Bayesian parameter estimation of ligament properties based on tibio-femoral kinematics during squatting

Laura Bartsoen[1], Matthias G.R. Faes[2], Michael Skipper Andersen[3], Roel Wirix-Speetjens[4], David Moens[1], Ilse Jonkers[5], and Jos Vander Sloten[1]

[1]*KU Leuven, Mechanical engineering, Leuven, Belgium*
[2]*TU Dortmund University, Chair for Reliability Engineering, Dortmund, Germany*
[3]*Aalborg university, Department of Materials and Production, Aalborg, Denmark*
[4]*Materialise N.V., Leuven, Belgium*
[5]*KU Leuven, Movement science department, Leuven, Belgium*

Abstract

The objective of this study is to estimate the, probably correlated, ligament material properties and attachment sites in a highly non-linear, musculoskeletal knee model based on kinematic data of a knee rig experiment for seven specific specimens. Bayesian parameter estimation is used to account for uncertainty in the limited experimental data by optimization of a high dimensional input parameter space (50 parameters) consistent with all probable solutions. The set of solutions accounts for physiologically relevant ligament strain ($\epsilon < 6\%$). The transitional Markov Chain Monte Carlo algorithm was used. Alterations to the algorithm were introduced in order to avoid premature convergence. To perform the parameter estimation with feasible computational cost, a surrogate model of the knee model was trained. Results show that there is a large intra- and inter-specimen variability in ligament properties, and that multiple sets of ligament properties fit the experimentally measured tibio-femoral kinematics. Although all parameters were allowed to vary significantly, large interdependence is only found between the reference strain and attachment sites. The large variation between specimens and interdependence between reference strain and attachment sites within one specimen, show the inability to identify a small range of ligament properties representative for the patient population. To limit ligament properties uncertainty in clinical applications, research will need to invest in establishing patient-specific uncertainty ranges and/or accurate in vivo measuring methods of the attachment sites and reference strain and/or alternative (combinations of) movements that would allow identifying a unique solution.

Keywords: Musculoskeletal knee model, ligament properties, Bayesian parameter estimation

1 Introduction

Musculoskeletal models are computational models consisting of rigid bodies (bones) connected by joints that are restrained by ligaments and actuated by muscles. The potential applications of musculoskeletal knee models (MSKM) in a clinical setting are numerous, e.g. pre-operative planning of and surgical navigation during total joint replacement surgery. Such applications require a MSKM that provides an accurate prediction of the model outputs, like kinematics and ligament strains. Similar to most biomechanical models, MSKM however rely on a set of input parameters like ligament material properties and attachment sites. In rigid body musculoskeletal models, the ligaments are typically assigned a material model [1] with a linear elastic, only tensile, behavior with a quadratic toe region that are ideally tuned to the individual patient. This model is given in eq. 1, where F [N] is the force, ϵ is the strain, k [N] is the ligament stiffness and ϵ_l is

an experimentally fitted parameter equal to 0.03. The strain ϵ can be computed through eq. 2, where L [m] is the strand length and L_r [m] and ϵ_r are respectively the strand length and strain at reference position. The reference position is defined with the knee in full extension.

$$F = \begin{cases} 0 & if \quad \epsilon < 0 \\ \frac{1}{4}\frac{k\epsilon^2}{\epsilon_l} & if \quad 0 \leq \epsilon \leq 2\epsilon_l \\ k(\epsilon - \epsilon_l) & if \quad \epsilon > 2\epsilon_l \end{cases} \tag{1}$$

$$\epsilon = \frac{(1+\epsilon_r)L - L_r}{L_r} \tag{2}$$

Measurements of the ligament material properties, linear stiffness k and reference strain ϵ_r, are scarce in the literature. There are some studies that determined the linear stiffness of the knee ligaments through in vitro experiments. Most studies [2], [3] determine the E-modulus. Converting this to the linear stiffness however requires an accurate extraction of the ligament geometry from MRI, which is not feasible due to the limited visibility of the ligaments. Some studies [4]–[9] did measure the linear stiffness K [N/mm]. Note that this is a different definition of the linear stiffness compared to the linear stiffness k [N] used in eq. 1, more specifically K [N/mm] is related to k [N] as given in eq. 3.

$$K = \frac{k}{L_r} \tag{3}$$

Table 1 gives the linear stiffness of the main knee ligaments as found in literature. The main ligaments are the Medial Collateral Ligament (MCL) with a deep and superficial bundle, Lateral Collateral Ligament (LCL), AnteroLateral ligament (ALL), PopliteoFibular ligament (PFL), Posterior Capsule (PC), Anterior Cruciate ligament (ACL) and Posterior Cruciate ligament (PCL). From this table it can be seen that there is a large variability in average stiffness between the different studies, showing a large variance between specimens. One could contradict that this is due to different measurement protocols between studies. The study of LaPrade *et al.* [8] however determined the variation in the tested population, which confirms the large variance between specimens.

Table 1: Linear stiffness K [N/mm] according to literature.

Linear stiffness K	deepMCL	supMCL	LCL	ALL	PFL	ACL	PCL
Sugita *et al.* [4]			58.1		43.6		
Chandrashekar *et al.* [5]						250	
Woo *et al.* [6]							220
Robinson *et al.* [7]	42	80					
LaPrade *et al.* [8]			33.5 ± 13.4		28.6 ± 13.6		
Trent *et al.* [9]		72	61			141	183

As the reference strain ϵ_r is a computational model parameter, in vitro measurements are non-existent. The studies of Blankevoort *et al.* [1] and Beidokhti *et al.* [10] performed an optimization of the ligament material properties (linear stiffness and reference strain) by minimizing the difference in kinematics predicted by their model to the experimentally measured movement. The study of Blankevoort *et al.* performed this optimization for only one specimen. Nevertheless the values found by their study are often used in studies to model the ligament behavior in other specimens [11]–[14]. The study of Beidokhti *et al.* performed this optimization based on laxity tests performed on three cadaver specimens. Even though the amount of

data is limited, the results show that there is a large variation between the specimens. Nevertheless, the number of specimens for which experimental data is available is too limited to determine a range of possible ligament properties for new specimens. Availability of such a range would allow to assess uncertainty caused by ligament properties in applications that require an accurate knee model, like pre-operative planning for Total Knee Arthroplasty (TKA). The studies of Blankevoort *et al.* and Beidokhti *et al.* also do not enforce a physiological ligament strain after optimization, meaning that the optimization does not account for ligament damage or joint instability which could lead to unrealistic combinations of ligament properties. It is known from literature [15], [16] that ligament damage starts occurring from 6% ligament strain. In this case damage means fiber rupture that leads to permanent elongation of the ligament. Furthermore, the studies of Blankevoort *et al.* and Beidokhti *et al.* also do not account for inaccuracies in experimental data and the possibility of multiple solutions.

The above mentioned studies are based on in vitro experiments and use methodologies that are infeasible in clinical settings i.e. in vivo conditions. As a result, often average values representative for a specific population are used in musculoskeletal models - if at all available. There are studies that investigated in vivo estimation of these material properties. The study of Slane *et al.* [17] for example measured the collateral ligament strains through ultrasound. They however came across several challenges, like the limitations of using a 2D imaging modality to characterize 3D motion, and the small size and complex anatomy of the collateral ligaments. The study of Pedersen et al. [18] described the development of an in vivo, non-invasive knee laxity measurement device. By combining different sets of movements actuated with the developed device, it might be possible to identify patient-specific ligament properties [19].

The in vitro studies show a large variability in ligament material properties between specimens, where research has responded by investing in identifying patient-specific ligament properties to increase accuracy of patient-specific knee models. It is however not clear if accurate ligament properties are required for application of the knee model. To be able to assess the uncertainty on model outputs caused by ligament properties, a population range of probable ligament properties has to be identified. This range can be established using inverse uncertainty quantification. To our knowledge, inverse uncertainty quantification on a MSKM has only been performed in the past in the study of Andersen *et al.* [19]. They identified uncertainty in ligament material properties due to experimentally introduced noise in laxity trials. The inverse uncertainty quantification was performed by drawing 20 samples from the experimental data to which uniformly distributed noise was applied. For each sample, the ligament material properties were optimized. A set of only 20 samples as well as the exclusion of uncertainty in the attachment sites might however underestimate the uncertainty in the ligament material properties.

The goal of this study is to investigate if a range of ligament material properties can be established allowing better estimation of uncertainty caused by unknown ligament properties. The developed framework should perform the Bayesian parameter estimation (BPE) with realistic computational cost. This study therefore opts to perform a BPE on a highly non-linear computational model. The model is a MSKM of which several, probably correlated, ligament properties (linear stiffness, reference strain and attachment sites) affect the model outcome (Tibio-Femoral (TF) kinematics and ligament strain). A statistical distribution is estimated on the ligament parameters based on the experimental measurement error of the TF-kinematics extracted from an experimentally performed squat movement, resulting in a high dimensional problem as in total 50 parameters have to be estimated. This allows assessment of the variation in the optimized material parameters between the cadaver specimens. If a small range and/or the existence of correlations between ligament properties can be established, a better estimation of the material properties could be done in a setting where in vivo measurement of the TF-kinematics is infeasible.

The contributions of this study are not limited to knee modeling. To perform the BPE, we apply an algorithm that is currently still being improved in literature on a real world, highly non-linear computational model. This results in a high-dimensional BPE problem, where 50 parameters have to be estimated. We show that Transitional Markov Chain Monte Carlo (TMCMC) as introduced [20] and improved [21] in literature,

cannot be directly applied to the computational knee model. A modification of the algorithm is proposed allowing BPE on the knee model.

2 Materials and methods

In brief, we performed a BPE of the ligament material properties and attachment sites on seven specimens. The ligament parameters were optimized (50 parameters) in order for the TF-kinematics to match the experimentally measured kinematics of a squat motion (Sec. 2.1). The knee model performs a squatting motion from 20° to 120° of flexion. The same squatting motion was experimentally simulated using an in vitro knee simulator [22]. The MSKM is described in detail in Section 2.2. To make the parameter estimation computationally efficient, the analysis was not performed on the knee model directly but on a surrogate model. This method is illustrated in Figure 1. As the surrogate model can be evaluated much faster compared to the MSKM, the parameter estimation process is possible with acceptable computational cost. As a surrogate modeling method an artificial neural network is applied (Sec. 2.3). The surrogate model still requires training, but generally this can be done with considerably less model evaluations compared to the parameter estimation itself. Section 2.4 describes the BPE of the ligament properties. This section closes with the description of the statistical analysis performed on the results of the BPE (Sec. 2.5).

Figure 1: Overview of the designed method for optimization towards the ligament properties.
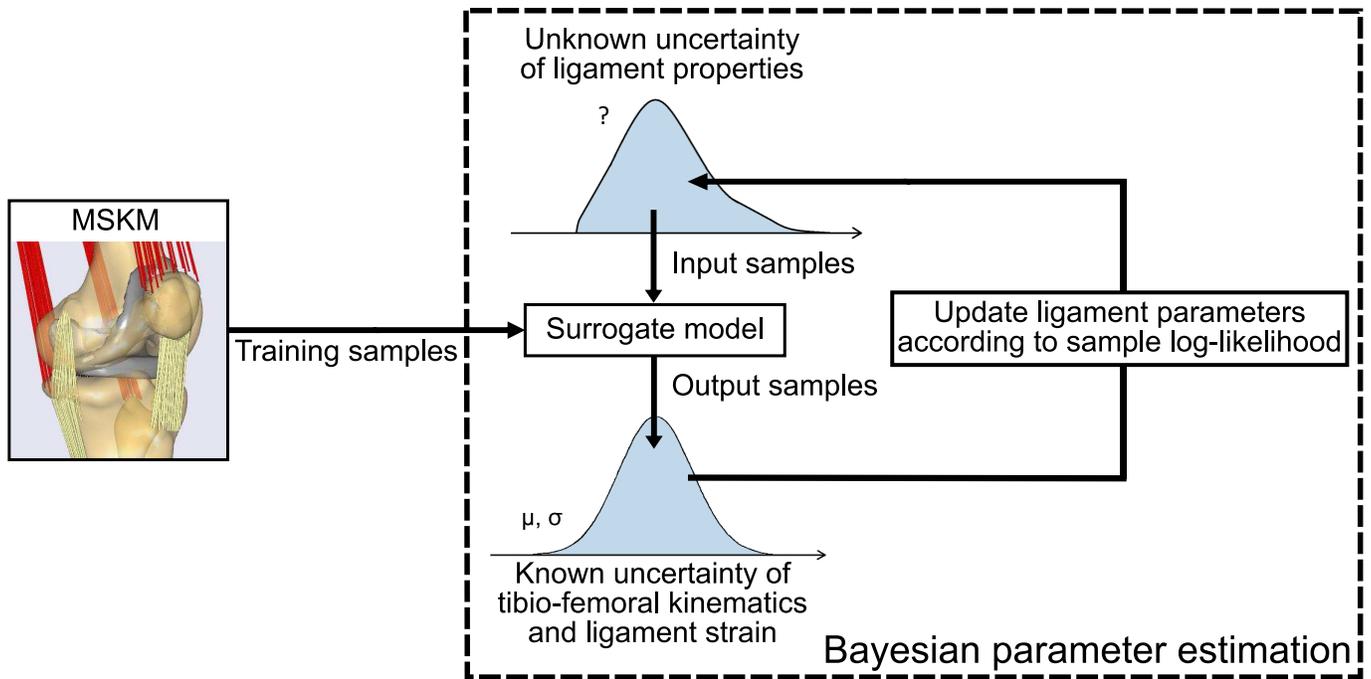

2.1 Experimental kinematics

The experimental kinematics were measured on cadaver specimens. Data from four specimens was adopted from the study of Vanheule *et al.* [11]. Data collection was approved by "Medical ethics committee - Faculty of Medicine - KU Leuven" with number "NH019 2015-05-01". Another three specimens were adopted of which data was collected during the study of Pedersen *et al.* [18]. Data collection was approved by the "Belgian National Council for Bioethics" with number "NH019 2016-10-01-KR09". Both data sets were gathered using the same experimental protocol.

A dynamic knee simulator system similar to the Oxford rig was used [22]. This set-up is designed in such way that it allows 6DOFs for both the tibio-femoral and patello-femoral joints. The force on the quadriceps is varied to keep the vertical force in the ankle at 111N. The force on the hamstrings muscles is kept constant at 50N each. The hip joint is allowed to slide up/down and flex/extend. The ankle joint is allowed to slide medio-laterally and rotate in all directions (FE, VV and IE).

2.2   Knee model

The MSKM is a rigid body model, implemented into the AnyBody modeling system 7.3.0 (AnyBody Technology A/S, Aalborg, Denmark), based on the knee model described in the study of Vanheule *et al.* [11]. The model (fig. 2) simulates a squat motion. The model is validated using three cadaver specimens, on which the squat is applied using a dynamic knee simulator system as described in the study of Vanheule *et al.* [11]. The same DOFs and forces as in the experiment are applied to the knee model. The magnitude of the quadriceps force is applied according to the magnitude measured during the experiment (Sec. 2.1). The flexion/extension (FE) angle is imposed and varies from 20° to 120° of flexion. The secondary kinematics, namely Medial/Lateral (ML), Anterior/Posterior (AP), Proximal/Distal (PD), Varus/Valgus (VV) and Internal/External (IE), are computed using Force Dependent Kinematics (FDK).

Figure 2: The MSKM with indication of the different DOFs.
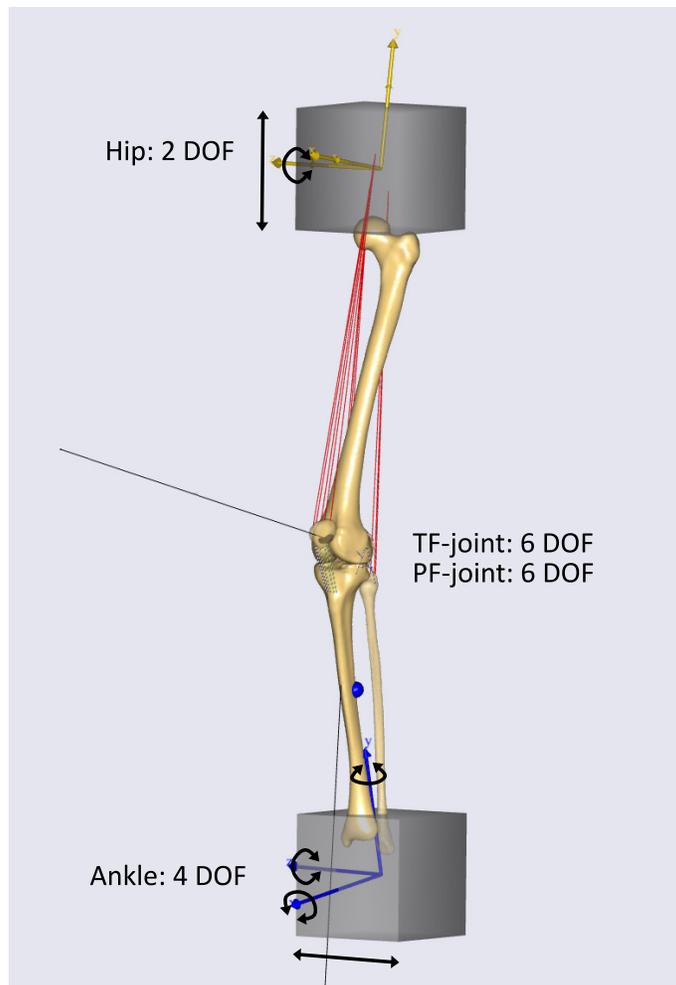
FDK computes forces (e.g. contact, muscle and ligament forces) and secondary kinematics simultaneously and is implemented in the AnyBody modeling system. The method is capable of computing kinematics while accounting for the complex bone geometry and the sometimes highly nonlinear force-strain behavior of the surrounding soft tissues. As the secondary joint motions are small, the dynamics are negligible. For each time step, the secondary joint motions are computed by solving a set of nonlinear equations describing the force equilibrium. For further details on the method we refer to the study of Andersen *et al.* [23]. As the flexion/extension motion of the tibio-femoral joint is imposed, eleven DOFs are computed with FDK, five from the tibio-femoral joint and six from the patello-femoral joint. This is in contrast with the original model of Vanheule *et al.* [11] where the patella tendon length was kept constant resulting in the computation of five DOFs of the patello-femoral joint with FDK. The contact is modeled using an elastic foundation, STL-based contact model. The contact force (F_c) between two objects is linearly related to the volume of the intersection between the STL-surfaces (V_p). The relationship is given in eq. 4, where P is the constant pressure module. The pressure module is specific for the contact object [24]. The pressure module for the tibio-femoral and patello-femoral contacts are equal to $2.3e10 N/m^3$.

$$F_c = P \cdot V_p \tag{4}$$

The model definition is personalized to the extent it would currently be feasible using imaging modalities currently available in clinics. More specifically, the bone and cartilage geometry is segmented from MR images using Mimics 17.0 (Materialise N.V., Leuven, Belgium). The ligament attachment sites are not clearly visible on MRI especially the medial and lateral ligaments as they are surrounded by several other soft tissues shown in the same gray value on MRI. Therefore, the attachments are estimated based on literature and perturbed in the parameter estimation. The deepMCL, supMCL, LCL, ALL, PFL, PC, ACL with an Anteromedial (ACL0) and Posterolateral (ACL1) bundle and PCL with an Anterolateral (PCL0) and Posteromedial (PCL1) bundle are modeled. The patello-femoral ligaments are replaced by artificial springs with linear elastic properties between the rigid bodies, in order to reduce computation time without considerable effects on tibio-femoral kinematics.

2.3 Artificial neural network

A computationally efficient optimization is performed using a surrogate model of the MSKM with as inputs the ligament stiffness, reference strain and attachment positions, and as outputs the TF-kinematics and ligament strains. As surrogate modeling technique, an artificial neural network (ANN) is used, which is implemented using Tensorflow 2.4.1 [25]. For further details, we refer to the study of Bartsoen *et al.* [26]. The ranges of the ligament material properties for which the ANN is trained are given in Table 2. The training ranges for both stiffness and reference strain are taken rather broad as limited experimental measurement information is available.

The attachment sites have to be varied in a way that enforces contact of the attachment sites to the bone surface. To achieve this, a raster is projected on the bone surface. Each of the points on the raster is assigned an x- and y-coordinate. The x-axis is parallel to the AP-axis of the knee joint and oriented in anterior direction. The direction of the y-axis is chosen in order to have an orthogonal, right turning coordinate system with the z-axis coincident with the normal directed away from the bone surface. Both parameters vary with a range given in Table 2 with respect to the initial estimation of the attachment sites based on literature. The variation of the attachment sites is based on two aspects. The first aspect looks at the identifiability of the bony landmark. For example, the fibular attachments of PFL and LCL do not vary largely as the space they have to vary on the bone is rather small. Another aspect considers the mechanical effect of the variation. The tibial attachments of deepMCL, supMCL, LCL and ALL only vary in AP direction as the PD direction will have a negligible effect on the ligament force moment arm

Table 2: Training bounds of the artificial neural network for the ligament material properties.

Ligaments	Stiffness [N]	Reference strain []	Femur [mm]		Tibia [mm]	
			x_F	y_F	x_T	y_T
deep MCL	[2000, 9000]	[-0.2, 0.3]	[-10, 10]	[-8, 8]	[-10, 10]	
superficial MCL	[2000, 9000]	[-0.2, 0.2]	[-10, 10]	[-8, 8]	[-10, 10]	
LCL	[2000, 9000]	[-0.2, 0.2]	[-8, 8]	[-8, 8]	[-5, 5]	
ALL	[2000, 9000]	[-0.3, 0.2]	[-8, 8]	[-8, 8]	[-10, 10]	
PFL	[2000, 9000]	[-0.3, 0.2]	[-10, 10]	[-6, 6]		
ACL0	[4000, 10000]	[-0.2, 0.4]	[-6, 6]	[-6, 6]	[-8, 8]	[-8, 8]
ACL1	[4000, 10000]	[-0.2, 0.4]	[-6, 6]	[-6, 6]	[-8, 8]	[-8, 8]
PCL0	[4000, 12000]	[-0.5, 0.4]	[-6, 6]	[-6, 6]	[-8, 8]	[-8, 8]
PCL1	[4000, 12000]	[-0.5, 0.4]	[-6, 6]	[-6, 6]	[-8, 8]	[-8, 8]
PC	[5000, 10000]	[0.05, 0.2]				

and are therefore not considered as variable parameters. The projection of the raster on the bone surface is illustrated in Figure 3.

Samples are collected for training and validation. The normalized Mean Absolute Error (nMAE), Mean absolute error (MAE) and 90% Absolute Error (90% AE) are evaluated. The nMAE means that the MAE is computed for the validation data when the output has been normalized to [0,1]. Comparison of nMAE and MAE allows to distinguish between errors due to a large variance of the output within the sampling space and a bad fit with the ANN. The 90% AE gives the 90th percentile of the absolute errors of the validation samples. It allows the evaluation of the ANN accuracy for the specific application.As literature [15], [16] shows that ligament damage starts occurring from 6 % strain, the 90% AE of the validation samples of the ANN needs to be considerably lower. The amount of validation samples is equal to 2 000 for each specimen. The amount of training samples is dependent on the specimen as enough samples are collected to get the 90% AE of the ligament strain below 3 % and of kinematics below 2 mm or 2 °. They are gathered according to the Sobol sequence [27]. The network has 51 inputs which are the ones given in Table 2 and the flexion angle. There are 13 outputs, which are the 5 DOFs of the TF-kinematics and the ligament strain in deepMCL, supMCL, LCL, ALL, PFL, ACL, PCL and PC. The architecture of the network is [51:256:512:1024:512:256:128:64:13] with activation function Softplus (eq. 5). The Huber loss is used as a loss function (eq. 6) with $d = 0.02$, which behaves similar to the mean squared error (MSE) for small errors and similar to the mean absolute error (MAE) for large errors.

$$a(x) = log(exp(x) + 1) \tag{5}$$

$$x = y_{true} - y_{predicted}$$
$$loss = \begin{cases} 0.5x^2 & if \ |x| \leq d \\ 0.5d^2 + d(|x| - d) & if \ |x| > d \end{cases} \tag{6}$$

The network is trained using the Adam [28] optimizer with a learning rate that decays upon convergence of the training loss. The initial learning rate is 0.001. The batch size is 32. To prevent overfitting, l2 kernel regularization along with dropout is applied. To further improve network accuracy a weighted ensemble averaging of six networks is performed.

7

Figure 3: Projected raster to vary the ligament attachment sites. On the right side in detail for the femoral attachment of the MCL with indication of the x- and y-axes.
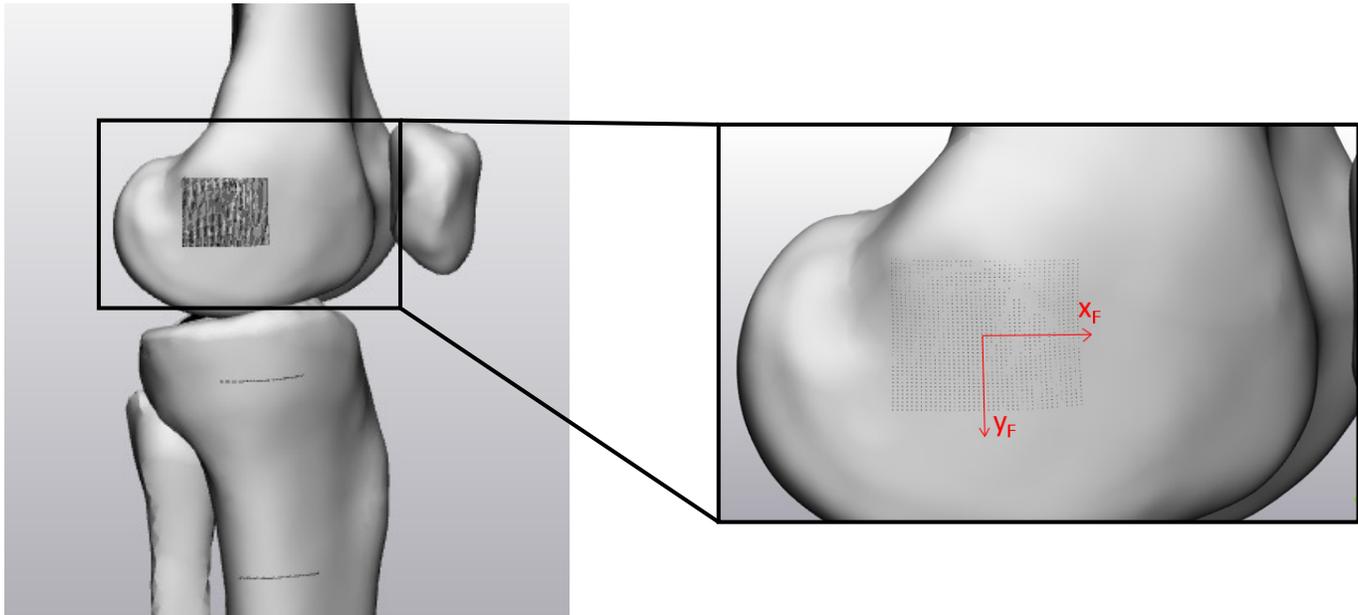
2.4 Ligament properties estimation

The ligament properties (linear stiffness, reference strain and attachment sites) are estimated using BPE. This means that a statistical distribution is estimated on the ligament parameters based on the error of the model outputs, more specifically TF-kinematics and ligament strain. Particularly for the problem described in this study, the statistical distribution of the TF-kinematics is defined as normally distributed with mean equal to the measured kinematics and standard deviation of 0.5 mm or 0.5 ° for ML- and AP-translation and IE-rotation, and 1.0 mm or 1.0 ° for PD-translation and VV-rotation. This error is chosen based on a repeatability analysis on the knee simulator, reported by Victor *et al.* [29] and a larger error is allowed on PD-translation and VV-rotation as these are largely influenced by the geometry of the contact surfaces. The constraint on the ligament strain is added to avoid ligament damage. As ligament damage starts occurring from 6% strain, the maximal ligament strain is limited to 6 % throughout the squat motion. As the squatting motion starts from 20 ° of flexion and highly slack ligaments (negative strain) throughout the entire squat would omit the use of the ligaments in other movements as well, the lower bound of the maximal ligament strain constraint is chosen slightly negative (-2 %). The maximal ligament strain throughout the squat is thus limited between -2 % and 6 % [15], [16]. These bounds are chosen as ligament damage is infeasible in a native knee squat movement and very small ligament strain throughout the entire squat would omit the use of the ligaments in other movements as well.

We apply the TMCMC algorithm [20], as also applied by Faes, Broggi et al. [30] in the context of high-dimensional Bayesian updating problems. This algorithm was chosen as it is suited for problems with complex probability distributions that can be both flat, peaked or multimodal [31]. The algorithm requires a prior distribution for the input parameters and gradually transforms the sampled distribution (set of chains) such that it approaches the posterior distribution. The amount of chains was set to 10 000. The prior was generated using uniform sampling over the ranges of the ANN (Table 2).

The original algorithm of Ching and Chen [20], was adapted with modifications 1 and 3 as introduced by the study of Betz *et al.* [21]. Modification 1 adapts the algorithm by recomputing the plausibility weight of each chain after each sub-iteration. This way the weight is adapted each time the chain moves. Modification

3 adapts the original algorithm by modifying the scaling parameter β after each iteration. The authors of the original TMCMC algorithm suggested $\beta = 0.2$ to be a good value for each problem, but Betz *et al.* show that for several problems this value is far from ideal resulting in a large correlation between the chains. They propose to adapt β based on the acceptance rate with a target acceptance rate of $t_{acr} = 0.23$. The flowchart given in Figure 4 shows the steps of the algorithm [20] with the modifications 1 and 3 from the study of Betz *et al.* [21]. For this study an extra modification was made by only performing re-sampling through selection of the current samples in the main loop and not as part of a sub-iteration. This avoids clustering of the samples at one specific solution resulting in premature convergence. As only a small subset of the ligament properties leads to a solution within the set distributions of the outputs, the acceptance rate of the first iteration in the main loop is very low leading to almost no MCMC. The resampling step however would quickly replace all samples with one most likely chain, leading to the premature convergence that is mentioned. This is illustrated in appendix B with results of a BPE using the TMCMC algorithm without the introduced adaptation on specimen 1. Moving the re-sampling step to the main loop however, allows enough iterations to sufficiently adapt the scaling parameter β during the initial iterations so that it reaches and keeps an acceptance rate of ± 23 % as suggested by the study of Betz *et al.* [21]. A detailed description of the algorithm can be found in appendix A.

The TMCMC algorithm uses the log-likelihood of a sample to perform the parameter estimation. The log-likelihood functions are defined according to a normal distribution for the error on the TF-kinematics. As an example, the log-likelihood is described for ML-translation at flexion angle ϕ in eq. 7, with $x^{exp}_{ML(\phi)}$ the experimentally measured ML-translation at flexion angle ϕ and $\sigma_{ML} = 0.5\ mm$.

$$log(f(D^{ML(\phi)}_{(j,k)}|M^{ML(\phi)}, \boldsymbol{\theta}_{(j,k)})) = -log(\sigma_{ML}\sqrt{2\pi}) - \frac{1}{2(\sigma_{ML})^2}(x_{ML(\phi)} - x^{exp}_{ML(\phi)})^2 \tag{7}$$

The ligament strains have to be optimized between -2 % and 6 % without preference for a specific value. A uniform distribution is thus assigned to the ligament strains. Using the log-likelihood of a uniform distribution would however lead to infinite values in case the strain is not within the allowed range, making the optimization infeasible. An adjustment to the log-likelihood is made to prevent this issue by using eq. 8 to compute the log-likelihood of the strain in a ligament, where $\epsilon_{ligament}$ is the maximal strain of the evaluated ligament throughout the squatting motion.

$$log(f(D^{ligament}_{(j,k)}|M^{ligament}, \boldsymbol{\theta}_{(j,k)})) = \begin{cases} -10 - 15 \cdot (\epsilon_{ligament} + 2.0)^2 & if\ -2.0\% < \epsilon_{ligament} \\ 0.0 & if\ -2.0\% \leq \epsilon_{ligament} \leq 6.0\% \\ -10 - 15 \cdot (\epsilon_{ligament} - 6.0)^2 & if\ \epsilon_{ligament} > 6.0\% \end{cases} \tag{8}$$

2.5   Statistical analysis

To describe the variation of the ligament properties, the absolute variation and variation relative to the ANN trained ranges are computed. The relative variation is computed for easy comparison across ligament parameters. For each ligament, the absolute variation is computed by subtracting the 95th percentile from the 5th percentile. The relative variation is computed by dividing the absolute variation by the ANN trained range.

A second analysis was performed to assess dependence between ligament parameters. This analysis consists out of a linear regression analysis between the reference strain and the attachments (eq. 9), the linear stiffness and the attachments (eq. 10), and the linear stiffness and the reference strain (eq. 11). The correlation coefficients, intercepts and coefficients of determination (R^2) are reported. The analysis is performed using Scikit-learn 0.22 [32].

9

Figure 4: Flowchart of the TMCMC algorithm as applied in this study. Modification 1 is indicated in blue, modification 3 in red and the modification introduced in the current study is indicated in green with * indicating the original position of the steps indicated in green.
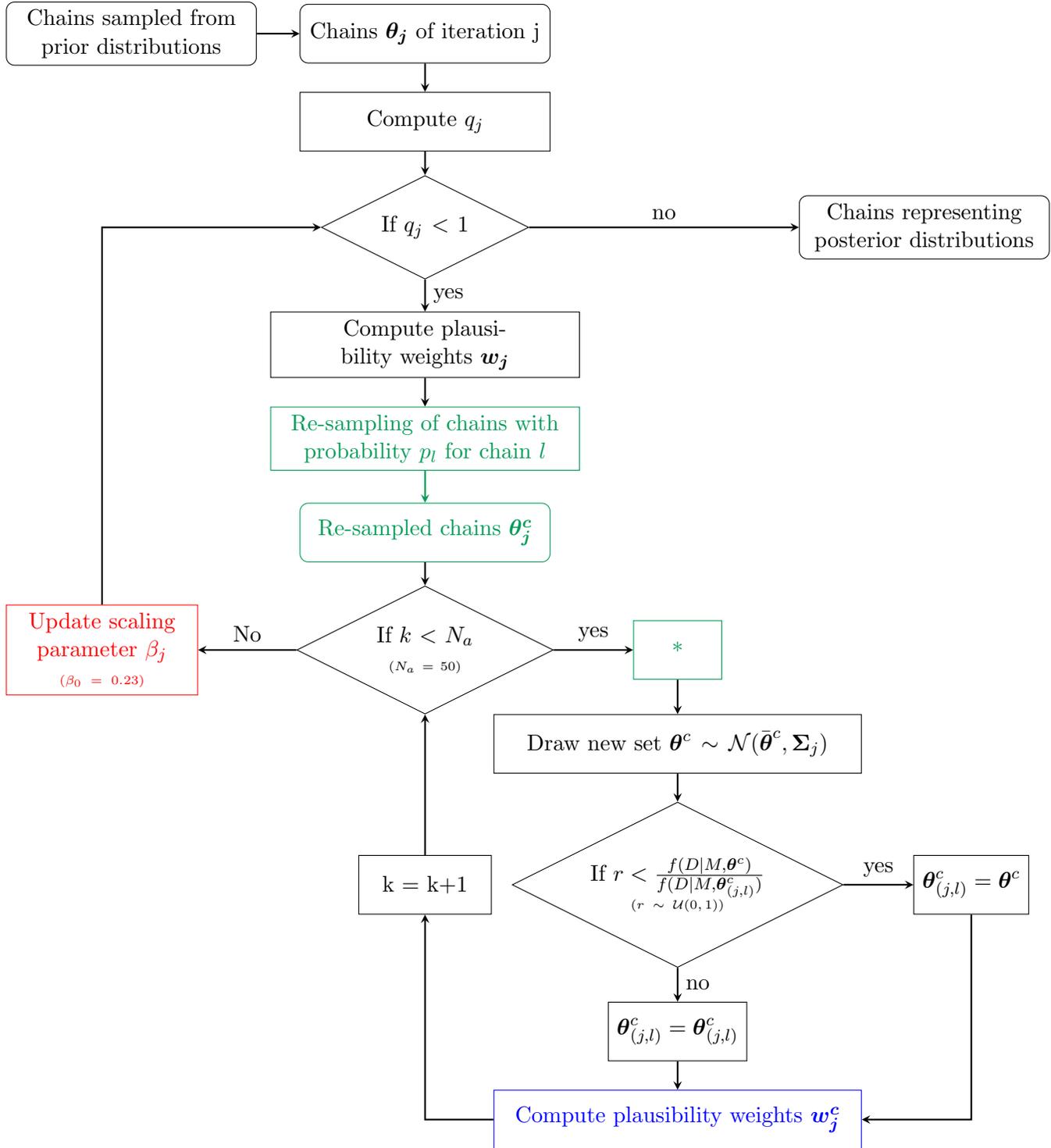
$\boldsymbol{\theta_j}$ are the chains of the current iteration. q_j is a parameter given by eq. 12. $\boldsymbol{w_j}$ are plausibility weights given by eq. 13. p_l is the resampling probability given by eq. 17. N_a is the total amount of iterations in the sub-loop. β_j is the scaling parameter and can be computed with eq . 19. $f(x)$ is the probability density function of x. $\boldsymbol{\Sigma}_j$ is the covariance matrix of the Gaussian proposal distribution given in eq. 15. D are the model outputs consistent with the current chains. M is the target probability distribution of the model outputs. $\mathcal{N}$ and $\mathcal{U}$ are respectively normal and uniform distributions.

$$\epsilon_r = a \cdot x_F + b \cdot y_F + c \cdot x_T + d \cdot y_T + e \tag{9}$$

$$k = a \cdot x_F + b \cdot y_F + c \cdot x_T + d \cdot y_T + e \tag{10}$$

$$k = a \cdot \epsilon_r + b \tag{11}$$

3 Results

The results of the ANN, the ligament properties estimation, the statistical analysis and the computational efficiency are given in respectively the Sections 3.1, 3.2, 3.3 and 3.4.

3.1 ANN accuracy

The validation error of the network is given in Table 3, indicated with the average, minimum and maximum error over the seven specimens. The 90% AE was aimed to be below 3% for the ligament strains and below 2 mm or 2 ° for TF-kinematics, which is the case for each output. The strain in the PCL has on average the largest MAE (1.20) and 90% AE (2.61) over the evaluated specimens. This is to be expected as it is the ligament with the largest variance in strain throughout the sampling space as it does not have the largest nMAE but it has the largest MAE and 90% AE. The amount of samples required for the ANN to reach the accuracy target, was different for each specimen with an average of 54 624 samples and ranging between 37 948 for specimen 7 and 70 027 for specimen 2.

Table 3: Accuracy of ANN for all outputs. The average [minimum, maximum] are given over the seven specimens. The 90% AE was aimed to be below 3 % for the ligament strains and below 2 mm or 2 ° for TF-kinematics.

	Output	nMAE [%]	MAE	90% AE
Kinematics	ML [mm]	0.76 [0.69,0.89]	0.21 [0.13,0.27]	0.44 [0.28,0.58]
	AP [mm]	0.52 [0.50,0.56]	0.30 [0.25,0.36]	0.63 [0.53,0.75]
	DC [mm]	0.33 [0.24,0.47]	0.10 [0.08,0.15]	0.18 [0.14,0.25]
	VV [°]	0.35 [0.29,0.48]	0.22 [0.14,0.37]	0.40 [0.25,0.63]
	IE [°]	0.67 [0.53,0.80]	0.78 [0.47,0.99]	1.53 [0.97,1.84]
Ligament strain	deepMCL [%]	0.39 [0.36,0.42]	0.58 [0.42,0.66]	1.19 [0.85,1.40]
	supMCL [%]	0.27 [0.24,0.30]	0.22 [0.18,0.26]	0.44 [0.36,0.50]
	LCL [%]	0.38 [0.34,0.45]	0.36 [0.28,0.52]	0.71 [0.55,1.01]
	PFL [%]	0.49 [0.44,0.58]	0.62 [0.43,0.89]	1.26 [0.91,1.77]
	ALL [%]	0.39 [0.33,0.50]	0.64 [0.50,0.78]	1.31 [1.02,1.52]
	ACL [%]	0.61 [0.50,0.75]	1.07 [0.97,1.22]	2.37 [2.11,2.77]
	PCL [%]	0.53 [0.46,0.70]	1.20 [1.01,1.31]	2.61 [2.18,2.97]
	PC [%]	0.67 [0.48,0.86]	0.57 [0.37,0.74]	1.19 [0.76,1.51]

3.2 Parameter estimation

Table 4 shows the errors on the TF-kinematics after optimization. The largest errors are on ML-translation for specimen 7. It can be seen that all maximal MAE are around 1 mm or 1 °, except for specimen 7. This means that the allowed error can be achieved for each specimen, except specimen 7. The ligament strains are optimized within the ranges -2 % and 6 % or slightly outside with a maximum strain of 7.0 % for the PCL of specimen 7.

Table 4: Accuracy of the optimization of the seven specimens over the experimental range of flexion. Average MAE [minimal MAE, maximal MAE] over the different chains of the TMCMC optimization.

Specimen	MAE on optimized output				
	ML [mm]	AP [mm]	DC [mm]	VV [°]	IE [°]
1	0.3 [0.2,0.6]	0.3 [0.1,0.6]	0.4 [0.3,0.5]	0.8 [0.6,1.1]	0.3 [0.1,0.6]
2	0.3 [0.1,0.6]	0.3 [0.1,0.6]	0.3 [0.2,0.3]	0.4 [0.2,0.5]	0.3 [0.1,0.6]
3	0.4 [0.2,0.7]	0.2 [0.1,0.5]	0.7 [0.6,0.8]	0.9 [0.7,1.1]	0.4 [0.2,0.7]
4	0.6 [0.4,1.0]	0.3 [0.1,0.7]	0.3 [0.2,0.4]	0.5 [0.4,0.7]	0.4 [0.2,0.7]
5	0.5 [0.3,0.7]	0.6 [0.4,0.8]	0.6 [0.5,0.7]	0.4 [0.3,0.8]	0.4 [0.2,0.6]
6	0.4 [0.3,0.7]	0.2 [0.1,0.5]	0.5 [0.4,0.5]	0.5 [0.3,0.8]	0.2 [0.1,0.6]
7	1.4 [1.2,1.7]	0.8 [0.5,1.1]	0.8 [0.7,0.9]	0.7 [0.6,1.0]	0.8 [0.5,1.3]

Figures 5 and 6 respectively show the reference strain and linear stiffness after optimization towards the experimental kinematics. The posterior distributions differ significantly from the prior distributions which were chosen uniformly distributed over the plotted ranges. It can be seen that for specimen 5 multiple solutions are found. This is clearly visible for ϵ_r of PCL1. For specimen 7, several parameters (e.g. ϵ_r for LCL, PCL0 and PC) are optimized towards the training edges of the neural network.

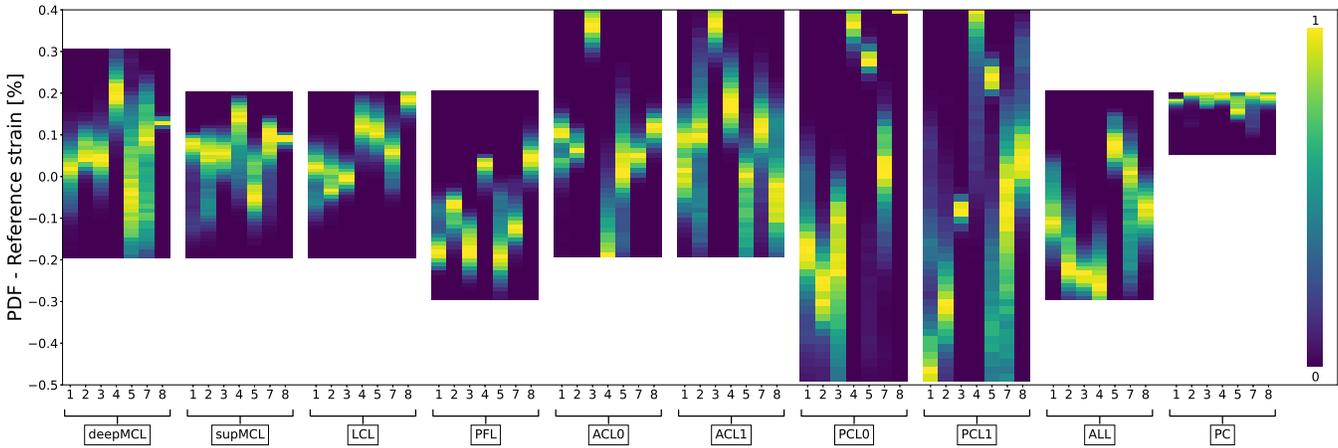Figure 5: Results of BPE: reference strain. Normalized histograms of the posterior distribution samples.

Figure 7 shows the root mean squared difference between the cumulative distribution functions of successive iterations of the parameter estimation of specimen 4. It can be seen that differences are largest in the first few iterations and convergence is achieved after about 22 iterations. This shows that regardless of the limited amount of experimental data, the posterior distributions are significantly different from the prior

Figure 6: Results of BPE: linear stiffness. Normalized histograms of the posterior distribution samples.
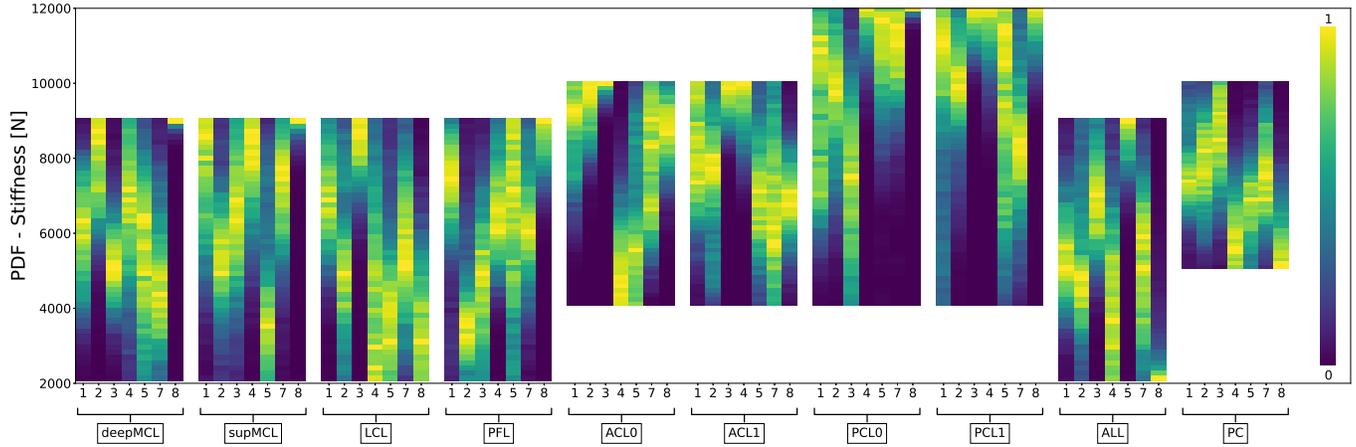
distributions. Similar performance is observed for the other specimens.

As specimen 7 did not fully fit the experimental data due to the too narrow training ranges of the ANN, it will be excluded from further analysis, reducing the test population to n=6.

3.3 Statistical analysis

Table 5 gives the difference between the 5th and 95th percentile of the variation of the input parameters. We can see that on average a smaller variation is allowed on the reference strain compared to the linear stiffness and attachment sites.

Table 5: Difference between 5th and 95th percentile of the variation in the input parameters as an average over the 6 specimens. The absolute and relative variation are given with average [minimum, maximum] over the 6 specimens.

Ligaments	Reference strain		Stiffness		Attachments	
	relative [%]	absolute	relative [%]	absolute [N]	relative [%]	absolute [mm]
deepMCL	37 [4,70]	0.19 [0.02,0.35]	57 [6,81]	4004 [467,5723]	53 [4,89]	10.0 [0.9,17.8]
supMCL	39 [8,55]	0.16 [0.03,0.22]	61 [13,83]	4322 [975,5815]	61 [6,88]	11.6 [1.3,17.7]
LCL	27 [13,37]	0.11 [0.05,0.15]	69 [36,86]	4840 [2554,6062]	69 [48,92]	9.7 [4.9,14.8]
PFL	25 [13,40]	0.13 [0.07,0.20]	66 [29,88]	4656 [2058,6221]	57 [25,86]	9.3 [3.6,15.4]
ACL0	22 [12,53]	0.14 [0.07,0.32]	52 [12,85]	3132 [754,5108]	51 [4,87]	8.3 [0.8,14.1]
ACL1	42 [17,69]	0.26 [0.10,0.42]	62 [25,86]	3758 [1507,5181]	66 [17,88]	10.7 [2.9,14.2]
PCL0	25 [1,45]	0.23 [0.01,0.41]	48 [5,77]	3864 [458,6232]	54 [1,95]	9.0 [0.3,17.1]
PCL1	48 [7,79]	0.43 [0.07,0.72]	55 [19,88]	4419 [1539,7077]	58 [9,96]	9.9 [1.6,17.4]
ALL	35 [20,62]	0.18 [0.10,0.31]	60 [28,81]	4268 [1984,5728]	57 [14,84]	10.0 [2.3,16.8]
PC	27 [7,56]	0.04 [0.01,0.08]	66 [55,75]	3346 [2798,3771]		

The coefficients of determination of the linear regression analysis are given in Table 6. It can be seen that the reference strain and the attachment sites are highly correlated while the linear stiffness and reference

13

Figure 7: Convergence of the BPE for reference strain, stiffness and maximal over the 4 DOFs of the attachment sites for specimen 4. The root mean squared error (RMSE) between Cumulative distribution functions of successive iterations is given.
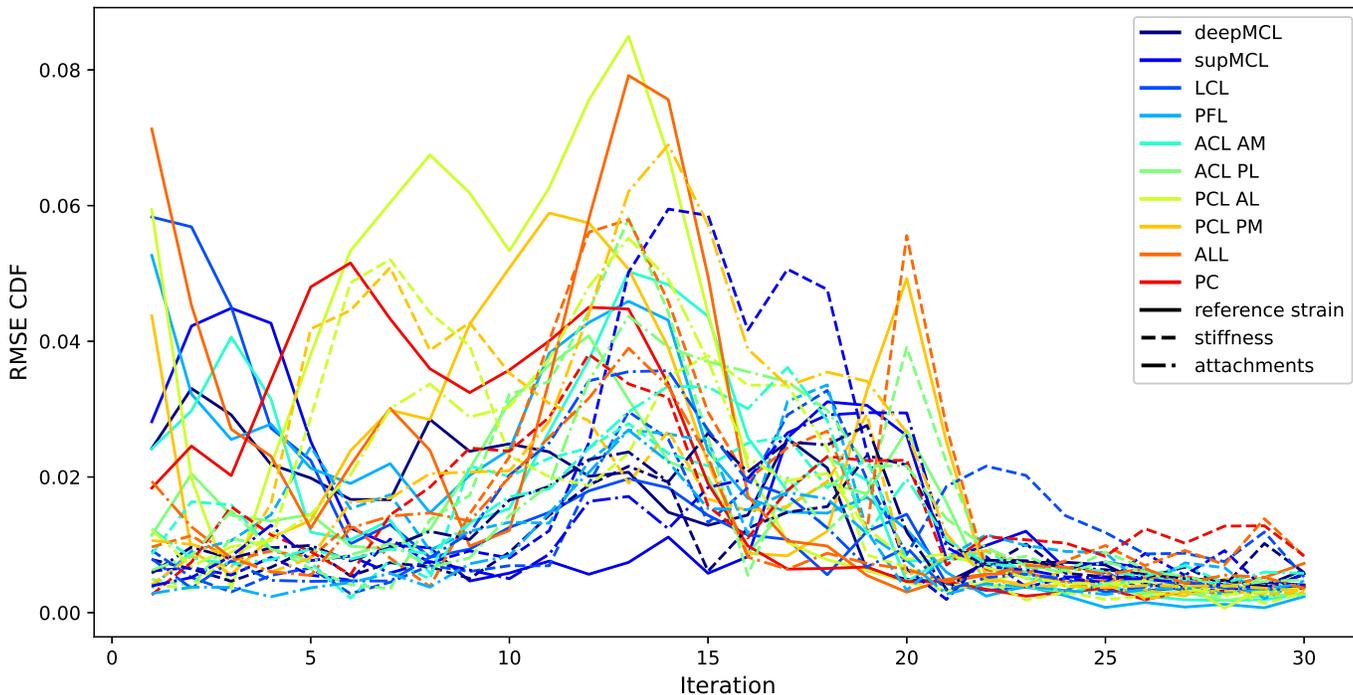
strain are not, nor are the linear stiffness and attachment sites. The coefficients and intercept of the linear regression functions between reference strain and attachments are given in Table 7. It can be seen that the coefficients do not vary largely between different specimens especially for the medial and lateral ligaments, while the intercept varies significantly.

3.4 Computational efficiency

The parameter estimation requires on average 20 iterations with 500 000 log-likelihood function evaluations per iteration to converge. The amount of iterations is dependent on the specimen and varies between 17 and 26. Each input sample is evaluated for 12 flexion angles between 20 ° and 120 °. This leads to a total of 6 000 000 knee model evaluations per iteration. Using the surrogate model this can be executed in 76 hours on average. Using the knee model directly this would take approximately 114 years per specimen with 8 simulations running in parallel. Each evaluation of the ANN takes a few milliseconds while an evaluation of the knee model takes ±4 min. By using the ANN to perform the parameter estimation, the amount of model evaluations can be reduced to an average of 54 624 samples required for training of the ANN.

4 Discussion

This study developed a framework that executes a BPE on a complex MSKM with feasible computational cost. By using an ANN as a surrogate model, the computational cost can be considerably reduced. It can be seen that - across specimens - there is a large variation in amount of samples to reach the required accuracy. This can be explained by the larger amount of knee model evaluation errors. For some specimens it is more likely that an unstable solution is found, leading to sparse regions in the sampling space caused by a

Table 6: Results of linear regression analysis. The average [minimum, maximum] coefficient of determination R^2 over the 6 specimens is given for each of the three combinations of ligament parameters.

Ligaments	ϵ_r - Attachments	k - Attachments	k - ϵ_r
deepMCL	0.85 [0.66,0.93]	0.14 [0.03,0.37]	0.02 [0.00,0.08]
supMCL	0.87 [0.67,0.94]	0.13 [0.03,0.39]	0.07 [0.00,0.23]
LCL	0.65 [0.49,0.77]	0.18 [0.00,0.36]	0.07 [0.00,0.19]
PFL	0.87 [0.72,0.96]	0.07 [0.01,0.16]	0.03 [0.00,0.11]
ACL0	0.61 [0.03,0.91]	0.16 [0.02,0.30]	0.05 [0.00,0.15]
ACL1	0.65 [0.16,0.94]	0.16 [0.05,0.36]	0.08 [0.01,0.30]
PCL0	0.64 [0.20,0.96]	0.16 [0.03,0.25]	0.06 [0.00,0.24]
PCL1	0.72 [0.41,0.96]	0.20 [0.04,0.41]	0.07 [0.00,0.23]
ALL	0.91 [0.87,0.97]	0.08 [0.02,0.12]	0.02 [0.00,0.05]
PC			0.02 [0.00,0.04]

Table 7: Results of linear regression analysis between reference strain and attachment sites (eq. 9). The average [minimum, maximum] of the coefficients and intercept of the linear regression model over the six specimens. With the reference strain expressed in %.

Ligaments	a	b	c	d	e
deepMCL	1.81 [1.16, 2.59]	-0.82 [-1.69, 0.09]	-0.05 [-0.44, 0.22]		1.96 [-3.48,14.1]
supMCL	0.98 [0.63, 1.25]	-0.52 [-0.76,-0.30]	0.05 [-0.04, 0.20]		0.69 [-1.82, 8.29]
LCL	0.56 [0.34, 0.73]	0.46 [0.02, 1.24]	0.36 [0.24, 0.68]		5.02 [-1.58,11.33]
PFL	-0.35 [-0.54,-0.11]	1.63 [1.39, 1.95]	0.00 [0.00, 0.00]		-11.87 [-17.67,-6.76]
ACL0	1.01 [-0.06, 2.81]	0.44 [-0.21, 2.02]	0.35 [-0.53, 0.95]	-0.19 [-0.56, 0.20]	4.31 [-5.88,12.79]
ACL1	1.00 [-0.05, 4.28]	0.64 [-1.33, 1.68]	0.51 [0.03, 1.33]	-0.76 [-1.65, 0.34]	11.89 [-8.57,40.40]
PCL0	1.27 [-0.45, 3.23]	2.69 [1.10, 4.18]	-0.03 [-0.80, 1.12]	-0.02 [-1.78, 1.71]	-20.88 [-33.32,-5.42]
PCL1	1.02 [0.11, 2.81]	3.27 [1.77, 5.51]	0.16 [-1.01, 1.19]	0.07 [-2.12, 1.89]	-21.95 [-38.98,-2.49]
ALL	1.39 [1.06, 1.89]	0.86 [-0.07, 1.55]	0.93 [0.66, 1.29]		-17.52 [-26.94,-5.62]

larger number of knee model evaluation errors. This problem could possibly have been tackled using another surrogate modeling technique like Kriging where a variance based sampling can be performed in order to choose the samples based on the areas in the sampling space with the largest variance [33], [34]. Exploring other types of surrogate modeling techniques (e.g. Support Vector regression, Polynomial Chaos expansion, Kriging), could possibly reduce the required amount of MSKM samples to reach the required accuracy. The study of Heidenreich *et al.* [35] shows that Polynomial Chaos expansion proved most effective in a BPE for an application involving Scatterometry. The study of Bartsoen *et al.* [26] however showed that ANN is the most effective method as a surrogate for the MSKM with respect to Kriging and Support Vector regression. Polynomial Chaos expansion was however not evaluated.

We chose a Bayesian method instead of an interval method as we wanted - in addition to identifying the possible ligament properties per specimen - also search for multi-modal output distributions and correlations between parameters. The interval methods also become inefficient for high dimensional problems. We do only have one measurement of the squat motion with an estimate on the error from a repeatability study

on the same experimental device. According to the study of Faes and Broggi *et al.* [30] a small amount of experimental data could lead to a posterior distribution influenced by the prior. The resulting distributions however do differ significantly from the prior distributions as was shown by monitoring the convergence. Regarding the BPE algorithm TMCMC, we found that the modifications 1 and 3 of the study of Betz *et al.* [21] significantly improved the results. Especially the update of the scaling parameter β was of a large importance. However, even the adaptations introduced by Betz *et al.* [21] are shown to be insufficient for successful BPE on our knee model. A modification was required consisting of performing re-sampling every main iteration as opposed to every sub-iteration. This avoided clustering of the samples at one specific solution resulting in premature convergence. This approach allowed the scaling parameter β to adapt sufficiently to increase to and keep an acceptance rate of ± 23 % as suggested by the study of Betz *et al.* [21]. This adaptation was most likely required as this specific problem required several iterations to optimize the set of ligament properties before actual MCMC sampling takes effect. If not included, the first few iterations are dominated by sample replacement causing a too fast increase in q-value, which results in premature convergence. This is supported by the review of Lye *et al.* [31], where they state that a burn-in period is generally unnecessary with TMCMC except when sampling from complex, high dimensional posteriors is performed. This is especially required when the posterior distribution consists of only a small area of the entire sampling space, which is the case for the reference strains.

By using a surrogate model, the computational cost of the BPE was considerably reduced. The parameter optimization however still takes 80 hours per specimen, due to the computational inefficiency of the TMCMC algorithm. The computational cost could potentially be further reduced by implementation of for example X-TMCMC [36] of which the authors claim it can reduce computational cost with an order of magnitude, M-TMCMC [37], Transitional Ensemble MCMC [38] or AK-MCMC [39].

Regarding the ligament properties, this study shows that there is a large variation between different specimens. This is consistent with the results of the study of Beidokhti *et al.* [10]. Between the three specimens on which Beidokhti *et al.* performed ligament properties optimization, the largest difference in stiffness is 7917 N between specimen 1 and 2 on the posterior bundle of the PCL. The largest difference in reference strain is 0.38 between specimens 1 and 3 on the anterior bundle of the ACL. Our results show these magnitudes of variance between different specimens but also within one specimen. Our results also show that a large variance is allowed on reference strain, linear stiffness and attachment sites, but that the reference strain and attachment sites are highly correlated. This is consistent with the study of Andersen et al. [19] where the effect of measurement noise on four different sets of laxity trials was studied. They found a large difference in the ligament stiffness with on average over the four sets a range of 6.54 kN and relatively small changes in ligament strain with on average 2.28 %. As they did not vary the attachment sites, the variance in the reference strain was much more limited.

Our results also show that for some specimens multiple distinct solutions can be found as well as large variations. This proves that solely measuring a squat movement is insufficient for proper estimation of the ligament material properties and attachment sites. Other types of movements will have to be added to eliminate the possibility of multiple solutions, like combinations of different laxity trials presented in the study of Andersen *et al.* [19]. Especially with in vivo set-ups, combinations of movements will be required, as larger measurement errors can be expected.

The linear regression analysis shows that even though a large variation is allowed on the reference strain and attachment sites, they are largely interdependent, thereby confirming them as critical parameters in the knee model. The linear stiffness is (almost) uncorrelated to the other parameters and is thus not critical to the knee model outputs that were studied (TF-kinematics and ligament strains).

The optimized parameters vary significantly between the different specimens. This means that a narrow range of ligament material properties cannot be identified. This will lead to high uncertainty in the ligament properties if used in applications like pre-operative planning. This study furthermore identifies that the reference strain and attachment sites are strongly correlated. This is highly relevant in clinical practice since

failing to account for this correlation in the uncertainty quantification of the knee model outputs used in pre-operative planning (TF kinematics and ligament strains) will largely overestimate the effect of the ligament properties. However, tracking of active motions, like the squat motion used in this study, is impractical in a clinical setting. To account for the parameter correlations when estimating the effect of uncertainty in the ligament properties on pre-operative planning for application in clinical practice, we suggest to define a patient-specific collection of sets of ligament properties representing their uncertainty based on the native knee simulations. The same methodology as introduced in this study can be applied to identify this patient-specific collection of sets of ligament properties, but with the difference that TF-kinematics are unknown. The feasible sets of ligament properties will have to be identified by solely constraining the ligament strain within the physiological limits. This will limit the possible outcomes that should be accounted for during pre-operative planning. If a future study using the proposed methodology shows that the ligament properties introduce too much uncertainty into the pre-operative planning process, it could be possible to use the same methodology as applied in this study on other (combinations of) motions. This could allow the identification of a set of motions to further limit uncertainty in the ligament properties up to the extent that the effect on pre-operative planning is limited enough for the reliable use of a computational knee model. These motions could be applied either pre-operatively through for example the device described by the study of Pedersen *et al.* [18] or intra-operatively using surgical navigation through for example the device described by the study of Fucentese *et al.* [40].

This study needs to be considered in view of the following limitations. First, only seven specimens are used for this study. We do however see a large difference between the different specimens. With this sample size it is impossible to identify possible outliers. Our results however confirm that it is impossible to establish a small range of ligament properties that represents the patient population. A second limitation is that the applied knee model does not include the menisci. Future research will need to investigate the effects of the presence of menisci. Uncertainties in the muscle parameters were not included in the study. In the experiment, a constant hamstrings force (50 N each) was imposed but this might slightly vary due to experimental noise. The quadriceps force is imposed during the experiment using a force feedback loop with the vertical ankle force. In the simulation, the Quadriceps force was imposed as a function varying with the flexion angle and averaged over the specimens as measured during the experiment. Per specimen this thus deviates from the experimental imposed quadriceps load. A final limitation is in the ANN training bounds for specimen 7. Some of the optimized parameters are pushed towards the limits which causes the kinematics for specimen 7 to not fully reach the imposed statistical distribution. This leads to a few different solutions with limited variation. This however does not influence the conclusions of the study. If slightly larger training ranges were taken, the TF-kinematics of specimen 7 could have been optimized properly with probably larger solution variation as is the case with the other specimens.

5   Conclusions

We developed a framework that allows parameter estimation - based on limited experimental data - of a high dimensional, non-linear knee model with feasible computational cost. Application of the TMCMC algorithm resulted in premature convergence. The implemented alteration through reduction of the amount of re-sampling, allowed successful parameter estimation. Despite the use of inverse uncertainty quantification, we see that no unique set of ligament properties can be found with solely measurement of TF-kinematics of a squat motion as multiple solutions are found for some of the specimens and a large variation is allowed on each of the ligament properties. We do however see a large correlation between ligament reference strain and attachment sites and no correlation with linear stiffness. To limit ligament properties uncertainty in applications like pre-operative planning, research will need to invest in establishing patient-specific uncertainty ranges and/or techniques for accurate in vivo determination of attachment sites and reference strains

and/or alternative (combinations of) movements that would lead to a unique solution.

Acknowledgment(s)

References

[1] L. Blankevoort and R. Huiskes, "Ligament-bone interaction in a three-dimensional model of the knee," *Journal of Biomechanical Engineering: Transactions of the ASME*, vol. 113, no. 3, pp. 263–269, 1991.

[2] K. Smeets *et al.*, "Mechanical analysis of extra-articular knee ligaments. part one: Native knee ligaments," *The Knee*, vol. 24, no. 5, pp. 949–956, 2017.

[3] D. L. Butler, M. D. Kay, and D. C. Stouffer, "Comparison of material properties in fascicle-bone units from human patellar tendon and knee ligaments," *Journal of biomechanics*, vol. 19, no. 6, pp. 425–432, 1986.

[4] T. Sugita and A. A. Amis, "Anatomic and biomechanical study of the lateral collateral and popliteofibular ligaments," *The American journal of sports medicine*, vol. 29, no. 4, pp. 466–472, 2001.

[5] N. Chandrashekar *et al.*, "Sex-based differences in the tensile properties of the human anterior cruciate ligament," *Journal of biomechanics*, vol. 39, no. 16, pp. 2943–2950, 2006.

[6] S. L.-Y. Woo *et al.*, "Tensile properties of the human femur-anterior cruciate ligament-tibia complex: The effects of specimen age and orientation," *The American journal of sports medicine*, vol. 19, no. 3, pp. 217–225, 1991.

[7] J. R. Robinson, A. M. Bull, and A. A. Amis, "Structural properties of the medial collateral ligament complex of the human knee," *Journal of biomechanics*, vol. 38, no. 5, pp. 1067–1074, 2005.

[8] R. F. LaPrade *et al.*, "Mechanical properties of the posterolateral structures of the knee," *The American journal of sports medicine*, vol. 33, no. 9, pp. 1386–1391, 2005.

[9] P. S. Trent, P. S. Walker, and B. Wolf, "Ligament length patterns, strength, and rotational axes of the knee joint.," *Clinical orthopaedics and related research*, no. 117, pp. 263–270, 1976.

[10] H. N. Beidokhti *et al.*, "The influence of ligament modelling strategies on the predictive capability of finite element models of the human knee joint," *Journal of biomechanics*, vol. 65, pp. 1–11, 2017.

[11] V. Vanheule *et al.*, "Evaluation of predicted knee function for component malrotation in total knee arthroplasty," *Medical engineering & physics*, vol. 40, pp. 56–64, 2017.

[12] D. G. Thelen, K. Won Choi, and A. M. Schmitz, "Co-simulation of neuromuscular dynamics and knee mechanics during human walking," *Journal of biomechanical engineering*, vol. 136, no. 2, 2014.

[13] C. R. Smith, S. C. Brandon, and D. G. Thelen, "Can altered neuromuscular coordination restore soft tissue loading patterns in anterior cruciate ligament and menisci deficient knees during walking?" *Journal of Biomechanics*, vol. 82, pp. 124–133, 2019.

[14] K. Kang *et al.*, "The effects of posterior cruciate ligament deficiency on posterolateral corner structures under gait-and squat-loading conditions: A computational knee model," *Bone & joint research*, vol. 6, no. 1, pp. 31–42, 2017.

[15] Z. Guo *et al.*, "Quantification of strain induced damage in medial collateral ligaments," *Journal of Biomechanical Engineering*, vol. 137, no. 7, 2015.

[16] P. P. Provenzano *et al.*, "Subfailure damage in ligament: A structural and cellular evaluation," *Journal of Applied Physiology*, vol. 92, no. 1, pp. 362–371, 2002.

[17] L. C. Slane *et al.*, "The challenges of measuring in vivo knee collateral ligament strains using ultrasound," *Journal of biomechanics*, vol. 61, pp. 258–262, 2017.

[18] D. Pedersen *et al.*, "A novel non-invasive method for measuring knee joint laxity in four dof: In vitro proof-of-concept and validation," *Journal of biomechanics*, vol. 82, pp. 62–69, 2019.

[19] M. S. Andersen *et al.*, "A methodology to evaluate the effects of kinematic measurement uncertainties on knee ligament properties estimated from laxity measurements," *Journal of Biomechanical Engineering*, 2021.

[20] J. Ching and Y.-C. Chen, "Transitional markov chain monte carlo method for bayesian model updating, model class selection, and model averaging," *Journal of engineering mechanics*, vol. 133, no. 7, pp. 816–832, 2007.

[21] W. Betz, I. Papaioannou, and D. Straub, "Transitional markov chain monte carlo: Observations and improvements," *Journal of Engineering Mechanics*, vol. 142, no. 5, p. 04 016 016, 2016.

[22] A. B. Zavatsky, "A kinematic-freedom analysis of a flexed-knee-stance testing rig," *Journal of biomechanics*, vol. 30, no. 3, pp. 277–280, 1997.

[23] M. Skipper Andersen *et al.*, "Introduction to force-dependent kinematics: Theory and application to mandible modeling," *Journal of biomechanical engineering*, vol. 139, no. 9, 2017.

[24] AnyBody Technology, *AnyForceSurfaceContact*, accessed May 12, 2021 from `https://anyscript.org/reference-manual/reference/classes/AnyForceSurfaceContact.html`, 2021.

[25] M. Abadi *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

[26] L. Bartsoen *et al.*, "Computationally efficient optimization method to quantify the required surgical accuracy for a ligament balanced tka," *IEEE Transactions on Biomedical Engineering*, 2021.

[27] I. M. Sobol', "On the distribution of points in a cube and the approximate evaluation of integrals," *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, vol. 7, no. 4, pp. 784–802, 1967.

[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[29] J. Victor, "A comparative study on the biomechanics of the native human knee joint and total knee arthroplasty," Ph.D. dissertation, KU Leuven, 2009.

[30] M. Faes *et al.*, "A multivariate interval approach for inverse uncertainty quantification with limited experimental data," *Mechanical Systems and Signal Processing*, vol. 118, pp. 534–548, 2019.

[31] A. Lye, A. Cicirello, and E. Patelli, "Sampling methods for solving bayesian model updating problems: A tutorial," *Mechanical Systems and Signal Processing*, vol. 159, p. 107 760, 2021.

[32] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[33] M. Faes *et al.*, "On the robust estimation of small failure probabilities for strong nonlinear models," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 5, no. 4, 2019.

[34] B. Echard, N. Gayton, and M. Lemaire, "Ak-mcs: An active learning reliability method combining kriging and monte carlo simulation," *Structural Safety*, vol. 33, no. 2, pp. 145–154, 2011.

[35] S. Heidenreich, H. Gross, and M. Bar, "Bayesian approach to the statistical inverse problem of scatterometry: Comparison of three surrogate models," *International Journal for Uncertainty Quantification*, vol. 5, no. 6, 2015.

[36] P. Angelikopoulos, C. Papadimitriou, and P. Koumoutsakos, "X-tmcmc: Adaptive kriging for bayesian inverse modeling," *Computer Methods in Applied Mechanics and Engineering*, vol. 289, pp. 409–428, 2015.

[37] H. Jensen, D. Jerez, and M. Valdebenito, "An adaptive scheme for reliability-based global design optimization: A markov chain monte carlo approach," *Mechanical Systems and Signal Processing*, vol. 143, p. 106 836, 2020.

[38] A. Lye, A. Cicirello, and E. Patelli, "An efficient and robust sampler for bayesian inference: Transitional ensemble markov chain monte carlo," *Mechanical Systems and Signal Processing*, vol. 167, p. 108 471, 2022.

[39] P. Wei, C. Tang, and Y. Yang, "Structural reliability and reliability sensitivity analysis of extremely rare failure events by combining sampling and surrogate model methods," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of risk and reliability*, vol. 233, no. 6, pp. 943–957, 2019.

[40] S. F. Fucentese and P. P. Koch, "A novel augmented reality-based surgical guidance system for total knee arthroplasty," *Archives of Orthopaedic and Trauma Surgery*, vol. 141, no. 12, pp. 2227–2233, 2021.

Appendices

A   TMCMC algorithm with modifications

The following steps describe the algorithm [20] with the modifications 1 and 3 from the study of Betz *et al.* [21]. For this study an extra modification was made by only performing re-sampling through selection of the

current samples in the main loop and not as part of a sub-iteration. This avoids clustering of the samples at one specific solution resulting in premature convergence.

At the initial iteration $j = 0$, all N_s chains $(\boldsymbol{\theta}_{0,1}, ..., \boldsymbol{\theta}_{0,N_s})$ are drawn from the prior Probability Distribution Function (PDF) $f_0(\boldsymbol{\theta}_0) = f(\boldsymbol{\theta}_0|M)$ with $\boldsymbol{\theta}$ the current set of input samples and M the aimed statistical distribution of the model output parameters.

1. q_j is found by solving eq. 12. If $q_j > 1$, then set $q_j = 1$.

$$q_{j+1} = \arg\min_q(|CV_j(q) - v_t|) \tag{12}$$

with CV_j the coefficient of variation of the set $\{f(D|M, \boldsymbol{\theta}_{(j,k)})^{(q-q_j)}\}_{k=1}^{N_s}$ and $v_t = 1.0$, with $q \in [q_j, 1]$.

2. Compute the plausibility weights (eq. 13) for all chains and the mean of the weighting coefficients S_j according to eq. 14.

$$w(\boldsymbol{\theta}_{(j,k)}) = f(D|M, \boldsymbol{\theta}_{(j,k)})^{q_{j+1}-q_j} \tag{13}$$

with D the output parameters based on the current input sample $\boldsymbol{\theta}_{(j,k)}$

$$S_j = \frac{\sum_{k=1}^{N_j} w(\boldsymbol{\theta}_{(j,k)})}{N_j} \tag{14}$$

3. Compute the covariance matrix of the Gaussian proposal distribution

$$\Sigma_j = \beta_j^2 \cdot \sum_{k=1}^{N_s} [\frac{w_{(j,k)}}{S_j \cdot N_s} \cdot (\boldsymbol{\theta}_{(j-1,k)} - \bar{\boldsymbol{\theta}}_j) \cdot (\boldsymbol{\theta}_{(j-1,k)} - \bar{\boldsymbol{\theta}}_j)^T] \tag{15}$$

with

$$\bar{\boldsymbol{\theta}}_j = \frac{\sum_{l=1}^{N_s} w_{(j,l)} \cdot \boldsymbol{\theta}_{(j-1,l)}}{\sum_{l=1}^{N_s} w_{(j,l)}} \tag{16}$$

where β_j is the current scaling coefficient, which is initialized as $\beta_0 = 2.4/\sqrt{M}$ with M the amount of input parameters equal to 50 for this study.

4. Perform a re-sampling from the current set of samples where each chain l has a probability p_l

$$p_l = \frac{w_{(j,l)}}{\sum_{n=1}^{N_s} w_{(j,n)}} \tag{17}$$

5. Repeat $N_a = 50$ times:

   (a) Draw a new set of samples θ^c from a normal distribution centered at the currently re-sampled set of samples $(\theta_{(j,l)}^c)$ with covariance matrix Σ_j

   (b) Set $\theta_{(j,l)}^c = \theta^c$ if $r < \frac{f(D|M,\theta^c)^{q_j}}{f(D|M,\theta_{(j,l)}^c)^{q_j}}$ with r drawn from a uniform distribution with bounds $[0, 1]$. otherwise the chain remains stationary.

(c) Recompute weights: (modification 1)

$$w(\boldsymbol{\theta}_{j,l}^c) = f(D|M, \boldsymbol{\theta}_{(j,l)}^c)^{q_j - q_{j-1}} \tag{18}$$

6. Adapt the scaling parameter β as follows: (modification 3)

$$\beta_{j+1} = \beta_j \cdot e^{\frac{p_{acr} - t_{acr}}{\sqrt{N_a}}} \tag{19}$$

where p_{acr} is the current acceptance rate and t_{acr} is the target acceptance rate equal to 0.23.

7. if $q_j = 1$, stop iteration; otherwise set $j = j + 1$ and continue with 1.

B   Results original TMCMC

To illustrate that the TMCMC algorithm without the adaptation proposed by this study leads to premature convergence, the original algorithm was executed for each of the specimens. As can be seen from Figures 8 and 9, all chains are concentrated around one solution.

Figure 8: Results of the original TMCMC: reference strain. Normalized histograms of the posterior distribution samples.
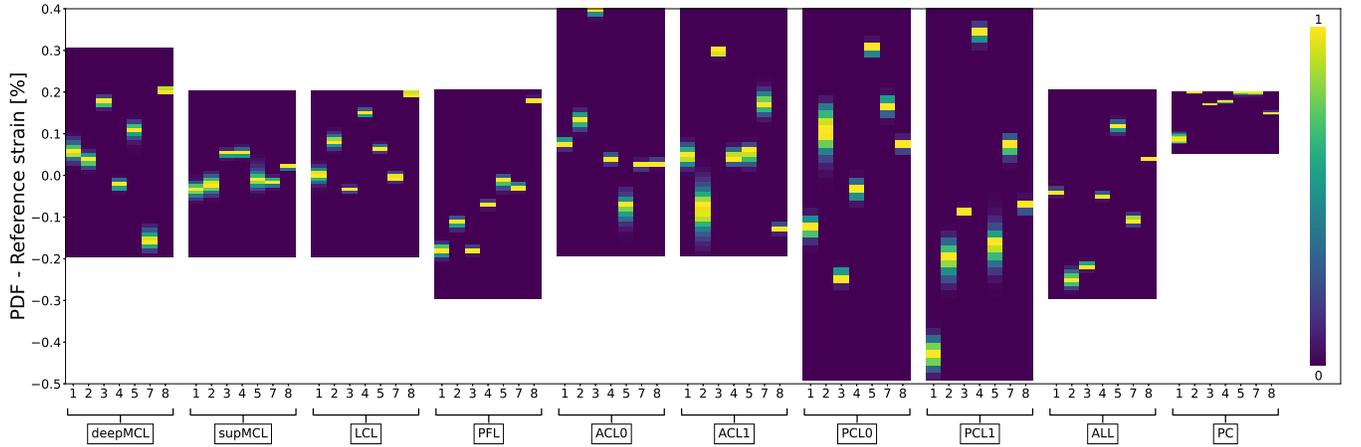

21

Figure 9: Results of the original TMCMC: linear stiffness. Normalized histograms of the posterior distribution samples.