the spatial predictions and uncertainties. The metamodel is trained from a reduced set of simulation data and the acquisition criterion is maximized to propose new configurations to be simulated at the next iteration. When the acquisition criterion is the expected improvement (EI), as first introduced in Mockus et al. [1978], the BO algorithm is often called EGO (Efficient Global Optimization, Jones et al. [1998]). EGO is currently a state-of-the-art approach to medium size, continuous and costly optimization problems, both from an empirical Le Riche and Picheny [2021] and a theoretical point of view Vazquez and Bect [2010].

However, in realistic settings, some of the decision variables are categorical. In structural design for example, the type of material, the number of components, the choice between alternative technologies lead to discrete variables with no obvious distance between them. The combination of continuous and categorical variables is called a mixed optimization problem. In non-costly cases, mixed optimization problems can be approached by Mixed-Integer NonLinear Programming Belotti et al. [2013] (when the discrete variables are

integers), by sampling based techniques such as evolutionary optimization Cao et al. [2000], Emmerich et al. [2008], Ocenasek and Schwarz [2002] or by alternating mixed programming Audet and Dennis Jr [2001].

When the objective function is costly, mixed optimization problems remain challenging and a topic for research. Bartz-Beielstein and Zaefferer [2017] provides an overview of metamodels that have or can be used in optimization when the variables are continuous or discrete. Bayesian optimization methods have already been extended to mixed problems. It was made possible by creating GP kernels (covariance functions) in mixed variables as a combination of continuous and discrete kernels. The acquisition function is defined over the same space as the objective function. Therefore maximizing the acquisition function is also a mixed variables problem.

To the best of our knowledge, the first EGO-like algorithm for mixed variables has been proposed in Hutter et al. [2011]. In this article, the mixed kernel is a product of continuous and discrete Gaussian kernels, and random forests constitute an alternative choice of mixed metamodel. More precisely, the discrete kernel is based on the hamming (also known as Gower) distance for ordinal or nominal variables, respectively. In Hutter et al. [2011], the expected improvement is first optimized with a multi-start local search for both continuous and discrete variables (thus a neighborhood for the discrete variables is defined) which is then complemented by a random search. This work was continued with the REMBO method in Wang et al. [2016], where a random linear embedding is introduced to tackle high-dimensional problems. Discrete variables were relaxed into continuous variables thanks to a mapping function. The optimization of the acquisition function was made with a combination of the DIRECT and CMA-ES continuous global optimizers. Both Hutter et al. [2011] and Wang et al. [2016] have been motivated by applications to the automatic configuration of algorithms. The goal of reaching very high dimensions (millions) probably forced the authors to use isotropic kernels.

A Bayesian mixed optimizer is presented in Pelamatti et al. [2019]. The GP kernels are products of continuous and discrete kernels. Different discrete kernels are compared, namely the homo- and hetero-scedastic hypersphere decomposition and the compound symmetric kernels. The optimization of the acquisition function is performed with a genetic algorithm in mixed variables. A similar BO with mixed kernel is described in Zuniga and Sinoquet [2020], but the expected improvement is optimized with the mixed version of the MADS algorithm Audet and Dennis Jr [2001] and a neighborhood for the categorical variables is defined through a probabilistic model. Random forests can replace the kriging model in BO with mixed inputs as they natively have a measure of prediction uncertainty. Such an implementation, first done in Hutter et al. [2011], is part of the `mlrMBO` `R` package Bischl et al. [2018], in conjunction with several acquisition criteria that can be optimized with a "focus-search" algorithm. The focus-search algorithm hierarchically samples the search space of the chosen acquisition criterion.

Recent developments in metamodels involving mixed variables assume that it is possible

to map categorical variables into quantitative un-observed continuous *latent variables* Zhang et al. [2019]. Whenever it is possible to write a model of the studied system, there often exist quantitative latent variables that describe the effects of the categorical variables. Typically, there are more latent variables than categorical ones. The existence of continuous latent variables can sometimes be established from the physics of the considered phenomena, e.g. in material science Zhang et al. [2020]. In structural mechanics for example, if the categorical variable describes the shape and the material of an element load in flexion, its bending moment of inertia is a candidate latent variable. Latent variables can emulate the properties of the original categorical variables, in particular within the metamodel, and open the way to reasoning with continuous quantities: the kernels of the Gaussian processes can be taken as continuous, gradients and neighborhoods are naturally defined during the optimization. On the contrary, categorical variables and their inherent lack of distance definition is the cause of complications in the kernel definition and in the optimization.

In this thesis we present a new Bayesian optimization algorithm for mixed variables called LV-EGO (for Latent Variable EGO). Our contribution with respect to Zhang et al. [2020] is that the continuity of latent variables is also considered during the optimization of the acquisition criterion. This implies that categorical variables must be recovered from the continuous latent variables, which creates a new "pre-image" problem that we tackle with a novel methodology based on augmented Lagrangian.

As optimization techniques have been recently used in inverse problems Ye et al. [2019], Kunze et al. [2021], we also address the possibility of applying our LV-EGO methodology for inverse problems with mixed variables and involving expensive simulators. Motivated by an industrial application we consider a bi-linear and expensive to evaluate function. That kind of problem seems not to be investigated much in the literature. More precisely, consider a continuous bi-linear inverse problem $y = mf(x) + \eta$, with $y \in \mathbb{R}^p$, unknowns $m \in \mathbb{R}$, and $x \in \mathbb{R}^d$, where f is an expensive-to-evaluate function. This specific product form $mf(x)$ makes the inverse problem ill-posed. This type of ill-posedness, as the product is scalar-vector, differs with classical bi-linear and self-calibration forms already studied in the literature like image blind deconvolution, compressed sensing and other applications Idier and Blanc-Féraud [2008] Ling [2017].

This particular setup appears as a calibration formulation for a radionuclide quantification application in Clement et al. [2018], where the aim is to quantify the radionuclide mass m inside a nuclear waste container defined by mixed source properties x, u via a non-destructive gamma spectrometry technique Guillot [2015], Dyrcz et al. [2021], Máduar and Miranda Junior [2007]. Figure 1.1 illustrates this process. The activity of a specific radionuclide is measured for a set of known energy levels $E_1, \ldots, E_p$ where $p \approxeq 6$. Then to obtain the calibration coefficient ϵ, reliable simulations can be performed with a set of simplified environmental setting involving continuous variables: Distance, Density, Eq-Surface, Eq-Thickness; and 2 categorical variables: 3 shapes

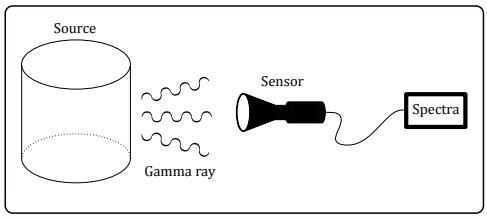


Figure 1.1: Diagram of a gamma spectrometry setup (Guillot [2015]).

(Parallelepiped, Sphere, Cylinder) and 4 materials (Iron, Vinyl, Chlorine, Plumb) as defined in Clement et al. [2018].

Motivated by this application, we study different deterministic and Bayesian approaches in order to address this specific type of inverse problem. To do that, we define different strategies for two key scenarios in a continuous setting, towards extending the proposed methodologies for a complete mixed-variable scenario.

1.1 Structure of the manuscript

This document is organized as follows. In Chapter 2 we present a brief overview of optimization of expensive functions and surrogate-based optimization, with a focus on Bayesian optimization (BO). Chapter 3 reviews the most common approaches to deal with continuous inverse problems in general. Then in Chapter 4 we present the LV-EGO strategy for optimization in presence of qualitative and quantitative variables. We define different variations of the main methodology as a way to improve its performance. We evaluate them on a set of test cases and an application in mechanics. In Chapter 5, we consider the application of LV-EGO methodology to inverse problems. Finally in Chapter 6 we discuss several future research lines, both theoretical and concerning the real mixed applications at CEA.

1.2 Scientific contributions

Results throughout this thesis are based on scientific contributions including one publication in an international journal, a technical report and communications in conferences.

Publications in international journals

1. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. A comparison of mixed-variables Bayesian optimization approaches. *Advanced Modeling and Simulation in Engineering Sciences*, special Issue on *Efficient Strategies*

for Surrogate-Based Optimization Including Multifidelity and Reduced-Order Models. Published - June 2022.

Technical Report

2. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Inversion of a costly multivariate function in presence of categorical variables. *Applied Inverse Problems Conference*, Grenoble, France, 2019. https://hal.archives-ouvertes.fr/hal-02273738.

Oral Presentations

3. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Optimization of a computationally expensive simulator with quantitative and qualitative inputs. OQUAIDO scientific days, May 2019, Nov 2019, Jun 2020, Dec 2020. Oral presentations.

4. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Optimization of a computationally expensive simulator with quantitative and qualitative inputs. CIROQUO scientific days, Jun 2021, Nov 2021. Oral presentations.

5. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Optimization of a computationally expensive simulator with quantitative and qualitative inputs. CEA-LETI PhD days, 2021. Oral Presentation.

Poster presentations

6. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Latent Variable Efficient Global Optimization for Qualitative and Quantitative Inputs. *Modeling and Numerical Methods for Uncertainty Quantification*. 2019 https://www.sigma-clermont.fr/en/mnmuq2019

7. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Optimization of a computationally expensive simulator with quantitative and qualitative inputs. CEA-LETI PhD days, 2019.

8. Cuesta-Ramirez J., Le Riche R., Roustant O., Perrin G., Durantin C., Glière A. Bayesian optimization for mixed continuous and categorical variables: A latent variable approach. MASCOT PhD student 2020 Meeting, Grenoble- France, 2020. https://www.gdr-mascotnum.fr/mascotphd20.html.

Chapter 2

Global Continuous Optimization of Expensive Functions

Contents

2.1 Expensive Functions and Global Optimum 7

2.2 Bayesian Optimization 9

2.2.1 Space-filling techniques 9

2.2.2 Gaussian Process Metamodel 10

2.2.3 Acquisition Criterion 11

The main goal of this chapter is to introduce the concepts behind the common terminology while finding the global optimum of an expensive-to-evaluate function. In particular we will present the population-based and surrogate-based frameworks as two possible choices when trying to tackle this challenge. Finally we review Bayesian Optimization (BO) as a selection of techniques to perform surrogate-based optimization in continuous inputs.

2.1 Expensive Functions and Global Optimum

In the framework of continuous global optimization, the main goal is to find, if possible, the best evaluation $x^\star$ also known as *global optimum*, from a set of feasible points $\mathcal{X}$ of a function $f : \mathbb{R}^{n_c} \to \mathbb{R}$. This is written as:

$$x^\star = \text{argmin}_{x \in \mathcal{X} \subset \mathbb{R}^{n_c}} y(x),$$

where $y = -f$ if we want to find the maximum or $y = f$ if we want to find the minimum. In the case of the *expensive functions* that will be considered in this document, f may not have explicit formula (non available derivatives) and it is represented by point-wise evaluations $(x_i, y_i), \quad i = \{1, \ldots, n\}$ defined on a space $\mathcal{X}$, typically$\mathcal{X} \in [0, 1]^{n_c}$ Frazier [2018]. In addition, further evaluations are controlled by a fixed budget t.

Under this setting, direct search techniques that often converge to *local optima* such as the Nelder-Mead method Nelder and Mead [1965] could not be applied. However, in the literature, there are different techniques that ensure convergence to the global optimum and could be divided in *population-based*, such as Covariance Matrix Adaptation Hansen [2006] and *surrogate-based*, such as Efficient Global Optimization Jones et al. [1998].

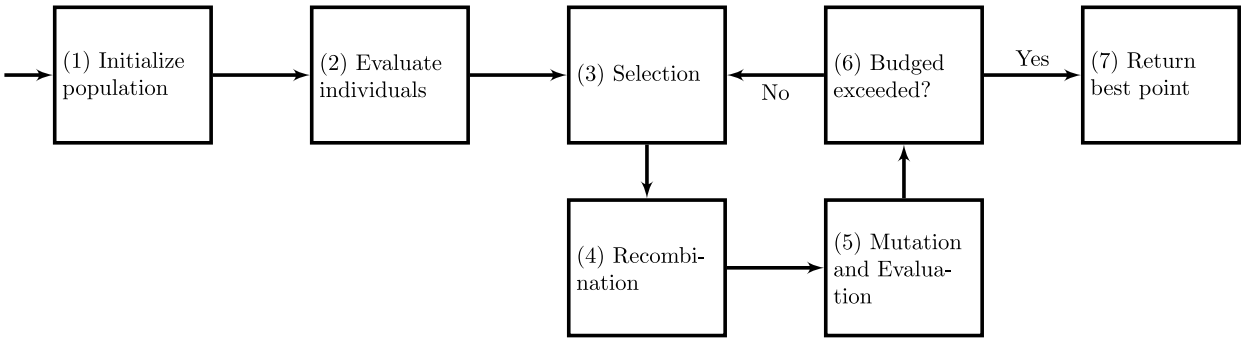


Figure 2.1: Diagram of population-based optimization.

On one hand, a *population-based* strategy proposes a set of steps inspired by biology, where each individual $x_i^{(t)}$ has the capability of evolution due to his adaptability, thereby producing better solutions $x_i^{(t+1)}$. Figure 2.1 shows the basic steps of evolutionary algorithms, where it starts by *selecting* individuals within a population of size n, as parents for the reproduction (*recombination*) step. Later, each newborn individual (or its parents) could experience a random *mutation* before its further *fitness* evaluation is made. This process is repeated until the budget t is consumed. At the end, the best individual among all the different iterations $x^\star$ and its evaluation $y^\star$ are returned.

Evolution strategies can be applied to any type of variable (continuous, discrete) and the implementation is very simple. As a drawback, they require a larger number of evaluations Elsawy et al. [2019] compared with surrogate-based methods. It is important to remark that there exist many different evolutionary strategies that vary in the way of dealing with the steps previously mentioned. For a review of population-based evolutionary algorithms we refer to Slowik and Kwasnicka [2020], and Boussaïd et al. [2013] for more general meta-heuristic methods.

On the other hand, *surrogate-based* strategies propose a way to mimic and further replace the current expensive function y, by a cheaper-to-evaluate model $Y(x)$. After executing the *first* step of its sequence (see Figure 2.2), we obtain a pair $(\mathbf{X}, \mathbf{Y})$ also called Design of Experiment (DoE) that corresponds to a set of initial informative points of size N_{DoE} already evaluated with the expensive function f. This set of points $\mathbf{X}$ can be chosen at random (e.g. from the uniform distribution) or by using a space-filling technique such as Latin Hypercube Sampling (LHS) McKay et al. [1979].

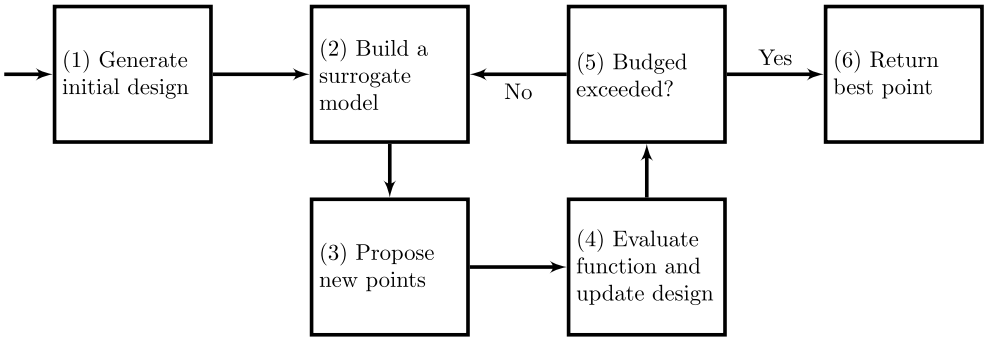


Figure 2.2: Diagram of Surrogate-based Optimization.

In the *second* step, a metamodel strategy is chosen to train the surrogate model Y with the DoE $(\mathbf{X}, \mathbf{Y})$. Then the *third* step makes use of an acquisition criteria, often a function that will propose a candidate x^t, to be evaluated through y in step 4. This process will continue until t is exhausted, and will return the best point found $x^\star$ in step 6.

Surrogate-based strategies are often more difficult to implement than population-based ones. However, as an advantage the surrogate Y can explore faster the solution set $\mathcal{X}$ which has already been exploited to improve the speed of convergence Jones et al. [1998]. When a Gaussian process is used as a surrogate, surrogate-based optimization is called Bayesian Optimization, that we now describe.

2.2 Bayesian Optimization

Bayesian Optimization (BO) is a common choice in engineering, when dealing with expensive black-box functions. Its three main ingredients are the space filling technique, the metamodel and the acquisition criterion that we now present.

2.2.1 Space-filling techniques

Optimization using expensive simulators requires generating an initial DoE $(\mathbf{X}, \mathbf{Y})$ to train Y. A well designed space-filling DoE is beneficial not only to the metamodel accuracy but also to the overall search effectiveness Tenne [2015]. In the context of BO, the black-box function may be non linear and a space filling design is used as a DoE. Furthermore, a LHS is often considered due to its good sampling properties with respect to marginals. Basically the idea of LHS is to sample points in a hypercube such that each point is the only one in each axis-aligned hyperplane containing it. Combining the two ideas, optimized versions of LHS such as the *maximin* distance criteria, where the sampled points are uniformly spaced, have been proposed in the literature Johnson et al. [1990].

Other ways to optimize LHS are the Translational Propagation LHS proposed by Viana et al. [2010], where the maximin distance can vary accordingly to the dimension n_c; and a Basic General Extension, where it is possible to construct an optimized LHS based on a previously constructed one during an iterative optimization process. The compatibility of those techniques in terms of convergence has not been studied under the BO framework, therefore are out of the scope of this document.

2.2.2 Gaussian Process Metamodel

GP based models are widely used in literature as the surrogate in BO. They are one of the most flexible statistical models available. Furthermore, in sequential designs, the uncertainty does not depend on the outputs and will be reduced by adding new points (See Eq. 2.1).

Formally, a Gaussian Process (GP) is a possible infinite collection of random variables, where any finite set of them has a joint Gaussian distribution Rasmussen and Williams [2006] Lawrence [2003]

$$Y \sim \mathcal{GP}(m(x), k(x, x')).$$

A GP is completely specified by a mean function (or trend) $m(x)$ and a covariance function (or kernel) $k(x, x')$ that represents the spatial dependence. The trend could be any function, and the kernel should fulfill the semi-definite positiveness (SDP) property:

$$\begin{gathered}\forall n \in \mathbb{N}^*, \forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}, \\ \forall x_1, x_2, \cdots, x_n \in \mathcal{X} \subset \mathbb{R}^{n_c} \\ \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j k(x_i, x_j) \geq 0\end{gathered}$$

Usually for the mean function a constant value is selected (often zero). For the covariance function there are a lot of possible choices depending on the input space properties (dimension, type of variables, etc). Usual choices are: Gaussian or Matérn kernels, but there are a lot of valid functions and combinations. More examples can be found in Rasmussen and Williams [2006].

In Gaussian process regression, the goal is to predict the response $y^{\text{new}} = Y(x^\star)$ of a new input $x^\star$, conditionally on a given training set $(\mathbf{X}, \mathbf{Y})$. These conditional distributions are Gaussian and given for centered GP's by:

$$\begin{aligned}\hat{m}(x^\star) &= k(x^\star, \mathbf{X}) k(\mathbf{X}, \mathbf{X})^{-1} \mathbf{Y} \\ \hat{\sigma}^2(x^\star) &= k(x^\star, x^\star) - k(x^\star, \mathbf{X}) k(\mathbf{X}, \mathbf{X})^{-1} k(\mathbf{X}, x^\star),\end{aligned} \tag{2.1}$$

where $k(\mathbf{X}, x^\star)$ is the $N_{\text{DoE}} \times 1$ covariance vector, $k(x^\star, \mathbf{X}) = k(\mathbf{X}, x^\star)^\top$, $k(\mathbf{X}, \mathbf{X})$ is $N_{\text{DoE}} \times N_{\text{DoE}}$, and each x_i is of dimension n_c.

Finally, the parameters for the trend, and the variance and kernel function are estimated via Maximum Likelihood Estimation (MLE). The log negative likelihood is given by:

$$-2\log p(\mathbf{Y}|\mathbf{X},\boldsymbol{\theta}) = n\log 2\pi + \mathbf{Y}^\top K^{-1}\mathbf{Y} + \log|K|,$$

where, $K = k(\mathbf{X},\mathbf{X})$ is the Gram matrix, and $\boldsymbol{\theta}$ is the vector of model parameters.

2.2.3 Acquisition Criterion

The last ingredient of BO is an acquisition criterion which is used to provide new design points. We focus on the expected improvement (EI), as an example. The basic idea of this acquisition criteria is to quantify an unvisited point $x^\star$ regarding how much its evaluation $Y(x^\star)$ is expected to be better, which in a minimization problem means smaller, than the current best evaluation $y_{\min} = \min(y(x_1),\dots,y(x_n))$. More formally EI is defined as:

$$\mathrm{EI}(x^\star) = \mathbb{E}[(y_{\min} - Y(x^\star))^+ | Y(x_1),\dots,Y(x_n)].$$

In the case of a GP Y, the EI has a closed form and can be written as a function of the trend and variance from Eq. 2.1 as :

$$\mathrm{EI}(x^\star) = \begin{cases} \hat{\sigma}(x^\star)[\mathrm{u}(x^\star)\Phi(\mathrm{u}(x^\star)) + \phi(\mathrm{u}(x^\star))] & \text{if } \hat{\sigma}(x^\star) > 0 \\ 0 & \text{if } \hat{\sigma}(x^\star) = 0, \end{cases}$$

where ϕ is the pdf of the standard normal distribution $\mathcal{N}(0,1)$, Φ its cdf, and

$$\mathrm{u}(x^\star) = \frac{y_{\min} - \hat{m}(x^\star)}{\hat{\sigma}(x^\star)}.$$

A new candidate point to be evaluated $x^{(i+1)}$ can be obtained by maximizing EI from $x^{(i+1)} \in \underset{x\in\mathcal{X}}{\operatorname{argmax}}\ \mathrm{EI}(x)$. This optimization can lead to either exploiting promising areas when $\hat{m}(x)$ is small or exploring new areas when $\hat{\sigma}(x)$ is large. More information on different acquisition functions can be found in Agnihotri and Batra [2020] and an optimized version for GP metamodel are proposed in Wilson et al. [2018].

Chapter 3

Solution of continuous ill-posed inverse problems

Contents

3.1 Deterministic Approach 13

3.2 Bayesian Approach 14

3.2.1 Stochastic Sampling 15

In this chapter we are going to review the most common approaches when dealing with an ill-posed inverse problem, where the approach to take will depend on its own properties. The general statement of an Inverse Problem suggests to find the set of inputs $x^\star \in \mathcal{X}$ to a mathematical model M, given an observation or set of observations $y \in \mathcal{Y}$. This is

$$y = M(x), \tag{3.1}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces. Frequently the inverse problems are ill-posed, this means that there is no solution, or the solution may not be unique, and may be sensitive to variations in the observations y. Among the different approaches to tackle inverse problems, the most common are the Deterministic and Bayesian Regularization ones.

3.1 Deterministic Approach

Under the *classical regularization* framework, the solution of any inverse problem from equation 3.1 can be written as the least square problem:

$$x^\star = \underset{x \in \mathcal{X}}{\operatorname{argmin}} \quad ||y - M(x)||^2_{\mathcal{Y}}, \tag{3.2}$$

which should suffice to find $x^\star$ when the problem is not ill-posed. Otherwise, finding a solution will require the addition of an auxiliary term depending on a point or center x_0 in the input space. This can be written as:

$$x^\star = \underset{x\in\mathcal{X}}{\operatorname{argmin}} \quad \left(\frac{1}{2}||y - M(x)||^2_{\mathcal{Y}} + \rho\right), \tag{3.3}$$

where ρ is a penalty term to be defined. Tikhonov Regularization Tarantola [2004] and the Ridge Regression Murphy [2012] are examples where the definition of ρ admits a closed-form solution for $x^\star$.

The main advantage of Classical regularization is the availability to obtain a deterministic solution to the inverse problem by using classical global optimization algorithms. This solution will depend on how the optimizer can deal with multiple local optima that appears often from equation 3.3.

3.2 Bayesian Approach

The Bayesian framework, instead of using partial information about the variable uncertainties, as is some of the classical approaches, it makes make use of all the possible probabilistic content available for the unknown to estimate. Here, the inverse problem is now written as

$$y = M(x) + \eta$$

where η is the observational noise. Then, using the Bayes theorem, the solution is written as a posterior distribution is

$$\pi(x|y) = \frac{\pi(y|x)\pi(x)}{\int_{\mathbb{R}^{nc}} \pi(y|x')\pi(x')dx'},$$

where $\pi(y|x)$ is called the likelihood to recover observation knowing the input parameters, $\pi(x)$ is the prior and the denominator or evidence $\pi(y)$ is often intractable. Under the Bayesian framework, by defining a prior $\pi(x) \sim \mathcal{N}(x_0, \Sigma_0)$ and if the noise is Gaussian $\eta \sim \mathcal{N}(0, \Gamma)$, the posterior takes the form

$$\pi(x|y) \propto \exp\{-\frac{1}{2}||y - M(x)||^2_\Gamma - \frac{1}{2}||x - x_0||^2_{\Sigma_0}\}, \tag{3.4}$$

where $||y||^2_\Gamma = y^T\Gamma^{-1}y$, $||x||^2_{\Sigma_0} = x^T\Sigma_0^{-1}x$. Under this framework and in the case of M being linear regarding x, the posterior distribution will be also Gaussian, and its Maximum a Posteriori,(or the value of x that maximizes its posterior density $\pi(x|y)$), correspond to the regularized least squares solution on equation 3.3.

Bayesian Regularization can be seen as a more general strategy that accounts for the uncertainty on $x^\star$, this at the cost of the evidence often being intractable or difficult to approximate when the posterior cannot be written as a Gaussian Distribution. For a more rigorous review on Bayesian Inverse problems, check Stuart [2010] or Dashti and Stuart [2016].

3.2.1 Stochastic Sampling

As is in general it is hard to obtain information from a probability measure where is not possible to express all the probabilities in terms of Gaussian distributions or the model M is nonlinear, or either in high dimensions Stuart [2010]; a commonly used method to extract information from a probability distribution is sampling. This means, generating a set of points $\{x_n\}_{n=1}^{N}$ that is distributed according to $\pi(x|y)$. The most common techniques of stochastic sampling are the Markov Chain Monte Carlo algorithms (MCMC) and its variants, where Metropolis-Hastings Hastings [1970] offers a framework to find an universal solution to the construction of an appropriate Markov chain, this means to generate a sequence that starts from a density $\tilde{\pi}(x)$, that converges to the target density $\pi(x)$ and explores all its support in a finite number of steps Robert and Changye [2020].

The basic idea of MH, also known as random walk, is to sample candidates x' from a a *proposal* or jumping distribution $\xi(x'|x^{(t)})$ (usually a normalized pdf), then accept or reject some of the proposed candidates. The acceptance probability is computed using the likelihood ratio of the previous and new candidates values:

$$\alpha(x', x^{(t)}) = \begin{cases} \min\{\frac{\tilde{\pi}(x')\xi(x'|x^{(t)})}{\tilde{\pi}(x^{(t)})\xi(x^{(t)}|x')}, 1\} & \text{if } \tilde{\pi}(x^{(t)})\xi(x^{(t)}|x') > 0 \\ 1, & \text{Otherwise} \end{cases}$$

If this probability is greater than a draw according to an uniform $U(0,1)$ distribution, then the candidate is accepted. The algorithm consists in defining a proposal distribution that accepts a large number of candidates in order to converge quickly to the target distribution but without increasing the autocorrelation too much. In the case of selecting a symmetric jumping distribution $\xi(x^{(t)}|x') = \xi(x'|x^{(t)})$ (usually Gaussian) the algorithm simplifies to the following steps:

Algorithm 1 Metropolis-Hastings Algorithm Hastings [1970]

Initialize $x^0 \sim \xi(x)$
for iteration $t = 0, 2, \ldots$ **do**
  Propose: $x' \sim \xi(x'|x^{(t)})$
  Acceptance probability:
    $\alpha = \min\left\{1, \frac{\tilde{\pi}(x')}{\tilde{\pi}(x^{(t)})}\right\}$
  $\nu \sim$ uniform $(0, 1)$
  if $\nu < \alpha$ **then**
    Accept the proposal: $x^{(t+1)} \leftarrow x'$
  else
    Reject the proposal: $x^{(t+1)} \leftarrow x^{(t)}$
  end if
end for

Common adaptations to the original MH are the adaptive covariance version where the variance of the proposal is modified in order to satisfy an acceptance rate Roberts and

Rosenthal [2009]; and the Metropolis-Within Gibbs algorithm which is commonly used on multiple dimensions as:

Algorithm 2 Metropolis-Hastings Within Gibbs Algorithm Ghirmai [2015]

Let be $x \in \mathbb{R}^{n_c}$
Initialize $x^{(0)} \sim \xi(x)$
for iteration $t = 0, 2, \ldots$ **do**
Propose and accept with MH: $x_1{}' \sim \xi_1(x_1{}'|x_2{}^{(t)}, \ldots, x_{n_c}{}^{(t)})$
Propose and accept with MH: $x_2{}' \sim \xi_2(x_2{}'|x_1{}', x_3{}^{(t)}, \ldots, x_{n_c}{}^{(t)})$
$\vdots$
Propose and accept with MH: $x_{n_c-1}{}' \sim \xi_{n_c-1}(x_{n_c-1}{}'|x_2{}', x_3{}^{(t)}, \ldots, x_{n_c-2}{}^{(t)}, x_{n_c}{}^{(t)})$
Propose and accept with MH: $x_{n_c}{}' \sim \xi_{n_c}(x_{n_c}{}'|x_2{}', x_3{}^{(t)}, \ldots, x_{n_c-1}{}^{(t)})$
end for

where it is possible to define different proposals ξ_i for each variable Ghirmai [2015].

MCMC convergence criteria

In order to set up and interpret a MCMC sampling routine, we need first to recall that the goal of MCMC sampling is to generate enough independent samples from the target distribution $\tilde{\pi}$, this means, being able to represent the true distribution while exploring all the support of it. To achieve that, it is necessary to identify when the chain or the sequence of sampled points has reached its stationary distribution and when the consecutive observations can be considered independent. The former is related to *Convergence* metrics and the later is to the *Mixing* metrics for any chain. Even that may exist different techniques to ensure convergence or mixing, it is recommended to apply a variety of them Givens and Hoeting [2005]

The group regarding *Mixing* metrics of the chain can be analyzed in a simple way graphically by using either the *sample path* or the *autocorrelation plot*. The first one corresponds to a figure relating the iteration number versus the realizations x. As shown in figures 3.1 and 3.2, a chain is *mixing well* when it moves from its starting value and continuously oscillates rapidly regarding the number of iterations in the support of the target distribution. On the other hand a *poor mixture*, is a representation of slower oscillations that are not compacted in the total number of iterations, this indicates that a longer chain may be required to observe a good mixture.

In the case of the autocorrelation plot, it summarizes the correlation in the sequence at different intervals of time or *lags*, this means, that autocorrelation at lag k is the correlation between iterates that are t iterations apart. As shown by figures 3.3 and 3.4 we expect that a good chain will show a faster decay as the lag increases and it will remain close to zero for higher lags k.

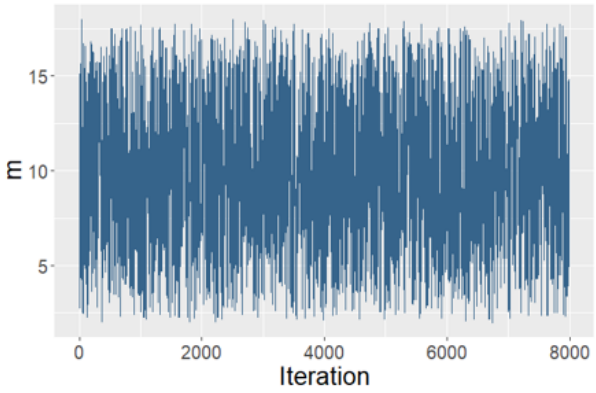


Figure 3.1: Sample path representing a good mixture.

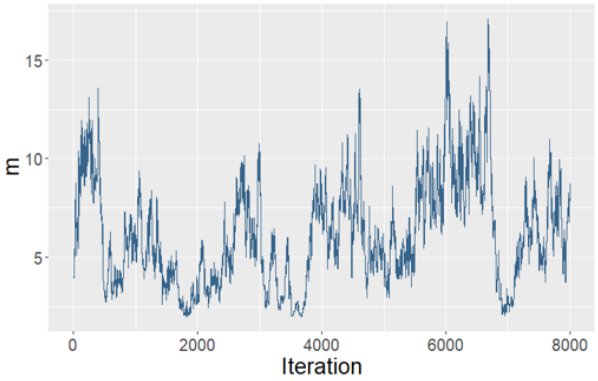


Figure 3.2: Sample path representing bad mixture.

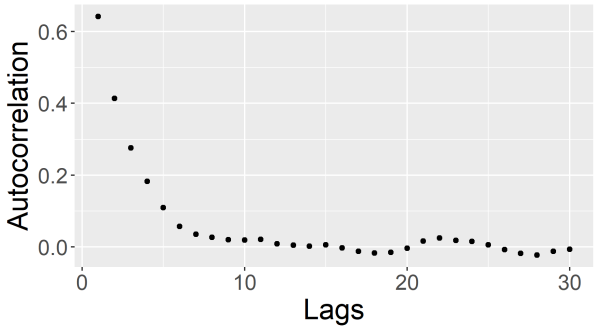


Figure 3.3: Sample path representing a good autocorrelation.

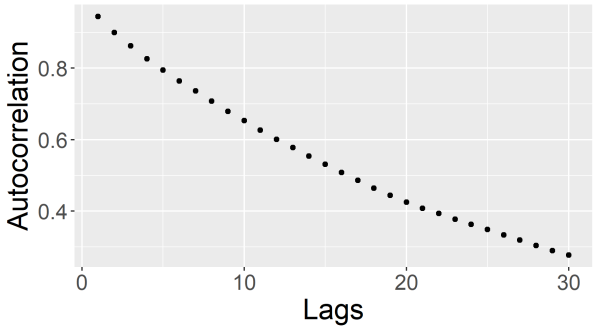


Figure 3.4: Sample path representing bad autocorrelation.

In the case of *convergence* the key considerations are the burn-in and the effective sample size. The *burn-in* phase allows us to adjust the proposal distribution in order to obtain a suitable acceptance rate. Candidates obtained during this phase are not necessarily in the zones of statistical content (high value of the likelihood) and are then discarded. Typically the number of samples to discard D is fixed to a few hundred or thousand values Givens and Hoeting [2005]. On the other hand the effective sample size (ESS) of the chain is the size of an i.i.d. sample that would contain the same information (e.g. mean and standard deviation). This value can be computed by:

$$ESS = \frac{L}{\hat{\tau}}$$

$$\tau = 1 + 2\sum_{k=1}^{\infty} r(k),$$

where $r(k)$ is the autocorrelation with lag k and L are the iterations after the burn-in period. Commonly, τ is estimated by truncating the summation when $\hat{r}(k) < 0.1$ For a

fixed number of iterations an MCMC algorithm with a larger ESS is likely to converge more quickly, this means that ESS can be used to compare the efficiency of different MCMC techniques Gong and Flegal [2016].

The **Geweke-test** is another common metric used to analyze whether the mean estimates have converged. This metric proposes to compare the mean values from the early and latter part of the Markov chain Geweke [1995]. Its interpretation is related to the a Z-test for the equality of means, where if $|Z_i| > 1.96$ implies that the means are different and the chain did not converge.

Another set of metrics used in tandem to evaluate the stopping criteria of a chain are the **Heidelberger-Welch Stationary and Half-Width Tests** . The former tests if the chain is already stationary and the latter if the sample size is adequate to meet the accuracy for the mean estimate. If the first one fails, it could indicate that a longer chain is required. Then, If the first one succeeds, it is the failure of the second that could indicate a longer chain is required Heidelberger and Welch [1983].

Finally the **Minimum effective sample size** is a formula that provides an estimator of the minimum iterations for your MCMC algorithm given the dimensions and desired confidence interval on the estimation of the sequence meanGong and Flegal [2016].

Chapter 4

Optimization of an expensive simulator with mixed variables

Contents

4.1 Latent Variable EGO 22
4.1.1 The vanilla LV-EGO algorithm 23
4.1.2 LV-EGO algorithms with Augmented Lagrangian . . . 25
4.2 Description of the numerical experiments 30
4.2.1 Test cases 32
4.2.2 Experiments setup and metrics 34
4.3 Results and discussion 34
4.3.1 Analytical test functions 35
4.3.2 Beam bending application 38
4.3.3 Summarized results 41
4.4 Complementary heuristics 43
4.4.1 Static exponential penalty 44
4.4.2 Adaptive exponential penalty 46
4.4.3 Updated results 48
4.4.4 Re-visiting the Beam Bending Problem 49
4.5 Conclusions 50

In this chapter we present LV-EGO as a novel strategy to perform global optimization in the presence of mixed variables. This model proposes a relaxation scheme from the mixed set of variables w to a full continuous one, therefore allowing the use of the classical BO algorithm as introduced in chapter 2. We define different variations of the algorithm as a way to improve its performance through comparisons on a group of test cases (including a mechanical application) and versus different related mixed optimizers. Most of the following contents correspond to our publication in the Journal *Advanced Modeling and Simulation in Engineering Sciences*, special Issue on *Efficient Strategies for*

Surrogate-Based Optimization Including Multifidelity and Reduced-Order Models.

We consider the problem of minimizing a function $y(x,u)$ depending on a vector of continuous variables $x=(x_1,\ldots,x_{n_c})$ and a vector of discrete variables $u=(u_1,\ldots,u_{n_d})$, where each u_i has m_i levels encoded $1,\ldots,m_i$. We denote $\mathcal{X}$ the domain of definition for the continuous inputs, typically, after rescaling, the hypercubic domain $[0,1]^{n_c}$. Similarly, we denote $\mathcal{U}=\prod_{j=1}^{n_d}\{1,\ldots,m_j\}$ the domain of definition for the discrete inputs. We also denote $w=(x,u)$ and $\mathcal{W}=\mathcal{X}\times\mathcal{U}$. For simplicity, the definition of y is overloaded in the following, and we will not make a distinction between $(x,u)\mapsto y(x,u)$ and $w\mapsto y(w)$. We focus on costly functions, meaning that each evaluation of y is time-consuming, and we aim at minimizing y with a tiny budget of evaluations. In this context, minimizing directly y is hardly possible. An alternative is to use Bayesian optimization (BO). In BO approaches, there are two main ingredients: a Gaussian process (GP) serving as a fast proxy, often called metamodel, built from the current learning set, and a sampling criterion, often called acquisition criterion, used to update the learning set with a new data point computed with y. A famous acquisition criterion is the expected improvement (EI). In that case, the BO approach is often called Efficient Global Optimization (EGO) algorithm.

To be more precise, let $\mathbf{W}=\{w^{(1)},\ldots,w^{(t)}\}\in\mathcal{W}^t$ be a design of experiments (DoE), and $y_i=y(w^{(i)})$ be the corresponding function evaluations ($i=1,\ldots,t$). Let $y_{\min}=\min(y_1,\ldots,y_t)$ be the current minimum. Let us now assume that y is a particular realization of the GP Y defined on $\mathcal{W}$. In that case, the EI criterion is defined by

$$\mathrm{EI}(w)=\mathrm{E}\left[\max(y_{\min}-Y^t(w),0)\right],\ w\in\mathcal{W},$$

where Y^t is the conditional GP knowing the observations:

$$Y^t:=Y\mid\{Y(w^{(1)})=y_1,\ldots,Y(w^{(t)})=y_t\}.$$

Notice that $\mathrm{EI}(w)$ is large when exploiting interesting area, that is to say when there is a good chance that $Y^t(w)$ is smaller than $y_{\min}$. This may occur when $\mathrm{E}[Y^t(w)]$ is close to $y_{\min}$, or when exploring unvisited areas, i.e. when the variance of $Y^t(w)$ is large compared to $(\mathrm{E}[Y^t(w)]-y_{\min})^2$. The idea of EGO is to evaluate y at a new point maximizing the EI criterion until a stopping criterion is reached. See Algorithm 3 for a synthetic description of the EGO algorithm when the stopping criterion is a maximum number of evaluations of y, noted budget.

Algorithm 3 EGO algorithm on a generic space

Generate the initial DoE of size N_{DoE}, $\mathbf{W}$, and calculate $\mathbf{Y} = (y_1, \ldots, y_{N_{\text{DoE}}})$, $t \leftarrow N_{\text{DoE}}$.
while $t \leq$ budget **do**
  Estimate the GP Y^t from the learning set formed by $\mathbf{W}$ and $\mathbf{Y}$.
  Look for the current minimum $y_{\min}$ and maximize $w \mapsto \text{EI}(w)$ on $\mathcal{W}$: $w^{t+1} \in \text{argmax}_{w \in \mathcal{W}} \text{EI}(w)$.
  Evaluate y at w^{t+1}, $y^{t+1} = y(w^{t+1})$.
  Update the learning set: $\mathbf{W} \leftarrow \mathbf{W} \cup \{w^{t+1}\}$, $\mathbf{Y} \leftarrow \mathbf{Y} \cup \{y^{t+1}\}$.
  $t \leftarrow t + 1$
end while
$w^\star = \arg\min_{w \in \mathbf{W}} y(w)$, $y^\star = y(w^\star)$
return $(w^\star, y^\star)$

This EGO algorithm has been intensively studied to minimize nonlinear functions that are expensive to be evaluated in the case $\mathcal{W} = \mathcal{X}$, i.e. when all input variables are continuous (see Le Riche and Picheny [2021] for numerical illustrations of its efficiency). The application of this algorithm in the presence of categorical variables is much less documented (see e.g. Pelamatti et al. [2019], Zuniga and Sinoquet [2020]), which can be explained by two main difficulties. The first one is related to the difficult estimation of covariance kernels on mixed spaces. Indeed, multi-dimensional covariance functions are often built by combination of one-dimensional ones. Therefore, covariance functions on $\mathcal{W}$ can be obtained by combining covariance functions on $\mathcal{X}$ and $\mathcal{U}$, so that, for all $w = (x, u)$ and $w' = (x', u')$ in $\mathcal{W}$:

$$\text{Cov}(Y(w), Y(w')) = k_1^x(x_1, x_1') * \cdots * k_{n_c}^x(x_{n_c}, x_{n_c}') * k_1^u(u_1, u_1') * \cdots * k_{n_d}^u(u_{n_d}, u_{n_d}'), \quad (4.1)$$

where $k_1^x, \ldots, k_{n_c}^x, k_1^u, \ldots, k_{n_d}^u$ are covariance functions and $*$ is an operation that preserves positive definiteness, such as sum or product. If we focus on the single categorical variable u_j with levels $1, \ldots, m_j$, we can identify the covariance function k_j^u to a $(m_j \times m_j)$-dimensional positive semidefinite matrix $\mathbf{T}$, such that for all $1 \leq k, \ell \leq m_j$,

$$(\mathbf{T})_{k\ell} = k_j^u(k, \ell). \quad (4.2)$$

This means that $\sum_{j=1}^{n_d} m_j(m_j + 1)/2$ coefficients need to be estimated to determine a covariance on $\mathcal{U}$ in the general case. That number can be large when m is large, which very often makes this estimation very difficult in practice. Furthermore, the optimization problem is often harder than the box-constrained one met with continuous variables. Indeed it is either constrained by the positive definiteness of $\mathbf{T}$, which is non-linear, or defined on a manifold if $\mathbf{T}$ is parameterized in spherical coordinates. We refer to Roustant et al. [2020] for more details and other parsimonious representations of k_j^u, which can reduce but not totally fix these issues. The second reason that can explain the few number of direct applications of EGO algorithm on mixed space is related to the difficult

maximization of the expected improvement, i.e. the search of the new input points where to call the function y, which are solutions of:

$$\max_{x,u \in \mathcal{X} \times \mathcal{U}} \mathrm{EI}(x,u) \ . \tag{4.3}$$

Indeed, classical optimization algorithms on continuous spaces usually try to exploit information related to the gradient of the function to be maximized, as well as notions of proximity in the space of the inputs. However, these two notions are difficult to exploit when dealing with categorical inputs, i.e. without any a priori ordering between the input instances. To circumvent this difficulty, a naive approach of resolution would consist in no longer considering a single maximization problem on $\mathcal{W}$, but the resolution in parallel of $\prod_{j=1}^{n_d} m_j$ maximization problems on $\mathcal{X}$, i.e. one problem per combination of instances of the categorical inputs u. Such an approach is not tractable when the number of optimization problems to be solved becomes large, which has motivated the definition of heuristics, such as evolutionary algorithms Li et al. [2013a], Cao et al. [2000], Lin et al. [2018], which seek to concentrate the searches only on the interesting instances of u. However, these approaches still rely on a large number of calls to the function to be optimized, and their convergence is not always easy to quantify.

Because mixed optimization problems are difficult, an alternative approach is proposed in the rest of this manuscript. It is based on the possibility to relax the discrete variables into continuous latent variables, therefore benefiting from the more efficient search mechanisms that exist in continuous spaces (e.g. gradients).

4.1 Latent Variable EGO

For an easier handling of categorical inputs, it was proposed in Zhang et al. [2019] to replace each categorical input u_j by a vector of $q_j \geq 1$ continuous inputs with values in $\mathbb{R}^{q_j}$, noted ℓ_j. To give an intuition of the underlying idea in the automotive domain, a category of lubricant may be determined by physical continuous features such as boiling temperature, viscosity, etc that act as latent variables. In structural mechanics, the shape of a load carrying structure, which is categorical, has underlying continuous flexural and membrane moments that drive its behavior. This amounts to associating to the Gaussian process (GP) Y a new GP $\widetilde{Y}$, such that for each instance u of the categorical inputs there exists a particular value of $\ell := (\ell_1, \ldots, \ell_{n_d}) \in \mathcal{L} \subset \mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_{n_d}}$, which is called *latent variable*, allowing us to write:

$$Y(x,u) \overset{\text{in law}}{=} \widetilde{Y}(x,\ell), \ x \in \mathcal{X}. \tag{4.4}$$

An important point is that the values of ℓ are unobserved and therefore $\widetilde{Y}$ is unknown. Nevertheless, in order to replace the EI maximization problem on $\mathcal{X} \times \mathcal{U}$ by a new optimization problem on $\mathcal{X} \times \mathcal{L}$, a precise knowledge of $\widetilde{Y}$ is not necessary. Indeed, assuming that kernels for mixed inputs are built by combining 1-dimensional ones as in

(4.1), it is sufficient to identify the mappings ϕ_j from $\{1, \ldots, m_j\}$ to $\mathbb{R}^{q_j}$ to each variable u_j such that

$$k_j^u(u_j, u_j') \approx k_j(\phi_j(u_j), \phi_j(u_j')), \tag{4.5}$$

where k_j is a continuous kernel on $\mathbb{R}^{q_j} \times \mathbb{R}^{q_j}$. Thus, it is not so much the values of $\phi_j(u_j)$ that are important, but their relative positions in $\mathbb{R}^{q_j}$ in order to allow a reasonable reconstruction of the dependency structure between $Y(x, u)$ and $Y(x', u')$.

According to the works achieved in Zhang et al. [2019], it appears that interesting mappings can be obtained by likelihood maximization and that relatively small values of q_j can give a satisfying reconstruction. Following their recommendations, q_j can be chosen equal to 1 if $m_j \leq 3$ and to 2 otherwise, which will be the values chosen in the rest of this manuscript. We denote by $n_\ell = \sum_{j=1}^{n_d} q_j$ the total number of latent variables. Following Roustant et al. [2020], the continuous kernel k_j associated to the latent variables was chosen as the dot product kernel $k_j(t, t') = \langle t, t' \rangle$. The corresponding covariance matrix is then low-rank, and provided better performances than the Gaussian kernel in the examples considered in the latter reference.

This new parameterization leads us to the following adaptation of the EI maximization problem defined by Eq. (4.3), which we name *acquisition problem* as it allows to acquire a new point to evaluate:

$$\begin{aligned} &\max_{x, \ell \in \mathcal{X} \times \mathcal{L} \subset \mathbb{R}^{n_c + n_\ell}} \mathrm{EI}^{(t)}(x, \ell) \\ &\text{such that } \exists u \in \mathcal{U} \text{ with } \ell = \phi^{(t)}(u). \end{aligned} \tag{4.6}$$

Here, $\mathrm{EI}^{(t)}(x, \ell)$ is the expected improvement associated with GP $\widetilde{Y}$ at iteration t, $\phi^{(t)} = (\phi_1^{(t)}, \ldots, \phi_{n_d}^{(t)})$ is the vector-valued mapping from $\prod_{j=1}^{n_d} \{1, \ldots, m_j\}$ to $\mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_{n_d}}$ at iteration t, and the constraint on the values of ℓ is driven by the fact that the values of the latent variables at the new point have to remain compatible with the current mapping functions.

We follow two paths to solve this acquisition problem. In the vanilla LV-EGO approach, which will be described soon, the EI maximization and the latent-discrete compatibility constraint are addressed one after each other. Alternatively, with the augmented Lagrangian approaches, which will be described in Section 4.1.2, the full constrained optimization problem is treated.

4.1.1 The vanilla LV-EGO algorithm

At each iteration, the vanilla LV-EGO algorithm first maximizes EI in a relaxed, fully continuous, formulation where the discrete variables are replaced by relaxed continuous

latent variables. Then, a pre-image problem is solved where EI is maximized over the discrete variables only, the continuous variables being fixed at their value of the relaxed problem. The LV-EGO methodology is summarized in Algorithm 4.

Algorithm 4 Vanilla LV-EGO with mixed inputs

1: **Generate** the initial DoE of size N_{DoE}: $\mathbf{X}$, $\mathbf{U}$
2: **Costly function evaluations** $y(x^i, u^i)$, $i = 1, \ldots, N_{\text{DoE}}$, $t \leftarrow N_{\text{DoE}}$
3: **while** $t \leq$ budget **do**
4: **Estimate** the latent variable mappings $\phi^{(t)}$ and the parameters of the continuous GP $\widetilde{Y}$.
5: **Perform** one EGO iteration in the *relaxed continuous* space : $(x^{t+1}, \ell^{t+1}) = \arg\max_{x,\ell \in \mathcal{X} \times \mathcal{L} \subset \mathbb{R}^{n_c+n_\ell}} \text{EI}^{(t)}(x, \ell)$.
6: **Recover** the *discrete pre-image* component u^{t+1} as: $u^{t+1} = \arg\max_{u \in \mathcal{U}} \text{EI}^{(t)}(x^{t+1}, \phi^{(t)}(u))$.
7: **Update** the DoE with (x^{t+1}, u^{t+1}) with output value $y(x^{t+1}, u^{t+1})$.
8: $t \leftarrow t + 1$
9: **end while**
10: **Return** $(x^\star, u^\star) = \arg\min_{x^t, u^t \in (\mathbf{X}, \mathbf{U})} y(x^t, u^t)$

The main difference with the generic Bayesian algorithm 3 is the new discrete pre-image problem in line 6. Notice that the pre-image is formulated in terms of the EI objective, as opposed to a more arbitrary distance like $\|\ell^{t+1} - \phi^{(t)}(u)\|$.

In terms of implementation, the EI maximization (line 5) is done with the COBYLA algorithm, a gradient free non-linear optimization technique Powell [1994]. Since COBYLA is a local optimizer and the EI is a multimodal function, the maximization is repeated (10 times) from randomly chosen initial points and the best result is kept. An exhaustive search is carried out for the EI maximization of the pre-image problem (line 6).

A comparison of the numerical complexities of the vanilla LV-EGO (Algorithm 4) and the generic EGO (Algorithm 3) shows that the cost of the latent variables is limited. Let us consider that the discrete space can be searched essentially by enumeration in $\mathcal{O}(\text{card}\,\mathcal{U}) = \mathcal{O}(\prod_{i=1}^{n_d} m_i)$ operations (where m_i is the number of levels per discrete variable) while a continuous space can be searched more efficiently in linear time. At each iteration, the Bayesian algorithms of this manuscript have three steps: first a GP is learned, then an acquisition criterion (EI for now and an augmented Lagrangian later) is maximized and finally a pre-image problem is solved. In the vanilla LV-EGO algorithm, these steps take place at lines 4, 5 and 6 of Algorithm 4, respectively. Table 4.1 summarizes the number of operations per step. The number of operations for learning the GPs is proportional to the cube of the number of points evaluated (t) because of the inversions of the covariance matrices, times the number of (continuous) parameters of the GP for the likelihood maximization.

The two other steps, the acquisition and the pre-image, imply predictions by the GP in t^2 operations times a number of operations that depends on the specific algorithm.

	Mixed space search (Alg. 3)	Vanilla LV-EGO (Alg. 4)	ALV-EGO-g (Alg. 5+6)	ALV-EGO-l (Alg. 5+7)
GP learning	$(n_c + \sum_{i=1}^{n_d} m_i) \times t^3$	$(n_c + q \times \sum_{i=1}^{n_d} m_i) \times t^3$	$(n_c + q \times \sum_{i=1}^{n_d} m_i) \times t^3$	$(n_c + q \times \sum_{i=1}^{n_d} m_i) \times t^3$
max acquisition	$(\prod_{i=1}^{n_d} m_i) \times n_c \times t^2$	$(n_c + q \times \sum_{i=1}^{n_d} m_i) \times t^2$	$(N'_{\text{DoE}} + n_c + q \times \sum_{i=1}^{n_d} m_i) \times t^2$	$(n_c + q \times \sum_{i=1}^{n_d} m_i) \times t^2$
pre-image	0	$(\prod_{i=1}^{n_d} m_i) \times t^2$	$(\prod_{i=1}^{n_d} m_i) \times t^2$	$(\prod_{i=1}^{n_d} m_i) \times t^2$

Table 4.1: Numerical complexities of the algorithms compared at each iteration (for a given t).

Comparing in Table 4.1 the column of the generic EGO with that of the vanilla LV-EGO, and assuming that for all i $m_i = m$ to keep the discussion simple, it can be seen that the latent variables induce a slight extra cost to be learnt. When $q = 2$, which is our default here, this extra cost is $n_d \times m_i \times t^3$ operations. $q = 1$ would not add any cost to the learning. An advantage, which comes from the sequential resolution of the mixed problem, occurs in the maximization of the acquisition criterion when $n_c + q \times n_d \times m < m^{n_d} \times n_c$, at the cost of an additional pre-image problem to solve. Thus, LV-EGO will be faster than a mixed EGO once the latent variables are estimated if $m^{n_d} + n_c + q \times m \times n_d < m^{n_d} \times n_c$, which happens frequently (take for example $n_c = 4, n_d = 2, m = 10, q = 2$).

4.1.2 LV-EGO algorithms with Augmented Lagrangian

A possible pitfall of the vanilla LV-EGO detailed in Algorithm 4.1.1 is that the link between the discrete variables u and their relaxed continuous counterparts ℓ is lost when maximizing $\text{EI}^{(t)}(x, \ell)$ in line 5. Recovering it during the discrete pre-image problem where x is fixed to a value optimal in the relaxed formulation but possibly non-optimal with respect to the mixed problem (4.3) may yield a sub-optimal solution. For this reason, we now propose LV-EGO algorithms that account for the discreteness constraint during the optimization using augmented Lagrangians.

In that prospect, notice that problem (4.6) can be approximated as an optimization problem with an inequality constraint:

$$\begin{aligned} &\min_{x,\ell \in \mathcal{X} \times \mathcal{L} \subset \mathbb{R}^{n_c+n_\ell}} f^{(t)}(x, \ell) := -\log(1 + \text{EI}^{(t)}(x, \ell)) \\ &\text{such that } g^{(t)}(\ell) := \min_{u \in \mathcal{U}} \|\ell - \phi^{(t)}(u)\| - \epsilon \ \leq \ 0 \end{aligned} \tag{4.7}$$

where ϵ is a small positive relaxation constant and $\|\cdot\|$ the Euclidean norm. In this reformulation, called *relaxed acquisition problem*, notice the log scaling of the EI which does not change the solution but improves the conditioning of the problem. Two values of

ϵ will be discussed in the sequel, $\epsilon = 0$ in which case the constraint becomes an equality constraint, $\min_{u\in\mathcal{U}}\|\ell - \phi^{(t)}(u)\| = 0$, and $\epsilon > 0$ but small which corresponds to a relaxation of the equality. In the sequel, ϵ is normalized with respect to the size of the vector of latent variables and set to $\epsilon = 0.01$.

The constrained optimization problem (4.7) is solved through an augmented Lagrangian approach Minoux [1986], Nocedal and Wright [2006]. The augmented Lagrangian is that of Rockafellar Rockafellar [1993] which, specified for Problem (4.7), is,

$$L_A^{(t)}(x,\ell;\lambda,\rho) = \begin{cases} f^{(t)}(x,\ell) - \frac{\lambda^2}{2\rho} & \text{, if } g^{(t)}(\ell) \leq \frac{-\lambda}{\rho} \\ f^{(t)}(x,\ell) + \lambda g^{(t)}(\ell) + \frac{\rho}{2} g^{(t)}(\ell)^2 & \text{, otherwise} \end{cases} \tag{4.8}$$

When $\epsilon = 0$, the constraint $g^{(t)}(\ell) \leq 0$ becomes an equality constraint, $g^{(t)}(\ell) = 0$. In this case, the augmented Lagrangian connected to that of Rockafellar is that of Hestenes Hestenes [1969] and takes the form

$$L_A^{(t)}(x,\ell;\lambda,\rho) = f^{(t)}(x,\ell) + \lambda g^{(t)}(\ell) + \frac{\rho}{2} g^{(t)}(\ell)^2 \tag{4.9}$$

Complementary explanations about the augmented Lagrangians are given in Appendix A.1.

Augmented Lagrangians require to specify the values of the Lagrange multiplier, λ, and of the penalty parameter, ρ. The general principle to fix them is to calculate the generalized Lagrange multiplier with a dual formulation Minoux [1986]: the dual function $D^{(t)}$ is maximized with respect to the multiplier λ while the penalty parameter ρ should take the smallest value that allows to find feasible solutions,

$$\begin{aligned} &\rho_t = \arg\min_{\rho\geq 0} \rho \;\text{ such that }\; g(\ell^t) \leq 0 \\ &\text{where } \lambda_t = \arg\max_{\lambda\geq 0} D^{(t)}(\lambda,\rho) \;, \\ &\quad D^{(t)}(\lambda,\rho) = \min_{x,\ell\in\mathcal{X}\times\mathcal{L}\subset\mathbb{R}^{n_c+n_\ell}} L_A^{(t)}(x,\ell;\lambda,\rho) \;, \\ &\quad \text{and } (x^t,\ell^t) \in \arg\min_{x,\ell\in\mathcal{X}\times\mathcal{L}\subset\mathbb{R}^{n_c+n_\ell}} L_A^{(t)}(x,\ell;\lambda,\rho) \;. \end{aligned} \tag{4.10}$$

There are two logics to solve Problem (4.10), both of which have been investigated in this study. Following an idea presented in Le Riche and Guyon [2002] for classical Lagrangians, we first propose to approximate the dual function $D()$ as the lower front of the augmented Lagrangians of a finite set of calculated points. The approximated dual is

$$\widehat{D}(\lambda,\rho) = \min_{(x,\ell)\in(\mathbf{X}',\mathbf{L}')} L_A^{(t)}(x,\ell;\lambda,\rho) \tag{4.11}$$

where $(\mathbf{X}',\mathbf{L}')$ is a DoE that should not be mistaken for $(\mathbf{X},\mathbf{U})$, the DoE of the original expensive problem. $(\lambda_t,\rho_t,x^t,\ell^t)$ comes from solving Problem (4.10) with minimizations over the finite set $(\mathbf{X}',\mathbf{L}')$ instead of the initial $\mathcal{X}\times\mathcal{L}$. The functions in Problem (4.7) are not costly, $(\mathbf{X}',\mathbf{L}')$ can be quite large.

This approach is called *global dual* as a global approximation to the dual function is built and maximized. It applies to very general functions, e.g., non differentiable functions. Another advantage of this approach is to allow large changes in the dual space. Figure A.1 provides an illustration of the approximated dual function and the effect of ρ on the dual problem. The sketch is done for an inequality constraint, yet it also stands with marginal changes for an equality (cf. Appendix A.1 and the caption to the Figure). Under the non-restrictive hypothesis that there is a ρ beyond which the solution to the primal problem (4.7) maximizes the dual function, maximizing the dual function preserves the global aspect of the search. However, the accuracy of the obtained (λ_t, ρ_t)'s will depend on the DoE. Because there is only one constraint in the current problem and evaluating it does not require calling the costly function, the maximization on λ and ρ is done by enumeration on a 100×20 grid and $(\mathbf{X}', \mathbf{L}')$ is a 100 LHS sample.

The other path to updating the multiplier is to progressively change them based on the minimizers of the augmented Lagrangian at the current step. This updating can be seen as a step in the dual space which makes it general, although it is usually proved by analogy with the Karush Kuhn and Tucker optimality conditions Nocedal and Wright [2006] which add unnecessary conditions (like differentiability), cf. Appendix A.1. Let (x^t, ℓ^t) be a solution to

$$\min_{x,\ell \in \mathcal{X} \times \mathcal{L} \subset \mathbb{R}^{n_c+n_\ell}} L_A^{(t)}(x, \ell; \lambda_t, \rho_t) \tag{4.12}$$

The update formula reads

$$\lambda_{t+1} = \lambda_t + \rho_t \left(g^{(t)}(\ell^t) + \max(0, \frac{-\lambda_t}{\rho_t} - g^{(t)}(\ell^t)) \right) \tag{4.13}$$

As in Picheny et al. [2016], the penalty parameter ρ is simply increased if the constraint is not satisfied,

$$\rho_{t+1} = \begin{cases} \rho_t & \text{if} \quad g^{(t)}(\ell^t) \leq 0 \\ 2\rho_t & \text{otherwise} \end{cases} \tag{4.14}$$

The update scheme based on equations (4.13) and (4.14) is called *local dual* as a local step in the dual (λ, ρ) space is taken.

Algorithm 5 Augmented Lagrangian Latent Variables EGO with global or local dual scheme (ALV-EGO-g or ALV-EGO-l)

1: **generate** the initial DoE of size N_{DoE} for $(\mathbf{X}, \mathbf{U})$
2: **costly function evaluations** $y(x^i, u^i)$, $i = 1, \ldots, N_{\text{DoE}}$, $t \leftarrow N_{\text{DoE}}$
3: **initialize** budget, ϵ
4: **while** $t \leq$ budget **do**
5: **estimate** the latent variables $\phi^{(t)}$ and the GP parameters from current DoE.
6: {approximately solve the relaxed acquisition problem (4.7) with $f^{(t)}(\cdot) = -\log(1 + \text{EI}^{(t)}(\cdot))$}
$(x^{t+1}, \ell^{t+1}) = \arg\min_{x,\ell} f^{(t)}(x, \ell)$ s.t. $g^{(t)}(\ell^{t+1}) = \min_{u \in \mathcal{U}} \|\ell - \phi^{(t)}(u)\| - \epsilon \leq 0$,
ALV-EGO-g variant: with the global dual scheme, cf. Algorithm 6
ALV-EGO-l variant: with the local dual scheme, cf. Algorithm 7
7: **recover** the *discrete pre-image* component u^{t+1} as: $u^{t+1} = \arg\max_{u \in \mathcal{U}} \text{EI}^{(t)}(x^{t+1}, \phi^{(t)}(u))$
8: **update DoE**: add (x^{t+1}, u^{t+1}) and its costly evaluation $y(x^{t+1}, u^{t+1})$ to the DoE $(\mathbf{X}, \mathbf{U})$.
9: $t \leftarrow t + 1$
10: **end while**
11: **return** $(x^\star, u^\star) = \arg\min_{(\mathbf{X},\mathbf{U})} y(x, u)$

Algorithm 5 gathers all these changes and is called ALV-EGO. The essential difference between this ALV-EGO algorithm and the vanilla counterpart (Algorithm 4) is that the EI maximization step is constrained so that the link between the discrete variables and the relaxed latent variables (hence the continuous x) is not lost and left to the pre-image step. The coupling between the continuous and the discrete variables is better accounted for. However, a pre-image step (line 7) is still necessary to fully recover a discrete solution in cases when the constraint is relaxed ($\epsilon > 0$). In ALV-EGO like in the vanilla LV-EGO, there are $q = 2$ continuous latent variable per discrete variable.

The global and local dual schemes are further detailed in Algorithms 6 and 7. The continuous minimizations of the Augmented Lagrangians once the Lagrange multipliers are set are always done with 10 random restarts of the COBYLA algorithm Powell [1994]. They occur in Algorithm 6, line 4 and Algorithm 7 line 5. To allow comparisons, this implementation is identical to the EI maximization of the vanilla LV-EGO (step 5 of Algorithm 4).

Algorithm 6 Global dual scheme (makes ALV-EGO-g when used in Algorithm 5)

Ensure: An estimation of the solution to the relaxed acquisition problem (4.7)
Require: $f^{(t)}()$, an objective function, $g^{(t)}()$, a constraint
N'_{DoE}, N_λ, $N_\rho > 0$
1: Calculate a DoE $(\mathbf{X}', \mathbf{L}') \in (\mathcal{X}, \mathcal{L})^{N'_{\text{DoE}}}$.
Half of the points are feasible by *i)* sampling a $u \in \mathcal{U}$ and *ii)* setting $\ell' = \phi^{(t)}(u)$
2: Create a grid of Lagrange multipliers and penalty parameters, $(\boldsymbol{\lambda}, \boldsymbol{\rho}) = \{\lambda_1, \ldots, \lambda_{N_\lambda}\} \times \{\rho_1, \ldots, \rho_{N_\rho}\}$, with $\lambda_i \geq 0$ and $\rho_j \geq 0$ for all i, j
3: Approximately solve the dual problem by enumeration:
ρ_t smallest $\rho \in \boldsymbol{\rho}$ that yields a feasible solution, $g(\ell^t) \leq 0$ where
$(\lambda_t, x', \ell') = \arg\max_{\lambda \in \boldsymbol{\lambda}} \min_{(x,\ell) \in (\mathbf{X}', \mathbf{L}')} L_A^{(t)}(x, \ell; \lambda, \rho)$
4: Fine tune the next candidate: $(x^{t+1}, \ell^{t+1}) = \arg\min_{(x,\ell) \in (\mathcal{X}, \mathcal{L})} L_A^{(t)}(x, \ell; \lambda_t, \rho_t)$
5: **return** x^{t+1}, ℓ^{t+1}

Algorithm 7 Local dual scheme (makes in ALV-EGO-l when used in Algorithm 5)

Ensure: An estimation of the solution to the relaxed acquisition problem (4.7)
Require: $f^{(t)}()$, an objective function, $g^{(t)}()$, a constraint
initial values of the Lagrange multiplier and penalty, $\lambda_{N_{\text{DoE}}} = 0$ and $\rho_{N_{\text{DoE}}} = 1$, t
1: **if** $t > N_{\text{DoE}}$ **then**
2: {when $t = N_{\text{DoE}}$ the initial $\lambda_{N_{\text{DoE}}}, \rho_{N_{\text{DoE}}}$ are used}
Update λ according to Eq. (4.13)
$\lambda_t = \lambda_{t-1} + \rho_{t-1} \left(g^{(t-1)}(\ell^t) + \max(0, \frac{-\lambda_{t-1}}{\rho_{t-1}} - g^{(t-1)}(\ell^t)) \right)$
3: Update ρ according to Eq. (4.14)
$\rho_t = \rho_{t-1}$ if $g^{(t-1)}(\ell^t) \leq 0$, $2\rho_{t-1}$ otherwise
4: **end if**
5: $(x^{t+1}, \ell^{t+1}) = \arg\min_{(x,\ell) \in (\mathcal{X}, \mathcal{L})} L_A^{(t)}(x, \ell; \lambda_t, \rho_t)$
6: **return** x^{t+1}, ℓ^{t+1}

While the local update of λ and ρ might seem less robust, it is the most common implementation and it might be sufficient for the constrained EI maximization. Indeed, between two iterations, the EI changes only locally around the current iterate. Providing the latent mapping functions do not change too much, a local update of λ and ρ seems appropriate. The numerical complexity of the ALV-EGO-g and -l algorithms is essentially the same as that of the vanilla LV-EGO, cf. Table 4.1. The global dual scheme has a slight extra-cost because of the search for the Lagrange multiplier and penalty parameter that require N'_{DoE} extra GP predictions.

Eventually, four variants of ALV-EGO are considered, ALV-EGO-ge or -gi or -le or -li where g stands for global, l for local, e for equality ($\epsilon = 0$) and i for inequality ($\epsilon > 0$).

name	formulation	metamodel	acq. crit.	optimizer of the acq. crit.
LV-EGO	LV	GP	EI	restarted COBYLA
LV-RFO	LV	`randomForest` toolbox	EI	focus-search (from `mlrMBO`)
ALV-EGO-ge or -gi	LV	GP	EI	DoE (for λ_t and ρ_t) and restarted COBYLA
ALV-EGO-le or -li	LV	GP	EI	restarted COBYLA
MS-RFO	MS	`randomForest` toolbox	EI	focus-search (from `mlrMBO`)
MS-ES	MS	none	$-y(x,u)$	evolution strategy (from Li et al. [2013a] in `CEGO` implementation Zaefferer [2014–2021])
MS-MKES	MS	GP (sym. compound disc. kernel)	EI	evolution strategy (from Li et al. [2013a] in `CEGO` implementation Zaefferer [2014–2021])

Table 4.2: Summary of the 9 algorithms tested: name, space over which it is defined (mixed versus continuous with latent variables), metamodel used, acquisition criterion, optimizer of the acquisition criterion.

4.2 Description of the numerical experiments

This section presents the different algorithms tested as well as the test-cases and applications used to compare their performance. The set of algorithms tested are summarized in the Table 4.2 which provides their names, the type of formulation for the mixed variables, the type of metamodel, the acquisition criterion and the technique to optimize the acquisition criterion. The two possible formulations for the mixed variables are either by searching in a mixed space (MS) or by a formulation in latent variables (LV). All Gaussian processes (GPs) are built with the `kerpg` package Deville et al. [2017–2021]. The meaning of the acronyms is: LV-EGO, Latent Variables EGO; LV-RFO, Latent Variables Random Forest Optimization; ALV-EGO-ge/-gi/-le/-li, Augmented Lagrangian Latent Variables global/local dual scheme with equality/inequality pre-image constraints; MS-RFO, Mixed Space search with Random Forest Optimization; MS-ES, Mixed Space search with Evolution Strategy; MS-MKES, Mixed Space search with Mixed Kriging metamodel and Evolution Strategy. The different algorithms will be tested on the suite

of test problems soon to be described. Before that, we provide a few more details about the evolution strategy and the mixed kriging model.

Mixed space evolution strategies

Among population-based techniques, the Evolution Strategy (ES) (μ^+, λ) is a stochastic optimization algorithm modified to solve problems with categorical and continuous inputs. As proposed in Li et al. [2013b], they extend the classical representation of the individuals by defining the space $\mathbb{I} = \mathcal{X} \times \mathcal{U}$, with $\mathcal{U} = \mathbb{Z} \times \mathbb{D}$, where $\mathcal{X}, \mathbb{Z}, \mathbb{D}$ denotes the continuous, integer and factor variables respectively that are sampled (which includes mutated) independently. The combined goal of the stochastic operators, the mutation and the recombination, and the selection of the best points, is to concentrate the search in interesting instances of the input variables. Given the mixed space of the individuals $\mathbb{I}$, the Evolution Strategy $(\mu \overset{+}{,} \lambda)$ proceeds as follows

Algorithm 8 (μ^+, λ) Evolution Strategy

$t \leftarrow 0$
Initialize population $P(t) \in \mathbb{I}$
Evaluate the μ initial individuals with objective function f
while termination criteria not fulfilled **do**
  for all $i \in \{i = 1, 2, \ldots, \lambda\}$ **do**
    chose uniform randomly parents c_{i_1}, c_{i_2} from $P(t)$(repetition is possible)
    $x_i \leftarrow$ mutate (recombine:c_{i_1}, c_{i_2})
    $Q(t) \leftarrow Q(t) \cup \{x_i\}$ (set of offspring individuals $Q(t)$)
  end for
  $P(t+1) \leftarrow \mu$ select individuals from: $(P \cup Q)$ ("+" version) or Q (","' version)
  $t \leftarrow t + 1$
end while

This version of the $(\mu \overset{+}{,} \lambda)$ Evolution Strategy Algorithm is used as the MS-ES technique defined in table 4.2 and will be combined with a mixed metamodel, based on a mixed variable kernel, that will be defined hereafter.

Mixed variable kernel

A mixed Gaussian Process $\{(x, u); (y(x, u))\}$ can be written as

$$Y \sim GP(\mathbf{0}, \mathrm{Cov}(Y(w), Y(w')))$$

where the covariance function can be a tensor product as in Equation 4.1:

$$\mathrm{Cov}(Y(w), Y(w')) = k_1^x(x_1, x_1') * \cdots * k_{n_c}^x(x_{n_c}, x_{n_c}') * k_1^u(u_1, u_1') * \cdots * k_{n_d}^u(u_{n_d}, u_{n_d}'),$$

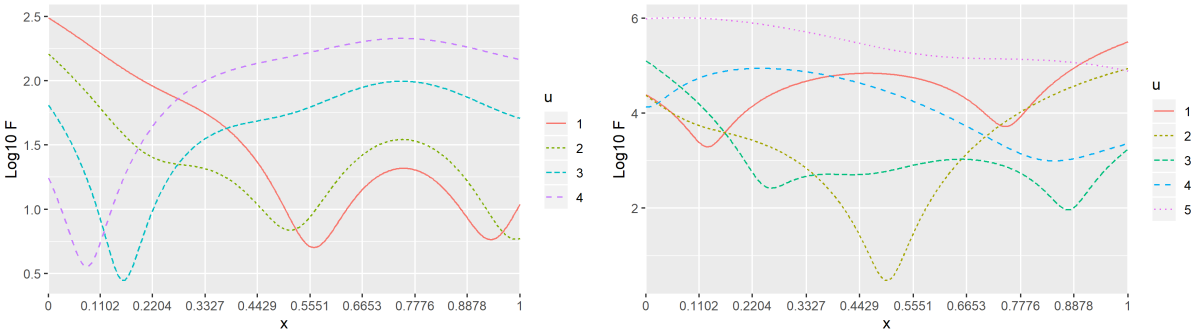
Under the strong assumption of a common correlation value for all the levels of each categorical variable, it is possible to reduce the effective number of parameters to estimate. Then each k_j^u is constructed as follows

$$k^u{}_j(u, u') = \begin{cases} \sigma_u^2, & \text{if } u = u' \\ c, & \text{if } u \neq u' \end{cases}$$

where c is a constant value. This corresponds to a Compound Symmetry (CS) kernel, for a more general mixed kernel representation see Roustant et al. [2020]. Under this setting, an hybrid MS-MKES is formed by a mixed GP with a CS kernel for the discrete variables optimized via the previously defined $(\mu \overset{+}{,} \lambda)$ ES. In MS-MKES, at each iteration, the EI acquisition criterion is maximized in terms of the mixed variables.

4.2.1 Test cases

There are 3 analytical test cases and a beam bending problem. The analytical test cases have all been designed by discretizing some of the variables of classical multimodal continuous test functions. The following notation is introduced to describe the discretization: if the continuous variable x_i is discretized with u_j that takes values in $\{1, \ldots, m_j\}$, then $u_j(k) = \beta$ means $x_i = \beta$ when $u_j = k$, β a scalar, $1 \leq k \leq m_j$.



(a) Discretized Branin-Hoo function (b) Discretized Goldstein-Price function

Figure 4.1: Two of the test functions with 1 discrete variable.

Test case 1: discretized Branin function. We modified the 2 dimensional *Branin-Hoo* function whose expression is

$$y(x_1, x_2) = (x'_2 - bx'^2_1 + cx'_1 - r)^2 + s(1-t)\cos(x'_1) + s,$$
$$x' = x'^{\min} + (x'^{\max} - x'^{\min}) \times x$$

where $b = 5/(4\pi^2), c = 5/\pi, r = 6, s = 10, t = 1/(8\pi)$, $x'^{\min} = [-5; 0], x'^{\max} = [10; 15]$ by keeping x_1 continuous in $[0; 1]$ and making x_2 discrete with 4 levels

$\{u(1) = 0; u(2) = 0.333; u(3) = 0.666; u(4) = 1\}$. The discretized Branin, which was already used in Zhang et al. [2020], has several local minima as shown in Figure 4.1a. The global optimum is located at $(x_1^\star, u^\star) = (0.182; u(3))$ with $y(x_1^\star, u^\star) = 2.791$.

Test case 2: discretized Goldstein function. As a second test case, the continuous Goldstein function

$$\begin{aligned} y(x_1, x_2) =& [1 + (x'_1 + x'_2 + 1)^2 (19 - 14x'_1 + 3x'^2_1 - 14x'_2 + 6x'_1 x'_2 + 3x'^2_2)] \\ & \times [30 + (2x'_1 - 3x'_2)^2 (18 - 32x'_1 + 12x'^2_1 + 48x'_2 - 36x'_1 x'_2 + 27x'^2_2)] \ , \\ x' =& x'^{\min} + (x'^{\max} - x'^{\min}) \times x \quad , \quad x'^{\min} = [-2, -2] \ , \ x'^{\max} = [2, 2] \end{aligned}$$

is partly discretized by replacing x_2 by u with 5 levels $\{u(1) = 0; u(2) = 1/2; u(3) = 1/2; u(4) = 3/4; u(5) = 1\}$. The discretized Goldstein, which has also been studied in Zhang et al. [2020], is drawn in Figure 4.1b. It has several local optima. The global optimum is located at $(x_1^\star, u^\star) = (0.5; u(2))$ with $y(x_1^\star, u^\star) = 3$.

Test case 3: discretized Hartman function. Two variables are discretized in the 6 dimensional Hartman function,

$$y(x) = -\sum_{i=1}^{4} \alpha_i \exp\left(-\sum_{j=1}^{d} A_{ij}(x_j - P_{ij})\right) ,$$

where $x \in [0, 1]^d$, $d = 6$, $\alpha = [1, 1.2, 3, 3.2]^\top$ and

$$A = \begin{pmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix} , \ P = 10^{-4} \begin{pmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4135 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{pmatrix} .$$

x_5 and x_6 are discretized with 5 and 4 levels respectively such that $\{u_1(1) = 0.350; u_1(2) = 0.257; u_1(3) = 0.477; u_1(4) = 0.312; u_1(5) = 0.657\}$ and $\{u_2(1) = 0.150; u_2(2) = 0.657; u_2(3) = 0.512; u_2(4) = 0.741\}$. Again, there are multiple local minima and the global optimum is located at $(x^\star, u^\star) = (0.202; 0.150; 0.477; 0.275; u_1(4), u_2(2))$ with $y(x^\star, u^\star) = -3.322$.

Euler-Bernoulli beam bending problem. This test case corresponds to an horizontal beam that is clamped at one end and subject to a vertical force at the other end. If the length of the beam is sufficiently long compared to the dimensions of its cross section,

and if it is operating within its linear elastic range, the final beam deflection y (to be minimized) is expressed as

$$D(L, S, \tilde{I}) = \frac{L^3}{3\, S^2\, \tilde{I}} \tag{4.15}$$

where $L \in [0, 1]$ is the horizontal length of the beam, $S \in [0, 1]$ is the cross-section area and $\tilde{I} = I/S^2, \in \{\tilde{I}(1), \tilde{I}(2), \ldots, \tilde{I}(12)\}$ is the normalized moment of inertia that can explicitly be derived for a given catalog of beam profiles. The 12 levels of the normalized moment of inertia are

$$\tilde{I} = \{0.083; 0.139; 0.380; 0.080; 0.133; 0.363; 0.086; 0.136; 0.360; 0.092; 0.138; 0.369\} \ . \tag{4.16}$$

We are interested in finding the best compromise between a minimization of the vertical deflection and the total weight, as expressed in the objective

$$y(x_1, x_2, u_1) = D(L, S, \tilde{I}) + \alpha L\, S \ , \tag{4.17}$$

$$\text{where } L = 10 + 10 \times x_1 \ , \ S = 1 + x_2 \ , \ u_1 = \tilde{I} \ , \ \alpha = 60 \tag{4.18}$$

$$\text{and } (x_1, x_2) \in [0, 1]^2 \ . \tag{4.19}$$

The solution is $(x_1^\star, x_2^\star, u_1^\star) = (0; 0.43; \tilde{I}(3))$ with output $y^\star = 1.287385 \times 10^3$.

4.2.2 Experiments setup and metrics

The optimization of each pair of algorithm and test case are repeated 50 times from different initial DoEs. The DoEs are generated by minimax Latin Hypercube Sampling. The size of the DoEs is $N_{\text{DoE}} = 4 \times n_c \times n_d \times \max(m_i)$ and a budget of $N_{\text{DoE}} + 50$ evaluations of the true objective function. Remember that the true objective function is supposed to be computationally intensive although it is not in these experiments so that runs can be repeated. The evolution strategies are stopped after $N_{\text{DoE}} + 50$ evaluations of the true function, like the other algorithms.

The internal local optimizer, COBYLA, is restarted 5 times during the likelihood maximization and 10 times during the maximization of the acquisition criterion. The focus-search algorithm has a sample size of 1000 with 5 boundary reduction iterations and 3 multi-starts, for a total of 3000 calls to the acquisition criterion.

A summary of the dimensions involved in the different examples is given in Table 4.3.

4.3 Results and discussion

In this section we analyze the obtained results by employing 4 main metrics. The performance of an algorithm is classically described by the median objective function over the 50 repeated runs, calculated at each iteration. The associated measure of dispersion of the performance is the interquartile over the repetitions as a function of the iteration. To

Name	n_c	n_d	m_i	N_{DoE}	n_{local}
Branin-Hoo	1	1	4	16	> 4
Goldstein	2	1	5	40	> 4
Hartmann	4	2	{5,4}	160	> 4
Beam Bending	2	1	12	96	NA

Table 4.3: Dimensions and DoE size of the test cases.

discriminate between methods that are rapid but provide rough solutions from the ones that take more time but yield better solutions, the two other metrics are based on the definition of targets. For each test case, a target is a given quantile of all the objectives functions found by all the algorithms throughout all the repetitions. A 10% target is difficult, while a 50% target is the median performance. The third metric is the iteration number at which the median objective function of a given algorithm reaches a given target. The fourth metric is the success rate (given a target), which is the percentage of the runs that do better than the target. The metrics associated to the quantile targets have the advantage that they are normalized with respect to the test cases: thanks to the quantiles, the definitions of an easy, a median or a hard target stands across the different functions to optimize. The target-based metrics will later be averaged over the different test cases.

Let us now review the performances of the algorithms on each test case.

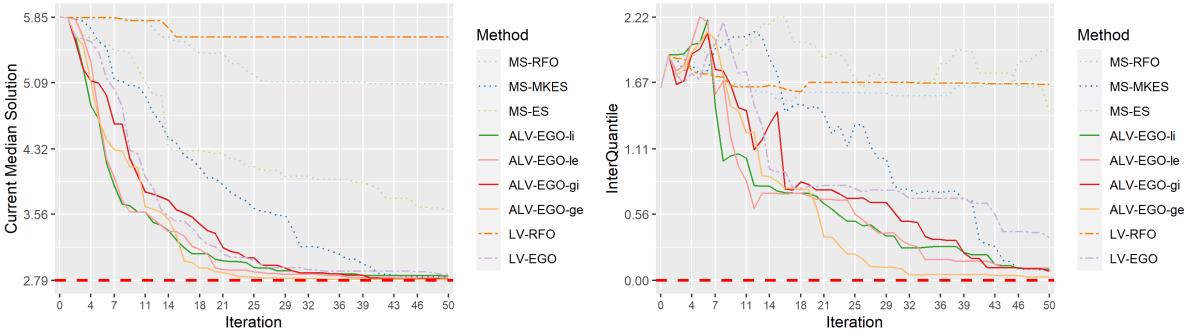
4.3.1 Analytical test functions

Branin function. Figure 4.2 presents the results for the Branin function with the four metrics. On the top left plot, showing the median value for the objective function, it is clear that the two methods that rely on the random forest metamodel (MS-RFO and LV-RFO) are overtaken by all other methods. This indicates that, whether in the mixed or in the latent-augmented space, random forests do not represent sufficiently well the Branin function in comparison to Gaussian processes. Looking at Figure 4.2b, it is observed that the fast methods typically have the lowest spread in performance and vice versa. This is expected as non converging runs may yield a wide range of performances. All methods involving the discrete constraint (i.e., the augmented Lagrangians) managed to improve over the LV-EGO performance; and including a mixed metamodel increased significantly the success rate and the median solution for the evolutionary strategy.

Regarding the success rate on Figure 4.2d, the methods MS-MKES, LV-EGO, ALV-EGO-li, -le, -ge and -gi were the most prominent, the latter being capable to reach success rates of about 20% for a 10% target. Notice that all these methods contain Gaussian processes. Indeed, the Branin function is easy to represent by a GP whether continuous or mixed. In the same vein, MS-MKES which differs from MS-ES by the use of a GP, clearly benefits from that metamodel.

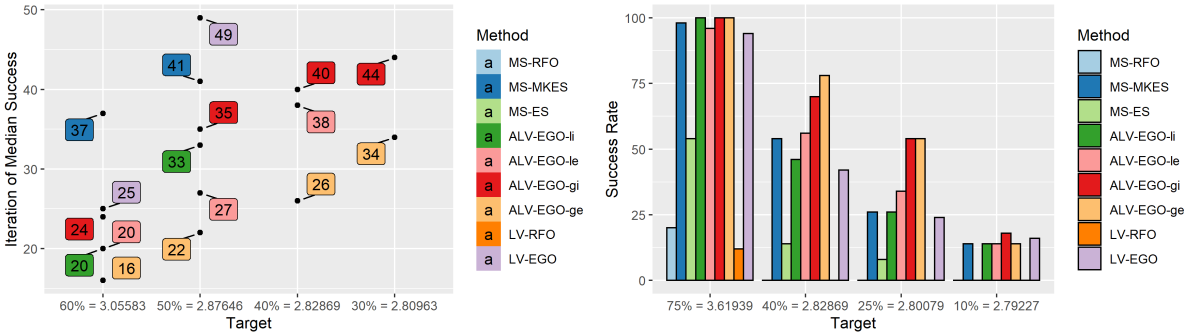
All ALV- methods, which account for the discrete constraint, obtained the best median performances. ALV-EGO-ge in particular found all targets, in the median sense, earlier than the other algorithms as can be seen from Figure 4.2c.

A last comment is necessary regarding the bottom of Figure 4.2: the plot on the left describes the median performance (in terms of targets reached) while the right plot counts the success rate at reaching a target over all runs. Therefore, some targets are reached on the right by some of the runs of a given algorithm, while they are never attained on the left by the median of the same algorithm. This comment stands accross all test cases.



(a) Median solution

(b) Interquartile range



(c) Iteration to median success

(d) Success rate

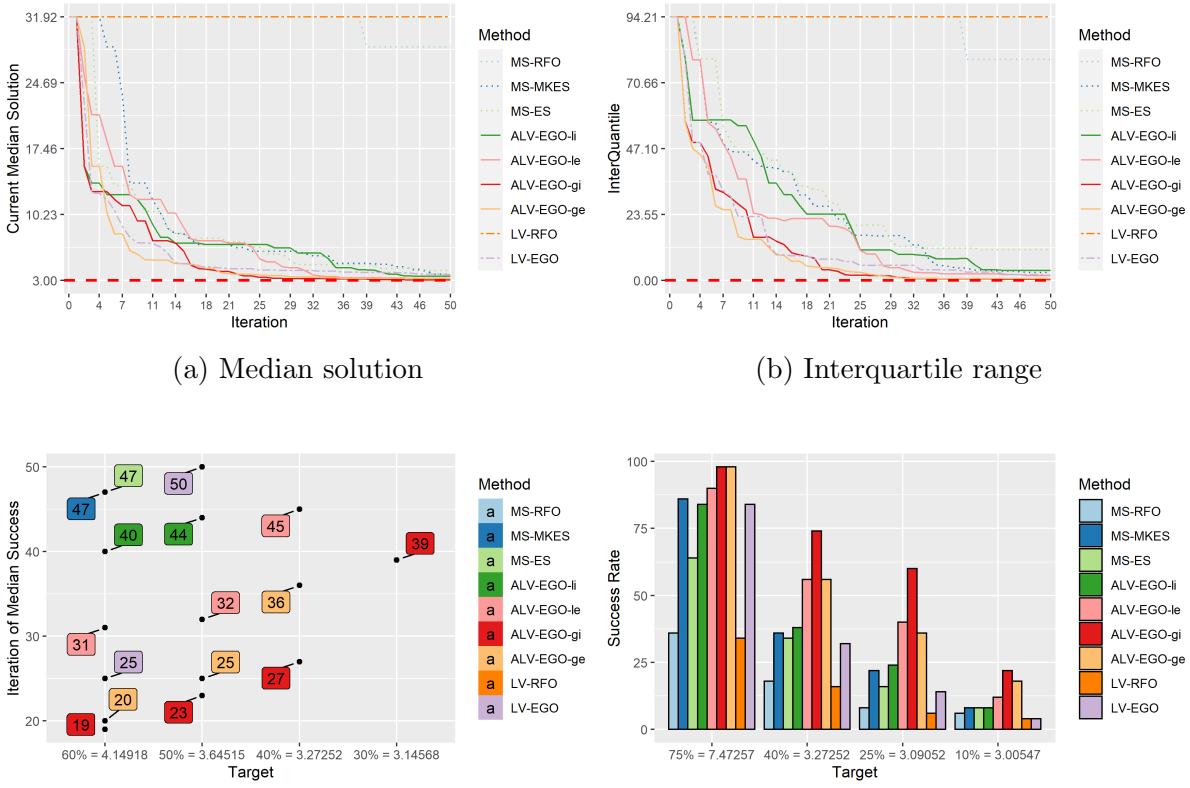
Figure 4.2: Comparison of all 9 algorithms on the Branin function. $y^\star = 2.79118$.

Goldstein function. The experiments done with the Goldstein test function are summed up in Figure 4.3. Like with the Branin function, algorithms relying on random forests (LV-RFO and MS-RFO) showed both poor performance (top left plot). The associated high constant interquartile (top right) is that of the best points in the initial designs, which remains unchanged since no better point is found by these algorithms.

Considering the success rates for all targets (bottom plots), it is seen that accounting for the discreteness through a constraint (which is the distinctive feature of ALV- methods) is useful with the Goldstein function: like with Branin, ALV-EGO-gi is the best performer, but the other ALV- follow and outperform LV-EGO. All ALV- strategies almost reach the absolute target of percentile 25% with a rate of 25% or higher.

The comparison of the plots 4.3c and 4.3d also shows that, behind the ALV- methods, LV-EGO has a good median performance (cf. Figure 4.3c) but more of the MS-MKES

searches manage to find difficult targets (the 25% and 10% quantiles).



(a) Median solution

(b) Interquartile range

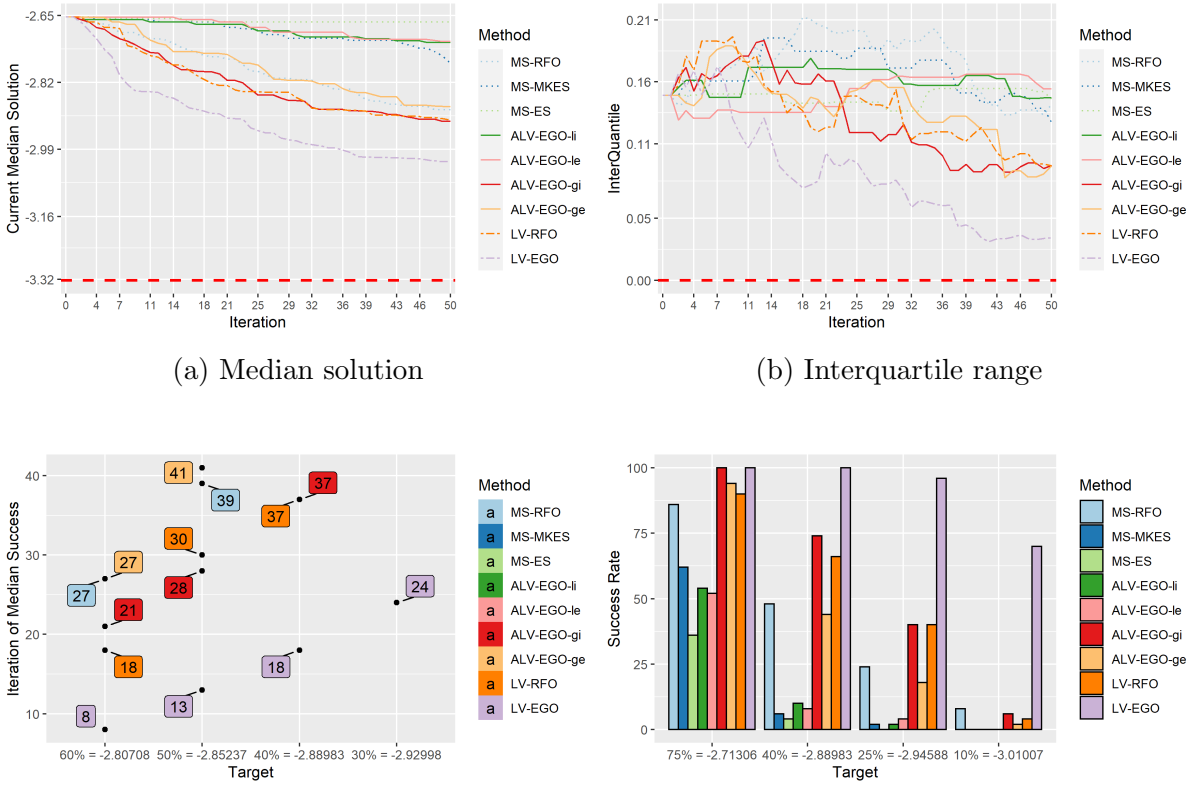
(c) Iteration to median success

(d) Success rate

Figure 4.3: Comparison of the 9 algorithms on the Goldstein function. $y^\star = 3$.

Hartmann function. Results on the Hartmann function which has 4 continuous and 2 discrete variables, with a total of 9 discrete levels, will be impacted by the sensitivity of the algorithms to an increase in dimension. These results are reported in Figure 4.4.

LV-EGO stands out as the best method with respect to all criteria for Hartmann. The next two best methods are LV-RFO and ALV-EGO-gi, followed by MS-RFO and ALV-EGO-ge. This time, LV-RFO and MS-RFO, which both rely on random forests, belong to the efficient methods: random forests gain in relative performance with respect to the GPs when the dimension and the size of the initial DoE increase. For Hartmann, LV-EGO consistently outperforms the ALV- implementations. The importance of keeping the coupling between discrete and latent variables during the optimization seems less crucial, and even somewhat detrimental, in the Hartmann case. We think that this is due to the very tight budget (50 iterations after the initial DoE) which does not allow the convergence of the optimizers, as can be seen in the Plot 4.4a where the global optimum is not reached. Because the optimum is not really found, constraints on discreteness are superfluous and their handling through the pre-image problem is sufficient. As in the other test cases, MS-ES was slower than the other methods.



(a) Median solution

(b) Interquartile range

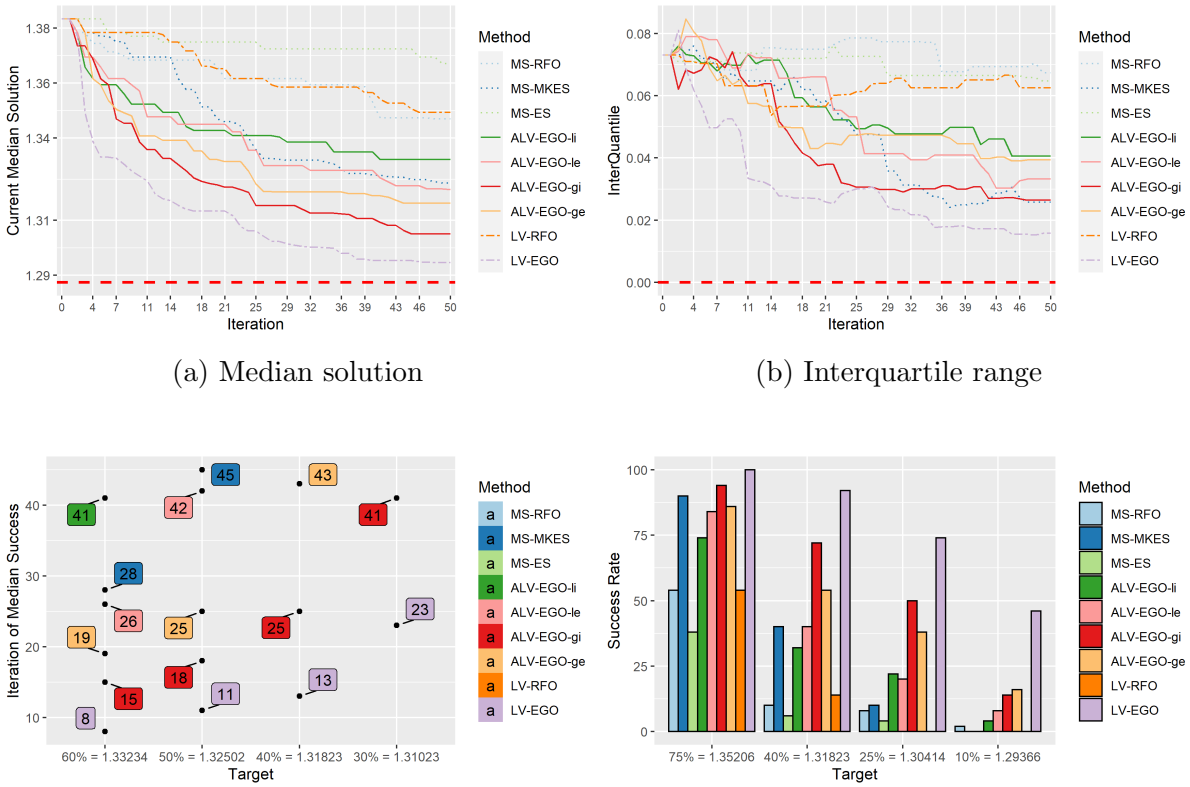
(c) Iteration to median success

(d) Success rate

Figure 4.4: Comparison of the 9 algorithms on the Hartmann function (for which $y^\star = -3.32237$).

4.3.2 Beam bending application

Optimization results. Figure 4.5 summarizes the 4 comparison metrics of all 9 algorithms in the bended beam test case. The ranking of the algorithms is similar to that obtained with the Branin and Goldstein functions. LV-EGO has the best convergence both in terms of median speed (cf. plots of the left column) and accuracy (bottom right plot). ALV-EGO-gi is the second most efficient method followed by ALV-EGO-ge. Again, the algorithms that resort to random forests, LV-RFO and MS-RFO, are the slowest and most inaccurate. They share this counter-performance with MS-ES.



(a) Median solution

(b) Interquartile range

(c) Iteration to median success

(d) Success rate

Figure 4.5: Comparison of all 9 algorithms on the beam design test case ($y^\star = 1.28738$).

Latent variables in the beam application. The beam subject to a bending load is a test case that allows to interpret the latent variables. Indeed, the normalized moment of inertia, $\tilde{I}$, is a candidate latent variable once it is allowed to take continuous values as it determines, with the continuous cross-section S and the length L, the output (the penalized beam deflection) y in Equation (4.19). The levels of $\tilde{I}$ (given in Equation (4.16)) correspond to 3 increasingly hollow profiles of 4 shapes, as illustrated in Figure 4.6. Because a relaxed $\tilde{I}$ is a possible latent variable, it is expected that the latent variables $\phi^{(t)}$ learned from the data will be grouped in the same way as $\tilde{I}$. Looking at $\tilde{I}$ values and at Figure 4.6, we thus expect, in the image space defined by latent variables, three groups of levels: those corresponding to solid forms (levels $\{1, 4, 7, 10\}$), medium-hollow forms (levels $\{2, 5, 8, 11\}$) and hollow forms (levels $\{3, 6, 9, 12\}$).

For the sake of interpretation, we select 1 run that found the global optimum with the Vanilla LV-EGO algorithm. In Figure 4.7, we represent in a color scale the estimated correlation matrix corresponding to the categorical kernel of Equation (4.5), at iterations $[1; 26; 49; 50]$.

At the beginning of the optimization, at iteration 1, we can see a block-structure which corresponds quite well to the three groups of forms described above. This structure

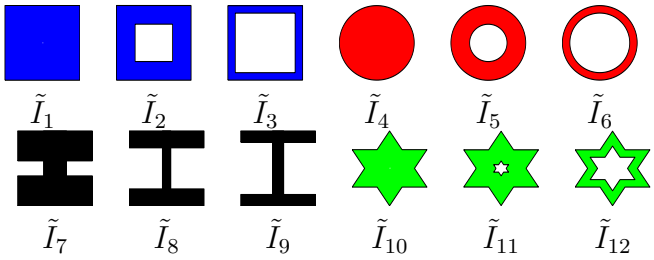
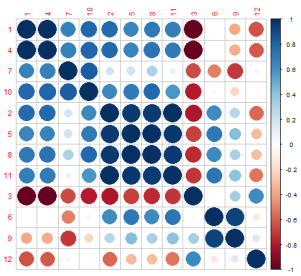
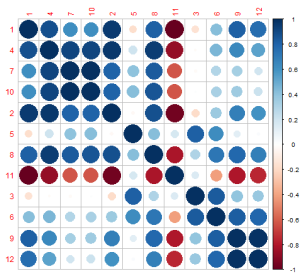


Figure 4.6: Shapes of the considered beam profiles. The scale differs from one picture to another, as the areas are supposed to be the same for each cross-section. From Roustant et al. [2020].

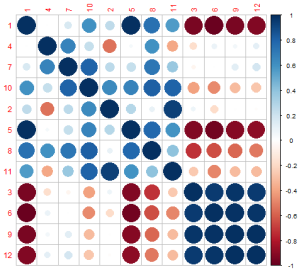
becomes less clear for the next iterations of the LV-EGO algorithm. This may be explained by the fact that the algorithm creates an unbalanced design, with more points in the promising areas according to the optimizers, so that all levels are no longer properly represented.



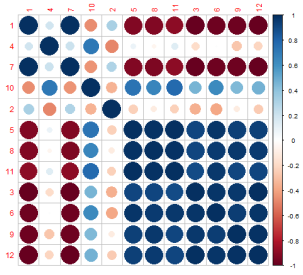
(a) Correlation of the latent variables at iteration #1



(b) Correlation of the latent variables at iteration #26



(c) Correlation of the latent variables at iteration #49



(d) Correlation of the latent variables at iteration #50

Figure 4.7: Representation of the correlation between the latent variables at various iterations t. The correlations correspond to the categorical kernel of Equation (4.5). The levels were grouped according to $\tilde{I}$: $\{1, 4, 7, 10\}$, $\{2, 5, 8, 11\}$, $\{3, 6, 9, 12\}$.

4.3.3 Summarized results

The results of all the previous test cases which are measured through targets can be averaged. For example, the success rate of an algorithm at 25% difficulty is the average of the rates for the 25% quantiles of all test cases. The average results are presented in Figure 4.8.

The three leading algorithms out of the 9 tested are ALV-EGO-gi, -ge and LV-EGO. Among them, LV-EGO is slightly better at locating difficult targets (10% quantile) while ALV-EGO-gi (closely followed by ALV-EGO-ge) is more robust at locating 50% targets as can be seen from the median success plot in Figure 4.8a. All three algorithms have in common to use latent variables. In particular, these algorithms outperformed MS-MKES which benefits from a Gaussian process but works only in the mixed space, i.e., MS-MKES does not imply latent variables. This shows that latent variables are useful to speed up a Bayesian search for mixed problems.

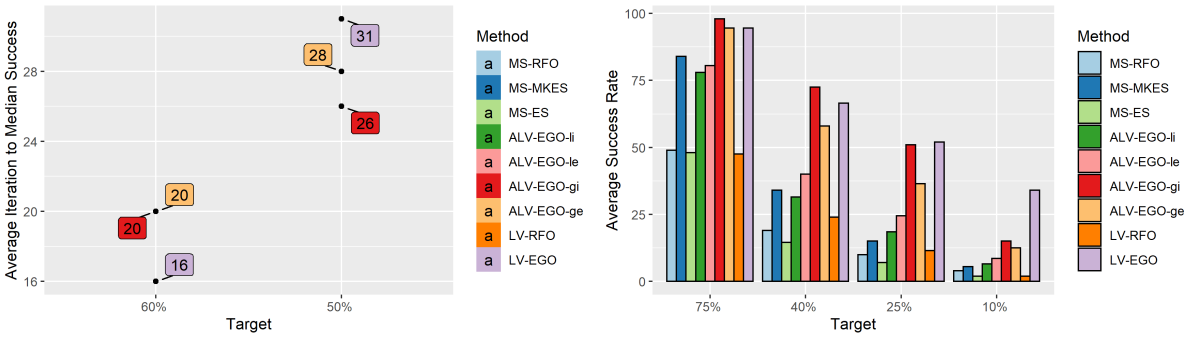
No clear advantage, on the average, was found for accounting for the discrete nature of the variables through constraints: LV-EGO, which ignores the link between latent variables and the discrete variables until the pre-image problem, is competitive with the best of the augmented Lagrangian ALV-EGO algorithms. We hypothesize that the constraint on latent variables, by creating disconnected feasibility islands around $\phi^{(t)}(u),\ u \in \mathcal{U}$, makes the optimization of the acquisition criterion almost as difficult to solve as it originally was in the mixed space, therefore not allowing to fully benefit from the continuity of the $\mathcal{X} \times \mathcal{L}$ space.

In our tests, the global updating of the Lagrange multipliers was always preferable to the local counterparts, ALV-EGO-gi and -ge eclipsing ALV-EGO-li and -le. The ALV-EGO-gi approach, where the discrete constraint is relaxed and turned into an inequality (Equation (4.7)), works better on the average than ALV-EGO-ge where the constraint is an equality. This illustrates the positive effect of the relaxation ϵ, that softens the phenomenon we mentioned above where the feasible domain is broken into disconnected regions.

MS-ES is consistently less efficient than the other algorithms. It was expected, because there is no metamodel to save calls to the function. Furthermore, the sampling is done in the mixed space. The optimizers based on random forests have also rather poor average performances, to the exception of the 6 dimensional Hartmann function. We believe the random forests need a sufficiently large initial DoE (which happened with a higher dimension) to fruitfully guide the search.

As a final comment, we discuss the necessity of re-estimating the latent variables at each iteration. The estimation of the latent variables has an important numerical cost of about $qt^3 \sum_{i=1}^{n_d} m_i$ operations at each iteration t (cf. Table 4.1).It was repeated at each iteration in the algorithms with latent variables considered so far. In the experiment reported in Figure 4.9, a version of the LV-EGO algorithm is considered where the latent variables are estimated once only, with the initial DoE, yielding the NR-LV-EGO algorithm (for Non Repeated estimation of $\phi()$).

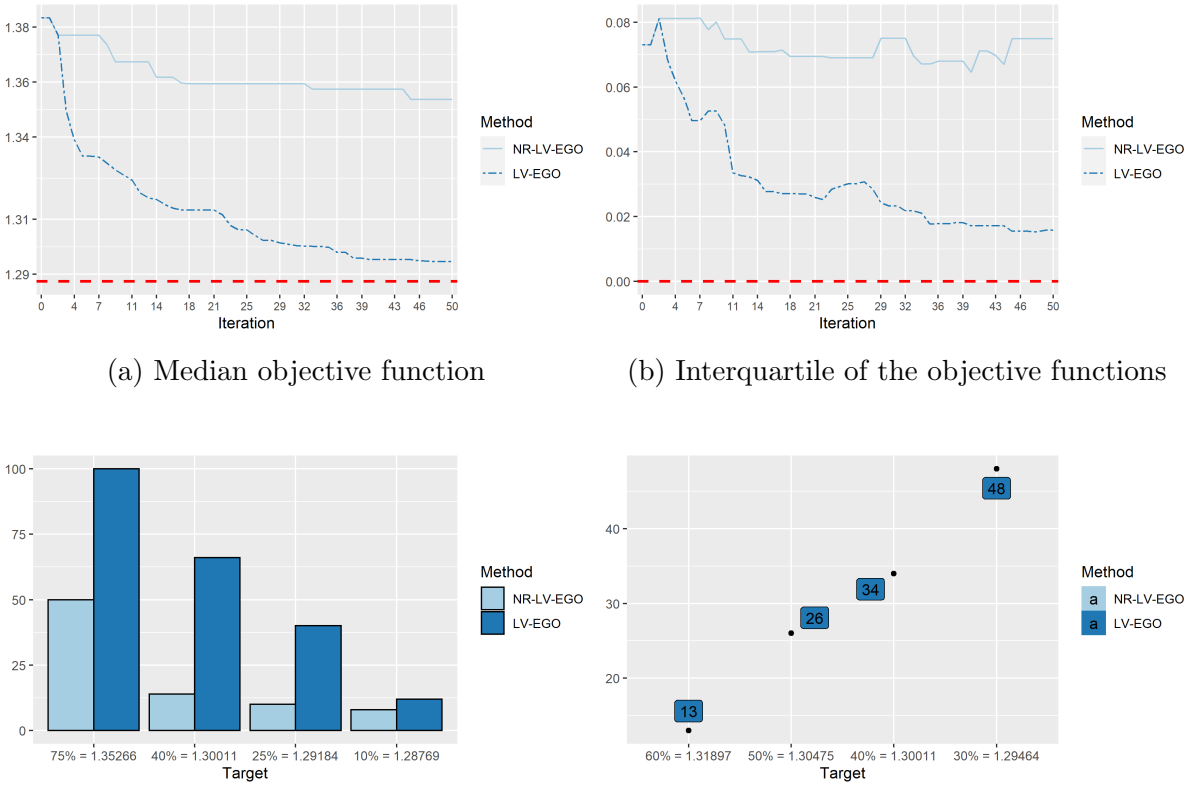
As can be seen in Figure 4.9 when comparing LV-EGO with NR-LV-EGO, the re-



(a) Average iteration to median success

(b) Average success rates

Figure 4.8: Comparison of the 9 algorithms tested with results averaged over all test cases.



(a) Median objective function

(b) Interquartile of the objective functions

(c) Success rate

(d) Iteration to median success

Figure 4.9: Comparison of LV-EGO with and without (NR-LV-EGO) a repeated estimation of the latent variables. Results for the beam design application.

estimation of the latent variables at each iteration, as implemented in the LV-EGO algorithm and its ALV-EGO variants, improves considerably its performance. An

accompanying result is the visualization of the correlation matrix of the discrete variable provided in Figure 4.7, where one notices that the correlation (hence the latent variables) evolves in time. Our experiments indicate that this evolution is beneficial to the optimization efficiency.

4.4 Complementary heuristics

In this section we present two different heuristics deduced from our empirical analysis of the LV-EGO performance with the different test cases.The formulations were inspired from the analysis of the performance of the augmented Lagrangian methods.

We have observed that ALV-EGO variants struggle to find suitable pairs (x^t, ℓ^t) that satisfies $g^{(t)}(\ell) \leq 0$ in the narrow region close to each latent variable $\phi^{(t)}$ (the q-ball of radius ϵ). As a direct consequence of this, in some test-cases like Hartmann and the Beam Bending applications, vanilla LV-EGO which first ignores the constraint and focuses on the EI maximization, manages to reach better and faster solutions than the rest of the methods.

One possible cause for this behavior is that as EI changes at each iteration, it does not guide the search. Also the connection between the recovery of the discrete variables and the optimization problems may not be strong enough. The basis for our heuristics is that the original problem

$$\begin{aligned} &\min_{x,\ell \in \mathcal{X}\times\mathcal{L}\subset\mathbb{R}^{n_c+n_\ell}} f^{(t)}(x,\ell) \coloneqq -\mathrm{EI}^{(t)}(x,\ell) \\ &\text{such that } g^{(t)}(\ell) \coloneqq \min_{u\in\mathcal{U}} \|\ell - \phi^{(t)}(u)\| - \epsilon \ \leq \ 0 \end{aligned} \tag{4.20}$$

can be handled through the penalization

$$\begin{aligned} &\min_{x,\ell \in \mathcal{X}\times\mathcal{L}\subset\mathbb{R}^{n_c+n_\ell}} P^{(t)}(x,\ell;\rho_t,\gamma_t) \\ &P^{(t)}(x,\ell;\rho_t,\gamma_t) \doteq -\mathrm{EI}^{(t)}(x,\ell)\exp\{-\rho_t(\gamma_t)h^{(t)}\}, \end{aligned} \tag{4.21}$$

ρ_t is the positive penalty parameter and $\exp\{-\rho_t(\gamma_t)h^{(t)}(\ell)\}$ is our penalty factor depending on the equality constraint

$$h^{(t)}(\ell) \coloneqq \min_{u\in\mathcal{U}} \|\ell - \phi^{(t)}(u)\| \tag{4.22}$$

i.e., $h^{(t)}(\ell) \equiv g^{(t)}(\ell)$ when $\epsilon = 0$. ρ_t is a function of

$$\gamma_t = \max_{u\in\mathcal{U}} \|\ell^t - \phi^{(t)}(u)\| \tag{4.23}$$

which is the largest distance of the best point obtained from the EGO iteration to any of the latent variables $\phi^{(t)}$. γ_t acts as a scaling quantity for the penalty factor and its role

will be detailed hereafter.

With this product by an exponential, we expect that feasible points with a low $\mathrm{EI}^{(t)}(x, \ell)$ will more often be proposed instead of infeasible points with better EI's. The exponential penalty may counterbalance well the rapid, also exponential, decrease in EI that is often observed in the neighborhood of visited points. Otherwise we expect that just feasible regions with radius γ_t should be taken into consideration.

Like any penalization approach, infeasible points may still be proposed and the inaccuracy in ℓ^{t+1} should be compensated by the very high expected improvement at ℓ^{t+1}, x^{t+1} which, in turns, expresses a good compromise between expected performance and gain of information. Under this setting we are going to evaluate two different ways to obtain ρ_t: we either assume that everything is independent at each iteration (meaning the ϕ changes at each iteration), or we assume that there is a dependency on the previous values (meaning that the ϕ converge). The first scenario leads to a **static exponential penalty** while the second leads to an **adaptive exponential penalty**.

The two approaches that we are describing are called heuristics because we did not mathematically prove that they provide a solution to Problem (4.6), as opposed to the augmented Lagrangians that have proofs partly included in Appendix A.1. However, as will be seen, they provide very good empirical results.

4.4.1 Static exponential penalty

The first idea tries to balance the contribution of $h^{(t)}$, which measures feasibility, and $\mathrm{EI}^{(t)}$ the expected improvement within the penalized function $P^{(t)}$ through the estimation of ρ_t. This can be done by defining a reference value $\mathrm{EI}_r^{(t)}$ that will help to identify how $P^{(t)}$ should behave. This can be summarized as

	$h^{(t)} < \gamma_t$	$h^{(t)} = \gamma_t$	$h^{(t)} > \gamma_t$
$\mathrm{EI}^{(t)}(x, \ell) > \mathrm{EI}_r^{(t)}$	$--$	$-$	$+-$
$\mathrm{EI}^{(t)}(x, \ell) = \mathrm{EI}_r^{(t)}$	$-$	1	$+$
$\mathrm{EI}^{(t)}(x, \ell) < \mathrm{EI}_r^{(t)}$	$+-$	$+$	$++$

Table 4.4: $P^{(t)}$ expected variations

The penalty parameter $\rho_t > 0$ is obtained from the middle scenario $1 = \mathrm{EI}_r^{(t)} \exp\{-\rho_t \gamma_t\}$, that is

$$\rho_t = \begin{cases} \frac{\log \mathrm{EI}_r^{(t)}}{\gamma_t}, & \text{if } \mathrm{EI}_r^{(t)} \geq 1 \\ \frac{-\log \mathrm{EI}_r^{(t)}}{\gamma_t}, & \text{otherwise,} \end{cases} \tag{4.24}$$

γ_t (cf. Equation 4.23), which measures the worst distance between a point and the feasible domain, is an indication of how difficult it can be to satisfy the discreteness

constraint. If γ_t is small, it is easy to solve the discrete pre-image problem and it does not harm to force an early convergence to the feasible domain with a strong penalty, hence the $1/\gamma_t$ term. Vice versa, a large γ_t is associated to a more progressive penalization so as to allow the optimizer to take short cuts through the infeasible domain.

The reference $\mathrm{EI}_r^{(t)}$ is obtained by evaluating $\mathrm{EI}^{(t)}(x,\ell)$ on a random DoE $(\mathbf{X}',\mathbf{L}') \in (\mathcal{X},\mathcal{L})^{N'_{\mathrm{DoE}}}$:

$$\mathrm{EI}_r^{(t)} = \begin{cases} \max_{(x,\ell)\in(\mathbf{X}',\mathbf{L}')} \mathrm{EI}^{(t)}(x,\ell), & \text{if } \max_{(x,\ell)\in(\mathbf{X}',\mathbf{L}')} \mathrm{EI}^{(t)}(x,\ell) \geq 1 \\ \mathrm{median}_{(x,\ell)\in(\mathbf{X}',\mathbf{L}')} \mathrm{EI}^{(t)}(x,\ell), & \text{otherwise.} \end{cases} \tag{4.25}$$

The overall penalization scheme (Equations (4.21), (4.24) and (4.25)) is dominated by situations where $\mathrm{EI}_r^{(t)} \ll 1$ where a negligible progress in objective is expected and the emphasis is shifted towards satisfying the discreteness constraint. This **L**atent **V**ariable **EGO** with **s**tatic exponential penalty or LV-EGO-s is summarized in the following algorithm

Algorithm 9 LV-EGO-s Algorithm

1: **generate** the initial DoE of size N_{DoE} for $(\mathbf{X},\mathbf{U})$
2: **costly function evaluations** $y(x^i,u^i)$, $i=1,\ldots,N_{\mathrm{DoE}}$, $t \leftarrow N_{\mathrm{DoE}}$
3: **initialize** budget, $\gamma_0 = 2\sqrt{q}$ (where q is the rank of the latent variables parameterization), define N'_{DoE}
4: **while** $t \leq$ budget **do**
5: **estimate** the latent variables $\phi^{(t)}$ and the GP parameters from current DoE.
6: Calculate a DoE $(\mathbf{X}',\mathbf{L}') \in (\mathcal{X},\mathcal{L})^{N'_{\mathrm{DoE}}}$.
7: Estimate $\mathrm{EI}_r^{(t)}$ according to equation 4.25
8: Estimate ρ_t according to equation 4.24
9: $(x^{t+1},\ell^{t+1}) = \arg\min_{(x,\ell)\in(\mathcal{X},\mathcal{L})} P^{(t)}(x,\ell;\rho_t,\gamma_t)$
10: **update** $\gamma_{t+1} = \max_{u\in\mathcal{U}} \|\ell^{t+1} - \phi^{(t)}(u)\|$
11: **Recover** the *discrete pre-image* component u^{t+1} as: $u^{t+1} = \arg\max_{u\in\mathcal{U}} \mathrm{EI}^{(t)}(x^{t+1},\phi^{(t)}(u))$
12: **update DoE**: add (x^{t+1},u^{t+1}) and its costly evaluation $y(x^{t+1},u^{t+1})$ to the DoE $(\mathbf{X},\mathbf{U})$.
13: $t \leftarrow t+1$
14: **end while**
15: **return** $(x^\star,u^\star) = \arg\min_{(\mathbf{X},\mathbf{U})} y(x,u)$

A numerical advantages of this technique is that it does not require internal loops to solve the dual sub-optimization problem like it was the case with the augmented Lagrangians (ALV-EGO-...) approaches.

4.4.2 Adaptive exponential penalty

The main idea in this second technique is to propose an iterative way to compute the penalty parameter ρ_t. To do that we introduce the quantity $P^{(t)}(x^{t+1}, \ell^{t+1}; \rho_t, \gamma_t) + \mathrm{EI}^{(t)}(x^{t+1}, \phi^{(t)}(u^{t+1}))$ as a measure of the discrepancy between the proposed EI and the penalized one. Note that we are using all the values obtained at the end of optimization and pre-image on iteration t. With this information it could be possible to detect how much of the EI is being really recovered. Then if this value is low, the recovered pair x^{t+1}, ℓ^{t+1} corresponds to a feasible or near feasible pair. We can define $0 < \alpha_{t+1} < 1$ as the normalized EI discrepancy factor

$$\alpha_{t+1} = \frac{P^{(t)}(x^{t+1}, \ell^{t+1}; \rho_t, \gamma_t) + \mathrm{EI}^{(t)}(x^{t+1}, \phi^{(t)}(u^{t+1}))}{\max\left(P^{(t)}(x^{t+1}, \ell^{t+1}; \rho_t, \gamma_t); \mathrm{EI}^{(t)}(x^{t+1}, \phi^{(t)}(u^{t+1}))\right)}, \tag{4.26}$$

where $\alpha_{t+1} \approx 0$ when the EI is similar between (x^{t+1}, ℓ^{t+1}) and (x^{t+1}, u^{t+1}) and $\alpha_{t+1} \approx 1$ when not. After the LV-EGO iteration, we want to obtain the next ρ_{t+1} such that the penalized function $P^{(t+1)}(\gamma_t, \dots)$ for the theoretical worst feasible point (x^γ, ℓ^γ) located at $h^{(t)} = \gamma_t$ will increase proportionally to how much α_{t+1} differs from 0. This rule is expressed with the following equations at the end of iteration t

$$\begin{aligned} P^{(t+1)}(x^\gamma, \ell^\gamma; \rho_{t+1}, \gamma_{t+1}) &= -\mathrm{EI}^{(t)}(x^\gamma, \ell^\gamma) \exp\{-\rho_{t+1}\gamma_{t+1}\} \\ &= -(1 - \alpha_{t+1})\mathrm{EI}^{(t)}(x^\gamma, \ell^\gamma) \exp\{-\rho_t\gamma_{t+1}\} \end{aligned}$$

From here we can solve for ρ_{t+1}

$$\rho_{t+1} = \rho_t - \frac{\log(1^+ - \alpha_{t+1})}{\gamma_{t+1}}, \tag{4.27}$$

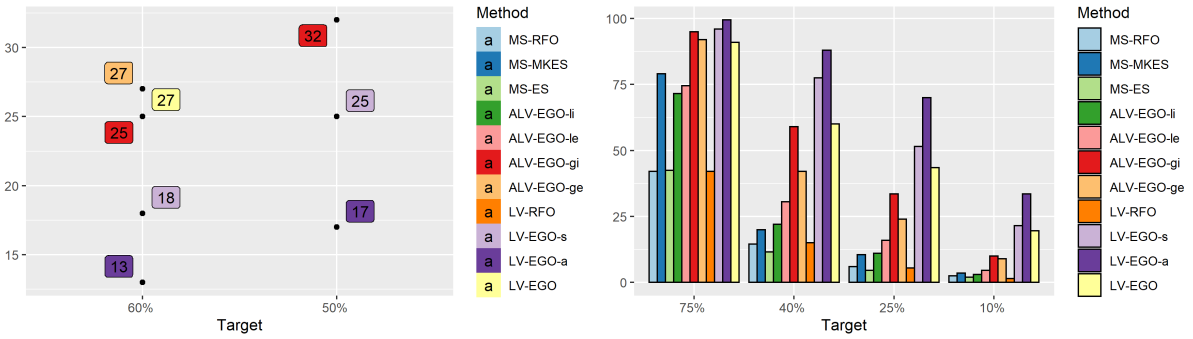
where 1^+ represents a slightly increase to 1 to avoid numerical issues when $\alpha_{t+1} \approx 1$, which may happen during the iterations. Since $(1^+ - \alpha_{t+1}) \leq 1$, ρ_t in Eq. (4.27) is an increasing sequence of t. With this we can construct the LV-EGO-a, where a stands for adaptive, algorithm.

Algorithm 10 LV-EGO-a Algorithm

generate the initial DoE of size N_{DoE} for $(\mathbf{X}, \mathbf{U})$
costly function evaluations $y(x^i, u^i)$, $i = 1, \ldots, N_{\text{DoE}}$, $t \leftarrow N_{\text{DoE}}$
initialize budget, $\gamma_0 = 2\sqrt{q}$ (where q is the rank of the latent variables parameterization), $\rho_0 = 0$.
while $t \leq$ budget **do**
  estimate the latent variables $\phi^{(t)}$ and the GP parameters from current DoE.
  $(x^{t+1}, \ell^{t+1}) = \arg\min_{(x,\ell)\in(\mathcal{X},\mathcal{L})} P^{(t)}(x, \ell; \rho_t, \gamma_t)$
  Recover the *discrete pre-image* component u^{t+1} as: $u^{t+1} = \arg\max_{u\in\mathcal{U}} \text{EI}^{(t)}(x^{t+1}, \phi^{(t)}(u))$
  update DoE: add (x^{t+1}, u^{t+1}) and its costly evaluation $y(x^{t+1}, u^{t+1})$ to the DoE $(\mathbf{X}, \mathbf{U})$.
  Estimate α_{t+1} according to equation 4.26
  update $\gamma_{t+1} = \max_{u\in\mathcal{U}} \|\ell^{t+1} - \phi^{(t)}(u)\|$
  Estimate ρ_t according to equation 4.27
  $t \leftarrow t + 1$
end while
return $(x^\star, u^\star) = \arg\min_{(\mathbf{X},\mathbf{U})} y(x, u)$

The main difference between the LV-EGO-a and LV-EGO-s heuristics is how in the former α_{t+1} will limit how much we can increase the penalty at each iteration, when in the latter ρ_t can increase or decrease more freely: in LV-EGO-s, ρ does not depend on any other optimization step or on the creation of an extra DoE. The LV-EGO-a approach can be seen as an approximated solution to the penalized problem (4.6) where the γ parameter could be related to the ρ proposed by Picheny et al. [2016].

4.4.3 Updated results



(a) Average iteration to median success

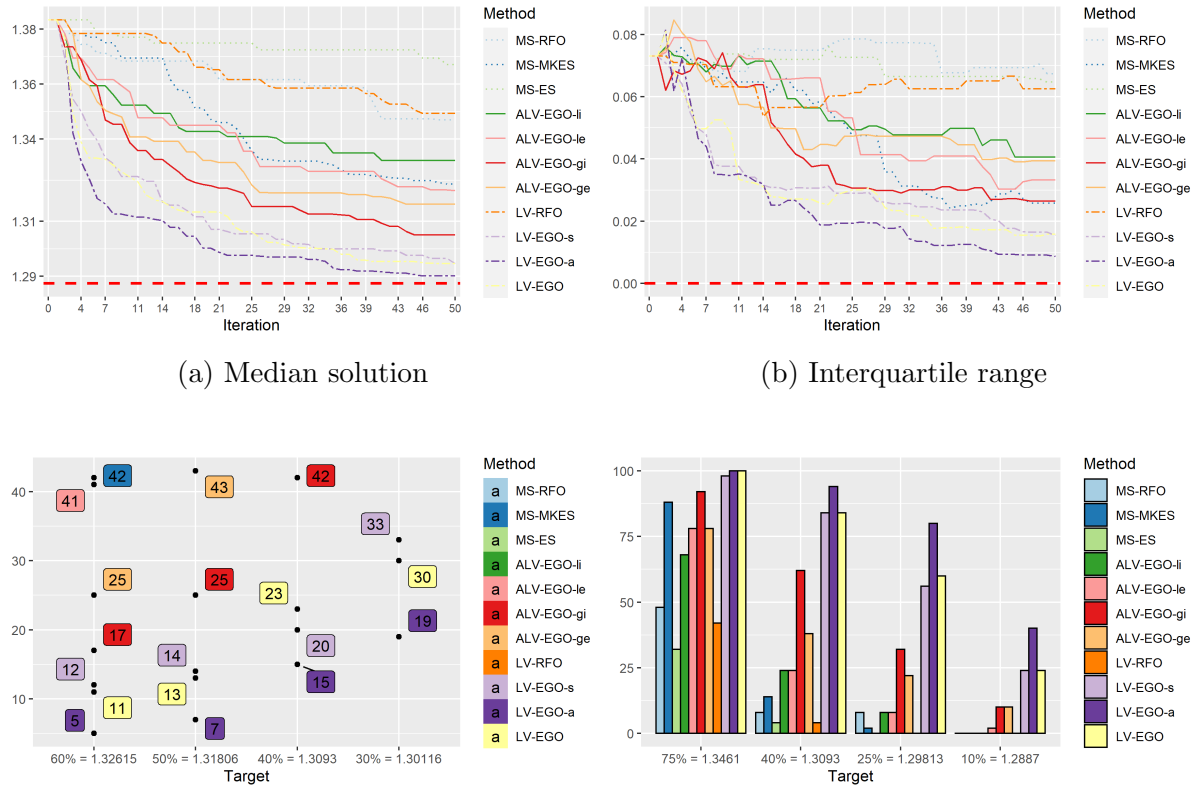
(b) Average success rates

Figure 4.10: Comparison of the 11 algorithms tested with results averaged over all test cases. By including LV-EGO-a and -s, the relative targets differ with respect to Figure 4.8 which explains changes in total performance.

Upon comparison of the averaged performance of all the algorithms for all the test cases and including the recently proposed heuristics, we observe on Figure 4.10 that both LV-EGO-s and -a formulations managed to outperform LV-EGO and all the ALV-EGO variants. It is the LV-EGO-a that clearly presents the fastest convergence and highest success rate.

More precisely LV-EGO-s managed to reach an average success rate of more than 50% for a 25% target and took less than half of the budget to reach a 50% median target improving the performance of LV-EGO and the . In the case of LV-EGO-a which performs even better than the -s formulation achieving the 50% target in one third of the budget and with a success rate close to 50% for the most difficult 10% target. This behavior can be explained by the combination of the exponential penalty that increases sufficiently fast to compensate for the rapid drop of the EI to 0, and the increasing penalty factor ρ_t that makes it possible to obtain high-performing non-feasible points at the beginning of the iterations. This clearly helps the iterative metamodel estimation and the overall performance of the LV-EGO.

4.4.4 Re-visiting the Beam Bending Problem



(a) Median solution

(b) Interquartile range

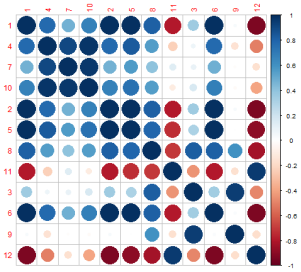
(c) Iteration to median success

(d) Success rate

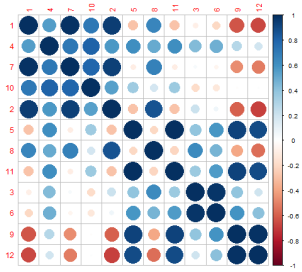
Figure 4.11: Comparison of all 11 algorithms on the beam design test case ($y^\star = 1.28738$). By including LV-EGO-a and -s, the relative targets differ with respect to Figure 4.5 which explains changes in total performance.

We now revisit the results regarding the Beam Bending application as presented in section 4.2.1 with the addition of the two heuristics, LV-EGO-s and LV-EGO-a. As shown in Figure 4.11 LV-EGO-a is the technique with the best performance, managing to improve all methods in overall and LV-EGO-s being slightly behind the vanilla LV-EGO algorithm.

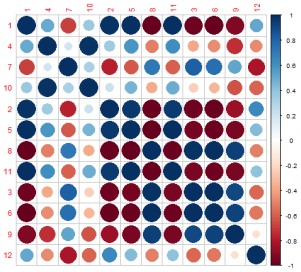
Figure 4.12 correspond to an updated representation of the correlation between the latent variables using LV-EGO-a. In this case is not possible to identify any of the original groups, neither a similar pattern of convergence, excepted, arguably, at the first iteration. This could either reinforce the unbalanced design hypothesis explained in section 4.3.2, or open a new discussion on how it is possible to obtain a correlation that would more clearly favor the apparition of different groups.



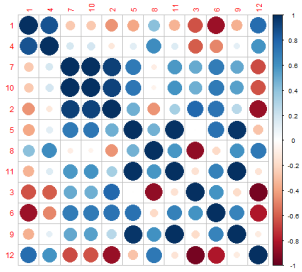
(a) Correlation of the latent variables at iteration #1



(b) Correlation of the latent variables at iteration #26



(c) Correlation of the latent variables at iteration #49



(d) Correlation of the latent variables at iteration #50

Figure 4.12: Representation of the correlation between the latent variables at various iterations t. The correlations correspond to the categorical kernel of Equation (4.5). The levels were grouped according to $\tilde{I}$: $\{1, 4, 7, 10\}$, $\{2, 5, 8, 11\}$, $\{3, 6, 9, 12\}$.

4.5 Conclusions

In this chapter, we have investigated five Bayesian optimization approaches to small and medium size mixed problems that hinged on latent variables. They differ in the way the coupling between the discrete variables and their relaxed pendants, the latent variables, is implemented. Algorithms involving latent variables were compared to other algorithms directly working in the mixed space and were found to consistently outperform them. LV-EGO and ALV-EGO-gi were more efficient (in terms of calls to the true objective function) than MS-MKES which also benefits from the Gaussian process. These first results show that latent variables provide a flexible way to handle mixed problems where the total number of levels and of variables is less or equal to about 10 variables and 10 levels in total.

Accounting for the discrete nature of some variables through a constraint during the relaxed optimization with augmented Lagrangians was not clearly found to further increase the performance of the search as LV-EGO competed equally and even sometimes

outperformed the ALV versions of the algorithms. It was also observed that expressing the discreteness as an inequality constraint by adding a tolerance was a better option than expressing it as an equality. The global updating strategy of the Lagrange multipliers, which to the best of our knowledge is original, improved over the more common local updating schemes. The random forests metamodels did not do as well as the Gaussian processes, whether in their continuous or mixed forms, within the Bayesian optimization algorithm.

Finally, we have introduced the concept of exponential penalization in the context of LV-EGO regularization, as a way to balance the sharp decrease in expected improvement during the search and to force convergence towards the feasible domain at the end. This inclusion proved to further increase the LV-EGO performance among all the test cases. We believe that introducing strong exponential penalties within a discrete constraint scenario does not harms the optimization because they instead of focusing in convergence, sample the discrete space which is an step towards feasibility in this setting.

Chapter 5

Inversion of a bi-linear black-box function

Contents

5.1 Scenario 1: $f(x) = X(x)\beta$ 54

5.1.1 Least-Squares approach 54

5.1.2 Gaussian approach 56

5.1.3 MCMC-based Bayesian approach 57

5.1.4 Computer Experiments 58

5.2 Scenario 2: $M(x)$, a non-linear black-box function . 67

5.2.1 Experiments MCMC-based Bayesian approach 68

5.3 Conclusions 69

This chapter focuses on studying different scenarios to solve an inverse problem for an application in radionuclide quantification. As presented in Clement et al. [2018], the inverse problem requires quantifying the mass of a radionuclide, that is then related to its activity measured by gamma spectrometry as

$$y = mf(x, u) + \eta, \tag{5.1}$$

where $y \in \mathbb{R}^p$ and small p (Usually $p \approxeq 6$). Here m, x, u are unknown quantities, with $m \in \mathbb{R}^+$, $x \in [0,1]^{n_c}$, $u \in \mathcal{U} = \prod_{j=1}^{n_d}\{1, \ldots, m_j\}$; and with an observation noise $\eta \sim \mathcal{N}(0, \gamma^2 I_p)$. This problem is also ill-posed, as reviewed in chapter 3 with the additions of a black-box function f that is potentially non-linear and can be expensive to evaluate and also could depend on mixed variables (x, u), and scarse data $p = 6$. Even though the mixed inputs property makes this problem able to be tackled by the LV-EGO methodology, the problem is already complex to solve for continuous variables. Therefore, in this chapter we will focus on defining a framework for the continuous approach, for a future mixed application of the LV-EGO methodology.

Towards tackling this continuous, potentially non linear, ill-posed inverse problem with scarse data, in this chapter we propose two global strategies for different scenarios of

increasing complexity for f. We start with the simplest case, assuming that f is an explicit continuous linear function $f(x) = x\beta$ depending on x only. This scenario still leads us to an ill-posed inverse problem that we analyze by defining classical and stochastic approaches (discussed in Chapter 3). In the second scenario, f is considered a non-linear black-box function, then we analyze the possible extension of the strategies proposed for the linear case.

It is important to remark that this particular problem, even in the first scenario, is classified as bi-linear due to the required estimation of the unknowns m and x Ling [2017]. Additionally, the product $mf(x)$ makes the problems immediately ill-posed, which differs from classical bi-linear self-calibration studies Ling [2017] Idier and Blanc-Féraud [2008], where the goal is to propose a low rank representation when m is a matrix, or a blind deconvolution where m is also a function of x. Therefore it corresponds to a problem with particularities that has not been studied before in the literature.

5.1 Scenario 1: $f(x) = X(x)\boldsymbol{\beta}$

Let be $y = mX(x)\ \boldsymbol{\beta} + \eta$ the bi-linear inverse problem with known vector $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_{n_c}]^\top$, $y \in \mathbb{R}^p$, $m \in \mathbb{R}^+$, η is the observation noise, and where $X \in \mathbb{R}^{p \times n_c}$ depends on the unknown $x = [x_1, \ldots x_{n_c}]^\top$ and is defined as

$$X(x) \triangleq \begin{bmatrix} x_1 & \ldots & x_{n_c} \\ \vdots & \vdots & \vdots \\ x_1 & \ldots & x_{n_c} \end{bmatrix}.$$

The objective is to obtain $\hat{m}$, an estimator for $m^\star$, which is the true but unknown value of m that explains y the best. Under this basic linear assumptions we study the possibility of finding an analytical expression, if possible for $\hat{m}$ by applying both Classical and Bayesian approaches.

5.1.1 Least-Squares approach

To solve this inverse problem from the Least-squares classical perspective, we proceed by computing the analytical expressions for $\hat{x}$ and $\hat{m}$ that minimize the least squares cost function $J(x, m) = ||y - mX(x)\boldsymbol{\beta}||^2$. This is done by solving the system of partial differential equations $\partial J/\partial x, \partial J/\partial m = 0$.

To understand the main difficulties of this approach and without loss of generality we consider the simpler bi-linear inverse problem

$$y = mx\beta + \eta,$$

where $x, y \in \mathbb{R}^p$, $\beta, m \in \mathbb{R}$. Here the cost function can be expressed in vectorial form as

$$\underset{\hat{x},\hat{m}}{\operatorname{argmin}} \quad (y - mx\beta)^T(y - mx\beta)$$

This problem is ill-posed, in the way that it has multiple possible solutions (m, x) (e.g. (m, x); $(m/2, 2x)$), and as we expressed in chapter 3 solving it would require adding a regularization term. Assuming a Gaussian prior for x, $\pi(x) \sim \mathcal{N}(\mu_x, \sigma_x^2 I_p)$, the least squares problem is now

$$\underset{\hat{x},\hat{m}}{\operatorname{argmin}} \quad (y - mx\beta)^T(y - mx\beta) + (x - \mu_x)^T\sigma_x^{-2}(x - \mu_x), \tag{5.2}$$

As we are interested in finding an estimator for $\hat{m}$, we first obtain $\hat{x}(m)$ from the partial differential equation $\partial J/\partial x = 0$ as

$$\hat{x}(m) = \frac{\mu_x + \sigma_x^2 m\beta y}{1 + \sigma_x^2 m^2\beta^2}.$$

Then replacing in J, we obtain

$$J(m) = \frac{(y - m\mu_x\beta)^T(y - m\mu_x\beta)}{1 + \sigma_x^2 m^2\beta^2},$$

which is quadratic on m in the numerator with a positive term in the denominator that will reduce while obtaining the derivative. This provides sufficient motivation about the existence of $\hat{m}$, which can be obtained from $J'(m) = 0$ as

$$m^2 + m\left(\frac{\omega^T\omega - y^Tyz}{\omega^Tyz}\right) + \frac{-y^T\omega}{\omega^Tyz} = 0$$

where $z = \beta^2\sigma_x^2$, $\omega = \beta\mu_x$. Then we can define the following theorem:

Theorem 5.1 *Let be $y = mx\beta + \eta$ the bi-linear inverse problem with $x \in \mathbb{R}^p$, $\beta \in \mathbb{R}, \beta \neq 0$, $y \in \mathbb{R}^p$ and $m \in \mathbb{R}^+$. It has a unique solution $\hat{m}$*

$$\hat{m} = \frac{\sigma_x^2|y|^2 - |\mu_x|^2 + \sqrt{(\sigma_x^2|y|^2 - |\mu_x|^2)^2 + 4\sigma_x^2(\mu_x^Ty)^2}}{2\beta\sigma_x^2(\mu_x^Ty)} \tag{5.3}$$

that satisfies $|\hat{m} - \hat{m}_{LS}| \leq \xi(\sigma_x^2)$ if and only if $\sigma_x > 0$ and $\hat{m}_{LS}$ corresponds to a Least Squares solution making $x = \mathrm{E}[x] = \mu_x$ as in equation B.6.

Under the regularized approach, meaning we have uncertainty on x, the condition $\sigma_x > 0$ is always satisfied, therefore equation 5.3 is a valid approximation. The complete procedure to obtain theorem 5.1 can be found in the appendix B.1 where we also evaluate the possibility to include a penalty term on m.

This solution could be extended to the more general case when $x \in \mathbb{R}^{n_c}$, by defining the matrix of continuous input $X \in \mathbb{R}^{p \times n_c}$ in which case the theorem 5.1 becomes:

Theorem 5.2 *Let $y = mX(x)\boldsymbol{\beta} + \eta$ be with $X \in \mathbb{R}^{p\times n_c}$ depends on the unknown $x = [x_1, \dots x_{n_c}]^\top$ and is defined as*

$$X(x) \triangleq \begin{bmatrix} x_1 & \dots & x_{n_c} \\ \vdots & \vdots & \vdots \\ x_1 & \dots & x_{n_c} \end{bmatrix},$$

the bi-linear inverse problem with known $\boldsymbol{\beta} = [\beta_1, \dots, \beta_{n_c}]^\top \neq 0$, $y \subset \mathbb{R}^p$ and $m \in \mathbb{R}^+$ that leads to the following regularization expression

$$\underset{m,X}{argmin} \quad J(X,m) = \ (y - mX\boldsymbol{\beta})^T(y - mX\boldsymbol{\beta}) + \sum_{i=1}^{n_c} \frac{(x_i - \mu_{x_i})^2}{\sigma_i^2},$$

has a unique positive solution $\hat{m}$

$$\hat{m} = \frac{\boldsymbol{\beta}^T\Sigma^{-1}\boldsymbol{\beta}|y|^2 - \boldsymbol{\beta}^T A^T A\boldsymbol{\beta} + \sqrt{(\boldsymbol{\beta}^T\Sigma^{-1}\boldsymbol{\beta}|y|^2 - \boldsymbol{\beta}^T A^T A\boldsymbol{\beta})^2 + 4(\boldsymbol{\beta}^T\Sigma^{-1}\boldsymbol{\beta})(\boldsymbol{\beta}^T A^T y)^2}}{2(\boldsymbol{\beta}^T A^T y)(\boldsymbol{\beta}^T\Sigma^{-1}\boldsymbol{\beta})} \tag{5.4}$$

that satisfies $|\hat{m} - \hat{m}_{LS}| \leq \xi(\Sigma)$ if and only if $\Sigma \succ 0$ is a positive definite matrix, where $A = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_{n_c}}]^\top$ and

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^{-2} & c_{1,2} & \dots & c_{1,n_c} \\ c_{2,1} & \sigma_{x_2}^{-2} & \dots & c_{2,n_c} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n_c,1} & c_{n_c,2} & \dots & \sigma_{x_{n_c}}^{-2} \end{pmatrix},$$

with off diagonal elements $c_{i,j} = c \in \mathbb{R} \forall i \neq j$; and $\hat{m}_{LS}$ corresponds to the Least Squares solution.

The complete deduction for this theorem can be found on the appendix B.3

Finally with the results obtained in equation 5.4, and trying to account for the uncertainty on x as under the Bayesian framework, we can generate n_x samples from $\pi(x)$, and construct an empirical distribution over m by evaluating on 5.4. This trick can be used in the next section when dealing with a continuous expensive black-box function.

5.1.2 Gaussian approach

Under the Bayesian framework, a classical approach consists in introducing a prior that can be conjugated with the available uncertainties. Here we recall the simpler bi-linear inverse problem

$$y = mx\beta + \eta,$$

where $x, y \in \mathbb{R}^p$, $\beta, m \in \mathbb{R}$. As η is assumed Gaussian $\eta \sim \mathcal{N}(0, \gamma^2\Gamma_0)$, a natural choice is to choose independent priors for m and x to be Gaussian. However by defining

$\pi(m) \sim \mathcal{N}(m_0, \sigma_m^2)$, $\pi(x) \sim \mathcal{N}(\mu_x, \sigma_x^2 I_{n_c})$, and assuming that the joint vector $z = (m, y, x)^T$ is Gaussian we cannot obtain $\pi(m|y) \sim \mathcal{N}(\mu^*, \Sigma^*)$ as a *conjugate* by employing Gaussian properties. This is due to the product of m x not leading to a Gaussian distribution on y, therefore this approach cannot be used even in the simplest case.

5.1.3 MCMC-based Bayesian approach

In this section, as we cannot derive a Gaussian estimator for m we can involve sampling methods, such as MCMC reviewed in Section 3.2.1. Under this framework, we do not require to specify $f(x)$ which can be either a deterministic formula or a black box function neither the priors $\pi(m), \pi(x)$. Therefore, we can avoid the problem of non Gaussian and non tractable distributions.

Starting from $y = mf(x) + \eta$ and applying Bayes' rule we obtain the approximate joint posterior $\pi(m, x|y)$ as

$$\pi\left(\begin{pmatrix} m \\ x \end{pmatrix} \middle| \; y\right) \propto \pi\left(y \middle| \begin{pmatrix} m \\ x \end{pmatrix}\right) \pi(m)\pi(x) \tag{5.5}$$

where $\pi(y|m, x)$ is the *likelihood* that can be non-linear in x, and we assume independent priors for m and x, $\pi(m)$ and $\pi(x)$ respectively. The goal again is to obtain $\hat{m}$, an estimator for $m^\star$, which is the value of m that explains y the best without incurring in strong assumptions. For that we define 3 different strategies based on 5.5: a pure Markov Chain Monte Carlo (MCMC) approach, a Monte Carlo integral within a MCMC and a point wise reconstruction $\pi(m|y)$ via Monte Carlo integral.

Strategy 1: Straightforward $n_c + 1$ MCMC This strategy relies on the fact that m and the components of $x \in \mathbb{R}^{n_c}$ are assumed a priori independent. Then we set up a MCMC within Gibbs algorithm for equation 5.5 following algorithm 1 and at the end of the sampling process, we extract the marginal samples of m to approximate the density $\pi(m|y)$, where we define $\hat{m} = MAP[\pi(m|y)]$ as the maximum a posteriori (MAP). The major advantage of this strategy is the flexibility to define dependent or independent proposals for each variable when both m and x are required to estimate.

Strategy 2: MC within MCMC In this case, we exploit the independence defined a priori between m and x and provide a more direct estimation of the density $\pi(m|y)$ by marginalizing x

$$\pi(m|y) \propto \int \pi\left(y \middle| \begin{pmatrix} m \\ x \end{pmatrix}\right) \pi(m)\pi(x)dx \tag{5.6}$$

In this approach, we require to generate n_x samples from $\pi(x)$ to set up a 1 dimensional MCMC algorithm which is considerably simpler than the $n_c + 1$ straightforward MCMC,

and we integrate by Monte-carlo. Then we approximate the density of $\pi(m|y)$ and apply $\hat{m}$ as the MAP.

Strategy 3: Point-Wise MC Another way to tackle Equation 5.6, given that $m \in \mathbb{R}$, is to define an interval of estimation for m between a $[m_{\mathrm{L}}, m_{\mathrm{U}}]$ from where for each discrete value within the interval m_i with $i = \{1, \ldots, n_m\}$ we point-wise evaluate the approximated posterior as

$$\pi(m_i|y) \propto \int \pi\left(y \,\middle|\, \begin{pmatrix} m_i \\ x \end{pmatrix}\right) \pi(m)\pi(x)dx \tag{5.7}$$

where for each m_i, the integral is approximated by Monte Carlo with the same number of samples n_x from $\pi(x)$. The main advantage of this method is that we are approximating directly the posterior density $\pi(m|y)$ where $\hat{m}$ is the MAP.

It is expected that while the pure MCMC approach provides more flexibility and is less dependent on strong assumptions on x, it is also, by definition, more costly computationally compared to its classical counterpart. Nevertheless in this particular case, as m is a scalar, strategies two and three are also considerably less costly.

5.1.4 Computer Experiments

In this section we perform a comparison between the classical approach and the MCMC-based strategies previously defined. As we consider that $f(x) = X\beta$ with $X \in \mathbb{R}^{p \times n_c}$ defined as

$$X(x) \triangleq \begin{bmatrix} x_1 & \ldots & x_{n_c} \\ \vdots & \vdots & \vdots \\ x_1 & \ldots & x_{n_c} \end{bmatrix}.$$

we generate $y \in \mathbb{R}^p$ by using

$$\begin{aligned} x^\star \in [0,1]^4 &= [0.2345, 0.7749, 0.9, 0.5643]^\top \\ \boldsymbol{\beta} &= [2.0293e-04, -1.1759e-04,\ 0.14258e-02, -8.890e-04]^\top, \end{aligned}$$

$m = 37.2349$, $\eta \sim \mathcal{N}(0, \gamma^2 I_p)$ with $\gamma^2 = 1e-04$. We analyze the impact of the prior parameters μ_x, σ_x on $m^\star$ for different numbers of available measurements $p = 100, 30, 6$ and under the following configurations

C1. Basic setup: all x_j are sampled with $\mu_x = [0.5, 0.5, 0.5, 0.5]^\top$ and $\sigma_x = [0.2, 0.2, 0.2, 0.2]^\top$ to satisfy $x \in [0,1]^4$.

C2. Small $\sigma_x = 0.05$, close mean: all x_j are sampled with $\mu_x = x^\star - \sigma_x$.

C3. Small $\sigma_x = 0.05$, far mean: all x_j are sampled with $\mu_x = x^\star - 2\sigma_x$.

C4. Big $\sigma_x = 0.1$, far mean: all x_j are sampled with $\mu_x = x^\star - 2\sigma_x$.

C5. Big $\sigma_x = 0.1$, close mean: all x_j are sampled with $\mu_x = x^\star - \sigma_x$.

C6. Small $\sigma_x = 0.05$, very far mean: all x_j are sampled with $\mu_x = x^\star - 3\sigma_x$.

To evaluate both classical and sampling approaches, we define the search interval for the mass $m \in [2, 100]$ with initial value $m_0 = 30$. For the classical approach we will apply equation 5.4, for a number 1000 times with a different value of x sampled from its Gaussian prior. To setup the sampling strategies 1 and 2, we computed the minimum effective sample sizes Gong and Flegal [2016] $mESS \geq 8605$ and $mESS \geq 6146$ at 95% of expected confidence and tolerance levels.

For the MCMC sampler, we used a Metropolis-Hastings Within Gibbs (MHWG) with Gaussian proposal for all the dimensions and adaptive variance for each individual dimension of 30% each 100 iterations for the first half of the chain and 10% for the rest of the sampling process, to keep the acceptance rate between $[0.45, 0.55]$ which is common in literature for faster convergence Ghirmai [2015]. This same configuration was used for the MC within MCMC strategy.

We defined a number of samples for pure MCMC of 20000 and 10000 for the MC within MCMC iterations including the corresponding discarded samples *(burn-in phase)* of 8000 and 4000 respectively. Those quantities were chosen given that the pure MCMC would require a higher number of iterations to converge given that it is on higher dimensions than the MC within MCMC. We also generate a total $n_x = 250$ samples from $\pi(x_j) \sim \mathcal{N}(\mu_{x_j}, \sigma_{x_j})$ to be used by the MC strategies. Finally for the point-wise MC we defined a total of 10000 points between the search interval $[m_\text{L}, m_\text{U}]$.

5.1.4.1 Experiments classical approach

	$p = 100$	$p = 30$	$p = 6$
C1, $\sigma_x = 0.2, \mu_x = 0.5$	88.95 ± 31.37	96.40 ± 66.48	155 ± 2402
C2, $\sigma_x = 0.05 \times 1$	40.44 ± **1.59**	42.29 ± **2.96**	44.59 ± **7.42**
C3, $\sigma_x = 0.05 \times 2$	41.76 ± 1.67	43.67 ± 3.13	46.03 ± 7.86
C4, $\sigma_x = 0.1 \times 2$	38.47 ± 2.87	40.33 ± 5.45	43.21 ± 14.50
C5, $\sigma_x = 0.1 \times 1$	**37.39** ± 2.71	**39.19** ± 5.15	**41.96** ± 13.68
C6, $\sigma_x = 0.05 \times 3$	44.61 ± 1.88	46.64 ± 3.52	49.15 ± 8.86

Table 5.1: Summarized results for the different configurations under 1000 samples from $\pi(x)$ and different measurement points p for a $m^\star = 37.2349$. For configuration one (C1) μ_x, σ_x represent the parameters of the prior. For the rest $Ci, = \sigma_x \times j$ means that configuration $i > 1$ uses a prior whose μ_x is deviated from the true value $x^\star$, j times σ_x.

As we know the deterministic formula from theorem 5.2 yields a single value for m, as we define different configurations for the prior of x we generated 1000 samples, estimated the empirical mean $\bar{m}$ and its confidence intervals using the notation $\hat{m} = \bar{m} \pm 3\sigma_m$ for each configuration. The results are summarized in Table 5.1, where we can observe three phenomena. The first is that higher values of p helps to reduce the uncertainty on m (σ_m). The second one is that the estimation of m is more sensitive to slight variations on σ_x than in μ_x. From the same table, by analyzing configurations 2 or 3 that share the similar proposed value of $\sigma_x = 0.05$, versus the configurations 4 or 5 respectively, where the $\sigma_x = 0.1$ is doubled, duplicates the uncertainty value for m but also increases the accuracy on $\hat{m}$.

The final phenomenon is related to the sensitivity of the estimation of m to the definition of $f(x)$, which in this case $f(x) = X(x)\beta$, benefits more from a $\pi(x)$ that provides diverse samples with a $\sigma_x = 0.1$, rather than $\sigma_x = 0.05$. This specific behavior points to the definition of the matrix X, as we generate x that are more similar, with the current number of parameters $\boldsymbol{\beta}$, it will become more difficult to reproduce variability among the different measurements p.

5.1.4.2 Experiments MCMC-based Bayesian approach

For this group of strategies, we consider important to analyze the most basic configuration possible (C0), this is, an uniform prior for both m and x which was not possible to include under the Gaussian prior assumption for the classical approach while defining Theorem 5.2.
The approximated distributions for the direct $n_c + 1$ MCMC strategy are presented in Figure 5.1. Here we observe how this strategy benefits more from a Gaussian rather than an Uniform uninformative prior while having its MAP estimate closer to the $m^\star$. This relative improvement, which requires no extra information on x is also reflected with their corresponding trace and autocorrelation plots from Figures 5.2 and 5.3 respectively. Specially seems beneficial for the case $p = 6$, where the Gaussian prior chain proved to be long enough in terms of its trace.

To analyze the performance for configurations from C2 to C6 we observe a relatively similar behavior in terms of density, trace and autocorrelation compared to C0 and C1. Here, we can also make use of the 90% confidence intervals presented as its lower value (CI-lower) and its length (CI-length) in Table 5.2. We observe that for identical $\sigma_x = 0.01$ configurations (C2, C3, C6) as μ_x deviates more from $x^\star$ it slightly displaces the CI-lower and also increases the CI-length which does not seem to impact the approximated distribution obtained. However, by comparing C2 and C4 (or C3 and C5) while σ_x doubles, the CI-length mimic this increment which is more impactful in terms of retrieving a value closer to the $m^\star$; finally by mixing this two variations we observe how the MAP estimate differs more from $x^\star$ while having the same p.

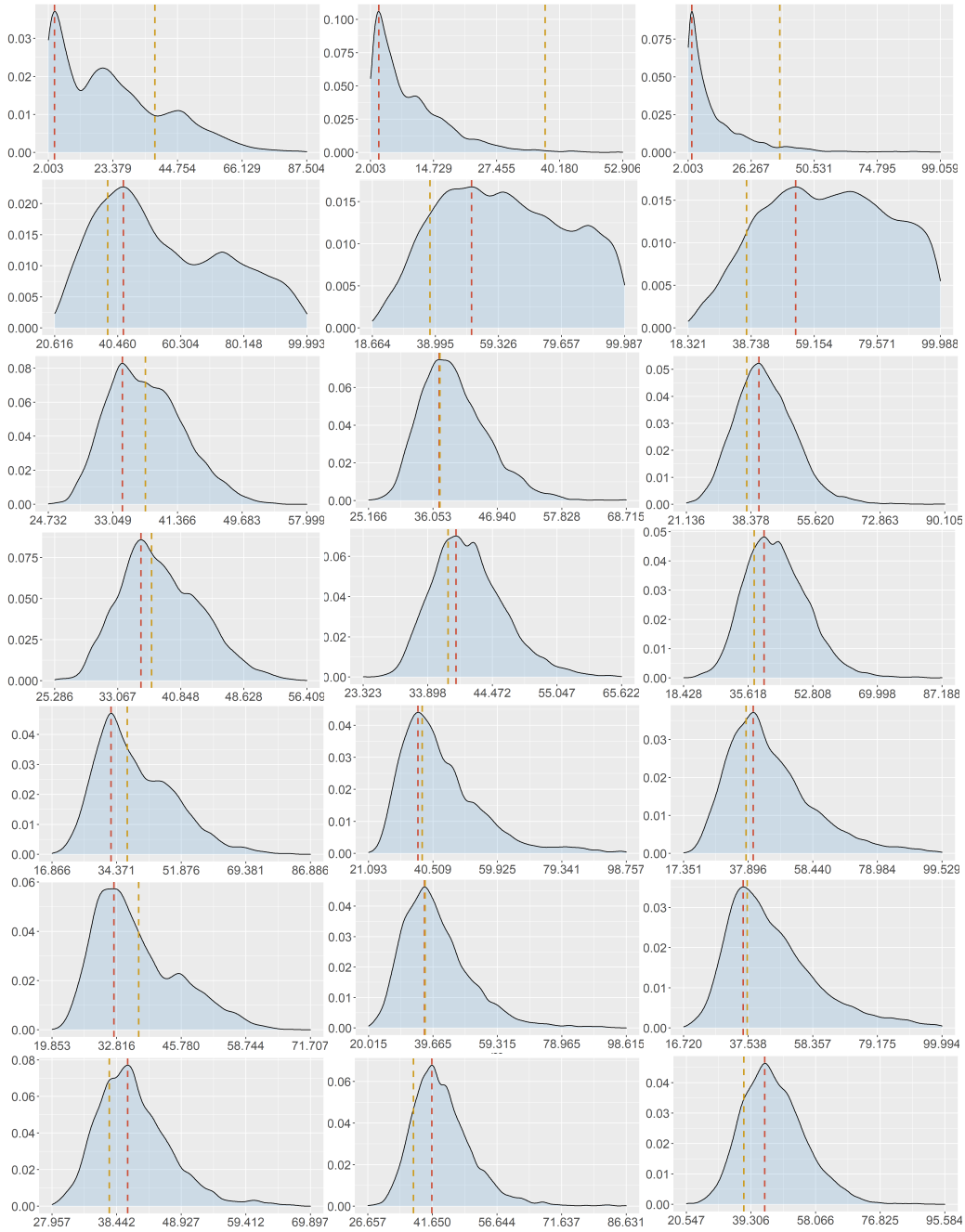


Figure 5.1: Results for the strategy 1 $n_c + 1$ MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

It is important to remark that for $p = 6$ the chains from C1 to C6 managed to converge

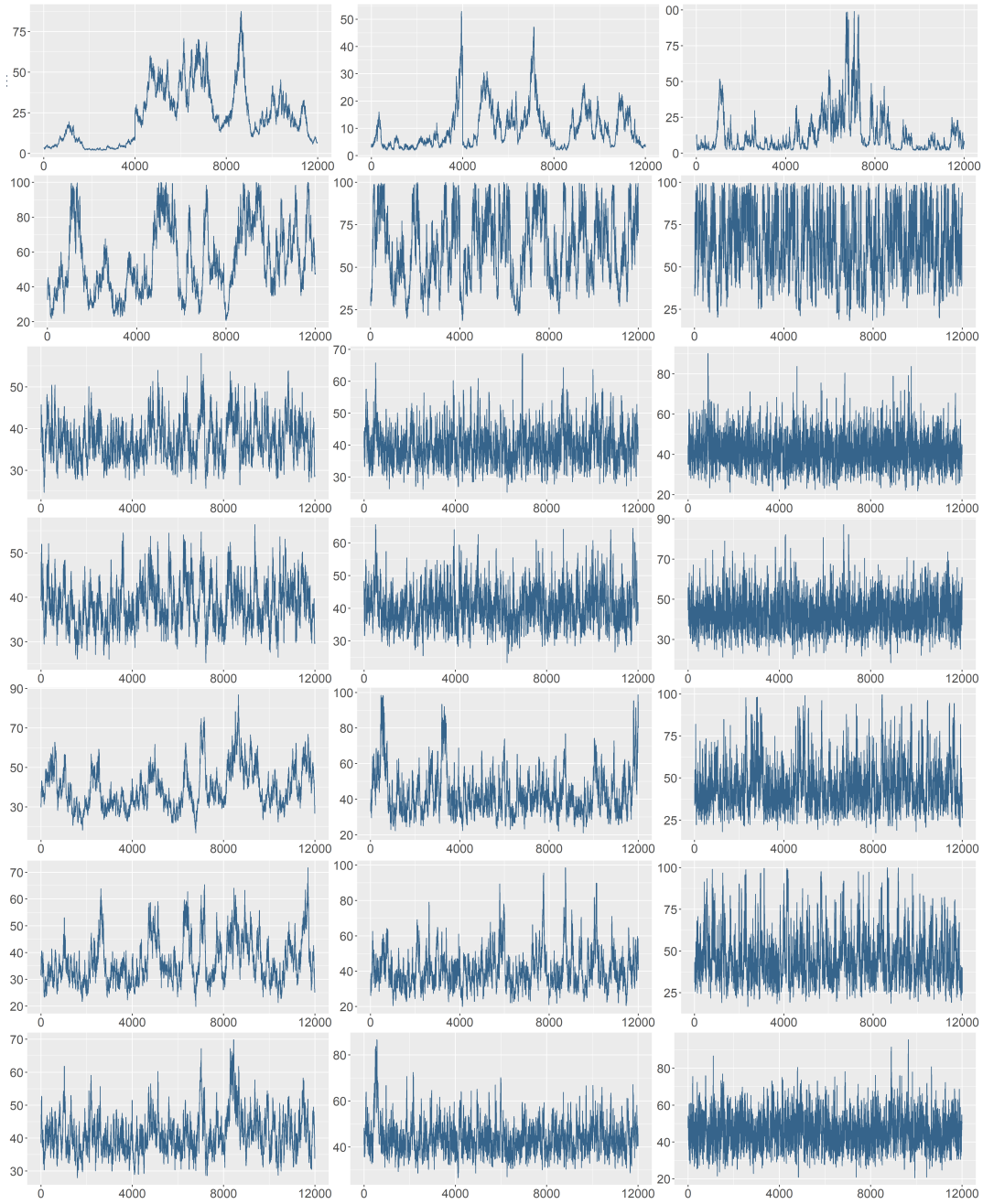


Figure 5.2: Trace plots for the strategy 1 $n_c + 1$ MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right).

even with a relatively poor prior as is C6. Also the fact that with $p = 100$ measurements this strategy found more difficult to find a chain without autocorrelated samples and with a proper trace plot with the provided priors.

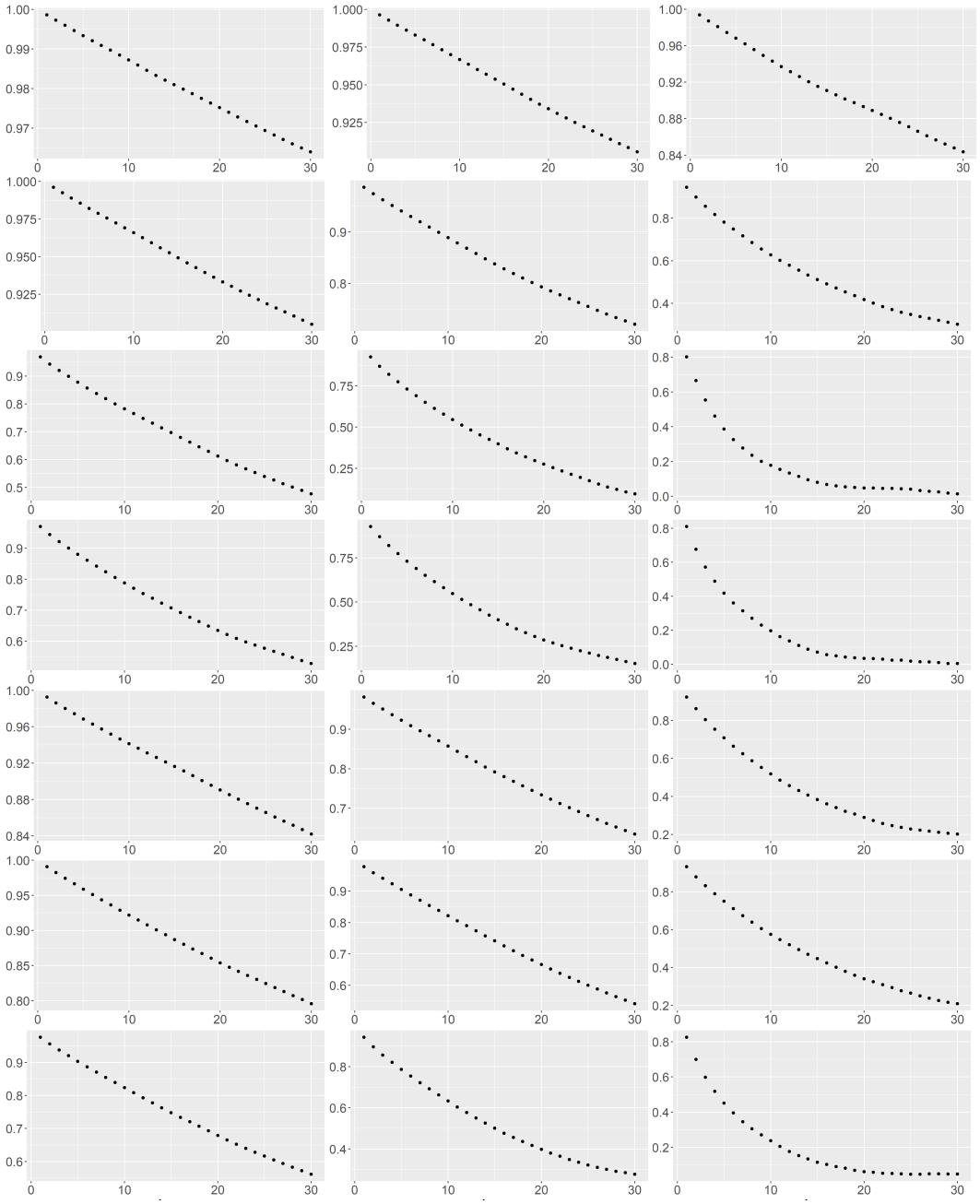


Figure 5.3: Autocorrelation plots for the strategy 1 $n_c + 1$ MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right).

In the case of strategy 2 MC within MCMC, differently from strategy 1, it benefits from a small chain requirement and managed to converge even under the C0 configuration as shown in Figures 5.4, 5.5 and 5.6. Also it presents the same behavior while comparing

	$p = 100$		$p = 30$		$p = 6$	
	CI-Lower	CI-Length	CI-Lower	CI-Length	CI-Lower	CI-Length
C0, Uniform	2.55	54.14	2.54	22.44	2.74	86.07
C1, $\sigma_x = 0.2, \mu_x = 0.5$	27.68	63.81	30.69	64.65	31.54	61.03
C2, $\sigma_x = 0.05 \times 1$	30.29	15.74	31.68	18.64	30.04	25.92
C3, $\sigma_x = 0.05 \times 2$	30.98	16.30	31.93	19.45	30.96	26.71
C4, $\sigma_x = 0.1 \times 2$	25.70	33.53	28.48	38.99	27.68	44.27
C5, $\sigma_x = 0.1 \times 1$	26.13	27.13	27.40	33.12	27.06	48.67
C6, $\sigma_x = 0.05 \times 3$	33.13	19.19	34.57	22.22	32.69	30.04

Table 5.2: Lower bound and length of the 90% confidence interval for the strategy $n_c + 1$ MCMC and different configurations. After configuration 1, configuration $Ci, = \sigma_x \times j$ means that configuration $i > 1$ uses a prior whose $\mu_x = x^\star - j\sigma_x$.

	$p = 100$		$p = 30$		$p = 6$	
	CI-Lower	CI-Length	CI-Lower	CI-Length	CI-Lower	CI-Length
C0, Uniform	23.21	70.64	24.95	67.95	26.13	69.16
C1, $\sigma_x = 0.2, \mu_x = 0.5$	31.00	64.88	32.19	62.46	29.71	66.84
C2, $\sigma_x = 0.05 \times 1$	31.15	15.15	31.55	18.23	29.57	23.94
C3, $\sigma_x = 0.05 \times 2$	31.93	17.17	32.22	19.10	30.99	26.60
C4, $\sigma_x = 0.1 \times 2$	27.32	36.90	28.47	38.14	28.23	50.16
C5, $\sigma_x = 0.1 \times 1$	26.78	35.85	27.77	35.68	27.93	49.83
C6, $\sigma_x = 0.05 \times 3$	33.71	18.99	34.21	21.10	33.31	28.88

Table 5.3: Lower bound and length of the 90% confidence interval for the strategy MC within 1 dimensional MCMC and different configurations. After configuration 1, configuration $Ci, = \sigma_x \times j$ means that configuration $i > 1$ uses a prior whose $\mu_x = x^\star - j\sigma_x$.

increments on μ_x and σ_x that strategy 1 as shown in Table 5.3. Additionally, it seems relatively more robust that strategy 1 to variations on the prior and its corresponding parameters which is always important when applying MCMC-based strategies.

In the case of strategy 3 point-wise MC as is derived from strategy 2, should present similar results as represented in Table 5.4. However as shown in Figure 5.7 we observe "wiggly" effects under configurations C0 and C1 which suggests that the prior of x is not informative enough for the MC integral. We also observe smoother versions of the approximated density for configurations C2 to C6 compared to the other strategies for $p = 30, 6$. In general, all MCMC-based approaches present desirable behaviors while approximating the m density even with non-Gaussian priors, being the MC within MCMC the most informative and robust of the 3.

As a final comparison between MCMC-based strategies and the deterministic approach for this specific scenario, we observe how both of them are viable strategies depending on the amount of information and precision required on the estimator of m versus the

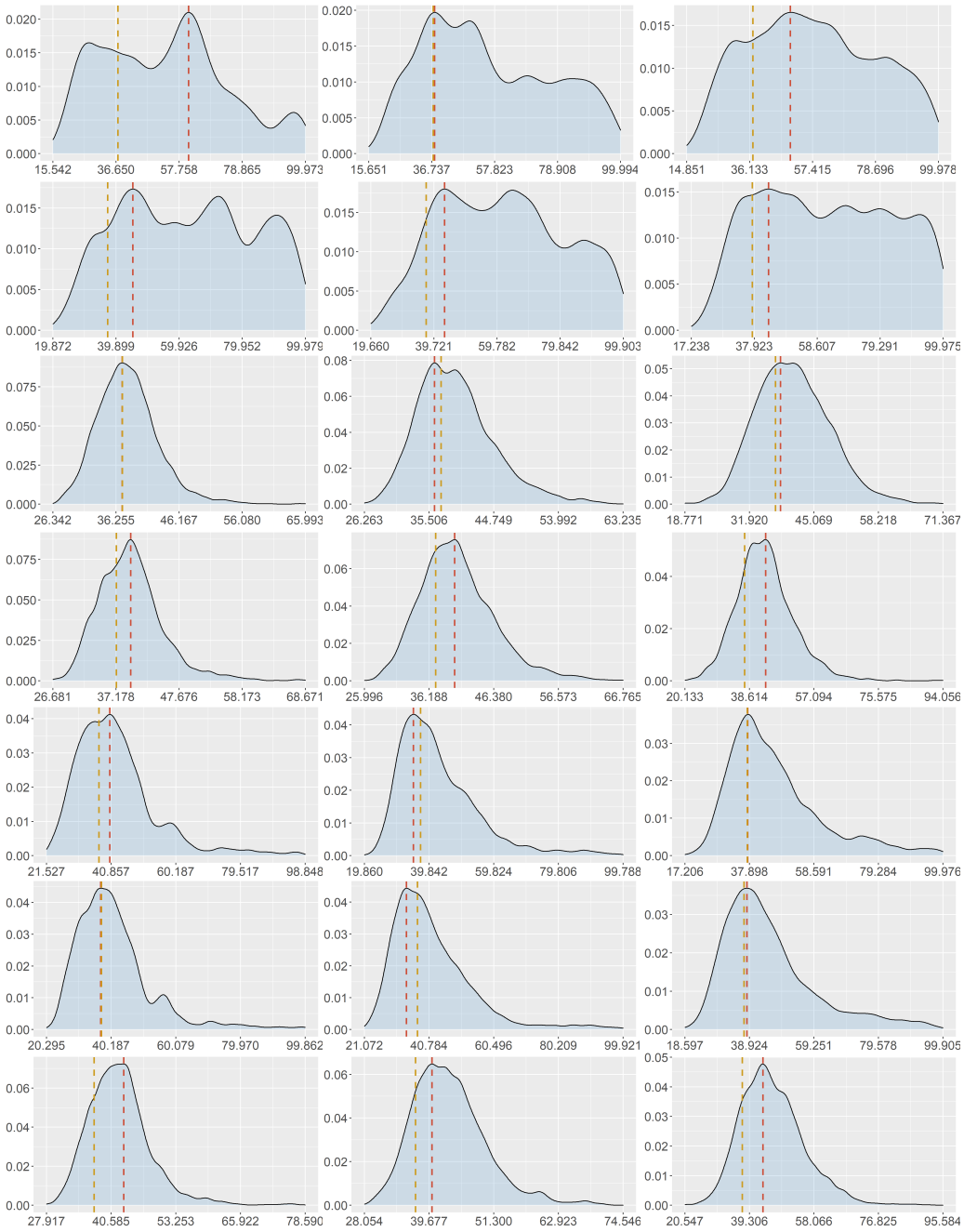


Figure 5.4: Results for the strategy 2 MC within MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

knowledge on the prior parameters.

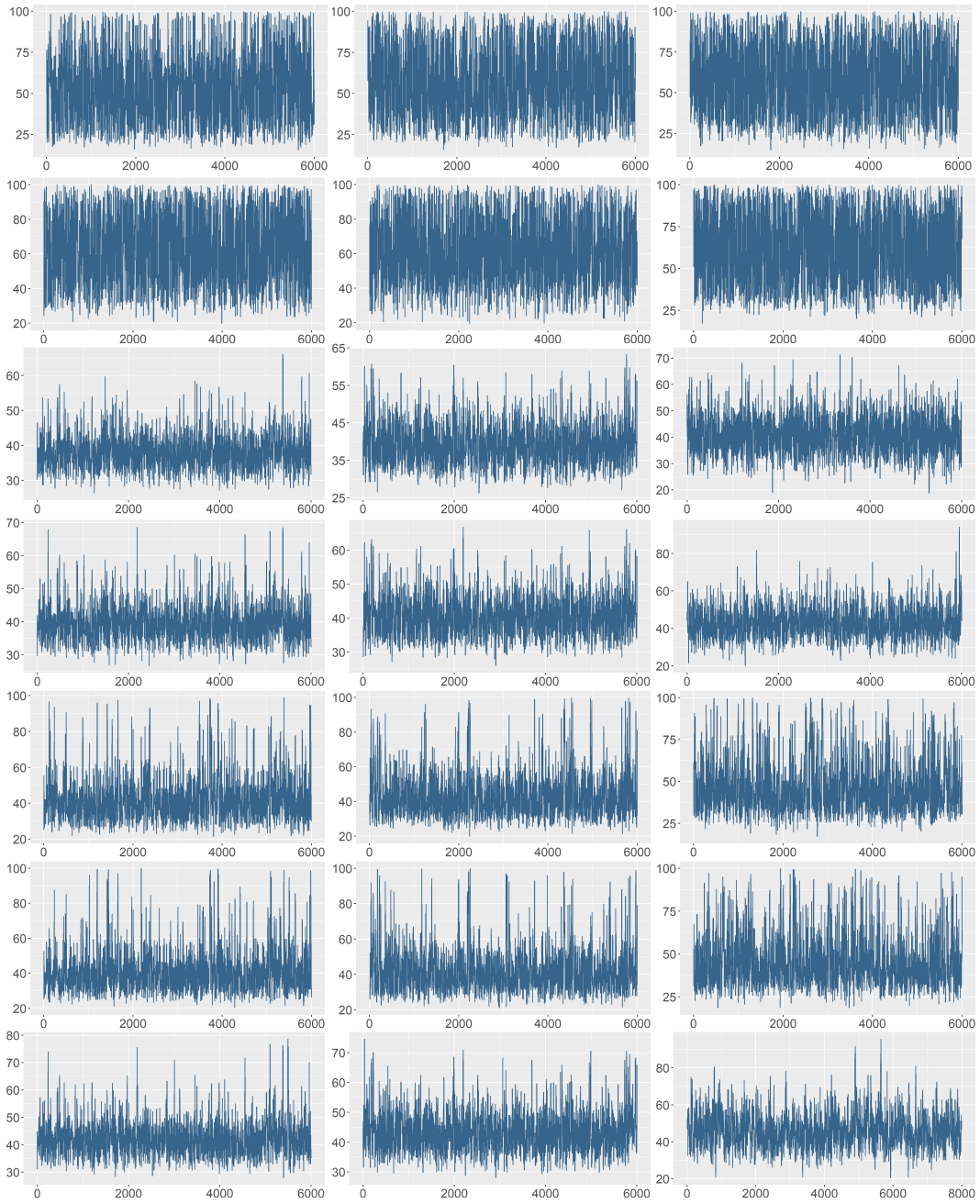


Figure 5.5: Trace plots for the strategy 2 MC within MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right).

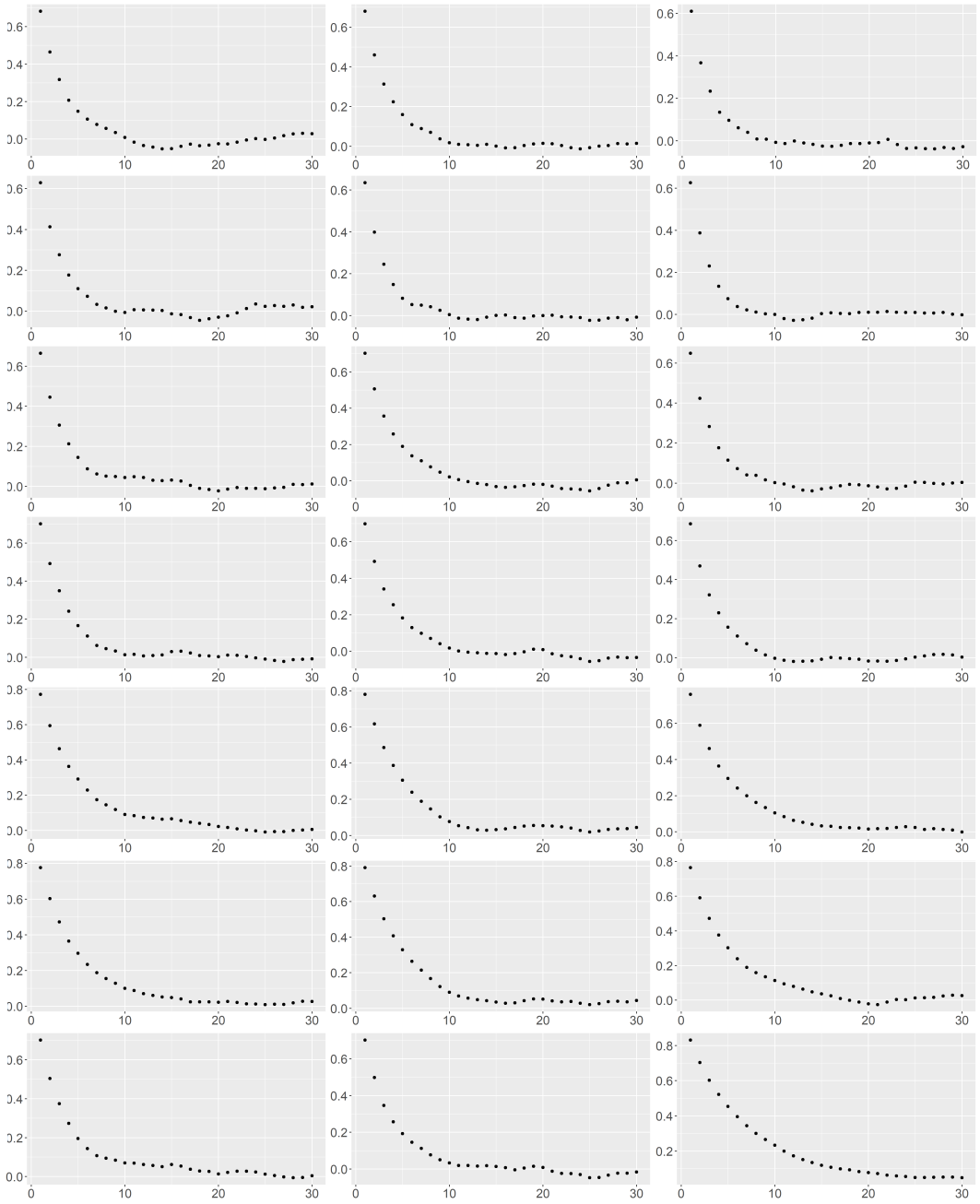


Figure 5.6: Autocorrelation plots for the strategy 2 MC within MCMC, for configurations C0 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right).

5.2 Scenario 2: $M(x)$, a non-linear black-box function

As we presented at the beginning of this section, we are inspired by an application for radionuclide quantification, where the available data is scarse, $p = 6$ and $f(x)$ is a

	$p = 100$		$p = 30$		$p = 6$	
	CI-Lower	CI-Length	CI-Lower	CI-Length	CI-Lower	CI-Length
C0, Uniform	23.34	70.61	24.77	68.45	25.15	70.57
C1, $\sigma_x = 0.2, \mu_x = 0.5$	30.89	65.35	31.79	62.83	32.75	62.92
C2, $\sigma_x = 0.05 \times 1$	31.05	16.52	31.40	17.40	30.01	24.70
C3, $\sigma_x = 0.05 \times 2$	31.87	17.60	32.22	18.40	31.05	28.28
C4, $\sigma_x = 0.1 \times 2$	27.50	37.34	27.97	36.14	26.60	45.16
C5, $\sigma_x = 0.1 \times 1$	26.93	36.36	27.38	34.85	27.66	50.73
C6, $\sigma_x = 0.05 \times 3$	33.62	20.09	34.03	20.68	32.93	28.87

Table 5.4: Lower bound and length of the 90% confidence interval for the strategy point-wise MC and different configurations. After configuration 1, configuration $Ci, = \sigma_x \times j$ means that configuration $i > 1$ uses a prior whose $\mu_x = x^\star - j\sigma_x$.

black-box function Clement et al. [2018]. Thus, for this scenario we treat $f(x)$ as a non-linear black-box function and $p = 6$ measurements are available for y.

To construct this function, we generated 1000 random values in the interval $[5.42e-09, 3.41e-05]$ as a valid range of values for a attenuation coefficient $f(x)$ Máduar and Miranda Junior [2007]. Then we trained a continuous Gaussian process $Y(x)$ conditioned on inputs $x \in \mathbb{R}^{n_c}$ in the interval [0,1], and a kernel of the class matérn Rasmussen and Williams [2006]. We selected $m = 10$ and

$$x^\star = [0.3511; 0.5391; 0.4306; 0.5193]$$
$$\eta = [3.4586; 2.3531; 3.2773; 3.9404; 4.0978; 4.3263] \times 10^{-07},$$

resulting in an output $y = [4.1079; 1.4581; 5.5302; 16.332; 19.354; 20.748] \times 10^{-06}$. The goal in this section is to analyze the applicability of the previously defined MCMC based approaches and its configurations (C0 to C6) under this more realistic scenario.

As an important note, in this section we discarded the possibility to apply a deterministic formula from Theorem 5.2, which required a parametric linear regression to estimate $\boldsymbol{\beta}$. This, due to the nature of f being a Gaussian process and the fact that a more suitable strategy can be a non-parametric approach like the kernelized ridge regression model for f Murphy [2012], which would lead to obtain a new theorem, and to create a dedicated section which is out of the scope of this document.

5.2.1 Experiments MCMC-based Bayesian approach

In this section, we are going to analyze the configurations C0, C1, C2 and a new configuration CT (true configuration) from where $\mu_x = x^\star$ and $\sigma_x = 0.0005$. This new configuration CT will be key to compare the performance of the different strategies towards a goal. As a reminder, C0 corresponds to an uniform prior for x, C1 a

non-informative Gaussian prior for x with $\mu_x = 0.5$ and $\sigma_x = 0.2$, and C2 is a more informative Gaussian prior with $\mu_x = x^\star - \sigma_x$, with $\sigma_x = 0.05$, as defined for the linear scenario in section 5.1.4. For this setup, for strategies 1 and 2 the starting point was set to $m_0 = 6$ in the same interval $[0, 100]$, the length of the chain, the number of burn-in discarded samples as well as, for strategy 3 the number of iterations and samples from $\pi(x)$ remain unchanged.

While trying to analyze C0, both strategies 1 and 2 struggle to obtain a likely value to start the chain, not even when a 0.5 value is relatively close to the true x. We observe this main difficulty in Figure 5.8 where a point-wise representation, for this case, leads to a "peak" distribution with very small "bumps" separated on the interval. This means that the true probability content lies on a small peak and that the prior information should be accurate enough not to displace the probability content further from the goal $m^\star$.

Now in Figure 5.9 we compare the performance for the configurations C1, C2 versus CT. We observe that the 3 strategies behave differently. $n_c + 1$ MCMC seems to approximate a distribution that seems correct according to the plot. However the information provided by Figure 5.10, shows that the chain is still far from convergence and that the samples are still highly autocorrelated. Thus, the apparent convergence for a uniform prior and the worse performance when introducing a more informative Gaussian prior, corresponds to a noisy chain based analysis.

For the second strategy, the behavior is different, even if Figure 5.11 suggests its convergence, the obtained density converges to a region where the $m^\star$ is not likely. A similar behavior is observed for the point-wise strategy, as it detects a principal peak that is far from the region where $m^\star$ is located. We believe that this behavior is due to the multi-modality induced by a non-linear $f(x)$ and the ill-posed product $mf(x)$.

After comparing with the configuration CT, we can deduce that the prior needs to be very precise, as the probability as the target distribution is very thin and presents multiple modes that are located far in the search space. This special multi-modality makes worse for the uninformative priors because it can lead to a lot of unlikely samples and most of the time to focus in a specific mode. It is also important to remark that this analysis is done by using the real η which is another variant to add when evaluating the likelihood of any of the MCMC-based strategies.

5.3 Conclusions

In this chapter, we have investigated and analyzed two different scenarios that helped us to understand the bi-linear inverse problem $y = mf(x) + \eta$ towards an application for radionuclide quantification. On the first scenario we provided a simple example where we managed to define a classical least-squares approach and 3 different MCMC-based

strategies. We were able to analyze the impact of different choices of prior on the estimation of the m value. As the deterministic approach provided a small confidence interval length, it was less robust for estimating m than its MCMC-based counterpart, the MC within MCMC strategy among the tested configurations. This strategy managed to avoid the necessity of a $n_c + 1$ chain while maintaining the convergence metrics. We also remark that any of the strategies can be used under particular scenarios depending on the prior information available, being a Gaussian type of prior more suitable for the techniques.

In the second scenario we created a continuous setting for a radionuclide application to analyze the performance of the MCMC-based strategies for different prior information available. Here we valued the importance of the trace and autocorrelation plots when analyzing the convergence of a chain. Furthermore we emphasize how this particular scenario requires a very precise and informative prior to be able to recover a distribution where $m^\star$ is likely while η is known. We believe that this is related to the non-linear properties of $f(x)$.

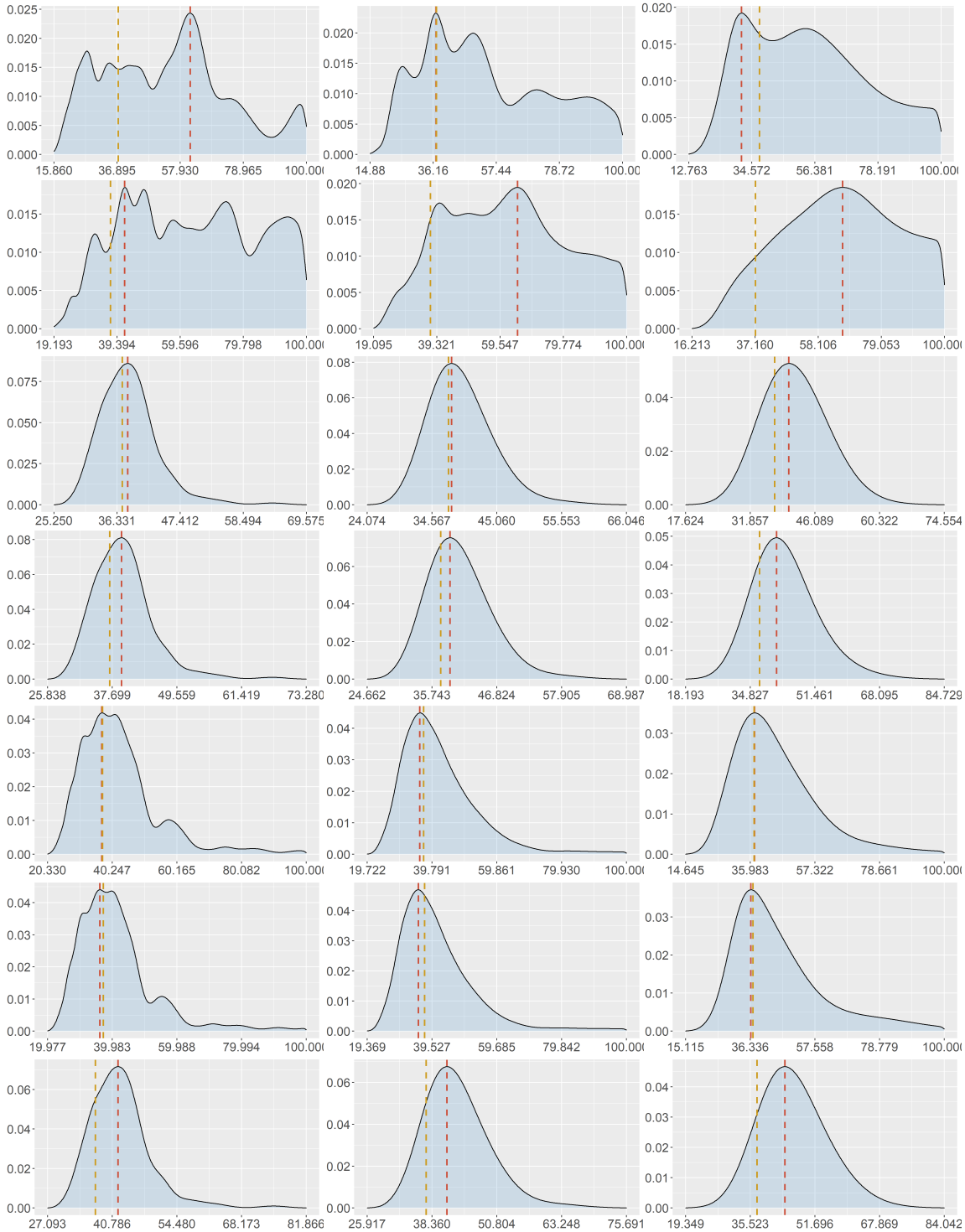


Figure 5.7: Results for the strategy 3 point-wise MC, for configurations C1 (top) to C6 (bottom)for different numbers of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$.

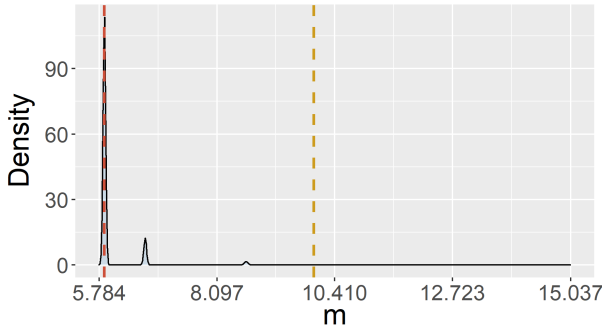


Figure 5.8: Obtained densities for the point-wise MC strategy. Red dashed line represents the MAP estimate. Golden dashed line represents $m^\star$

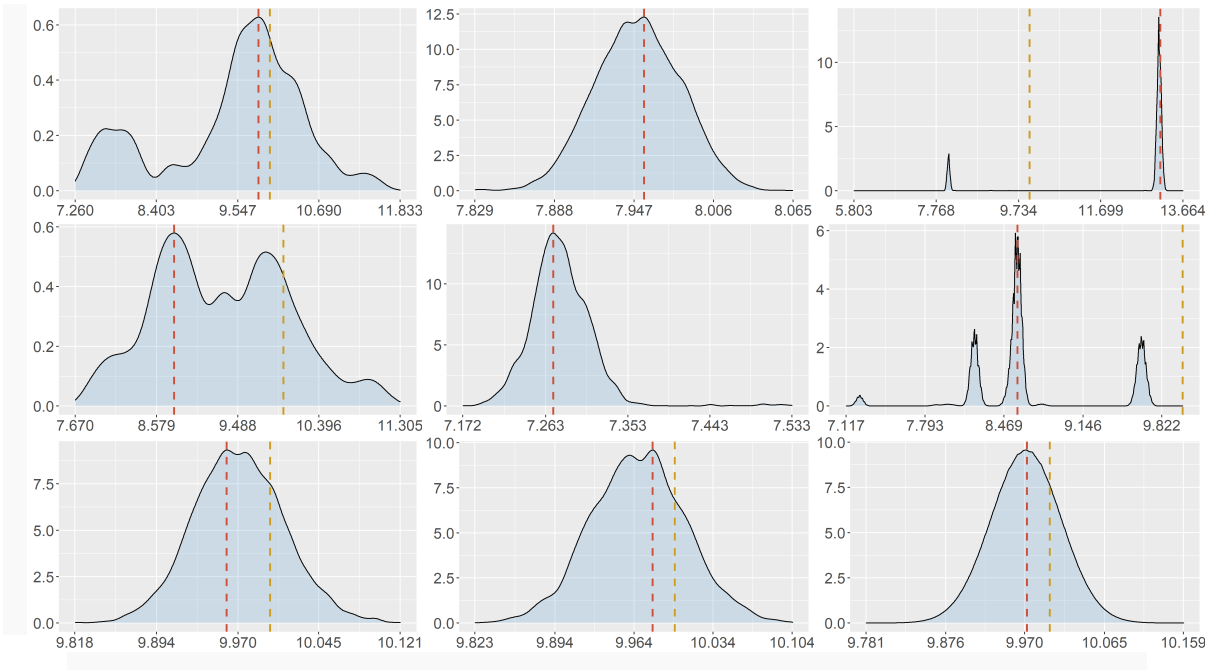


Figure 5.9: Obtained densities for the strategies MCMC (left), MC within MCMC (middle) and point-wise MC (right) for configurations C1 (top), C2 (middle) and CT (bottom). Red dashed line represents the MAP estimate. Golden dashed line represents $m^\star$

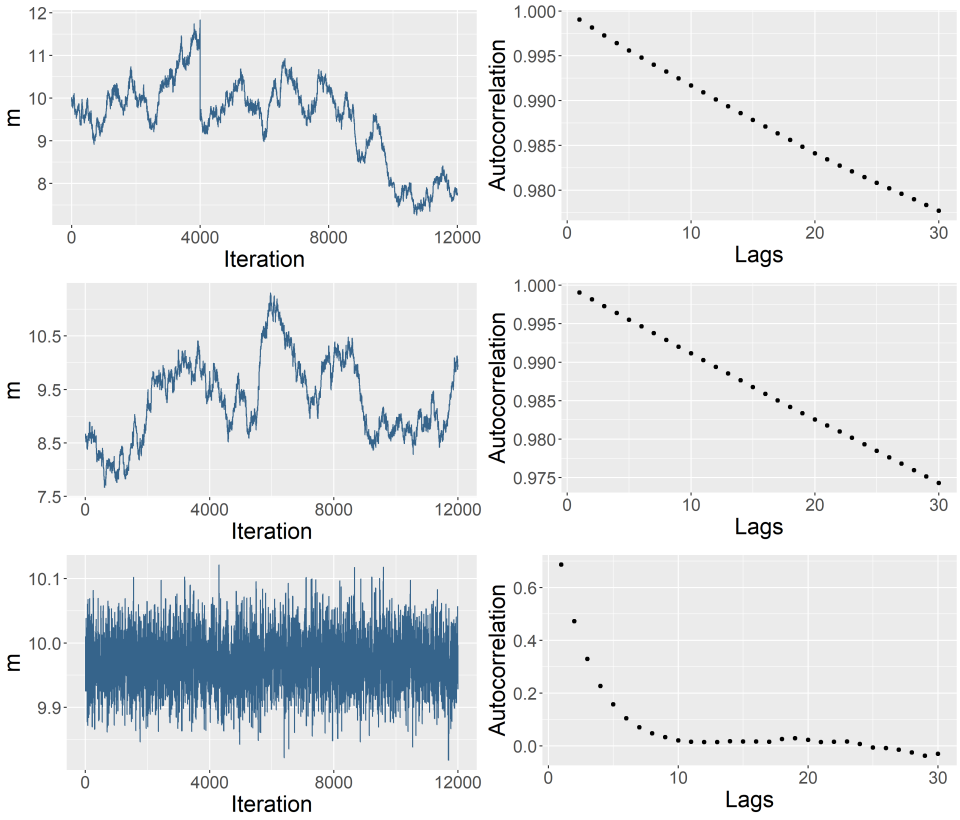


Figure 5.10: Obtained trace (left) and autocorrelation (right) plots for the MCMC strategy.

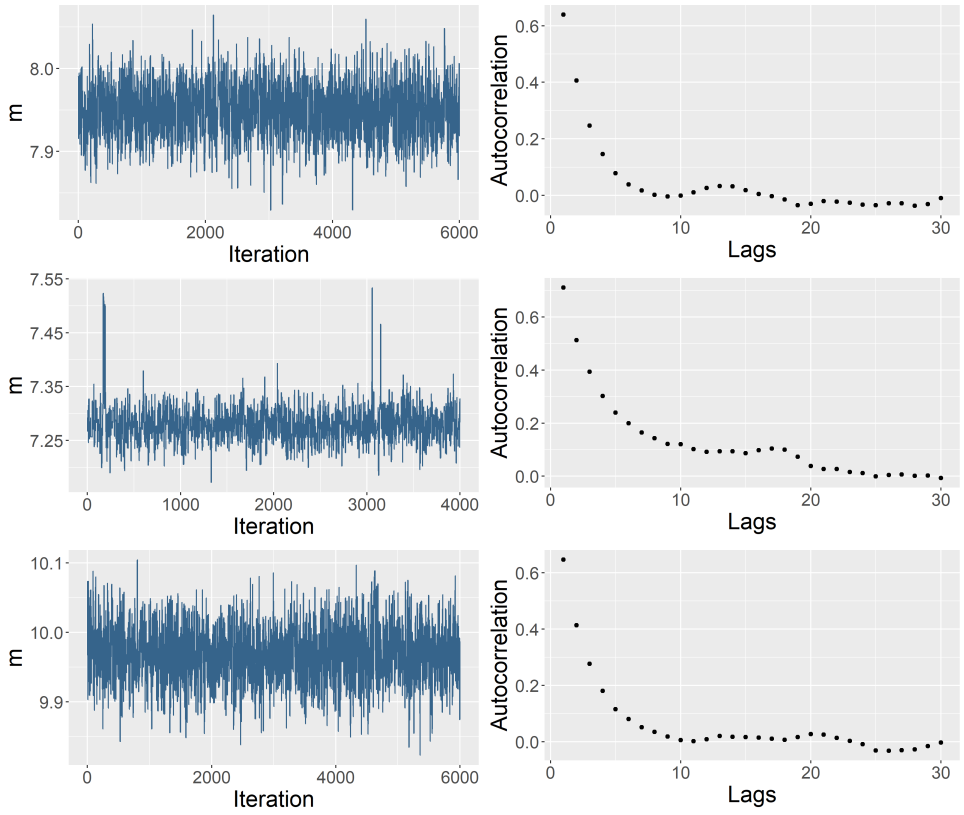


Figure 5.11: Obtained trace (left) and autocorrelation (right) plots for the MC within MCMC strategy.

Chapter 6

Conclusions and perspective

Summary of Contributions

In this document we investigated Bayesian approaches to solve mixed problems. We proposed the LV-EGO algorithm as a novel methodology that relaxes a problem from the mixed space to a continuous one. This was done by employing a latent variable mapping and preserving the link between the mapping and the categorical variables by inducing a constraint during the relaxed optimization. We also proposed different variations based on augmented Lagrangians (the ALV-EGO variants) to handle this constraint properly. Then we compared LV-EGO and its variants to algorithms working directly in the mixed space based on random forests, evolutionary strategies and mixed kernels.

Among the different algorithms, we found that the augmented Lagrangian accounting of the constraint as an inequality plus a tolerance value was better than the original equality form suggested by the constraint. However, during this relaxed optimization, it was not clear that the augmented Lagrangians increased the performance of the search versus the vanilla LV-EGO. We also introduced the concept of exponential penalization for LV-EGO regularization, as a strong penalty trying to force convergence on a feasible domain. This modification proved to further increase the performance of the LV-EGO among all the different test cases.

As an additional study, we proposed different scenarios to solve the specific inverse problem in a continuous setting $y = mf + \eta$, where $m \in \mathbb{R}$ is the variable of interest, η is the observation noise. We defined two different scenarios, one with f being a deterministic function and the second with f being a non-linear black box function.

For the first scenario, we considered both least-squares and three different MCMC based strategies in order to understand the impact of the uncertainty of the prior on the estimation of m. This helped us to understand the difficulties to apply classical approaches, and the advantages of using Monte-Carlo integral to simplify a higher dimensional MCMC. Then the second scenario showed us the difficulties of convergence for the different MCMC chains when there are few observations for different types of priors. Finally, we established a framework in continuous inputs, as a step of applying mixed methodologies

that translates the problem to a continuous inputs setting.

Perspectives

After studying Bayesian optimization and inversion in the context of mixed variables, we suggest possible lines of research that will complement the results presented in this document. For the *optimization* study, we propose three ways. Providing the *analytical gradients* for the acquisition function on the relaxed problem should increase the performance of the LV-EGO algorithms, as gradients are key elements during continuous optimization. Extending the *convergence results* of EGO to the scenario of mixed optimization through latent variables, which will increase the credibility of those methodologies. Developing a new parameterization for the discrete levels that can extend the application of LV-EGO to more than 10 levels.

For the *inversion* study, we also propose three different ways. Defining a more *general linear model* (e.g. kernel ridge regression) to derive analytical expressions that can be applied for different types of regularization and black-box functions.Further exploration of the impact of the prior parameters with other functions that exhibit different types of non-linearity seems to be a useful continuation. We also believe that hybrid strategies mixing deterministic and Bayesian approaches should be investigated as a path to cumulate the advantages of both methods.

Finally, as the contents in this document were motivated by different CEA applications, the following paragraphs present some of the developments related to tackling those specific problems.

Application to a mixed filter design

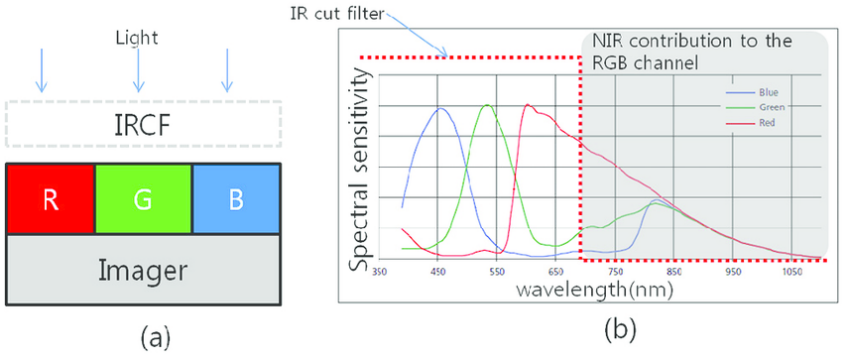


Figure 6.1: (a) Conventional camera system based on a color filter array (Bayer mosaic), an image sensor with the IR cut-off filter (IRCF). (b) Spectral sensitivity of the camera system. Source: Park and Kang [2016].

The near-infrared (NIR) is one of the regions closest in wavelength to the radiation detectable by the human eye. Therefore even though human eyes cannot detect NIR, silicon based filters can and are highly sensitive up to a wavelength (λ) of 1100nm as in Park and Kang [2016] and presented in Figure 6.1. Engineers at CEA-LETI are developing strategies to increase the response of an infrared cut filter (IRCF) in the NIR region. This can be done by designing an interference filter based on a stack of K multiple optical cavities of thickness e_k and optical index m_k with $k = 1, \dots, K$. Then the obtained Transmission spectrum $T(\lambda)$ should mimic the behavior presented in Figure 6.1. This design will require $T(\lambda)$ to satisfy

$$\begin{aligned} \alpha - T(\lambda_1) &\leq 0 \qquad && \forall \lambda_1 \in \mathcal{S}_1 \\ T(\lambda_2) - \beta &\leq 0 && \forall \lambda_2 \in \mathcal{S}_2, \end{aligned}$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ correspond to the visible and near infra-red regions respectively, $\alpha \in [0, 1]$ is the desired proportion of transmission and $\beta \in [0, 1]$ is the desired proportion of reflection. The values of $T(\lambda)$ will depend on the number of layers K, the thickness of each layer $\mathbf{e} = (e_1, \dots, e_K)$, and on the effective index N_k of the chosen material $\mathbf{m} = (m_1, \dots, m_K)$ corresponding to each layer. This yields, once again a mixed variables problem. Now for a fixed number of layers K we can create a surrogate of $T(\lambda)$, and optimize it in the LV-EGO fashion, through the Expected Improvement (EI) as:

$$\begin{aligned} &\max_{(\mathbf{e},\mathbf{m})\in\mathcal{S}} \; EI \\ \text{s.t.} \quad & g = \sum_{k=1}^{K} e_k \leq \; \tau, \end{aligned} \tag{6.1}$$

where g is a design constraint for the global thickness.

Application to radionuclide quantification

The inverse problem $y = mf(x, u) + \eta$ defined in chapter 5 can be reviewed in a new scenario where $f(x, u)$ is a mixed and expensive simulator. Under a Bayesian setup, we can directly rewrite the original inverse problem as a maximization problem of the mode of the posterior distribution $\pi(m, x, u|y)$, where the likelihood $\pi(y|m, x, u)$ will depend on the mixed surrogate definition (for example a mixed kernel surrogate). Then we can apply any of the LV-EGO variants from section 4.1. In this formulation, as in the continuous one proposed in Clement et al. [2018] we may require a mixed metamodel to account for the model error.

Appendices

Chapter A

Mixed Optimization

A.1 Complements on the augmented Lagrangians

Case of an equality constraint

Let us first consider an optimization problem with an equality constraint,

$$\begin{cases} \min_{x\in\mathcal{X}} f(x) \\ \text{such that } h(x) = 0 \end{cases} \tag{A.1}$$

At this point, $f()$ and $h()$ are very general functions on a d-dimensional general set $\mathcal{X}$. We only require that $\mathcal{X}$ is not empty, that $f()$ and $h()$ are bounded, and that there is at least one solution to (A.1), $x^\star \in \mathcal{X}$, which can be attained. $f()$ and $h()$ are not necessarily continuous, a fortiori not necessarily differentiable. With respect to the main body of the article, the notations are simplified in this Section: $\mathcal{X}$ stands for the cartesian product of $\mathcal{X}$ and $\mathcal{L}$, $f(x)$ generalizes $-\log(1+\mathrm{EI}^{(t)}(x,\ell))$ and $h(x)$ corresponds to $g^{(t)}(\ell)$ when $\epsilon = 0$. Note that $g^{(t)}()$, being made of the minimum distance to a discrete set of points (cf. Eq. (4.7)), is not differentiable. $g^{(t)}()$ is the only constraint in the article. This appendix considers one constraint too, but all the results given readily generalize to many constraints by replacing the products by vector scalar products.
Problem (A.1) can be equivalently reformulated as

$$\begin{cases} \min_{x\in\mathcal{X}} f(x) + \frac{1}{2}\rho h^2(x) \\ \text{such that } h(x) = 0 \end{cases} \tag{A.2}$$

where $\rho \geq 0$ is a penalty parameter. The two above formulations have the same solution $x^\star$ and the same value of optimal objective function since $x^\star$ is feasible, $h(x^\star) = 0$, therefore $f(x^\star) = f(x^\star) + \frac{1}{2}\rho h^2(x^\star)$. However, as proved in Minoux [1986] and sketched in Figure A.1, there is always a positive lower bound on the penalty parameters, $\rho \geq \rho^\star \geq 0$,

such that Problem (A.2) can be equivalently solved through the dual formulation,

$$\begin{aligned} &\max_{\lambda\in\mathbb{R}} D(\lambda,\rho) \\ &\text{where } D(\lambda,\rho) = \min_{x\in\mathcal{X}} L_A(x;\lambda,\rho) \\ &\text{and } L_A(x;\lambda,\rho) = f(x) + \lambda h(x) + \frac{1}{2}\rho h^2(x) \end{aligned} \tag{A.3}$$

In this way, the augmented Lagrangian of Hestenes [1969] is the classical Lagrangian of the penalized problem (A.2). We write $\lambda^\star, \rho^\star$ a solution to (A.3). $D(\lambda,\rho)$ is the lower front of all augmented Lagrangians for varying x at a given λ, ρ. The "global dual" update of (λ,ρ) comes from the resolution of (A.3) where the set $\mathcal{X}$ is approximated by the finite subset of samples $\mathbf{X}$.
Let us denote

$$x(\lambda,\rho) = \arg\min_{x\in\mathcal{X}} L_A(x;\lambda,\rho) \tag{A.4}$$

a solution at given multiplier and penalty parameter. The function $D(\lambda,\rho)$ is concave in λ and ρ and $h(x(\lambda,\rho))$ is a subgradient with respect to λ Minoux [1986]. This is at the root of updating strategies that we called "local dual" earlier and which consist in a gradient step in the dual space,

$$\lambda_{t+1} = \lambda_t + \alpha\partial_\lambda D(\lambda_t,\rho_t) = \lambda_t + \alpha h(x(\lambda_t,\rho_t)) \quad , \tag{A.5}$$

where $\alpha > 0$ is a step size factor.

More specific update strategies such as those given in Nocedal and Wright [2006], Picheny et al. [2016] stem from the Karush Kuhn and Tucker (KKT) optimality conditions and require the additional assumption that $\mathcal{X} \in \mathbb{R}^d$ and $f()$ and $h()$ are differentiable. At $x^\star$, since $h(x^\star) = 0$ and λ^{KKT} being the KKT multiplier[1], one has

$$\begin{aligned} \nabla f(x^\star) + \rho h(x^\star)\nabla h(x^\star) + \lambda^{KKT}\nabla h(x^\star) &= 0 \\ \Rightarrow \qquad \nabla f(x^\star) + \lambda^{KKT}\nabla h(x^\star) &= 0 \end{aligned} \tag{A.6}$$

At iteration t, the necessary conditions for $x^t = x(\lambda_t,\rho_t)$ to be the minimum of $L_A(;\lambda_t,\rho_t)$ are

$$\nabla f(x^t) + \rho_t h(x^t)\nabla h(x^t) + \lambda_t \nabla h(x^t) = \nabla f(x^t) + (\rho_t h(x^t) + \lambda_t)\nabla h(x^t) = 0 \tag{A.7}$$

Comparing equations (A.6) and (A.7), x^t can be driven to $x^\star$ if

$$\lambda_{t+1} = \lambda_t + \rho_t h(x^t) \tag{A.8}$$

The updates (A.5) and (A.8) have the same form, (A.8) is more restrictive since the KKT conditions must apply but the step size is known.

[1]The Lagrange multiplier that maximizes the dual function is equal to the KKT multiplier only when the functions are differentiable, the constraints qualification conditions apply, and there is a saddle point i.e., $min_x max_\lambda L_A(x;\lambda,\rho) = max_\lambda min_x L_A(x;\lambda,\rho)$.

The equality constraint of the article (Equation (4.7) with $\epsilon = 0$) is a minimum over distances. It has the additional feature that it is always positive or null, $\forall x \in \mathcal{X}$, $h(x) \geq 0$. Because of this, if h is locally differentiable around $x^\star$, $\nabla h(x^\star) = 0$ since h has a minimum at $x^\star$. The constraint qualification condition is not satisfied ($\nabla h(x^\star)$ does not span a non-empty set) and the KKT conditions do not apply. Another consequence is that the optimal Lagrange multiplier must be positive and the search for λ can be written $\max_{\lambda \geq 0} D(\lambda, \rho)$ in Problem (A.3), as in Problem (4.10).
Proof: Assume ρ is large enough for Problem (A.2) to have a saddle point at its optimum, $f(x^\star) \leq f(x) + \rho/2 h^2(x) + \lambda^\star h(x)$, $\forall x$ where $\lambda^\star$ is the optimum Lagrange multiplier. Since the optimization problem has an active constraint, there is a point x^I that is infeasible, $h(x^I) > 0$, and has a better objective function than the feasible solution (otherwise the constraint is useless), $f(x^I) + \frac{\rho}{2} h^2(x^I) \leq f(x^\star)$. If the optimum Lagrange multiplier is negative, $\lambda^\star < 0$, $f(x^I) + \frac{\rho}{2} h^2(x^I) + \lambda^\star h(x^I) < f(x^\star)$ which contradicts the fact that $x^\star$ is a solution to the dual problem. □

Inequality constraint

When $\epsilon > 0$, Problem (4.7) has an inequality constraint which we rewrite here more simply,

$$\begin{cases} \min_{x \in \mathcal{X}} f(x) \\ \text{such that } g(x) \leq 0 \end{cases} \tag{A.9}$$

The considerations on augmented Lagragian done above for equality constraints readily extend to inequality constraints by introducing a slack variable,

$$\begin{cases} \min_{x, s \in \mathcal{X} \times \mathbb{R}} f(x) \\ \text{such that } g(x) + s^2 = 0 \end{cases} \tag{A.10}$$

and the expression for the augmented Lagrangian (A.3) becomes

$$L_A(x, s; \lambda, \rho = f(x) + \lambda(g(x) + s^2) + \frac{1}{2}\rho(g(x) + s^2)^2 \tag{A.11}$$

The minimization of $L_A()$ on the slack variable s can be done analytically:

$$\frac{\partial L_A(x, s; \lambda, \rho)}{\partial s} = 0 \iff s^2 = -\frac{\lambda}{\rho} - g(x)$$

Since s^2 needs to be positive, all cases are summed up in

$$s^2 = \max\left(0, -\frac{\lambda}{\rho} - g(x)\right) \tag{A.12}$$

Reinjecting the expression of s^2 into the augmented Lagrangian yields

$$L_A(x; \lambda, \rho) = f(x) + \frac{1}{2\rho}\left[(\max(0, \lambda + \rho g(x)))^2 - \lambda^2\right] \tag{A.13}$$

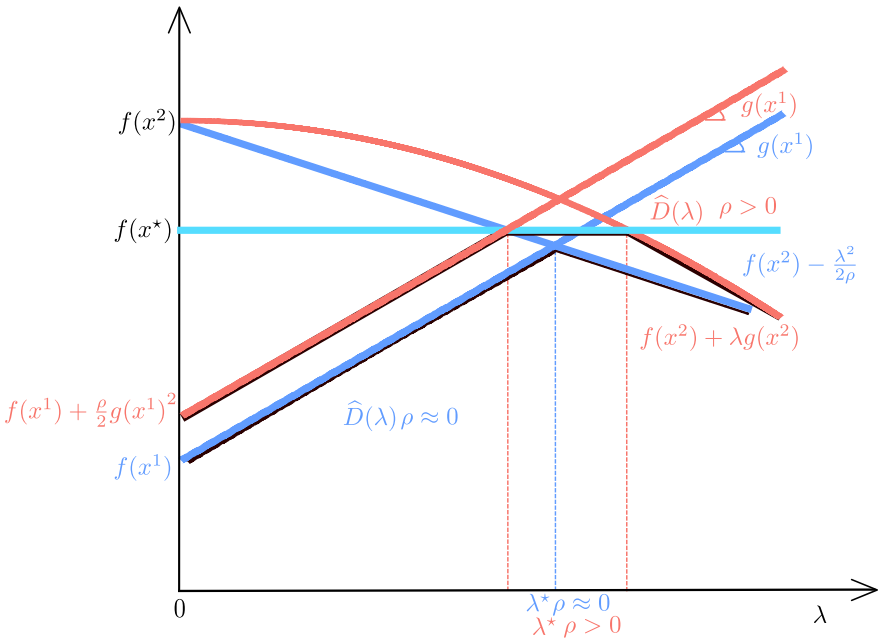


Figure A.1: Sketch of Rockafellar's augmented Lagrangian for $\rho \approx 0$ in blue and $\rho > 0$ in red. x^1 is infeasible, x^2 feasible (and $g(x^2) < -\lambda/\rho$) and $x^\star$ is an optimum with $g(x^\star) = 0$. The black highlighted curves are the approximation to the dual function, $\widehat{D}(\lambda)$ for $\mathbf{X} = \{x^1, x^2, x^\star\}$, for $\rho \approx 0$ and $\rho > 0$. There is no saddle point and a duality gap with the blue set of curves in that $x^\star \notin \arg\min_x L_A(x; \lambda^\star, \rho \approx 0)$ and $\widehat{D}(\lambda^\star) = min_x L_A(x; \lambda^\star, \rho \approx 0) < L_A(x^\star; \lambda^\star, \rho \approx 0)$, i.e., minimizing the augmented Lagrangian does not lead to the result of the problem. However, by increasing ρ, it is visible that the y-intercept of the infeasible points increase so that one always reaches a state where $x^\star = \arg\min_x L_A(x; \lambda^\star, \rho)$ as in the red set of curves. A similar illustration can be done with the augmented Lagrangian with equality constraint: $f(x) + \rho/2h^2(x)$ is the y-intercept and $h(x)$ is the slope of the augmented Lagrangian associated to x. The main difference is that all points contribute linearly in terms of λ to $L_A(x; \lambda, \rho)$.

which is equivalent to the expression of Rockafellar with the 2 cases given in Equation (4.8) (recall $-\log(1+EI)$ is $f(x)$).

The update equations for λ are the same as those for the equality case where the slack variable s^2 takes its optimal value. On the one hand, it is possible to solve the approximated dual problem as in (4.11). On the other hand, a step along a subgradient in the dual space can be taken,

$$\begin{aligned} \lambda_{t+1} &= \lambda_t + \alpha(g(x^t) + s_t^2) \\ \Rightarrow \lambda_{t+1} &= \lambda_t + \alpha\left(g(x^t) + \max(0, -\frac{\lambda}{\rho} - g(x^t))\right) \end{aligned} \tag{A.14}$$

where α is again a positive step factor. It has the same form as Equation (4.13). The update (4.13) is fully recovered from the KKT conditions as above for equalities, (A.8),

$$\begin{aligned} \lambda_{t+1} &= \lambda_t + \rho(g(x^t) + s_t^2) \\ \Rightarrow \lambda_{t+1} &= \lambda_t + \rho\left(g(x^t) + \max(0, -\frac{\lambda}{\rho} - g(x^t))\right) \end{aligned} \tag{A.15}$$

Equations (A.14) and (A.15) are the same but in the latest the step factor α is known and equal to ρ, which comes at the additional expense of the KKT validity conditions.

A.1.1 Random Forest Regression

The basic idea of Random Forest regression, as introduced by Breiman in Breiman [2001], is to combine a large collection of de-correlated tree predictors that are capable to capture complex structures in data with a relatively low bias and a bagging (bootstrap) sampling technique Tibshirani et al. [2009]. With this setting is possible to reduce the noise of an approximately unbiased model by averaging. Given a training set $(\mathbf{X}, \mathbf{Y})$ of size N_{DoE} and a **bagging** quantity B the basic setting of a random forest follows the algorithm:

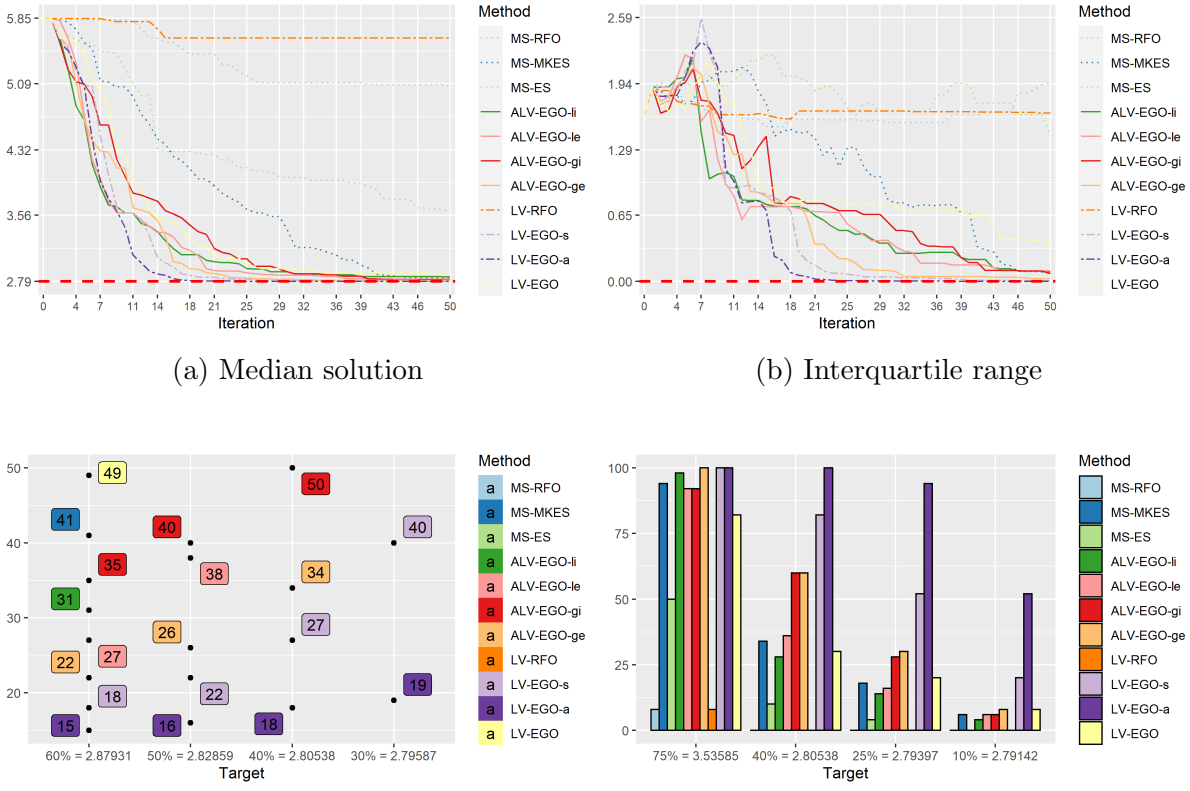
Algorithm 11 Random forest for Regression

1: **for** $b \in \{b = 1, 2, \ldots, B\}$ **do**
2: Sample with replacement N training points from $(\mathbf{X}, \mathbf{Y})$ and form the subset $(\mathbf{X}_b, \mathbf{Y}_b)$
3: Grow a random-forest tree T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until minimum node size $n_{\min}$ is reached.

- Select m variables at random from the p variables.
- Pick the best variable/split-point among the m.
- Split the node into two daughter nodes.

4: **end for**
5: Output the ensemble of trees $\{T_b\}_1^B$.
6: Predictions are made by averaging from all the B predictors $\hat{y}_{\mathrm{rf}}(x) = \dfrac{1}{B}\sum_{b=1}^{B} T_b(x)$

A.2 Supplementary results Including Heuristics



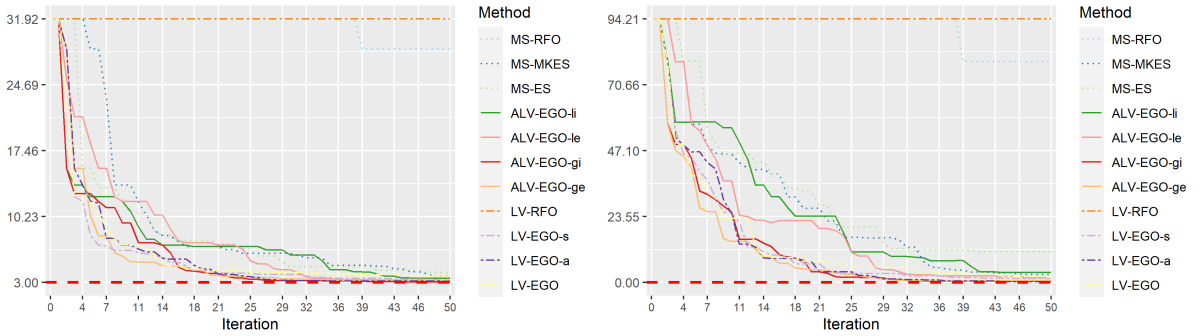
(a) Median solution

(b) Interquartile range

(c) Iteration to median success

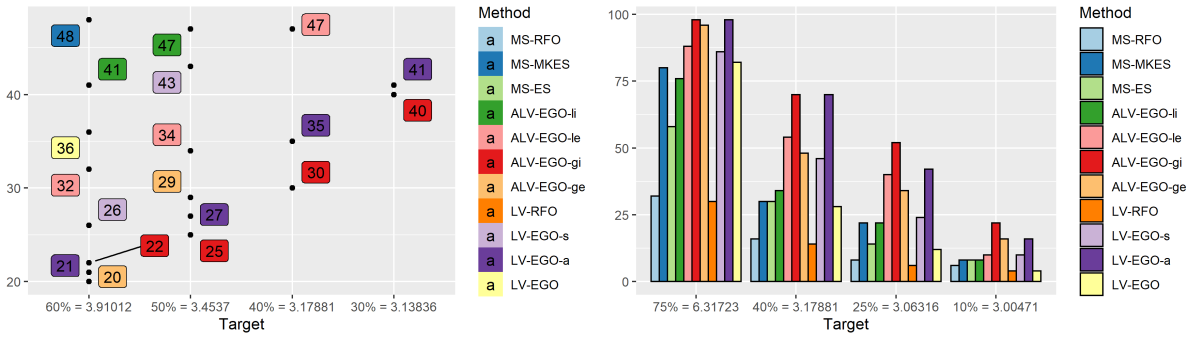
(d) Success rate

Figure A.2: Comparison of all 11 algorithms on the Branin function. $y^\star = 2.79118$.



(a) Median solution

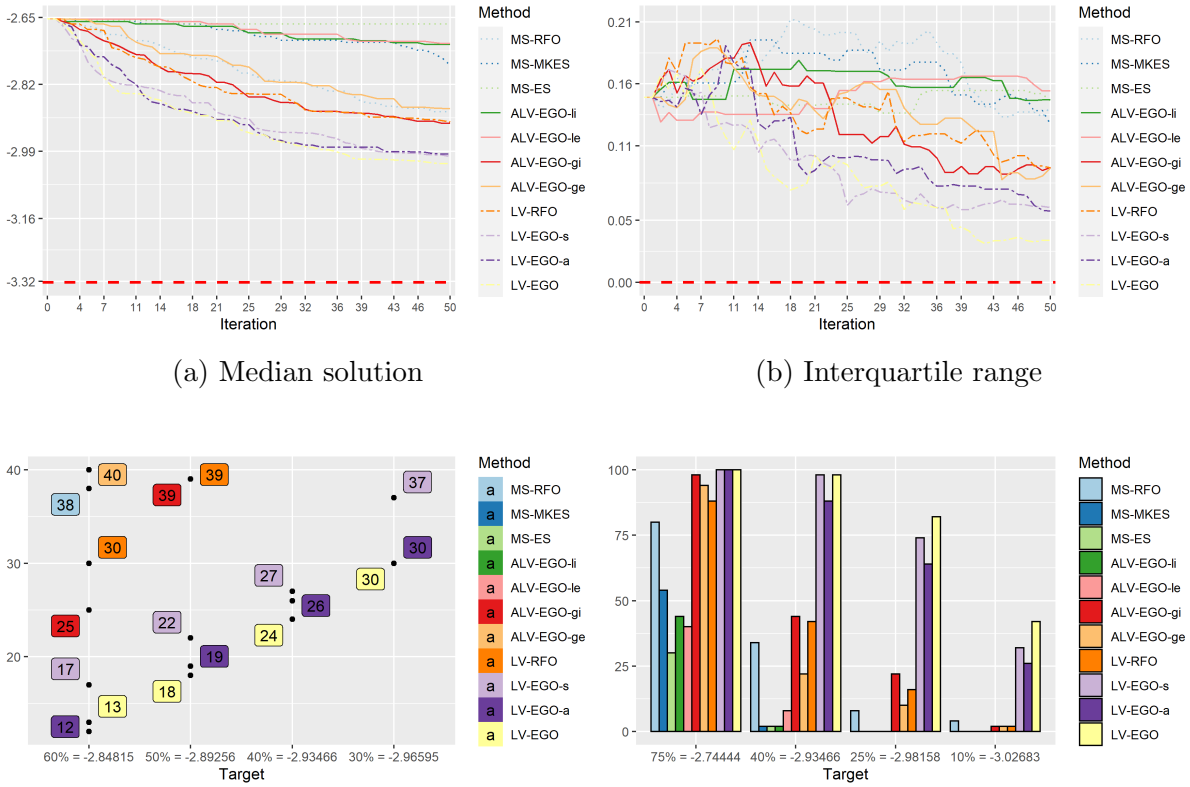
(b) Interquartile range



(c) Iteration to median success

(d) Success rate

Figure A.3: Comparison of the 11 algorithms on the Golstein function. $y^\star = 3$.



(a) Median solution

(b) Interquartile range

(c) Iteration to median success

(d) Success rate

Figure A.4: Comparison of the 11 algorithms on the Hartmann function (for which $y^\star = -3.32237$).

Chapter B

Mixed Inversion

B.1 Deterministic expression for m

In the first part of this section, we are going to add a ℓ_2 regularization term just for u and try to find an unique solution for $\hat{m}$. Later we will discuss about the necessity of adding a second penalty term related to the mass m and its implications.

For the first part, let $\mathcal{X} = \mathbb{R}^p$, $\mathcal{Y} = \mathbb{R}^p$ and assuming that u follows a multivariate Gaussian distribution $\pi(x) \sim \mathcal{N}(\mu_x, \sigma_x^2 I_p)$. The least squares problem is now

$$\underset{\hat{x},\hat{m}}{\text{argmin}} \quad (y - mx\beta)^T(y - mx\beta) + (x - \mu_x)^T\sigma_x^{-2}(x - \mu_x), \tag{B.1}$$

where the term in blue corresponds to the regularization term related to $\pi(x)$.

Again, we need to compute $\partial J/\partial x = 0$ to find $\hat{x}(m)$

$$\frac{\partial J}{\partial x} = -2m\beta y + 2m^2\beta^2 x - 2\sigma_x^{-2}\mu_x + 2\sigma_x^{-2}x = 0$$
$$\hat{x}(m) = \frac{\mu_x + \sigma_x^2 m\beta y}{1 + \sigma_x^2 m^2\beta^2},$$

now the updated function $J(m)$ will be

$$J(m) = \frac{(y - m\mu_x\beta)^T(y - m\mu_x\beta)}{1 + \sigma_x^2 m^2\beta^2},$$

which is a 1 variable real valued function. Figure B.1 show that $J(m)$, for real values of m is quadratic. This provides sufficient motivation to find an algebraic expression for real (positive) values of $\hat{m}$ as desired by the application.

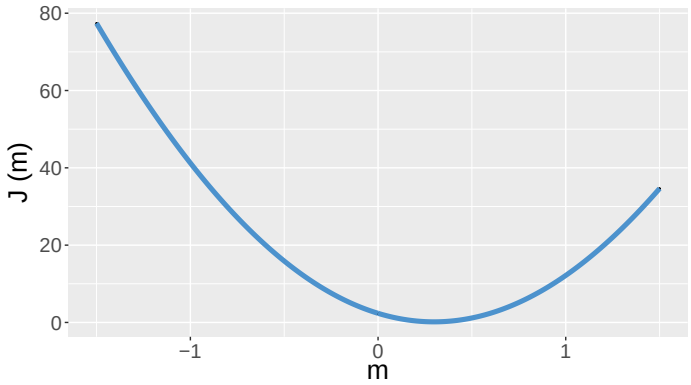


Figure B.1: Graphical representation of the function $J(m)$, for real values of m. This function reach its minimum value on the model estimator $\hat{m}$

Expression for $\hat{m}$: Re-defining useful constants $z = \beta^2\sigma_8 xx^2$, $\omega = \beta\mu_8 xx$

$$J(m) = \frac{(y - m\omega)^T(y - m\omega)}{1 + m^2 z}$$

Now computing $J'(m) = 0$

$$\begin{aligned}
J'(m) &= \frac{-2\omega(1 + m^2 z)(y - m\omega)^T - 2mz(y - m\omega)^T(y - m\omega)}{(1 + m^2 z)^2} \\
0 &= \frac{-2(y - m\omega)^T[\omega(1 + m^2 z) + mz(y - m\omega)]}{(1 + m^2 z)^2} \\
0 &= \frac{-2(y - m\omega)^T[\omega + \cancel{\omega m^2 z} + ymz - \cancel{\omega m^2 z}]}{(1 + m^2 z)^2} \\
0 &= \frac{-2(y - m\omega)^T(\omega + ymz)}{(1 + m^2 z)^2}
\end{aligned}$$

Now removing constants and positive values

$$\begin{aligned}
0 &= (y - m\omega)^T(\omega + ymz) \\
0 &= m^2(\omega^T yz) + m(|\omega|^2 - z|y|^2) - y^T\omega,
\end{aligned}$$

which corresponds to a second order polynomial of the form $Am^2 + Bm + C = 0$ where,

$$\begin{aligned}
A &= \omega^T yz = \beta^3\sigma_x^2\mu_x^T y \\
B &= |\omega|^2 - z|y|^2 = \beta^2(|\mu_x|^2 - \sigma_x^2|y|^2) \\
C &= -y^T\omega = -\beta\mu_x^T y
\end{aligned}$$

The test-case requires to find a unique value $m \geq 0$. To impose this constraint in the quadratic formula, we require both roots R_1, R_2 to be real ($\Delta = B^2 - 4AC > 0$) and with opposite sign ($R_1 R_2 < 0$).

We are going to find the scenarios when we could satisfy this constraint, for that, we are going to compare with the real roots polynomial $x^2 - x(R_1 + R_2) + R_1 R_2 = 0$. For that we are dividing by A to ensure positive sign on the second order term

$$\begin{aligned} m^2(\omega^T yz) + m\left(\omega^T\omega - y^T yz\right) - y^T\omega &= 0 \\ m^2 + m\left(\frac{B}{A}\right) + \frac{C}{A} &= 0 \\ m^2 + m\left(\frac{\omega^T\omega - y^T yz}{\omega^T yz}\right) + \frac{-y^T\omega}{\omega^T yz} &= 0 \end{aligned}$$

from where we could extract $R_1 R_2 = C/A = -1/z$. Forcing $R_1 R_2 < 0$, means that $-\frac{1}{z} < 0$. This could be possible if and only if $z > 0$, which recovering the original notation $z = \beta^2\sigma_x^2$, which will always satisfy $z > 0$, which also satisfies $R_1 R_2 < 0$ and $\Delta = B^2 - 4AC > 0$. This means, that we will always recover 1 positive and 1 negative root from the second order polynomial, ensuring that there is a single value of m that satisfies the constraints of the test-case.

Finally, we can find an analytical expression to $\hat{m}$

$$\hat{m} = \max \quad [m_1, m_2] \tag{B.2}$$

$$m_{1,2} = \frac{\beta^2\sigma_x^2|y|^2 - \beta^2|\mu_x|^2 \pm \sqrt{\beta^4|\mu_x|^4 - 2\beta^4\sigma_x^2|\mu_x|^2|y|^2 + \beta^4\sigma_x^4|y|^4 - 4\beta^4\sigma_x^2(\mu_x^T y)^2}}{2\beta^3\sigma_x^2(\mu_x^T y)} \tag{B.3}$$

$$m_{1,2} = \frac{\sigma_x^2|y|^2 - |\mu_x|^2 \pm \sqrt{|\mu_x|^4 - 2\sigma_x^2|\mu_x|^2|y|^2 + \sigma_x^4|y|^4 + 4\sigma_x^2(\mu_x^T y)^2}}{2\beta\sigma_x^2(\mu_x^T y)} \tag{B.4}$$

$$m_{1,2} = \frac{\sigma_x^2|y|^2 - |\mu_x|^2 \pm \sqrt{(\sigma_x^2|y|^2 - |\mu_x|^2)^2 + 4\sigma_x^2(\mu_x^T y)^2}}{2\beta\sigma_x^2(\mu_x^T y)} \tag{B.5}$$

Analyzing the expression for $\hat{m}$ Now we are interested in providing a mathematical interpretation of the results obtained in equation B.2. We know from the quadratic polynomial that we will obtain real roots with opposite sign if $\beta^2\sigma_x^2 > 0$, that could be easily satisfied having $\beta \neq 0$ (that will always be true, because if β is 0 we have no linear problem to solve) and $\sigma_x > 0^+$ (which also is always true because is a standard deviation

and by definition is a positive measure). Now, recalling the general quadratic solution

$$
\begin{aligned}
m_{1,2} &= \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \\
C &= -\beta \mu_x^T y \\
B &= \beta^2 |\mu_x|^2 - \beta^2 \sigma_x^2 |y|^2 \\
A &= -\beta^2 \sigma_x^2 C
\end{aligned}
$$

with $A \neq 0$ that forces $C = -\beta \mu_x^T y \neq 0$. Now applying $\beta^2 \sigma_x^2 > 0$, we will always get $AC < 0$, which means we will always get $\sqrt{B^2 + 4AC}$ and also $|\sqrt{B^2 + 4AC}| \geq |B|$.

With this information and knowing that the sign of $\beta \mu_x^T y$ will play a role for identifying

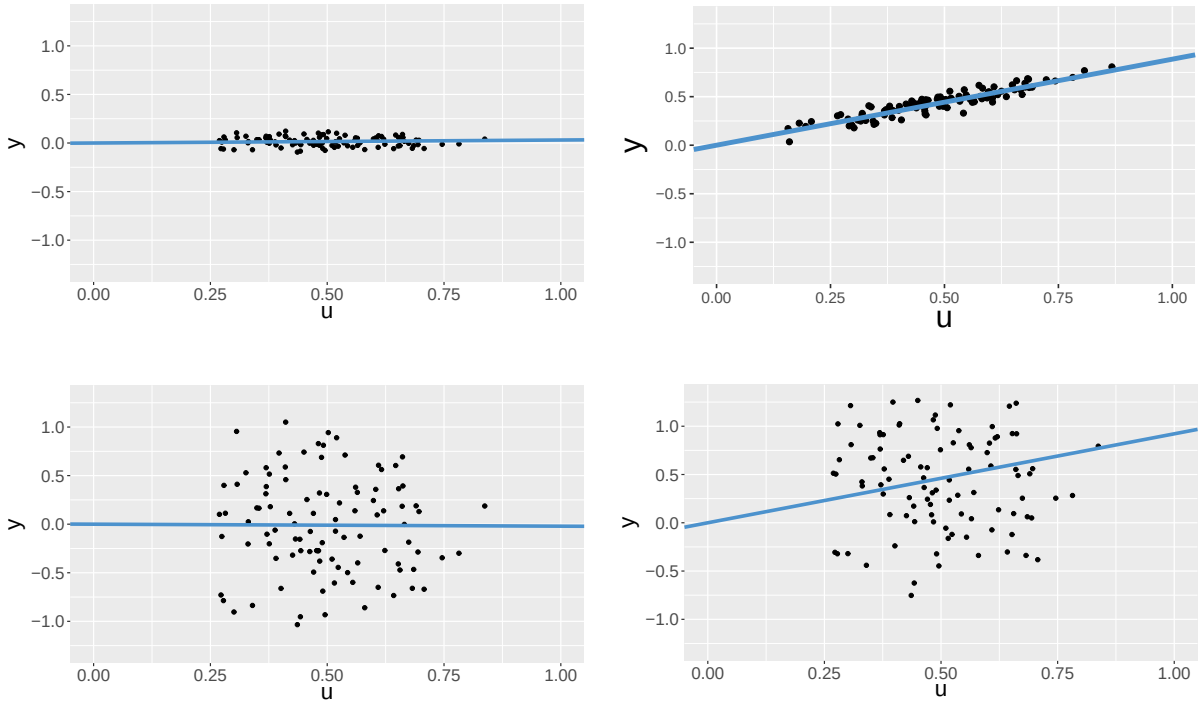


Figure B.2: Estimations of the m for different configurations. Top Left:$\gamma = 0.05$, $m = 0.3$, $\beta = 0.1$,$\hat{m} = 0.30824$. Top Right: $\gamma = 0.05$, $m = 0.3$, $\beta = 3$, $\hat{m} = 0.2991$. Bottom Left: $\gamma = 0.5$, $m = 0.3$, $\beta = -0.1$, $\hat{m} = 0.2045$. Bottom Right: $\gamma = 0.5$, $m = 0.3$, $\beta = 4$, $\hat{m} = 0.3071$.

the positive root of $m_{1,2}$, we will obtain 2 different scenarios

$$\text{Case 1: if} \quad \beta\mu_x^T y > 0$$

$$\hat{m} = \frac{+\sqrt{B^2 + 4AC} - B}{2A}$$

$$\text{Case 2: if} \quad \beta\mu_x^T y < 0$$

$$\hat{m} = \frac{+\sqrt{B^2 + 4AC} + B}{2|A|}$$

Recall that

$$m_{1,2}(\sigma_x^2) = \frac{-(|\mu_x|^2 - \sigma_x^2|y|^2) \pm \sqrt{(\sigma_x^2|y|^2 - |\mu_x|^2)^2 + 4\sigma_x^2(\mu_x^T y)^2}}{2\beta\sigma_x^2(\mu_x^T y)}$$

First, we are going to rewrite this expression in order to simplify. Defining $p = \sigma_x^2$, the expression inside the square root is

$$(p|y|^2 - |\mu_x|^2)^2 + 4p(\mu_x^T y)^2 = p^2|y|^4 - 2p|y|^2|\mu_x|^2 + |\mu_x|^4 + 4p(\mu_x^T y)^2$$

Defining $a = |y|^2$, $b = |\mu_x|^2$, $c = (\mu_x^T y)$

$$p^2|y|^4 - 2p|y|^2|\mu_x|^2 + |\mu_x|^4 + 4p(\mu_x^T y)^2 = p^2a^2 - 2pab + b^2 + 4pc^2$$

The new expression for $m_{1,2}(p)$

$$m_{1,2}(p) = \frac{-(b - pa) \pm (p^2a^2 - 2pab + b^2 + 4pc^2)^{1/2}}{2\beta pc}$$

Now we are going to check in both scenarios when $p \to 0^+$. First we are going to check Case 2

$$\lim_{p\to 0^+} \frac{+\sqrt{p^2a^2 - 2pab + b^2 + 4pc^2} + b - pa}{2\beta p|c|}$$

$$\lim_{p\to 0^+} \frac{2b}{0^+} \to +\infty$$

Given that $b = |\mu_x|^2 \neq 0$, we will fail to recover the mass.
Now for Case 1

$$\lim_{p\to 0^+} \frac{+\sqrt{p^2a^2 - 2pab + b^2 + 4pc^2} + pa - b}{2\beta pc}$$

$$\lim_{p\to 0^+} \frac{0^+}{0^+} = \text{undet..}$$

To solve this, we are going to approximate $(p^2a^2 - 2pab + b^2 + 4pc^2)^{1/2}$ using the Taylor expansion for $(1+x)^{1/2} = 1 + x/2 - x^2/4 + o(x^2)$, where $o(x^2)$ are negligible terms when $x \to 0$. Applying this expansion we obtain

$$(p^2a^2 - 2pab + b^2 + 4pc^2)^{1/2} = b\left(1 + p\left(\frac{a^2p - 2ab + 4c^2}{b^2}\right)\right)^{1/2}$$

$$b\left(1 + p\left(\frac{a^2p - 2ab + 4c^2}{b^2}\right)\right)^{1/2} = b[1 + \frac{p}{2b^2}(4c^2 - 2ab + a^2p) - \frac{p^2}{4b^4}(16c^4 - 16abc^2 + 4a^2b^2 - 4pa^3b + 8pa^2c^2 + p^2a^4) + o(p^3)]$$

When extracting the terms in red we will get $b - pa$, that will cancel with $+pa - b$ on the numerator of $\hat{m}$. Now neglecting p^3 terms we obtain

$$\hat{m}(p) = \frac{1}{2p\beta c}\left(p\frac{2c^2}{b} + \frac{p^2}{b}\left(\frac{a^2}{2} - a^2\right) + \frac{p^2}{b^2}(4c^2a) - \frac{p^2}{b^3}(4c^4) + o(p^3)\right)$$

Now applying the denominator we obtain

$$\hat{m}(p) = \frac{c}{\beta b} + \frac{p}{\beta b}\left(\frac{2ca}{b} - \frac{a^2}{4c} - \frac{2c^3}{b^2}\right) + o(p^2)$$

where this corresponds to a polynomial expansion of $\hat{m}(p)$ centered in 0. Now making $p \to 0$

$$\hat{m_{\text{LS}}}(p) = \frac{c}{\beta b} = \frac{\mu_x^T y}{\beta|\mu_x|^2} \tag{B.6}$$

which corresponds to the Least Squares solution, making $x = \mathrm{E}[x] = \mu_x$. Also we could use the term depending on p as an approximation error $\xi(p)$

$$\xi(p) \approxeq \frac{p}{\beta b}\left(\frac{2ca}{b} - \frac{a^2}{4c} - \frac{2c^3}{b^2}\right)$$

$$\xi(\sigma_x^2) \approxeq \frac{\sigma_x^2}{\beta|\mu_x|^2}\left(\frac{2(\mu_x^T y)|y|^2}{|\mu_x|^2} - \frac{|y|^4}{4\mu_x^T y} - \frac{2(\mu_x^T y)^3}{|\mu_x|^4}\right)$$

this means that $|\hat{m} - \hat{m}_{\text{LS}}| \leq \xi(\sigma_u^2)$ for this specific approach.

B.2 Including m penalty term

Lets try to include also a penalty term involving the uncertainty about m in $J(m,u)$ this means defining a prior $\pi(m) \sim \mathcal{N}(m_0, \sigma_m^2)$. Now the optimization problem is

$$\underset{\hat{u},\hat{m}}{\text{argmin}} \quad (y - mu\beta)^T(y - mu\beta) + (u - \mu_u)^T\sigma_u^{-2}(u - \mu_u) + \frac{(m - m_0)^2}{\sigma_m^2},$$

where is important to remark that if $\sigma_m \to \infty$ we recover the equation B.1. This new inclusion does not affect the value of $\hat{u}$, but it will change the form of $J(m)$ to

$$J(m) = \frac{(y - m\mu_u\beta)^T(y - m\mu_u\beta)}{1 + \sigma_u^2 m^2\beta^2} + \frac{(m - m_0)^2}{\sigma_m^2}$$

It could be proved that $J'(m) = 0$ will be no longer a second order polynomial (will include 5^{th} order terms into the equation) and will be not so obvious the possibility of finding an analytical expression for the test-case constraints, also to find the scenarios where that solution will be unique and positive. That is why, given the good experimental results with the approximation given by equation B.1, it will not be necessary for this test-case to include a penalty term involving m.

B.3 General expression for any $f(X) = X\beta$

We are interested in considering the more general case of $f(X) = X\beta$, here we are going to introduce the matrix of continuous input $X \in \mathbb{R}^{p\times n_c}$ and a expensive function $f(X) = X\beta$, then the inverse problem is

$$y = mX\beta + \eta,$$

where $y \in \mathbb{R}^p$, X is the design matrix with rows $X_i \triangleq x$, and the $\beta = [\beta_1, \ldots, \beta_j, \ldots \beta_{n_c}]^\top$ are learnt from data. Defining the same prior $\pi(x_j) \sim \mathcal{N}(\mu_{x_j}, \sigma_{x_j}^2 I_{n_c})$. Now, we formulate the regularization problem:

$$\begin{aligned} \underset{m,X}{\text{argmin}} \quad & J(X,m) = (y - mX\beta)^T(y - mX\beta) + \sum_{i=1}^{n_c} \frac{(x_i - \mu_{x_i})^2}{\sigma_i^2} \\ & J(X,m) = (y - mX\beta)^T(y - mX\beta) + \text{Tr}((X - A)^T(X - A)D^2) \end{aligned}$$

where:

$$A = [\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_{n_c}}]^\top$$

is a $p \times n_c$ design matrix with the mean values of the x_j stacked by columns, and D is a $n_c \times n_c$ matrix, that includes in its diagonal the inverse of the standard deviations of X.

$$D = \begin{pmatrix} \sigma_{x_1}^{-1} & d_{1,2} & \ldots & d_{1,n_c} \\ d_{2,1} & \sigma_{x_2}^{-1} & \ldots & d_{2,n_c} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n_c,1} & d_{n_c,2} & \ldots & \sigma_{x_{n_c}}^{-1} \end{pmatrix},$$

Again we compute the partial derivative $\partial J/\partial X = 0$, find $\hat{X}(m)$, and then update the expression of $J(\hat{X}(m), m)$ to find $\hat{m}$.

$$\begin{aligned} 0 &= -2m(y - mX\beta)\beta^T + 2(X - A)D^2 \\ (X - A)D^2 &= m(y - mX\beta)\beta^T \\ XD^2 + m^2X\beta\beta^T &= my\beta^T + AD^2 \\ X(m^2\beta\beta^T + D^2) &= AD^2 + my\beta^T \\ \hat{X} &= (AD^2 + my\beta^T)(D^2 + m^2\beta\beta^T)^{-1} \end{aligned}$$

where, $\hat{X}$ exists if $(D^2 + m^2\beta\beta^T)^{-1}$ exists. Now defining constants $\delta = m\beta$ and $D^2 = \Sigma$ and applying *Woodbory* inversion formula

$$\begin{aligned} (D^2 + m^2\beta\beta^T)^{-1} &= (\Sigma + \delta\delta^T)^{-1} \\ &= \Sigma^{-1} - \frac{\Sigma^{-1}\delta\delta^T\Sigma^{-1}}{1 + \delta^T\Sigma^{-1}\delta} \end{aligned}$$

now re defining the scalar product $z = \delta^T\Sigma^{-1}\delta$ and simplifying the denominator we obtain

$$(D^2 + m^2\beta\beta^T)^{-1} = \frac{\Sigma^{-1}}{1 + z}$$

Now, $\hat{X}m\beta$

$$\begin{aligned} \hat{X} &= (AD^2 + my\beta^T)(D^2 + m^2\beta\beta^T)^{-1}m\beta \\ &= (A\Sigma + y\delta^T)\left(\frac{\Sigma^{-1}}{1 + z}\right)\delta \\ &= \frac{yz + A\delta}{1 + z} \end{aligned}$$

Now, $y - \hat{X}m\beta$

$$\begin{aligned} y - m\hat{X}\beta &= y - \frac{yz + A\delta}{1 + z} \\ &= \frac{y - A\delta}{1 + z} \end{aligned}$$

Now computing $(X - A)^T$ and $(X - A)D^2$

$$\begin{aligned}
(X - A)D^2 &= (y - m\hat{X}\beta)m\beta^T \\
(X - A)D^2 &= \frac{(y - A\delta)}{1 + z}\delta^T \\
(X - A) &= \frac{(y - A\delta)}{1 + z}\delta^T\Sigma^{-1} \\
(X - A)^T &= \Sigma^{-1}\delta\frac{(y - A\delta)^T}{1 + z} \\
\mathrm{Tr}((X - A)^T(X - A)D^2) &= \mathrm{Tr}\left(\Sigma^{-1}\delta\frac{(y - A\delta)^T(y - A\delta)}{(1 + z)^2}\delta^T\right) \\
&= \mathrm{Tr}\left(\frac{(y - A\delta)^T(y - A\delta)}{(1 + x)^2}\delta^T\Sigma^{-1}\delta\right) \\
&= \mathrm{Tr}\left(\frac{(y - A\delta)^T(y - A\delta)}{(1 + z)^2}z\right)
\end{aligned}$$

The term inside the trace is scalar, so is possible to remove the trace operator and update $J(m, \hat{X})$

$$\begin{aligned}
J(m, \hat{X}) &= \frac{(y - A\delta)^T(y - A\delta)}{(1 + z)^2} + \frac{(y - A\delta)^T(y - A\delta)}{(1 + z)^2}z \\
&= \frac{\cancel{(1 + z)}(y - A\delta)^T(y - A\delta)}{(1 + z)^{\cancel{2}}} \\
&= \frac{(y - A\delta)^T(y - A\delta)}{(1 + z)} \\
J(m) &= \frac{(y - A\beta m)^T(y - A\beta m)}{(1 + m^2\beta^T\Sigma^{-1}\beta)}
\end{aligned}$$

Now, defining $w = \beta^T\Sigma^{-1}\beta$ and computing $J'(m) = 0$ we obtain

$$\begin{aligned}
0 &= \frac{1}{(1 + m^2w)^2}\left[-2(y - A\beta m)A\beta(1 + m^2w) - 2mz(y - A\beta m)^T(y - A\beta m)\right] \\
0 &= (y - A\beta m)^T\left(-A\beta - \cancel{m^2A\beta w} - ymw + \cancel{m^2A\beta w}\right) \\
0 &= (y^T - m\beta^TA^T)(-A\beta - ymw) \\
0 &= m^2(\beta^TA^Tyw) + m(\beta^TA^TA\beta - w|y|^2) - y^TA\beta
\end{aligned}$$

Again, we obtain a 2^{nd} order polynomial

$$\begin{aligned}
am^2 + bm + c &= 0 \\
a &= (\beta^TA^Ty)(\beta^T\Sigma^{-1}\beta) \\
b &= \beta^TA^TA\beta - \beta^T\Sigma^{-1}\beta|y|^2 \\
c &= -\beta^TA^Ty
\end{aligned}$$

where again we could ensure a single positive root if $a/c < 0$, this means $-1/(\beta^T\Sigma^{-1}\beta) < 0$ and $\beta^T\Sigma^{-1}\beta \succ 0$ which is possible if and only if $\Sigma \succ 0$ is a positive definite matrix, that also ensures the existence of $(D^2 + m^2\beta\beta^T)^{-1}$.

Expression for $\hat{m}$: Finally we could find an analytical expression for $\hat{m}$.

$$\hat{m} = \max \quad [m_1, m_2]$$
$$m_{1,2} = \frac{\beta^T\Sigma^{-1}\beta|y|^2 - \beta^T A^T A\beta \pm \sqrt{(\beta^T\Sigma^{-1}\beta|y|^2 - \beta^T A^T A\beta)^2 + 4(\beta^T\Sigma^{-1}\beta)(\beta^T A^T y)^2}}{2(\beta^T A^T y)(\beta^T\Sigma^{-1}\beta)}$$

B.4 Complete Results By Configuration

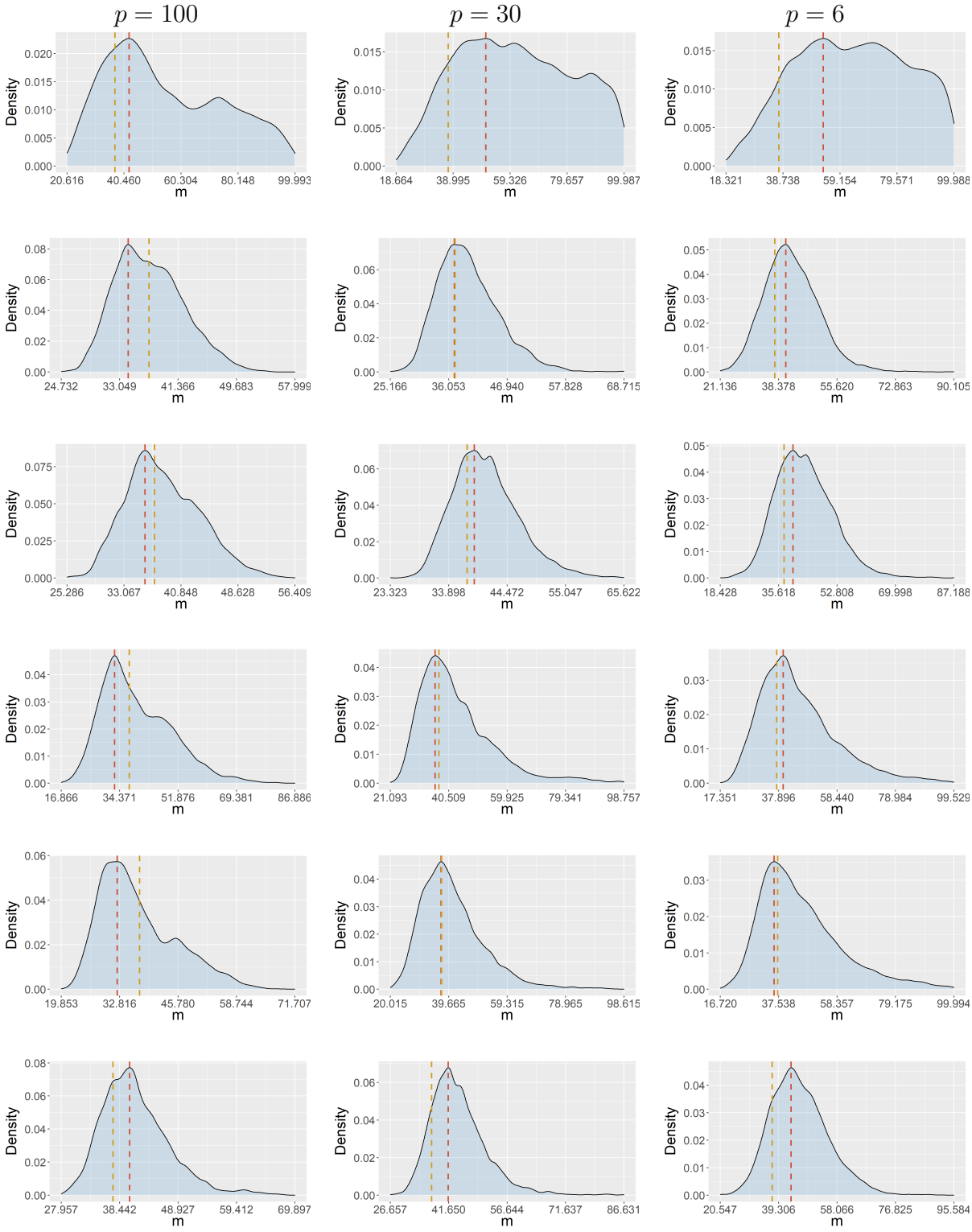


Figure B.3: Results for the strategy 1 $n_c + 1$ MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

Myles Saint-Etienne Guillermo Cuesta-Ramirez

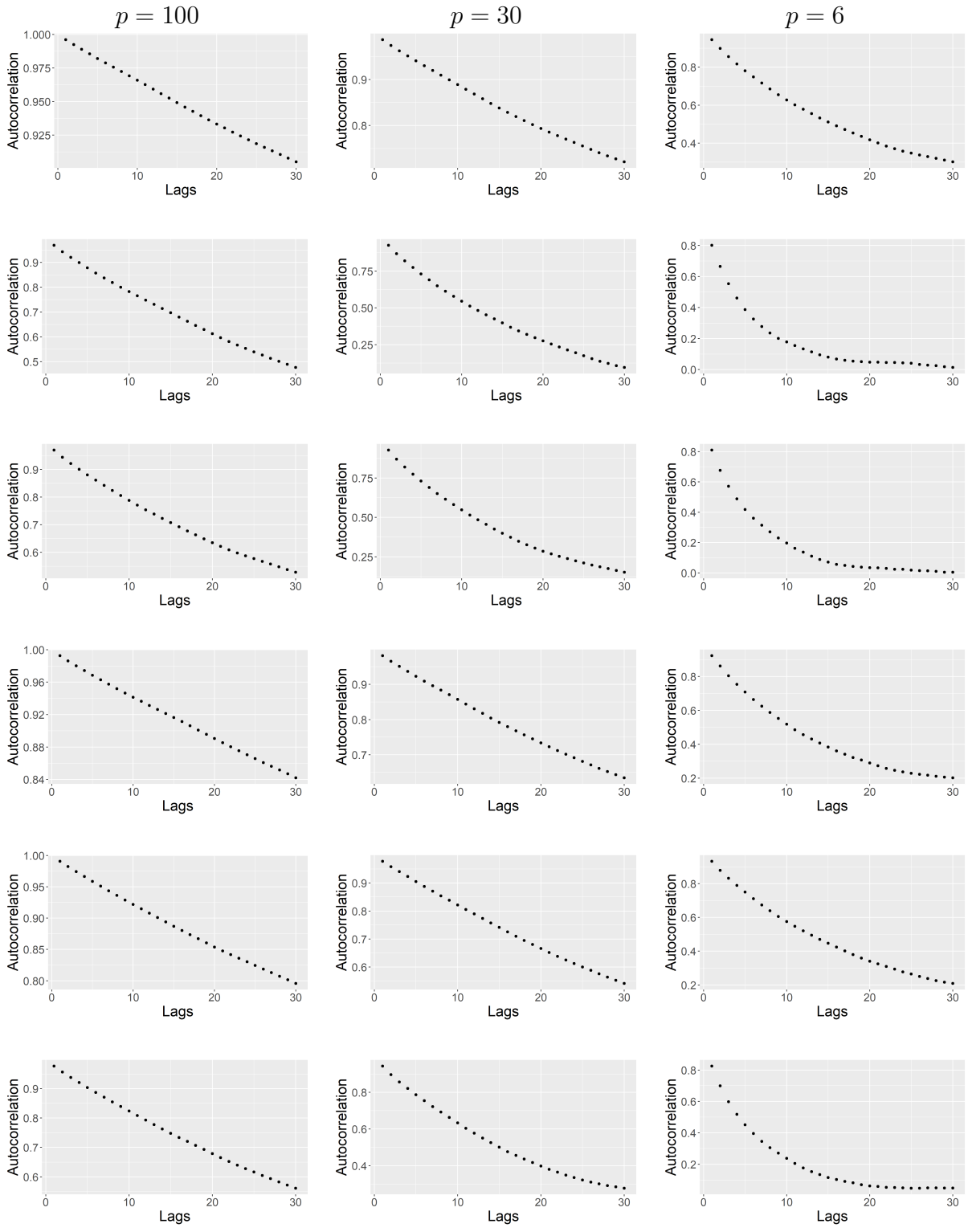


Figure B.4: Results for the strategy 1 $n_c + 1$ MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

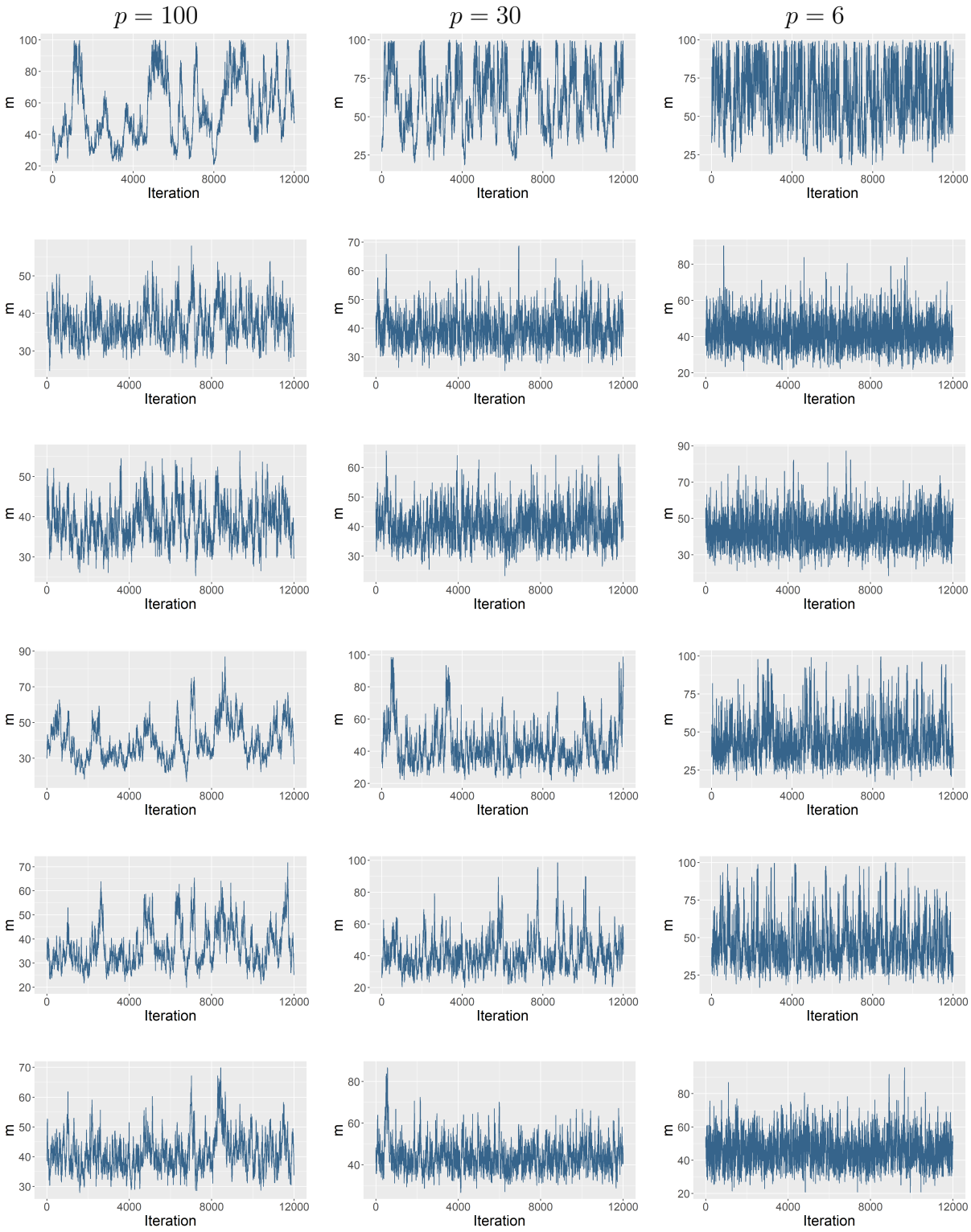


Figure B.5: Results for the strategy 1 $n_c + 1$ MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

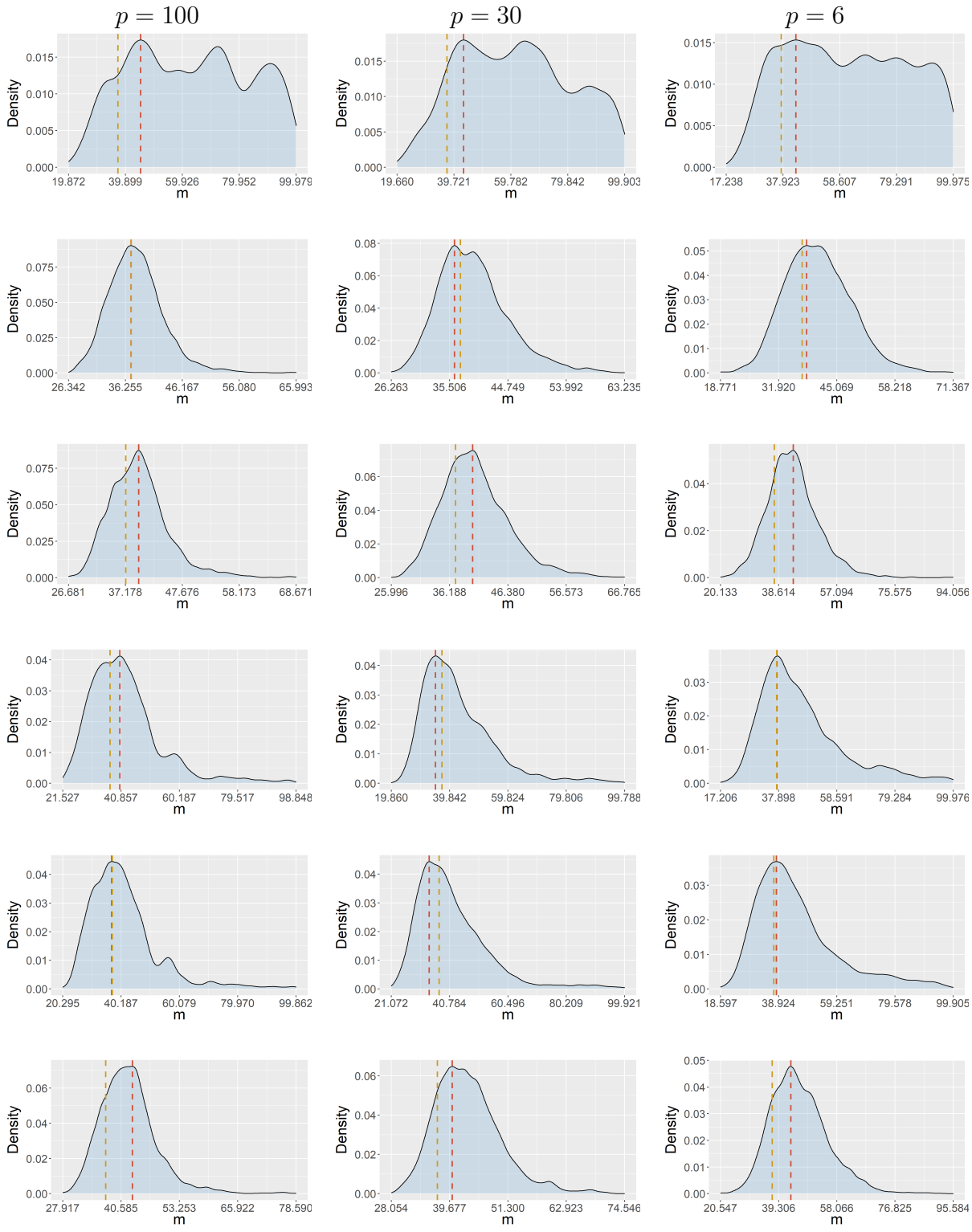


Figure B.6: Results for the strategy 2 MC within MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

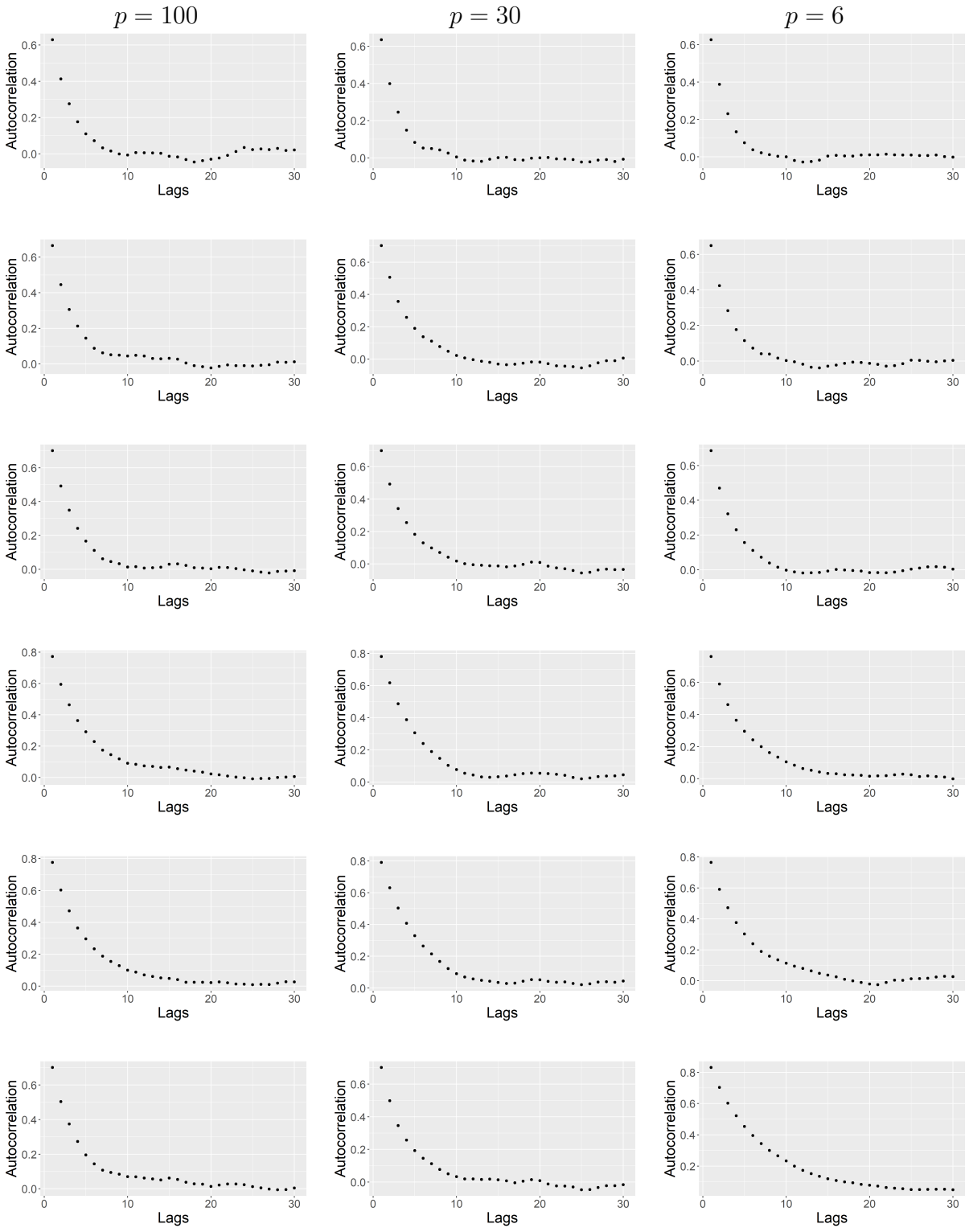


Figure B.7: Results for the strategy 2 MC within MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

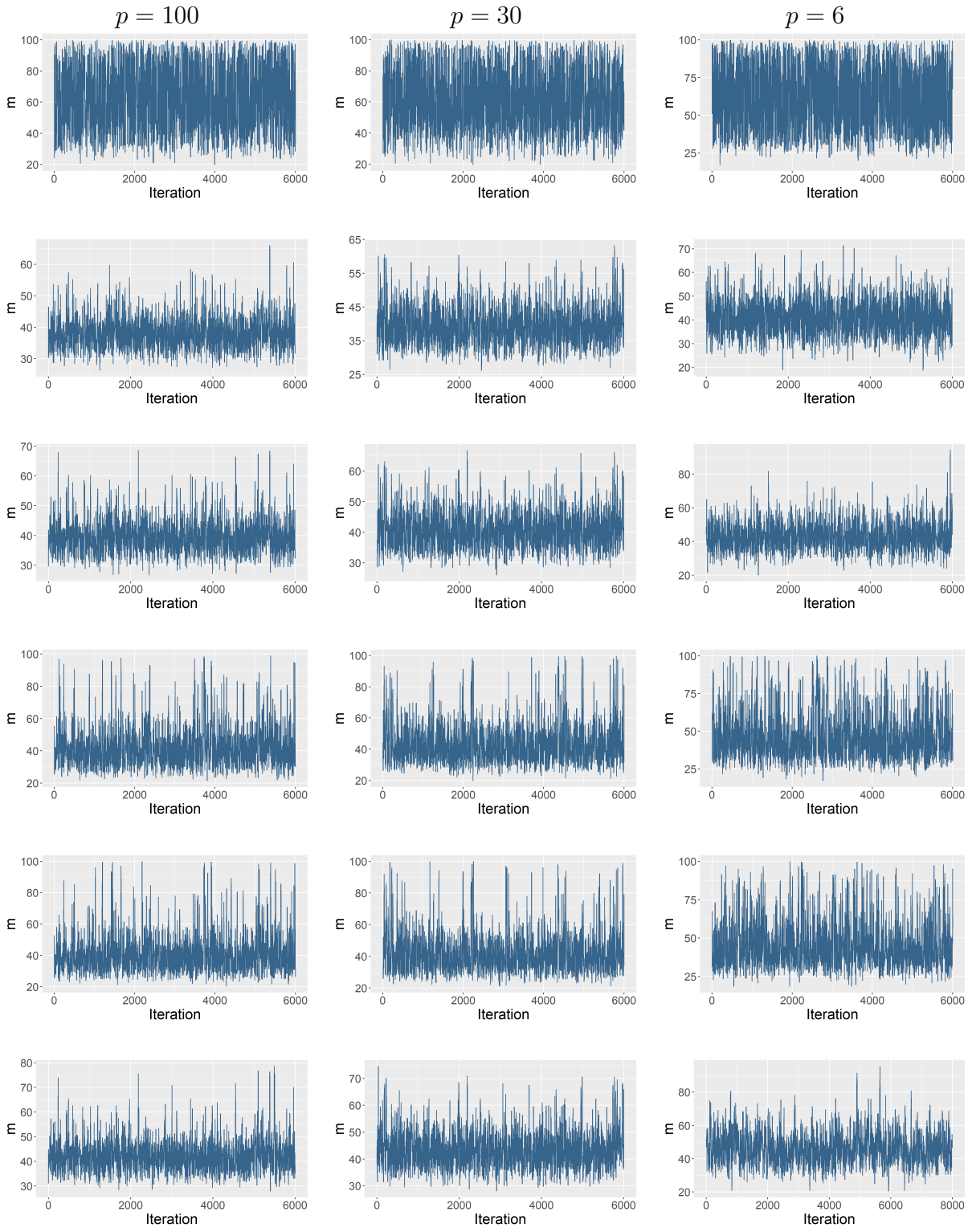


Figure B.8: Results for the strategy 2 MC within MCMC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

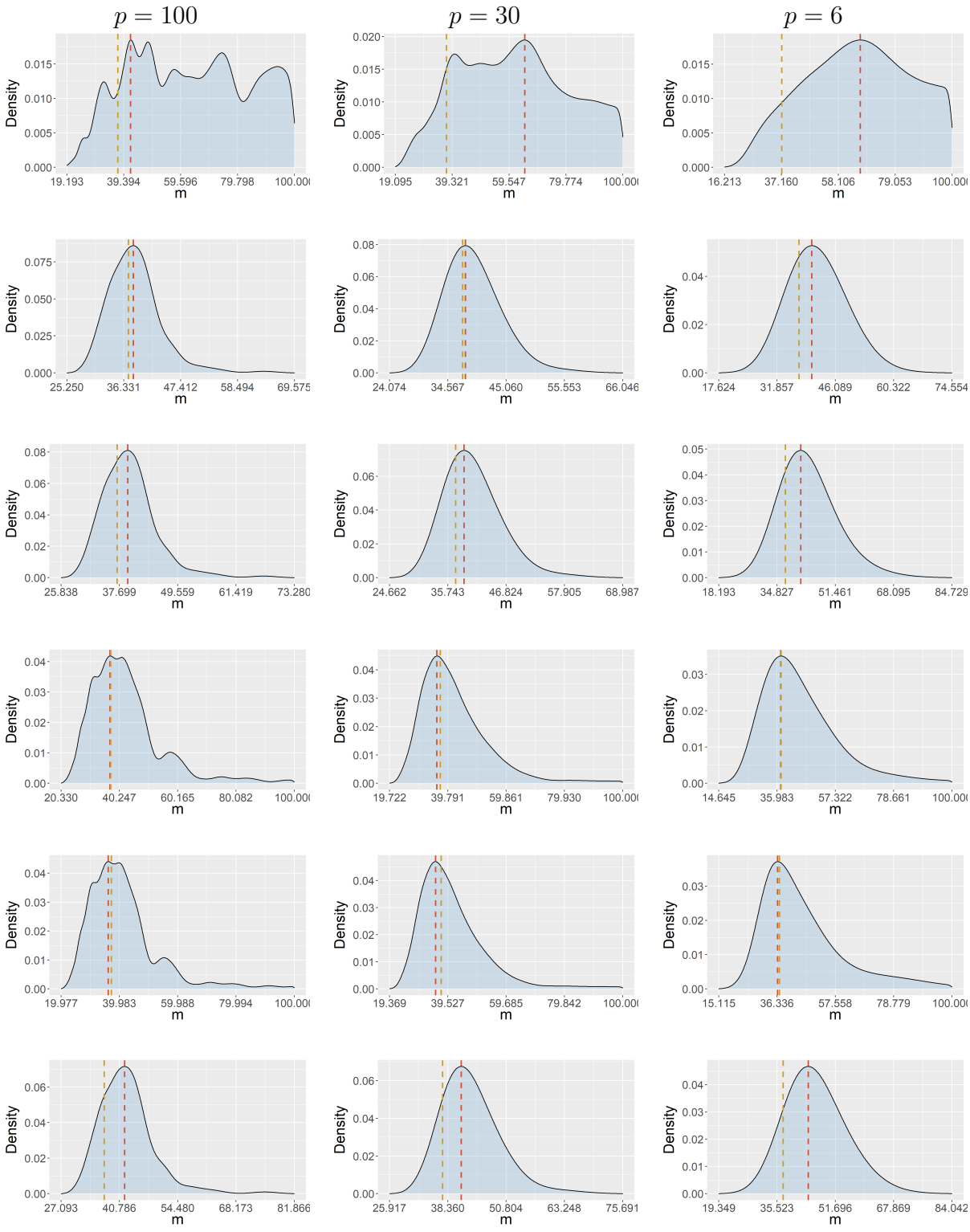


Figure B.9: Results for the strategy 3 point-wise MC, for configurations C1 (top) to C6 (bottom)for different number of observations $p = 100$ (left), $p = 30$ (mid), $p = 6$ (right). The red dashed line represents the MAP estimate and the gold dashed line the true value $m^\star$

References

Apoorv Agnihotri and Nipun Batra. Exploring bayesian optimization. *Distill*, 2020. doi: 10.23915/distill.00026. https://distill.pub/2020/bayesian-optimization.

Charles Audet and John E Dennis Jr. Pattern search algorithms for mixed variable programming. *SIAM Journal on Optimization*, 11(3):573–594, 2001.

T. Bartz-Beielstein, B. Filipič, P. Korošec, and E.G. Talbi. *High-Performance Simulation-Based Optimization*. Studies in Computational Intelligence. Springer International Publishing, 2019. ISBN 9783030187644. URL https://books.google.fr/books?id=8yGbDwAAQBAJ.

Thomas Bartz-Beielstein and Martin Zaefferer. Model-based methods for continuous and discrete global optimization. *Applied Soft Computing*, 55:154–167, 2017.

Pietro Belotti, Christian Kirches, Sven Leyffer, Jeff Linderoth, James Luedtke, and Ashutosh Mahajan. Mixed-integer nonlinear optimization. *Acta Numerica*, 22:1–131, 2013.

Bernd Bischl, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. mlrMBO: A modular framework for model-based optimization of expensive black-box functions, 2018.

Ilhem Boussaïd, Julien Lepagnot, and Patrick Siarry. A survey on optimization metaheuristics. *Information Sciences*, 237:82–117, July 2013. ISSN 00200255. doi: 10.1016/j.ins.2013.02.041. URL https://linkinghub.elsevier.com/retrieve/pii/S0020025513001588.

Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, October 2001. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324.

YJ Cao, L Jiang, and QH Wu. An evolutionary programming approach to mixed-variable optimization problems. *Applied Mathematical Modelling*, 24(12):931–942, 2000.

A. Clement, N. Saurel, and G. Perrin. Stochastic approach for radionuclides quantification. *EPJ Web of Conferences*, 170:06002, 2018. ISSN 2100-014X. doi: 10.1051/

epjconf/201817006002. URL https://www.epj-conferences.org/10.1051/epjconf/201817006002.

Masoumeh Dashti and Andrew M. Stuart. The Bayesian Approach to Inverse Problems. In Roger Ghanem, David Higdon, and Houman Owhadi, editors, *Handbook of Uncertainty Quantification*, pages 1–118. Springer International Publishing, Cham, 2016. ISBN 978-3-319-11259-6. doi: 10.1007/978-3-319-11259-6_7-1. URL https://doi.org/10.1007/978-3-319-11259-6_7-1.

Yves Deville, David Ginsbourger, Olivier Roustant, and Nicolas Durrande. kergp. https://cran.r-project.org/package=kergp, 2017–2021.

Patrycja Dyrcz, Thomas Frosio, Nabil Menaa, Matteo Magistris, and Chris Theis. Qualification of the activities measured by gamma spectrometry on unitary items of intermediate-level radioactive waste from particle accelerators. *Applied Radiation and Isotopes*, 167:109431, 2021. ISSN 0969-8043. doi: https://doi.org/10.1016/j.apradiso.2020.109431. URL https://www.sciencedirect.com/science/article/pii/S0969804320305777.

Mahmoud M. R. Elsawy, Stéphane Lanteri, Régis Duvigneau, Gauthier Brière, Mohamed Sabry Mohamed, and Patrice Genevet. Global optimization of metasurface designs using statistical learning methods. *Scientific Reports*, 9(1), December 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-53878-9. URL http://www.nature.com/articles/s41598-019-53878-9.

Michael Emmerich, A Zhang, R Li, I Flesch, and Peter J. Lucas. Mixed-integer bayesian optimization utilizing a-priori knowledge on parameter dependences. *Journal of Physical Chemistry A - J PHYS CHEM A*, pages 65–72, 01 2008.

Peter I. Frazier. A Tutorial on Bayesian Optimization. *arXiv e-prints*, page arXiv:1807.02811, July 2018.

John Geweke. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian Statistics*, 4, 11 1995.

Tadesse Ghirmai. Applying metropolis-hastings-within-gibbs algorithms for data detection in relay-based communication systems. In *2015 IEEE Signal Processing and Signal Processing Education Workshop (SP/SPE)*, pages 167–171, 2015. doi: 10.1109/DSP-SPE.2015.7369547.

Geof Givens and Jennifer Hoeting. *Computational Statistics: Second Edition*, volume 710. 02 2005. ISBN 0471461245. doi: 10.1002/9781118555552.

Lei Gong and James M. Flegal. A practical sequential stopping rule for high-dimensional markov chain monte carlo. *Journal of Computational and Graphical Statistics*, 25(3): 684–700, 2016. doi: 10.1080/10618600.2015.1044092.

Nicolas Guillot. *Gamma ray quantification by equivalent numerical modelling*. Theses, Université Blaise Pascal - Clermont-Ferrand II, 2015. URL https://tel.archives-ouvertes.fr/tel-01247249.

Nikolaus Hansen. *The CMA Evolution Strategy: A Comparing Review*, pages 75–102. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-32494-2. doi: 10.1007/3-540-32494-1_4. URL https://doi.org/10.1007/3-540-32494-1_4.

W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97–109, 1970.

P. Heidelberger and P. Welch. Simulation run length control in the presence of an initial transient. *Oper. Res.*, 31:1109–1144, 1983.

Magnus R Hestenes. Multiplier and gradient methods. *Journal of optimization theory and applications*, 4(5):303–320, 1969.

Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *International Conference on Learning and Intelligent Optimization*, pages 507–523. Springer, 2011.

Jérôme Idier and Laure Blanc-Féraud. *Deconvolution of Images*, chapter 6, pages 141–167. John Wiley & Sons, Ltd, 2008. ISBN 9780470611197. doi: https://doi.org/10.1002/9780470611197.ch6. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470611197.ch6.

M.E. Johnson, L.M. Moore, and D. Ylvisaker. Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, 26(2):131–148, 1990. ISSN 0378-3758. doi: https://doi.org/10.1016/0378-3758(90)90122-B. URL https://www.sciencedirect.com/science/article/pii/037837589090122B.

Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, Dec 1998. ISSN 1573-2916. doi: 10.1023/A:1008306431147. URL https://doi.org/10.1023/A:1008306431147.

H Kunze, D La Torre, and M Ruiz Galán. Optimization methods in inverse problems and applications to science and engineering. *Optimization and Engineering*, 22(4):2151–2158, December 2021.

Neil D. Lawrence. Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data. In *Proceedings of the 16th International Conference on Neural Information Processing Systems*, NIPS'03, pages 329–336, Cambridge, MA, USA, 2003. MIT Press. URL http://dl.acm.org/citation.cfm?id=2981345.2981387. event-place: Whistler, British Columbia, Canada.

Rodolphe Le Riche and Frédéric Guyon. Dual evolutionary optimization. *Lecture Notes in Computer Science*, (2310):281–294, 2002. selected papers of the 5th Int. Conf. Evolution Artificielle.

Rodolphe Le Riche and Victor Picheny. Revisiting bayesian optimization in the light of the coco benchmark. *Structural and MultiDisciplinary Optimization*, 2021. to appear.

Rui Li, Michael TM Emmerich, Jeroen Eggermont, Thomas Bäck, Martin Schütz, Jouke Dijkstra, and Johan HC Reiber. Mixed integer evolution strategies for parameter optimization. *Evolutionary computation*, 21(1):29–64, 2013a.

Rui Li, Michael T.M. Emmerich, Jeroen Eggermont, Thomas Bäck, M. Schütz, J. Dijkstra, and J.H.C. Reiber. Mixed Integer Evolution Strategies for Parameter Optimization. *Evolutionary Computation*, 21(1):29–64, 2013b. ISSN 1063-6560, 1530-9304. doi: 10.1162/EVCO_a_00059. URL http://www.mitpressjournals.org/doi/10.1162/EVCO_a_00059.

Ying Lin, Yu Liu, Wei-Neng Chen, and Jun Zhang. A hybrid differential evolution algorithm for mixed-variable optimization problems. *Information Sciences*, 466:170–188, 2018. ISSN 00200255. doi: 10.1016/j.ins.2018.07.035. URL https://linkinghub.elsevier.com/retrieve/pii/S0020025516318163.

Shuyang Ling. *Bilinear Inverse Problems: Theory, Algorithms, and Applications*. PhD thesis, 2017. URL https://www.proquest.com/dissertations-theses/bilinear-inverse-problems-theory-algorithms/docview/1949289070/se-2?accountid=28541. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Dernière mise à jour - 2021-05-14.

M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979. ISSN 00401706. URL http://www.jstor.org/stable/1268522.

M. Minoux. *Mathematical Programming: Theory and Algorithms*. A Wiley-Interscience publication. Wiley, 1986. ISBN 9780471901709. URL https://books.google.fr/books?id=5kDvAAAAMAAJ. translated by Vajda, S.

Jonas Mockus, Vytautas Tiesis, and Antanas Zilinskas. The application of bayesian methods for seeking the extremum. *Towards global optimization*, 2(117-129):2, 1978.

Kevin P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012. ISBN 0262018020. doi: 10.5555/2380985.

Marcelo Máduar and Pedro Miranda Junior. Gamma spectrometry in the determination of radionuclides comprised in radioactive series. 01 2007.

J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, January 1965. ISSN 0010-4620. doi: 10.1093/comjnl/7.4.308. URL https://doi.org/10.1093/comjnl/7.4.308.

Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer series in operations research. Springer, New York, 2nd ed edition, 2006. ISBN 978-0-387-30303-1. OCLC: ocm68629100.

Jiff Ocenasek and Josef Schwarz. Estimation of distribution algorithm for mixed continuous-discrete optimization problems. In *2nd Euro-International Symposium on Computational Intelligence*, pages 227–232. IOS Press Kosice, Slovakia, 2002.

Chulhee Park and Moon Kang. Color restoration of rgbn multispectral filter array sensor images based on spectral decomposition. *Sensors*, 16:719, 05 2016. doi: 10.3390/s16050719.

Julien Pelamatti, Loïc Brevault, Mathieu Balesdent, El-Ghazali Talbi, and Yannick Guerin. Efficient global optimization of constrained mixed variable problems. *Journal of Global Optimization*, 73(3):583–613, 2019. ISSN 0925-5001, 1573-2916. doi: 10.1007/s10898-018-0715-1. URL http://link.springer.com/10.1007/s10898-018-0715-1.

Victor Picheny, Robert B Gramacy, Stefan Wild, and Sebastien Le Digabel. Bayesian optimization under mixed constraints with a slack-variable augmented Lagrangian. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL https://proceedings.neurips.cc/paper/2016/file/31839b036f63806cba3f47b93af8ccb5-Paper.pdf.

M. J. D. Powell. *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*, pages 51–67. Springer Netherlands, Dordrecht, 1994. ISBN 978-94-015-8330-5. doi: 10.1007/978-94-015-8330-5_4. URL https://doi.org/10.1007/978-94-015-8330-5_4.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 2006. ISBN 978-0-262-18253-9. OCLC: ocm61285753.

Christian P. Robert and Wu Changye. Markov chain monte carlo methods, a survey with some frequent misunderstandings, 2020.

Gareth O. Roberts and Jeffrey S. Rosenthal. Examples of adaptive mcmc. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009. doi: 10.1198/jcgs.2009.06134. URL https://doi.org/10.1198/jcgs.2009.06134.

R. Tyrrell Rockafellar. Lagrange Multipliers and Optimality. *SIAM Review*, 35(2):183–238, 1993. URL http://www.jstor.org/stable/2133143.

Olivier Roustant, Espéran Padonou, Yves Deville, Aloïs Clément, Guillaume Perrin, Jean Giorla, and Henry Wynn. Group kernels for gaussian process metamodels with categorical inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):775–806, 2020. doi: 10.1137/18M1209386. URL https://doi.org/10.1137/18M1209386.

Adam Slowik and Halina Kwasnicka. Evolutionary algorithms and their applications to engineering problems. *Neural Computing and Applications*, March 2020. ISSN 0941-0643, 1433-3058. doi: 10.1007/s00521-020-04832-8. URL http://link.springer.com/10.1007/s00521-020-04832-8.

A. M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010. ISSN 0962-4929, 1474-0508. doi: 10.1017/S0962492910000061. URL https://www.cambridge.org/core/product/identifier/S0962492910000061/type/journal_article.

Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, USA, 2004. ISBN 0898715725. doi: 10.5555/1062404.

Yoel Tenne. Initial sampling methods in metamodel-assisted optimization. *Engineering with Computers*, 31(4):661–680, October 2015. ISSN 0177-0667, 1435-5663. doi: 10.1007/s00366-014-0372-z. URL http://link.springer.com/10.1007/s00366-014-0372-z.

H.A.L. Thi, H.M. Le, and T.P. Dinh. *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*. Advances in Intelligent Systems and Computing. Springer International Publishing, 2019. ISBN 9783030218034. URL https://books.google.fr/books?id=R46dDwAAQBAJ.

Robert Tibshirani, Jerome Friedman, and Trevor Hastie. *The Elements of Statistical Learning*. Springer Texts in Statistics. Springer, 2 edition, 2009. ISBN 0-387-84884-3 978-0-387-84884-6. URL https://web.stanford.edu/~hastie/ElemStatLearn/.

Emmanuel Vazquez and Julien Bect. Convergence properties of the expected improvement algorithm with fixed mean and covariance functions. *Journal of Statistical Planning and Inference*, 140(11):3088–3095, 2010. ISSN 03783758. doi: 10.1016/j.jspi.2010.04.018. URL https://linkinghub.elsevier.com/retrieve/pii/S0378375810001850.

Felipe A. C. Viana, Gerhard Venter, and Vladimir Balabanov. An algorithm for fast optimal Latin hypercube design of experiments: AN ALGORITHM FOR FAST OPTIMAL LHD. *International Journal for Numerical Methods in Engineering*, 82(2):135–156, April 2010. ISSN 00295981. doi: 10.1002/nme.2750. URL http://doi.wiley.com/10.1002/nme.2750.

Ziyu Wang, Frank Hutter, Masrour Zoghi, David Matheson, and Nando de Feitas. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.

James Wilson, Frank Hutter, and Marc Deisenroth. Maximizing acquisition functions for Bayesian optimization. *Nips*, page 12, 2018.

Nan Ye, Farbod Roosta-Khorasani, and Tiangang Cui. *Optimization Methods for Inverse Problems*, pages 121–140. Springer International Publishing, Cham, 2019. ISBN 978-3-030-04161-8. doi: 10.1007/978-3-030-04161-8_9. URL https://doi.org/10.1007/978-3-030-04161-8_9.

Martin Zaefferer. CEGO. https://cran.r-project.org/package=CEGO, 2014–2021.

Yichi Zhang, Siyu Tao, Wei Chen, and Daniel W. Apley. A Latent Variable Approach to Gaussian Process Modeling with Qualitative and Quantitative Factors. *Technometrics*, pages 1–12, 2019. ISSN 0040-1706, 1537-2723. doi: 10.1080/00401706.2019.1638834. URL https://www.tandfonline.com/doi/full/10.1080/00401706.2019.1638834.

Yichi Zhang, Daniel W. Apley, and Wei Chen. Bayesian Optimization for Materials Design with Mixed Quantitative and Qualitative Variables. *Scientific Reports*, 10 (1), December 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-60652-9. URL http://www.nature.com/articles/s41598-020-60652-9.

Miguel Munoz Zuniga and Delphine Sinoquet. Global optimization for mixed categorical-continuous variables based on gaussian process models with a randomized categorical space exploration step. *INFOR: Information Systems and Operational Research*, 58 (2):310–341, 2020. doi: 10.1080/03155986.2020.1730677. URL https://doi.org/10.1080/03155986.2020.1730677.

École Nationale Supérieure des Mines
de Saint-Étienne

NNT: 2022LYSEM010

Jhouben Janyk CUESTA RAMIREZ

Optimization of a computationally expensive simulator with quantitative and qualitative inputs.

Speciality: Applied Mathematics

Keywords: Mixed variable optimization, expensive optimization problem, Bayesian ptimization, mixed meta-model, Gaussian process, latent variables, augmented Lagrangian.

Abstract: In this thesis, costly mixed problems are approached through Gaussian processes where the discrete variables are relaxed into continuous latent variables. The continuous space is more easily harvested by classical Bayesian optimization techniques than a mixed space would. Discrete variables are recovered either subsequently to the continuous optimization, or simultaneously with an additional continuous-discrete compatibility constraint that is handled with augmented Lagrangians. Several possible implementations of such Bayesian mixed optimizers are compared. In particular, the reformulation of the problem with continuous latent variables is put in competition with searches working directly in the mixed space. Among the algorithms involving latent variables and an augmented Lagrangian, a particular attention is devoted to the Lagrange multipliers for which a local and a global estimation techniques are studied. The comparisons are based on the repeated optimization of three analytical functions and a mechanical application regarding a beam design. An additional study for applying a proposed mixed optimization strategy in the field of mixed self-calibration is made. This analysis was inspired in an application in radionuclide quantification, which defined an specific inverse function that required the study of its multiple properties in the continuous scenario. A proposition of different deterministic and Bayesian strategies was made towards a complete definition in a mixed variable setup.

École Nationale Supérieure des Mines
de Saint-Étienne

NNT : 2022LYSEM010

Jhouben Janyk CUESTA RAMIREZ

Optimisation de codes numériques coûteux en présence de variables quantitatives et qualitatives

Spécialité : Mathématiques Appliquées

Mots clefs : Optimisation en variables mixtes, problème d'optimisation coûteux, optimisation Bayésienne, méta-modèle mixte, processus Gaussien, variables latentes, Lagrangien augmenté..

Résumé :

Dans cette thèse, les problèmes d'optimisation mixtes coûteux sont abordés par le biais de processus gaussiens où les variables discrètes sont relaxées en variables latentes continues. L'espace continu est plus facilement exploité par les techniques classiques d'optimisation bayésienne que ne le serait un espace mixte. Les variables discrètes sont récupérées soit après l'optimisation continue, soit simultanément avec une contrainte supplémentaire de compatibilité continue-discrète qui est traitée avec des Lagrangiens augmentés. Plusieurs implémentations possibles de ces optimiseurs mixtes bayésiens sont comparées. En particulier, la reformulation du problème avec des variables latentes continues est mise en concurrence avec des recherches travaillant directement dans l'espace mixte. Parmi les algorithmes impliquant des variables latentes et un Lagrangien augmenté, une attention particulière est consacrée aux multiplicateurs de Lagrange pour lesquels des techniques d'estimation locale et globale sont étudiées. Les comparaisons sont basées sur l'optimisation répétée de trois fonctions analytiques et sur une application mécanique concernant la conception d'une poutre. Une étude supplémentaire analyse s'inspire d'une application de quantification des radionucléides, qui définit une fonction inverse spécifique nécessitant l'étude de ses multiples propriétés dans un scénario continu. Une proposition de différentes stratégies déterministes et bayésiennes a été faite en vue d'une définition complète dans un contexte de variables mixtes.