



Bayesian framework for solving transform invariant low-rank textures

Shihui Hu, Lei Yu, Menglei Zhang, Chengcheng Lv

► To cite this version:

Shihui Hu, Lei Yu, Menglei Zhang, Chengcheng Lv. Bayesian framework for solving transform invariant low-rank textures. ICIP, 2016, Phinex, United States. pp.3588 - 3592, 10.1109/ICIP.2016.7533028 . hal-01407365

HAL Id: hal-01407365

<https://inria.hal.science/hal-01407365>

Submitted on 2 Dec 2016

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.

BAYESIAN FRAMEWORK FOR SOLVING TRANSFORM INVARIANT LOW-RANK TEXTURES

Shihui Hu, Lei Yu, Menglei Zhang, Chengcheng Lv

School of Electronic and Information, Wuhan University, China

ABSTRACT

As a holistic image feature, Transform Invariant Low-Rank Textures (TILT) can effectively recover the rectification of user-specified patch with a rich class of low-rank textures in 2D images. Unlike conventional local image features, TILT isn't dependent on the extraction of points, corners or edges, which would bring inaccuracy and weak robustness. However, TILT is still rather rudimentary, and have some limitations in applications. In this paper, we proposed a novel algorithm for better solving TILT. Our method is based on the application of Bayesian framework in robust principal component analysis (RPCA), some missing entries of TILT's mathematical model are taken into account as well. Experimental results on both synthetic and real data indicate that our new algorithm outperforms the existing algorithm especially for the case with corruptions and occlusions.

Index Terms— Image feature, transform invariant low-rank textures, rectification, Bayesian framework, corruptions and occlusions

1. INTRODUCTION

Feature extraction in images is one of fundamental problems in computer vision. Most of existing image features, such as SIFT points [1], Harris corners [2] and Canny edges [3], are key to many high level computer vision applications over the past decades. However, the image features mentioned above, called “low level local features”, are always inaccurate and unrobust. Zhang et al. [4] proposed a holistic feature called “Transform Invariant Low-Rank Textures (TILT)” that can recover the rectification of user-specified patch in 2D images. Generally, this type of patch which can be recovered by TILT, has a global structure, such as full of regularity, symmetry, repetition, which can be measured by low-rankness. Hence, TILT has a good performance in accuracy and robustness.

Although TILT can work remarkably well for a very broad range of cases, it is still a new and rudimentary image feature. The original TILT paper [4] presented some examples of TILT where it failed, as shown in Fig.1. The first case shows that two incompatible dominant low rank structures (the facade and the shadow) overlapped result in an overall high rank region. The second case shows that the combined region containing



Fig. 1. Failure cases of TILT. *Left column* The original images, the user-specified patch specified by red rectangles, and the transforms found by TILT specified by green rectangles. *Right column* Rectified regions of interest.

two adjacent low rank regions each of which is distorted differently, might no longer be low rank. The third case shows that too many occlusions may lead to TILT's failure.

In this paper, the third case is considered. Although TILT is designed to be robust to corruptions and occlusions, it's important to note that an assumption is essentially necessary that the amount of corruptions and occlusions can not be very large. It inspires us to try to find an algorithm which is more robust to handle more complex situations in natural scenes.

Mathematically, TILT is modeled (approximately) as the robust principal component analysis (RPCA) problem presented in [5], it needs to find a rectified patch $\mathbf{I} \circ \tau$, which can be decomposed into a low-rank component \mathbf{I}^0 and a sparse component \mathbf{E} , namely $\mathbf{I} \circ \tau = \mathbf{I}^0 + \mathbf{E}$, meanwhile gets the appropriate transform τ . The existing solver for TILT is based on the alternating direction method (ADM) [6], one of critical

drawbacks of the existing solver is the simply modified ADM presented in [4] is not always guaranteed to converge to the optimal solution. The convergence of ADM has been well-studied and established under the case of *two variables*, while the inner loop of TILT problem has *three variables*. Ren et al. [7] proposed a improved algorithm for solving TILT problem, with a theoretical guarantee on the convergence of its inner loop. But another critical drawback of the existing solver, the weak robustness to large amount of corruptions and occlusions, has not been resolved.

In contrast with the convex optimization methods mentioned above, Bayesian method has a good performance in robustness to noises and perturbations. In this paper, based on the application of Bayesian framework in robust principal component analysis (RPCA) [8] [9], we propose a novel algorithm for solving TILT problem. We rewrite the mathematical model of TILT as $\mathbf{I} \circ \tau = \mathbf{I}^0 + \mathbf{E} + \mathbf{N}$, where \mathbf{N} is the measurement noise which exists everywhere and usually cannot be ignored in practice. Theoretically speaking, our algorithm outperforms the existing TILT solver in robustness to corruptions and occlusions. Meanwhile, it has been proved that Bayesian approach to sparsity (or low-rank) problems has fewer local minimizations [10]. In theory, our algorithm is with a guarantee of convergence to the optimal solution.

2. TRANSFORM INVARIANT LOW-RANK TEXTURES

In this section, we briefly review the mathematical model of TILT, as well as some implementation details.

2.1. Mathematical Model

Consider a 2D texture $\mathbf{I}^0 \in \mathbb{R}^{m' \times n'}$. If $r = \text{rank}(\mathbf{I}^0)$, with $r \ll \min(m', n')$, we can say \mathbf{I}^0 is a *low-rank texture*. In most real images, we almost can't find a perfectly low-rank texture, primarily because usually it undergoes a deformation and suffers from many types of corruptions or occlusions. As a result, a deformed texture \mathbf{I} with a proper inverse transform τ and then removed the corruption or occlusion \mathbf{E} , will be low-rank. So the mathematical model for TILT presented in [4] is:

$$\min_{\mathbf{I}^0, \mathbf{E}, \tau} \|\mathbf{I}^0\|_* + \zeta \|\mathbf{E}\|_1 \quad \text{s.t.} \quad \mathbf{I} \circ \tau = \mathbf{I}^0 + \mathbf{E}, \quad (1)$$

where $\|\cdot\|_*$ the nuclear norm to promote the low-rank property, $\|\cdot\|_1$ the ℓ_1 -norm to promote the sparsity, $\zeta > 0$ is a parameter to balance the two terms, $\tau: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ belongs to a certain group of transforms, e.g. affine transforms, perspective transforms.

Observing the objective function in (1) is convex, the constraint $\mathbf{I} \circ \tau = \mathbf{I}^0 + \mathbf{E}$ is nonlinear in τ , hence the problem is not convex. To linearize the constraint [11], the linear ap-

proximation can be obtained by Taylor expansion.

$$\min_{\mathbf{I}^0, \mathbf{E}, \tau} \|\mathbf{I}^0\|_* + \zeta \|\mathbf{E}\|_1 \quad \text{s.t.} \quad \mathbf{I} \circ \tau + \nabla \mathbf{I} \Delta \tau = \mathbf{I}^0 + \mathbf{E} \quad (2)$$

where $\nabla \mathbf{I}$ is the Jacobian matrix (derivatives of the image with respect to the transformation τ), the high-order terms $O(\nabla \mathbf{I} \Delta \tau)$ of Taylor expansion is ignored in TILT model.

2.2. Ambiguity of Low-Rank Textures Definition

By this definition of low-rank texture, obviously its rank is invariant under any scaling and translation transformations:

$$\text{rank}(\mathbf{I}^0(x, y)) = \text{rank}(c\mathbf{I}^0(ax + t_1, by + t_2)) \quad (3)$$

In [4], this ambiguity can be easily resolved by imposing additional constraints on the texture. To eliminate the ambiguity in translation, we fix the center x_0 of the initial patch after the transformation. That is, $\Delta \tau$ should be enforced a set of linear constraints as following:

$$A_t \Delta \tau = 0 \quad (4)$$

To eliminate the ambiguity in scaling, we enforce that the area and the ratio of edge length remain constant after the transformation. Notice that these equalities enforce non-linear constraints on the transformation τ , we linearize these non-linear constraints with respect to the transformation τ , just as (2). Then the another set of linear constraints is:

$$A_s \Delta \tau = 0 \quad (5)$$

By combining the two sets of linear constraints in (4)(5), we obtain the complete form of the constraints of transformation τ :

$$S \cdot \Delta \tau = 0 \quad (6)$$

3. BAYESIAN TILT FRAMEWORK

In this paper, a hierarchical Bayesian model will be employed to TILT problem, where all unknown quantities are treated as stochastic models. Then we obtain all expectations via variational Bayesian approach [12], where the posteriors are approximated by maximizing lower-bound of marginal likelihood.

As mentioned above, the transformed texture $\mathbf{I} \circ \tau \in \mathbb{R}^{m \times n}$, is the superposition of three parts: low-rank component $\mathbf{I}^0 \in \mathbb{R}^{m \times n}$, sparse component $\mathbf{E} \in \mathbb{R}^{m \times n}$, and measurement noise $\mathbf{N} \in \mathbb{R}^{m \times n}$ (in bayesian TILT model, we take the measurement noise \mathbf{N} into consideration). The mathematical model (2) of TILT can be rewritten as

$$\mathbf{Y} = \mathbf{I}^0 + \mathbf{E} + \Phi \Delta \tau + \mathbf{N} \quad (7)$$

where $\mathbf{Y} = \mathbf{I} \circ \tau$ with $\mathbf{Y} \in \mathbb{R}^{m \times n}$, $\Phi = -\nabla \mathbf{I}$ with $\Phi \Delta \tau \in \mathbb{R}^{m \times n}$, \mathbf{N} with $\mathbf{N} \in \mathbb{R}^{m \times n}$ contains measurement noise and the taylor series high-order terms $O(\nabla \mathbf{I} \Delta \tau)$ in (2), based on an assumption that $O(\nabla \mathbf{I} \Delta \tau)$ is stochastic quantity.

3.1. Hierarchical Bayesian Model

Suppose the low-rank component \mathbf{I}^0 with $r = \text{rank}(\mathbf{I}^0)$, it can be decomposed into column space $\mathbf{A} \in \mathbb{R}^{m \times r}$ and row space $\mathbf{B} \in \mathbb{R}^{n \times r}$,

$$\mathbf{I}^0 = \mathbf{AB}^T \quad (8)$$

each column of \mathbf{A} and \mathbf{B} is assumed to obey the Gamma-Gaussian model to introduce the column sparsity, i.e.

$$A_{\cdot i} \sim \mathcal{N}(0, \gamma_i^{-1}), \quad B_{\cdot i} \sim \mathcal{N}(0, \gamma_i^{-1}) \quad (9)$$

$$\gamma_i \sim \text{Gamma}(a_\gamma, b_\gamma) \quad (10)$$

with hyperparameters $a_\gamma, b_\gamma > 0$. when a_γ, b_γ are close to 0 will result in a low-rank promoting prior for the expectation of \mathbf{I}^0 .

The sparse component \mathbf{E} is injected the sparse constraint by employing independent Gamma-Gaussian model for each coefficient E_{ij} of the matrix \mathbf{E} , i.e.

$$E_{ij} \sim \mathcal{N}(0, \alpha_{ij}) \quad (11)$$

$$\alpha_{ij} \sim \text{Gamma}(a_\alpha, b_\alpha), \forall i, j \quad (12)$$

with hyperparameters $a_\alpha, b_\alpha > 0$.

Considering that we have imposed a set of linear constraints on transformation τ , the entry $S \cdot \Delta\tau$ in (6) is expected to be zero. In order to integrate $\Delta\tau$ into the whole bayesian framework, we treat $S \cdot \Delta\tau$ as a sparse variable, with obeying a hierarchical Gamma-Gaussian model:

$$S \cdot \Delta\tau \sim \mathcal{N}(0, \lambda^{-1}) \quad (13)$$

$$\lambda \sim \text{Gamma}(a_\lambda, b_\lambda) \quad (14)$$

hyperparameters $a_\lambda > 0$ and $b_\lambda > 0$.

Using the model proposed in (7), we follow the standard assumption that noise component \mathbf{N} is drawn from a Gaussian distribution, then incorporate the noise \mathbf{N} in the observations, that is:

$$\mathbf{Y} \sim \mathcal{N}(\mathbf{AB}^T + \mathbf{E} + \Phi\Delta\tau, \beta^{-1}\mathbf{I}_{mn}) \quad (15)$$

$$p(\beta) = \beta^{-1} \quad (16)$$

with β the noise precision, assigned with the Jeffrey's prior.

3.2. Variational Bayesian Approach

In bayesian framework, all variables of interest can be estimated by variational Bayesian approach after simple inference. The expectation of each variable is presented here as follows:

The update for \mathbf{A} and \mathbf{B} :

$$\langle \mathbf{A} \rangle = \langle \beta \rangle \langle \mathbf{Y} - \Phi\Delta\tau - \mathbf{E} \rangle \langle \mathbf{B} \rangle \Sigma^{\mathbf{A}}, \quad (17)$$

$$\langle \mathbf{B} \rangle = \langle \beta \rangle \langle \mathbf{Y} - \Phi\Delta\tau - \mathbf{E} \rangle \langle \mathbf{A} \rangle \Sigma^{\mathbf{B}} \quad (18)$$

with covariance:

$$\begin{aligned} \Sigma^{\mathbf{A}} &= (\langle \beta \rangle \langle \mathbf{B}^T \mathbf{B} \rangle + \text{diag}(\gamma))^{-1}, \\ \Sigma^{\mathbf{B}} &= (\langle \beta \rangle \langle \mathbf{A}^T \mathbf{A} \rangle + \text{diag}(\gamma))^{-1}. \end{aligned}$$

and the required expectations: $\langle \mathbf{A}^T \mathbf{A} \rangle = \langle \mathbf{A} \rangle^T \langle \mathbf{A} \rangle + m\Sigma^{\mathbf{A}}$, $\langle \mathbf{B}^T \mathbf{B} \rangle = \langle \mathbf{B} \rangle^T \langle \mathbf{B} \rangle + n\Sigma^{\mathbf{B}}$.

The update for \mathbf{E} :

$$\langle E_{ij} \rangle = \langle \beta \rangle \Sigma_{ij}^E (Y_{ij} - \langle A_{\cdot i} \rangle \langle B_{\cdot i} \rangle^T - \Phi_{ij} \langle \Delta\tau \rangle), \quad (19)$$

with covariance:

$$\Sigma_{ij}^E = \frac{1}{\langle \beta \rangle + \langle \alpha_{ij} \rangle}.$$

The update for $\Delta\tau$:

$$\langle \Delta\tau \rangle = \Sigma^{\Delta\tau} \langle \beta \rangle \Phi_{mn \times d}^T (\mathbf{Y} - \mathbf{AB}^T - \mathbf{E})_{mn \times 1}, \quad (20)$$

with covariance:

$$\Sigma^{\Delta\tau} = (\langle \beta \rangle \Phi_{mn \times d}^T \Phi_{mn \times d} + \lambda S^T S)^{-1}$$

where d is dimension of the transformation τ , the subscripts $(\cdot)_{mn \times d}$ and $(\cdot)_{mn \times 1}$ denote that the matrix (\cdot) is reshaped into a mn -by- d matrix or a mn -by-1 matrix.

The update for γ :

$$\langle \gamma_i \rangle = \frac{2a_\gamma + m + n}{2b_\gamma + \langle A_{\cdot i}^T A_{\cdot i} \rangle + \langle B_{\cdot i}^T B_{\cdot i} \rangle} \quad (21)$$

with the required expectations: $\langle A_{\cdot i}^T A_{\cdot i} \rangle = \langle A_{\cdot i} \rangle^T \langle A_{\cdot i} \rangle + m(\Sigma^{\mathbf{A}})_{ii}$, $\langle B_{\cdot i}^T B_{\cdot i} \rangle = \langle B_{\cdot i} \rangle^T \langle B_{\cdot i} \rangle + n(\Sigma^{\mathbf{B}})_{ii}$.

The update for α :

$$\langle \alpha_{ij} \rangle = \frac{2a_\alpha + 1}{2b_\alpha + \langle E_{ij}^2 \rangle} \quad (22)$$

with the required expectation: $\langle E_{ij}^2 \rangle = \langle E_{ij} \rangle^2 + \Sigma_{ij}^E$.

The update for λ :

$$\langle \lambda \rangle = \frac{2a_\lambda + 1}{2b_\lambda + \langle (S \cdot \Delta\tau)^2 \rangle} \quad (23)$$

with the required expectation: $\langle (S \cdot \Delta\tau)^2 \rangle = \langle S \cdot \Delta\tau \rangle^2 + \lambda^{-1}$.

The update for β :

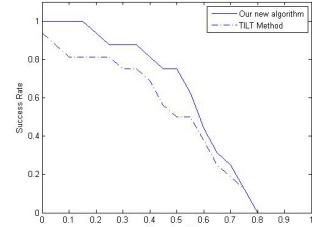
$$\langle \beta \rangle = \frac{mn}{\| \mathbf{Y} - \mathbf{AB}^T - \mathbf{E} - \Phi\Delta\tau \|_F^2} \quad (24)$$

with the required expectation:

$$\begin{aligned} \langle \| \mathbf{Y} - \mathbf{AB}^T - \mathbf{E} - \Phi\Delta\tau \|_F^2 \rangle &= \\ &\| \mathbf{Y} - \langle \mathbf{A} \rangle \langle \mathbf{B} \rangle^T - \langle \mathbf{E} \rangle - \Phi \langle \Delta\tau \rangle \|_F^2 + \sum_{i=1}^m \sum_{j=1}^n \Sigma_{ij}^E \\ &+ \text{Tr}(n \langle \mathbf{A} \rangle^T \langle \mathbf{A} \rangle \Sigma^{\mathbf{B}}) + \text{Tr}(m \langle \mathbf{B} \rangle^T \langle \mathbf{B} \rangle \Sigma^{\mathbf{A}}) \\ &+ \text{Tr}(mn \Sigma^{\mathbf{A}} \Sigma^{\mathbf{B}}) + \sum_{i=1}^m \sum_{j=1}^n \text{Tr}(\Phi_{ij}^T \Phi_{ij} \Sigma^{\Delta\tau}) \end{aligned}$$



(a)



(b)

Fig. 2. Robustness test of TILT method and our new algorithm on various low-rank patterns. (a) low-rank patterns (b) Success rate of TILT method and our new algorithm under different levels of corruption.

4. EXPERIMENTS

In this section, we conduct several experiments with both artificial synthetic images and real images to demonstrate the performance of our algorithm, the original TILT method proposed in [4] is presented here for comparison.

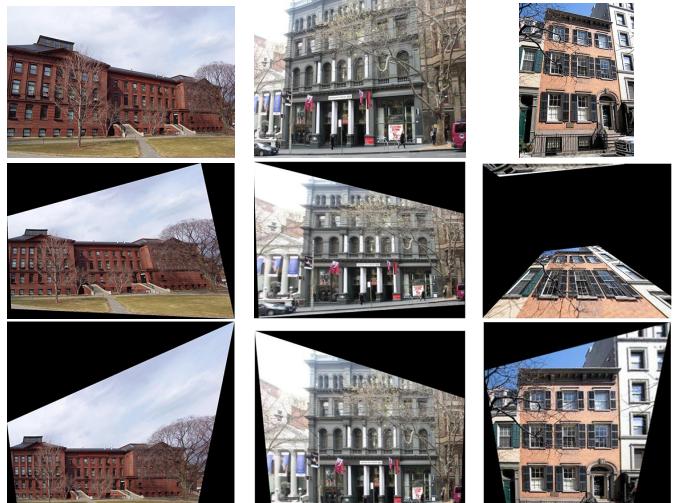
4.1. Synthetic Images

First, we use artificial synthetic images to demonstrate the superiority of our algorithm. Following the same setting in [4], we deform some typically low-rank patterns by an affine transformation defined by the form: $y = Ax + b$, where $x, y \in \mathbb{R}^2$. The affine matrix A is parameterized as $A(\theta, t) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \times \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$, where θ is the rotation angle and t is the skew value. We introduce a small deformation (say rotation by 12° and skew by 0.2), and examine whether the two methods can converge to the correct transform under different levels of random corruption.

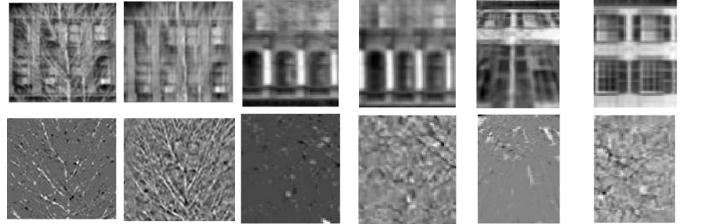
The comparison between two methods is shown in Fig.2(b). It is obvious that our algorithm outperforms TILT method when the percentage of corruption is less than 75%, under 75% or larger corruption the two methods have the same performance because of the over-corruption. The contrast experiments indicate that our algorithm is more robust than TILT method to corruption.

4.2. Real images

Now, we move on to discuss the experimental results on real images to verify the conclusion in the last subsection. Fig.3(a) shows some typical cases. When a regular patch of building is hidden from a tree or some other things, TILT method is not so robust to noise that recovers the correct transform from the deformed and corrupted image. Fig.3(b) shows the low-rank components and sparse error components of each experiment, it is worth highlighting that low-rank components in TILT method have a lot of “tree’s pixels” which should be in-



(a) Original images and Rectified images



(b) Comparisons for low-rank components and sparse components

Fig. 3. Largely corrupted cases to both TILT method and our new algorithm. (a) First row The Original images. Second row Rectified results by TILT. Third row Rectified results by our algorithm. (b) First row Low-rank components. Second row Sparse components. First, third and fifth columns Results by TILT. Second, fourth and sixth columns Results by our algorithm

corporated into sparse components while our new algorithm provides much better results of separation between low-rank components and sparse components. The impure low-rank component leads to a result that TILT method can not converge to a lower-rank solution, then obtains a relatively wrong transformation.

5. CONCLUSIONS

In this paper, based on bayesian framework, we propose a novel algorithm for solving TILT problem, some missing entries are considered in our model as well. Two critical drawbacks of the existing method of TILT, convergence can't be guaranteed and weak robustness to large corruptions and occlusions, are settled well in our algorithm. Experimental results show that our algorithm significantly outperforms the existing methods, both artificial synthetic images and real images are utilized in experiments.

6. REFERENCES

- [1] David G Lowe, “Distinctive image features from scale invariant keypoints,” *Intl Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.
- [2] Chris Harris and Mike Stephens, “A Combined Corner and Edge Detector,” *Proceedings of the Alvey Vision Conference 1988*, pp. 147–151, 1988.
- [3] John Canny, “A Computational Approach to Edge Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [4] Zhengdong Zhang, Arvind Ganesh, Xiao Liang, and Yi Ma, “TILT: Transform Invariant Low-Rank Textures,” *International Journal of Computer Vision*, vol. 99, no. 1, pp. 1–24, 2012.
- [5] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright, “Robust Principal Component Analysis?,” *CoRR*, vol. abs/0912.3, 2009.
- [6] Zhouchen Lin, Minming Chen, and Yi Ma, “The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices,” *arXiv:1009.5055*, p. 23, 2013.
- [7] Xiang Ren and Zhouchen Lin, “Linearized alternating direction method with adaptive penalty and warm starts for fast solving transform invariant low-rank textures,” *International Journal of Computer Vision*, vol. 104, pp. 1–14, 2013.
- [8] Xinghao Ding, Lihan He, and Lawrence Carin, “Bayesian robust principal component analysis,” *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3419–3430, 2011.
- [9] S. Derin Babacan, Martin Luessi, Rafael Molina, and Aggelos K. Katsaggelos, “Sparse bayesian methods for low-rank matrix estimation,” *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 3964–3977, 2012.
- [10] David Wipf, “Non-convex rank minimization via an empirical bayesian approach,” *arXiv preprint arXiv:1207.2440*, , no. 3, 2012.
- [11] Dimitri P Bertsekas, “Nonlinear programming,” 1999.
- [12] Matthew J Beal, *Variational algorithms for approximate bayesian inference*, Ph.D. thesis, University of London, London, May 2003.