Model-based three-material decomposition in dual-energy CT using the volume conservation constraint

Stephen Z. Liu[1], **Matthew Tivnan**[1], **Greg M. Osgood**[2], **Jeffrey H. Siewerdsen**[1], **J. Webster Stayman**[1], **Wojciech Zbijewski**[1,*]

[1]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205, USA

[2]Department of Orthopaedic Surgery, Johns Hopkins University, Baltimore, MD 21205, USA

Abstract

Objective: We develop a model-based optimization algorithm for "one-step" Dual-Energy (DE) CT decomposition of three materials directly from projection measurements.

Approach: Since the three-material problem is inherently undetermined, we incorporate the volume conservation principle (VCP) as a pair of equality and nonnegativity constraints into the objective function of the recently reported Model-Based Material Decomposition (MBMD). An optimization algorithm (Constrained MBMD, CMBMD) is derived that utilizes voxel-wise separability to partition the volume into a VCP-constrained region solved using interior-point iterations, and an unconstrained region (air surrounding the object, where VCP is violated) solved with conventional two-material MBMD. CMBMD is validated in simulations and experiments in application to bone composition measurements in the presence of metal hardware using DE Cone-Beam CT (CBCT). A kV-switching protocol with non-coinciding low- and high-energy (LE and HE) projections was assumed. CMBMD with decomposed base materials of cortical bone, fat, and metal (Titanium, Ti) is compared to MBMD with (i) fat-bone and (ii) fat-Ti bases.

Main Results: Three-material CMBMD exhibits a substantial reduction in metal artifacts relative to the two-material MBMD implementations. The accuracies of cortical bone volume fraction estimates are markedly improved using CMBMD, with ~5–10x lower Normalized Root Mean Squared Error (NRMSE) in simulations with anthropomorphic knee phantoms (depending on the complexity of the metal component) and ~2–2.5x lower in an experimental test-bench study.

Significance: In conclusion, we demonstrated one-step three-material decomposition of DE CT using volume conservation as an optimization constraint. The proposed method might be applicable to DE applications such as bone marrow edema imaging (fat-bone-water decomposition) or multi-contrast imaging, especially on CT/CBCT systems that do not provide coinciding LE and HE ray paths required for conventional projection-domain DE.

Keywords

cone-beam CT; dual-energy CT; constrained optimization; model-based reconstruction; material decomposition; metal artifacts; bone mineral density

---

*wzbijewski@jhu.edu .

1. Introduction

Dual-energy (DE) CT is of interest in many clinical imaging applications due to its ability to provide material discrimination (Johnson *et al* 2007, Nicolaou *et al* 2012, Patino *et al* 2016). Algorithms for DE material decomposition can be divided into three categories: i) image-domain decomposition (IDD), ii) projection-domain decomposition (PDD), and iii) one-step (i.e., optimization-based) decomposition. In IDD, material volume fractions are estimated (e.g., by voxel-wise inversion of a decomposition matrix) from individual reconstructions of the low-energy (LE) and high-energy (HE) channel data (Goodsitt *et al* 1987, Maaß *et al* 2009, Zbijewski *et al* 2014). The primary advantage of IDD is that it is applicable to most DE acquisition protocols; the potential disadvantage is that it requires additional beam hardening pre-correction (Joseph and Spital 1978) to achieve accurate material density maps. PDD addresses the beam hardening issue by the direct estimation of material line integrals from LE and HE projections, typically by means of pre-calibrated polyenergetic lookup tables (Alvarez and Macovski 1976, Zhao *et al* 2021). However, PDD cannot be applied to DE sampling patterns that do not provide geometrically matched ray paths between two spectral channels, e.g., dual/triple-source scanner configurations (Macari *et al* 2010, Kaza *et al* 2012, Liu *et al* 2019, 2020b) and single-source image acquisitions with slow kV-modulation (Szczykutowicz and Chen 2010, Rigie and La Riviere 2014, Sisniega *et al* 2014). One-step decomposition overcomes this limitation of PDD. Decomposition and reconstruction are performed simultaneously by optimizing a data discrepancy function between the measured DE data and a polyenergetic forward model (Long and Fessler 2014, Barber *et al* 2016, Mory *et al* 2018, Tilley *et al* 2019). The object is typically parameterized as the product of unknown material densities and their energy-dependent mass-attenuation coefficients. The forward model can capture arbitrary acquisition geometries, so matching ray paths are no longer needed.

In principle, the physics of DE imaging permits a maximum of two base materials. However, three-material decompositions are desirable in many clinical scenarios, such as contrast-enhanced DE imaging of the liver (Wortman *et al* 2018), bone mineral density (BMD) measurements (Goodsitt *et al* 1987, Wichmann *et al* 2014), and detection of bone marrow edema (Ali *et al* 2018, Müller *et al* 2020). Since direct estimation of three materials from DE data is inherently underdetermined, the Volume Conservation Principle (VCP, Goodsitt *et al* 1987, Liu *et al* 2009, Long and Fessler 2014) is often used as the constraint to enable such decomposition. (Note that even with more energy channels, three-material decomposition is ill-posed in the absence of K-edges, because most tissues belong to the span of two bases, namely photoelectric and Compton effects, Alvarez and Macovski 1976). VCP is typically posed as: i) an equality constraint that sums volume fractions of all base materials to one at each image voxel, and ii) a nonnegativity constraint for all fractions. As these constraints are specified in the reconstruction domain, VCP is incompatible with PDD. Therefore, three-material decomposition is typically achieved using either IDD (Granton *et al* 2008, Ding *et al* 2018, Jiang *et al* 2020) or a hybrid approach, whereby a preliminary two-material PDD is followed by image-domain basis change that applies VCP (Yu *et al* 2009, Liu *et al* 2022). Here, we propose an algorithm that incorporates VCP into the one-step decomposition framework to provide direct estimation of volume fractions of three materials from DE data.

This method aims to combine the benefits of VCP-constrained IDD, namely the ability to include a third material through VCP and the applicability to geometries with unmatched LE and HE ray paths, with the capacity to utilize the projection information obtained in two spectral channels to account for beam hardening, similar to PDD but without the potentially noise-amplifying multi-stage processing of the hybrid approach.

Our algorithm extends the previously reported Model-Based Material Decomposition (MBMD, Tilley *et al* 2019) by applying a constrained optimization framework to impose VCP on the Penalized-Weighted Least-Squares (PWLS) polyenergetic objective function of MBMD. Compared to the Poisson likelihood model used in other recently proposed constrained one-step methods for spectral CT (Long and Fessler 2014, Barber *et al* 2016), the PWLS objective permits straightforward inclusion of system blur and the associated noise correlations, potentially enabling simultaneous model-based deblurring and decomposition (Wang *et al* 2021). The proposed Constrained MBMD (CMBMD) is formulated as a partitioned optimization scheme, where each voxel receives either constrained or unconstrained updates depending on its location. This heuristic strategy was developed because any realistic object is likely to contain regions where the constraint needs to be relaxed because of the presence of materials – for example, air – that cannot be decomposed into the CMBMD three-material base set without violating the VCP (we assume that the region where VCP is applicable is known *a priori*.) Other proposed solutions to this challenge include: i) estimating best-matching material n-tuples in each voxel as part of the optimization objective (Long and Fessler 2014), or ii) replacing VCP with other constraints such as the nonnegativity of a virtual monochromatic image synthesized from the decomposition (Barber *et al* 2016).

An early version of CMBMD with initial validation results was previously presented in a conference publication (Liu *et al* 2021a). This paper provides a substantially expanded derivation of the algorithm and additional implementation details, and reports on new feasibility studies in simulations and physical experiments involving quantitative DE cone-beam CT (CBCT) imaging in the presence of metal implants. For smaller components made of relatively low-density metals, such as Titanium, photon starvation and noise amplification effects are minimal. In such cases, beam hardening is among the dominant sources of artifacts. One-step DE decomposition should be fairly robust to these distortions, because it inherently accounts for polyenergetic attenuation. However, since the objective of DE imaging is typically not to minimize metal artifacts, but rather to discriminate other materials (e.g., mineralized bone from bone marrow), it is often undesirable to use the metal as one of the two decomposition bases in an unconstrained DE inversion. We show that such scenarios can be efficiently handled by adding the metal as a third base enabled by VCP and solved using CMBMD. Our studies specifically address the application of DE CBCT to quantify bone mineral density (BMD).

2.   Algorithms

Table 1 summarizes the key variables used in algorithm derivation. Matrices are denoted in bold upright (non-italic) uppercase and vectors are in bold italic lowercase; their components

are represented by dropping the bold-face and adding subscripts. Constants are always in non-bold letters; among them, the dimensions of variables are written using script letters.

The algorithm derivations are presented below using a fairly general notation that does not explicitly restrict the number of energy channels and base materials. However, the two numbers are related to each other by the need to yield a well determined decomposition problem. Therefore, for a DE acquisition, the number of base materials $\mathcal{k}$ is 2 for the unconstrained MBMD and 3 for constrained MBMD with VCP. Typically, the total number of projections rays $\mathcal{i}$ is approximately equally split between LE and HE spectra.

The algorithmic methods are described as follows. Sec. 2.1 reviews the multi-energy imaging system model employed in the decomposition. Sec 2.2 introduces the unconstrained MBMD algorithm, including a new voxel-separable update term (in contrast to the voxel- and material-separable updates in Tilley *et al* 2019). Sec. 2.3 shows how MBMD can be expanded to include the volume conservation constraint. Finally, Sec. 2.4 provides a detailed outline of the proposed CMBMD framework, which integrates the constrained optimization in Sec. 2.3 with an optimization strategy that enables regional application of the constraints. (Note that since we anticipate that almost any practical application of constrained decomposition will require a partitioned approach similar to Sec 2.4, we apply the term "CMBMD" to this regional formulation instead of the general constrained algorithm of Sec. 2.3).

2.1. Polyenergetic Forward Model

Both the constrained and unconstrained algorithms employ a discretized polyenergetic forward model which assumes that the object consists of $\mathcal{k}$ base materials (Tilley *et al* 2019):

$$\bar{y}_i(\boldsymbol{\pi}) = \sum_{\epsilon = 1}^{e} s_i(\epsilon)\Delta\epsilon\exp\left(\sum_{\kappa = 1}^{\mathcal{k}} -\mathbf{A}_i\mu^{\{\kappa\}}(\epsilon)\boldsymbol{\pi}^{\{\kappa\}}\right) \tag{1}$$

The spatial distribution of volume fractions of the base material κ is encapsulated by vector $\boldsymbol{\pi}^{\{\kappa\}} \in \mathbb{R}^{\mathcal{j}}$. The intensity measured by the projection ray i is given by $\bar{y}_i(\boldsymbol{\pi})$. $\mathbf{A}_i \in \mathbb{R}^{1 \times \mathcal{j}}$ is the ith row of the projection matrix $\mathbf{A} \in \mathbb{R}^{\mathcal{i} \times \mathcal{j}}$. The x-ray spectrum and detector response for ray i at the discretized energy ϵ (ranging from 1 to e keV, where e is the tube potential) are included in $s_i(\epsilon)$. ϵ is the energy bin width. In a DE image acquisition, some of the rays are obtained using an HE spectrum and some using an LE spectrum; the spectra are appropriately included in $s_i(\epsilon)$ of each ray.

As shown by Tilley *et al* 2019, the forward model for an entire DE image acquisition of $\mathcal{i}$ rays $\bar{\boldsymbol{y}}(\boldsymbol{\pi}) \in \mathbb{R}^{\mathcal{i}}$ can be re-written analogous to a system with (linear) blur:

$$\bar{\boldsymbol{y}}(\boldsymbol{\pi}) = \begin{bmatrix} \bar{y}_1(\boldsymbol{\pi}) & \cdots & \bar{y}_{\ell}(\boldsymbol{\pi}) \end{bmatrix}^T$$

$$= \begin{bmatrix} \mathbf{B}_1 & & \\ & \ddots & \\ & & \mathbf{B}_{\ell} \end{bmatrix} \exp\left(- \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_{\ell} \end{bmatrix} \boldsymbol{\pi} \right) = \mathbf{B}\exp(-\mathbf{M}\boldsymbol{\pi}) \tag{2}$$

where $\boldsymbol{\pi} \in \mathbb{R}^{\jmath \hbar}$ is the column concatenation of the volume fractions of all $\hbar$ bases:

$$\boldsymbol{\pi} := \begin{bmatrix} \boldsymbol{\pi}^{\{1\}} \\ \vdots \\ \boldsymbol{\pi}^{\{\hbar\}} \end{bmatrix} = \begin{bmatrix} \left[\boldsymbol{\pi}_1^{\{1\}} & \cdots & \boldsymbol{\pi}_{\jmath}^{\{1\}} \right]^T \\ \vdots \\ \left[\boldsymbol{\pi}_1^{\{\hbar\}} & \cdots & \boldsymbol{\pi}_{\jmath}^{\{\hbar\}} \right]^T \end{bmatrix} \tag{3}$$

and $\mathbf{M}_i \in \mathbb{R}^{e \times \jmath \hbar}$ and $\mathbf{B}_i \in \mathbb{R}^{1 \times e}$ are respectively defined as (with zero entries omitted):

$$\mathbf{M}_i := \begin{bmatrix} \mu^{\{1\}}(\epsilon_1) & \cdots & \mu^{\{\hbar\}}(\epsilon_1) \\ \vdots & \vdots & \vdots \\ \mu^{\{1\}}(\epsilon_e) & \cdots & \mu^{\{\hbar\}}(\epsilon_e) \end{bmatrix} \begin{bmatrix} \mathbf{A}_i & & \\ & \ddots & \\ & & \mathbf{A}_i \end{bmatrix} \tag{4a}$$

$$\mathbf{B}_i := \begin{bmatrix} s_i(\epsilon_1) & \cdots & s_i(\epsilon_e) \end{bmatrix} \Delta \epsilon \tag{4b}$$

Essentially, the application of matrix $\mathbf{M} \in \mathbb{R}^{\ell e \times \jmath \hbar}$ computes the line integrals of the object for every projection ray i at each discretized beam energy ϵ. The matrix $\mathbf{B} \in \mathbb{R}^{\ell \times \ell e}$ applies the summation over the energy bins weighted by the system spectral response $s_i(\epsilon)$, equivalent to the outer energy integral in Eq. (1).

2.2. Unconstrained MBMD

This section briefly recapitulates the unconstrained MBMD algorithm. Importantly for the proposed CMBMD approach, we introduce a new optimization scheme that does not rely on material-wise separable surrogates originally employed in Tilley *et al* 2019.

In MBMD, the reconstruction of the unknown base material fraction maps $\boldsymbol{\pi}$ is formulated as the minimization of a PWLS objective function $\Phi(\boldsymbol{\pi})$ (Tilley *et al* 2017, 2019):

$$\hat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi} \geq \mathbf{0}} \Phi(\boldsymbol{\pi})$$

$$= \arg\min_{\boldsymbol{\pi} \geq \mathbf{0}} \frac{1}{2} \| \boldsymbol{y} - \mathbf{B}\exp(-\mathbf{M}\boldsymbol{\pi}) \|_{\Sigma_{\boldsymbol{y}}^{-1}}^2 + r(\boldsymbol{\pi}) \tag{5}$$

$$= \arg\min_{\boldsymbol{\pi} \geq \mathbf{0}} l(\boldsymbol{\pi}) + r(\boldsymbol{\pi})$$

where $l(\boldsymbol{\pi})$ is the data fidelity term and $r(\boldsymbol{\pi})$ is the regularization term. We assume that the measurements $\boldsymbol{y} \in \mathbb{R}^{\ell}$ follow an uncorrelated Gaussian distribution with mean $\bar{\boldsymbol{y}}(\boldsymbol{\pi})$ and covariance $\Sigma_{\boldsymbol{y}} \in \mathbb{R}^{\ell \times \ell}$, where $\Sigma_{\boldsymbol{y}}^{-1}$ is diagonal with entries equal to the reciprocal of $\boldsymbol{y}$. For $r(\boldsymbol{\pi})$, we use the following general form:

$$r(\boldsymbol{\pi}) = \sum_{\kappa = 1}^{\hbar} r^{\{\kappa\}}\left(\boldsymbol{\pi}^{\{\kappa\}}\right) = \sum_{\kappa = 1}^{\hbar} \beta^{\{\kappa\}}\psi\left(\boldsymbol{\pi}^{\{\kappa\}}\right) \tag{6}$$

Throughout this work, we apply the same smoothness penalty $\psi(\boldsymbol{\pi}^{\{\kappa\}})$ for each of the base material images; however, the penalty strength $\beta^{\{\kappa\}}$ is different for each basis κ.

We optimize $\Phi(\boldsymbol{\pi})$ using separable surrogates. For each iteration n, the current estimate $\boldsymbol{\pi}^{(n)}$ is updated by taking a Newton-Raphson (NR) step towards the optimum of an iteration-dependent surrogate function $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ that majorizes $\Phi(\boldsymbol{\pi})$. In Tilley *et al* 2019, $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ was designed such that the resulting update equation was separable in both voxels and materials (i.e., entry-wise separable). Here, however, we applied the additive convexity technique (Erdogan and Fessler 1999) to obtain $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ that is separable only in voxels, but not in materials (see derivation in Appendix I). This yields the following NR update equation:

$$\boldsymbol{\pi}^{(n+1)} = \boldsymbol{\pi}^{(n)} - \alpha_{\mathrm{NR}}{}^{(n)}\varDelta\boldsymbol{\pi}_{\mathrm{NR}}{}^{(n)} \tag{7a}$$

$$\mathscr{H}_{\varphi}{}^{(n)}\varDelta\boldsymbol{\pi}_{\mathrm{NR}}{}^{(n)} = \nabla\varphi\left(\boldsymbol{\pi}^{(n)}; \boldsymbol{\pi}^{(n)}\right) = \nabla\Phi\left(\boldsymbol{\pi}^{(n)}\right) \tag{7b}$$

where $\nabla\varphi\left(\boldsymbol{\pi}^{(n)}; \boldsymbol{\pi}^{(n)}\right) \in \mathbb{R}^{\mathscr{I}\hbar}$ and $\mathscr{H}_{\varphi}{}^{(n)} \in \mathbb{R}^{\mathscr{I}\hbar \times \mathscr{I}\hbar}$ denote the gradient and the Hessian of $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ evaluated at $\boldsymbol{\pi} = \boldsymbol{\pi}^{(n)}$, respectively. Importantly, due to the voxel-wise separability, $\nabla\varphi(\boldsymbol{\pi}^{(n)}; \boldsymbol{\pi}^{(n)})$ is equivalent to the gradient of the original objective $\nabla\Phi\left(\boldsymbol{\pi}^{(n)}\right) \in \mathbb{R}^{\mathscr{I}\hbar}$ evaluated at the current iteration; $\mathscr{H}_{\varphi}{}^{(n)}$ is block-diagonal and is equivalent to the Hessian of the original objective $\mathscr{H}_{\Phi}{}^{(n)} \in \mathbb{R}^{\mathscr{I}\hbar \times \mathscr{I}\hbar}$ but with all cross-voxel second-order derivatives set to zero (Tivnan *et al* 2020). Specifically:

$$\begin{aligned} \nabla\Phi\left(\boldsymbol{\pi}^{(n)}\right) &= \nabla l\left(\boldsymbol{\pi}^{(n)}\right) + \nabla r\left(\boldsymbol{\pi}^{(n)}\right) \\ &= \mathbf{M}^T\mathbf{D}\left\{\exp\left(-\mathbf{M}\boldsymbol{\pi}^{(n)}\right)\right\}^T\mathbf{B}^T\boldsymbol{\Sigma}_{\boldsymbol{y}}{}^{-1}\left(\boldsymbol{y} - \overline{\boldsymbol{y}}\left(\boldsymbol{\pi}^{(n)}\right)\right) + \nabla r\left(\boldsymbol{\pi}^{(n)}\right) \end{aligned} \tag{8a}$$

$$\begin{aligned} \mathscr{H}_{\Phi}{}^{(n)} &= \mathscr{H}_{l}{}^{(n)} + \mathscr{H}_{r}{}^{(n)} \\ &= \nabla^T \nabla l\left(\boldsymbol{\pi}^{(n)}\right) + \nabla^T \nabla r\left(\boldsymbol{\pi}^{(n)}\right) \\ &= \mathbf{M}^T\mathbf{D}\left\{\exp\left(-\mathbf{M}\boldsymbol{\pi}^{(n)}\right)\right\}^T\mathbf{B}^T\boldsymbol{\Sigma}_{\boldsymbol{y}}{}^{-1}\mathbf{B}\mathbf{D}\left\{\exp\left(-\mathbf{M}\boldsymbol{\pi}^{(n)}\right)\right\}\mathbf{M} + \nabla^T \nabla r\left(\boldsymbol{\pi}^{(n)}\right) \end{aligned} \tag{8b}$$

where $\nabla r\left(\boldsymbol{\pi}^{(n)}\right) \in \mathbb{R}^{\mathscr{I}\hbar}$ and $\mathscr{H}_{r}{}^{(n)} \in \mathbb{R}^{\mathscr{I}\hbar \times \mathscr{I}\hbar}$ represent the gradient and the Hessian of the regularization term evaluated at $\boldsymbol{\pi} = \boldsymbol{\pi}^{(n)}$, respectively. Procedurally, $\mathscr{H}_{\Phi}(n)$ is obtained by: 1) forward projecting a volume of ones for each material κ (can be precomputed before iterations), 2) applying weight terms computed using the current estimate $\boldsymbol{\pi}^{(n)}$, and 3) backprojecting the weighted projections.

During the optimization, the NR search direction for each iteration n, $\varDelta\boldsymbol{\pi}_{\mathrm{NR}}{}^{(n)} \in \mathbb{R}^{\mathscr{I}\hbar}$, is first obtained by solving the linear system in Eq. (7b). The voxel- and material-wise separability

used in Tilley *et al* 2019 led to a diagonal linear system for $\boldsymbol{\pi}_{\mathrm{NR}}^{(n)}$, which could be solved by simply finding the reciprocals of all diagonal entries of the Hessian (i.e., coordinate descent). Here, as we do not apply material-wise separability, Eq. (7b) is block-diagonal and thus requires solving a dense (yet small) $k \times k$ linear system at every voxel j. However, this additional computational burden is acceptable compared to the entry-wise separable surrogate in Tilley *et al* 2019, because of the resulting improvement in the convergence rate (Tivnan *et al* 2020). To further accelerate convergence, momentum-based updates (Nesterov 2005) and Ordered Subsets (OS, Hudson and Larkin 1994) can be easily incorporated. The resulting algorithm is referred to as OS-NR.

The search direction $\boldsymbol{\pi}_{\mathrm{NR}}^{(n)}$ obtained from Eq. (7b) is then applied to compute the new estimate $\boldsymbol{\pi}^{(n+1)}$ for the next iteration following Eq. (7a). The step length $a_{\mathrm{NR}}^{(n)} \in (0, 1]$ can be either pre-defined or obtained from standard line search (Nocedal and Wright 2006).

2.3. Interior-Point Algorithm for MBMD with Volume Conservation Constraint

We now introduce VCP as a pair of linear equality and nonnegativity constraints for the PWLS problem in Eq. (5):

$$\widehat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi}} \Phi(\boldsymbol{\pi}) \quad \text{s.t.} \quad \begin{cases} \displaystyle\sum_{\kappa=1}^{k} \boldsymbol{\pi}^{\{\kappa\}} - 1 = \mathbf{C}\boldsymbol{\pi} - 1 = 0 \\ \boldsymbol{\pi} \geq \mathbf{0} \end{cases} \tag{9a}$$

$$\mathbf{C} := \begin{bmatrix} \mathbf{1}^T & & \\ & \ddots & \\ & & \mathbf{1}^T \end{bmatrix} = \begin{bmatrix} [1 \ \cdots \ 1] & & \\ & \ddots & \\ & & [1 \ \cdots \ 1] \end{bmatrix} \tag{9b}$$

where $\mathbf{0}$ and $\mathbf{1}$ are vectors of zeros and ones, respectively. The first constraint in Eq. (9a) sums the base material fractions at each voxel to one (the summation matrix $\mathbf{C} \in \mathbb{R}^{j \times jk}$ is defined in Eq. (9b)); the second constraint enforces the nonnegativity of $\boldsymbol{\pi}$.

To derive the update term for Eq. (9a), we first integrate the nonnegativity constraint into the objective function using the barrier technique (Fiacco and McCormick 1990):

$$\widehat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi}} \Gamma_\tau(\boldsymbol{\pi}) = \arg\min_{\boldsymbol{\pi}} \Phi(\boldsymbol{\pi}) - o_\tau(\boldsymbol{\pi}) \quad \text{s.t.} \quad \mathbf{C}\boldsymbol{\pi} - 1 = 0 \tag{10a}$$

$$o_\tau(\boldsymbol{\pi}) := \tau \sum_{j=1}^{j} \sum_{\kappa=1}^{k} \ln\left(\boldsymbol{\pi}_j^{\{\kappa\}}\right) = \tau \mathbf{1}^T \ln(\boldsymbol{\pi}) \tag{10b}$$

where $\Gamma_\tau(\boldsymbol{\pi})$ is the new objective consisting of $\Phi(\boldsymbol{\pi})$ and the Fiacco-McCormick log-barrier function $o_\tau(\boldsymbol{\pi})$ defined in Eq. (10b). The barrier strength $\tau > 0$ is adapted per iteration using the Predictor-Corrector (PC) framework (Shanno and Vanderbei 2000, Nocedal *et al* 2009, Zanelli *et al* 2020) as explained in Sec. 2.5 such that the updates remain within a close neighborhood about the central path.

Following Nocedal *et al* 2009, we can derive the Lagrangian $\mathscr{L}_\tau(\boldsymbol{\pi}, \lambda)$ of Eq. (10a):

$$\mathscr{L}_\tau(\boldsymbol{\pi}, \lambda) = \Gamma_\tau(\boldsymbol{\pi}) - \lambda^T(\mathbf{C}\boldsymbol{\pi} - \mathbf{1}) = \Phi(\boldsymbol{\pi}) - o_\tau(\boldsymbol{\pi}) - \lambda^T(\mathbf{C}\boldsymbol{\pi} - \mathbf{1}) \tag{11}$$

and its associated first-order Karush-Kuhn-Tucker (KKT) optimality conditions that define the primal-dual feasibilities and complementarity:

$$\nabla_{\boldsymbol{\pi}}\mathscr{L}_\tau(\boldsymbol{\pi}, \lambda) = \nabla\Phi(\boldsymbol{\pi}) - \mathbf{C}^T\lambda - \tau\frac{1}{\boldsymbol{\pi}} = \nabla\Phi(\boldsymbol{\pi}) - \mathbf{C}^T\lambda - \theta = \mathbf{0} \tag{12a}$$

$$\nabla_\lambda\mathscr{L}_\tau(\boldsymbol{\pi}, \lambda) = \mathbf{C}\boldsymbol{\pi} - 1 = 0 \tag{12b}$$

$$\theta := \tau\frac{1}{\boldsymbol{\pi}} \rightarrow \mathbf{D}\{\boldsymbol{\pi}\}\mathbf{D}\{\theta\}\mathbf{1} - \tau\mathbf{1} = \mathbf{0} \tag{12c}$$

$$\boldsymbol{\pi} \geq \mathbf{0} \tag{12d}$$

$$\theta \geq \mathbf{0} \tag{12e}$$

where vectors $\lambda \in \mathbb{R}^{\mathscr{I}}$ and $\theta \in \mathbb{R}^{\mathscr{I}\mathscr{k}}$ are the KKT multipliers, and $(1/\boldsymbol{\pi}) \in \mathbb{R}^{\mathscr{I}\mathscr{k}}$ denotes the vector of element-wise reciprocals of $\boldsymbol{\pi}$. The first three KKT equalities in Eqs. (12a) – (12c) respectively represent zeroing the gradient of $\mathscr{L}_\tau$ with respect to the primal-dual variables $\boldsymbol{\pi}$, λ and θ. (Note that the third KKT condition in Eq. (12c) can be obtained by setting $\theta := \tau/\boldsymbol{\pi}$ in $o_\tau(\boldsymbol{\pi})$, and then zeroing $\nabla_\theta\mathscr{L}_\tau$).

To estimate the unknown $\boldsymbol{\pi}$ and the associated KKT multipliers λ and θ, we approximate the KKT equalities in Eqs. (12a) – (12c) using the nonlinear Interior-Point (IP) iterations (the nonnegativities in Eqs (12d) – (12e) are ascertained by the selection of τ described in Sec. 2.5):

$$\begin{bmatrix}\boldsymbol{\pi}^{(n+1)} \\ \theta^{(n+1)} \\ \lambda^{(n+1)}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\pi}^{(n)} \\ \theta^{(n)} \\ \lambda^{(n)}\end{bmatrix} - \alpha_{\mathrm{IP}}(n)\begin{bmatrix}\Delta\boldsymbol{\pi}_{\mathrm{IP}}^{(n)} \\ \Delta\theta_{\mathrm{IP}}^{(n)} \\ \Delta\lambda_{\mathrm{IP}}^{(n)}\end{bmatrix} \tag{13a}$$

$$\mathscr{J}\nabla\mathscr{L}_\tau^{(n)}\begin{bmatrix}\Delta\boldsymbol{\pi}_{\mathrm{IP}}^{(n)} \\ \Delta\theta_{\mathrm{IP}}^{(n)} \\ \Delta\lambda_{\mathrm{IP}}^{(n)}\end{bmatrix} = \nabla\mathscr{L}_\tau\left(\begin{bmatrix}\boldsymbol{\pi}^{(n)} \\ \theta^{(n)} \\ \lambda^{(n)}\end{bmatrix}\right) \tag{13b}$$

where the constant $\alpha_{\mathrm{IP}}^{(n)} \in (0, 1]$ denotes the step length, $\nabla\mathscr{L}_\tau\left(\begin{bmatrix}\boldsymbol{\pi}^{(n)}; \theta^{(n)}; \lambda^{(n)}\end{bmatrix}\right) \in \mathbb{R}^{(\mathscr{I}\mathscr{k} + \mathscr{I}\mathscr{k} + \mathscr{I})}$ is the gradient of $\mathscr{L}_\tau(\boldsymbol{\pi}, \theta, \lambda)$ at a given τ evaluated at the current iteration (i.e., the left-hand-sides of the KKT equalities in Eqs. (12a) – (12c)):

$$\nabla \mathscr{L}_\tau \left(\begin{bmatrix} \boldsymbol{\pi}^{(n)} \\ \boldsymbol{\theta}^{(n)} \\ \lambda^{(n)} \end{bmatrix} \right) = \begin{bmatrix} \nabla_{\boldsymbol{\pi}} \mathscr{L}_\tau \\ \nabla_{\boldsymbol{\theta}} \mathscr{L}_\tau \\ \nabla_\lambda \mathscr{L}_\tau \end{bmatrix} = \begin{bmatrix} \nabla \varphi \left(\boldsymbol{\pi}^{(n)}; \boldsymbol{\pi}^{(n)} \right) - \mathbf{C}^T \lambda^{(n)} - \boldsymbol{\theta}^{(n)} \\ \mathbf{D}\{\boldsymbol{\pi}^{(n)}\} \mathbf{D}\{\boldsymbol{\theta}^{(n)}\} \mathbf{1} - \tau \mathbf{1} \\ \mathbf{C}\boldsymbol{\pi}^{(n)} - \mathbf{1} \end{bmatrix} \tag{14}$$

and $\mathscr{J}_{\nabla \mathscr{L}_\tau}{}^{(n)} \in \mathbb{R}^{(\jmath\hbar + \jmath\hbar + \jmath) \times (\jmath\hbar + \jmath\hbar + \jmath)}$ is the corresponding Jacobian (independent of τ):

$$\mathscr{J}_{\nabla \mathscr{L}_\tau}{}^{(n)} = \nabla^T \nabla \mathscr{L}_\tau \left(\begin{bmatrix} \boldsymbol{\pi}^{(n)} \\ \boldsymbol{\theta}^{(n)} \\ \lambda^{(n)} \end{bmatrix} \right) = \begin{bmatrix} \mathscr{H}_\varphi{}^{(n)} & -\mathbf{C}^T & -\mathbb{I} \\ \mathbf{D}\{\boldsymbol{\theta}^{(n)}\} & \mathbf{0} & \mathbf{D}\{\boldsymbol{\pi}^{(n)}\} \\ \mathbf{C} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{15}$$

where $\mathbb{I} \in \mathbb{R}^{\jmath\hbar \times \jmath\hbar}$ is the identity matrix.

Similar to the procedures for the unconstrained MBMD, the primal-dual variables are simultaneously updated by applying the IP steps $\Delta\boldsymbol{\pi}_{\text{IP}}{}^{(n)} \in \mathbb{R}^{\mathcal{I}\hbar}$, $\Delta\theta_{\text{IP}}{}^{(n)} \in \mathbb{R}^{\mathcal{I}\hbar}$ and $\Delta\lambda_{\text{IP}}{}^{(n)} \in \mathbb{R}^{\mathcal{I}}$ that can be obtained by solving the linear system in Eq. (13b). We adopt the same iteration-dependent voxel-wise separable surrogate $\varphi(\boldsymbol{\pi}, \boldsymbol{\pi}^{(n)})$ as in the unconstrained MBMD to compute $\nabla \mathscr{L}_\tau \left(\left[\boldsymbol{\pi}^{(n)}; \theta^{(n)}; \lambda^{(n)} \right] \right)$ and $\mathscr{J}_{\nabla \mathscr{L}_\tau}{}^{(n)}$ associated with the primal-dual variables $\boldsymbol{\pi}$, $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$. In this manner, Eq. (13b) becomes block-diagonal per voxel after proper rearrangement (similar to Eq. (7b) for the unconstrained MBMD) such that $\boldsymbol{\pi}_{\text{IP}}{}^{(n)}$, $\boldsymbol{\theta}_{\text{IP}}{}^{(n)}$ and $\boldsymbol{\lambda}_{\text{IP}}{}^{(n)}$ can be obtained by solving a small $(\hbar + \hbar + 1) \times (\hbar + \hbar + 1)$ linear system at each voxel. In our implementation, row reduction provides further reduction of the size of the system in Eq. (13b).

2.4. CMBMD: Constrained Model-Based Decomposition with Regional Constraints

It is unlikely that VCP with one specific set of $\hbar$ materials will hold globally for a realistic object. Even if the base materials are chosen such that their attenuation curves span the entire range of tissue attenuations present in the object, the surrounding air – which is typically excluded from the base set – will violate the VCP constraints, as it should have voxel values of zero in all material images. Therefore, we propose a heuristic strategy whereby the volume is partitioned into a constrained region denoted $\boldsymbol{\pi}_{\mathscr{C}} \in \mathbb{R}^{\rho\hbar}$ (ρ voxels in total) and an unconstrained region $\boldsymbol{\pi}_{\mathscr{U}} \in \mathbb{R}^{\varphi\hbar}$ (φ voxels in total) based on some prior knowledge. Due to the application of VCP, one can estimate the volume fraction of one more material in $\boldsymbol{\pi}_{\mathscr{C}}$ than in $\boldsymbol{\pi}_{\mathscr{U}}$. We assume that the first $\hbar - 1$ bases are solved for in both regions, whereas the $\hbar^{\text{th}}$ material is only estimated in $\boldsymbol{\pi}_{\mathscr{C}}$ and set to zero for all voxels in $\boldsymbol{\pi}_{\mathscr{U}}$ and its gradient.

To jointly estimate $\boldsymbol{\pi}_{\mathscr{C}}$ and $\boldsymbol{\pi}_{\mathscr{U}}$, we exploit voxel-wise separability and update the material fractions in $\boldsymbol{\pi}_{\mathscr{C}}$ using the IP iterations of Eq. (13), whereas the NR iterations of conventional MBMD (Eq. (7)) are applied in $\boldsymbol{\pi}_{\mathscr{U}}$. Since both the constrained and unconstrained algorithm

employ the same surrogate gradient and Hessian, the two regions are coupled in the computation of these two terms. This involves forward- and backprojection of a composite volume π obtained as a union of $\pi_{\mathscr{C}}$ and $\pi_{\mathscr{U}}$.

The pseudocode of this locally constrained optimization approach is shown in Table 2. Each iteration begins by computing $\nabla \varphi(\pi^{(n)}; \pi^{(n)})$ and $\mathscr{H}_{\varphi}^{(n)}$ from the current solution for the entire volume $\pi^{(n)}$ (row 3a in Table 2). This is the only instance in the algorithm where global operations of forward and backprojection are performed. The remaining steps are applied to individual voxels, now partitioned into the constrained and unconstrained sets $\pi_{\mathscr{C}}^{(n)}$ and $\pi_{\mathscr{U}}^{(n)}$ by simple index selection within vector $\pi^{(n)}$ (row 3b in Table 2). The gradient and Hessian are partitioned in the same manner.

The voxels in $\pi_{\mathscr{U}}^{(n)}$ are updated using the NR steps of Eq. (7) (rows 3c and 3d in Table 2). It is again worth emphasizing that this involves solving a simple linear system of equations in each voxel; the equation coefficients are obtained from the gradient and Hessian of $\pi^{(n)}$ in that voxel.

The update of the constrained voxels $\pi_{\mathscr{C}}^{(n)}$ is more complex than for $\pi_{\mathscr{U}}^{(n)}$, because it involves iteration-dependent optimization of the barrier strength $\tau^{(n)}$ using the PC framework (Nocedal *et al* 2009, Zanelli *et al* 2020). For each voxel in $\pi_{\mathscr{C}}^{(n)}$, Eq. (13b) is first solved with $\tau = 0$ to obtain preliminary "predictor" directions $[\Delta\pi_{\mathscr{C}}*; \Delta\theta*; \Delta\lambda*] \in \mathbb{R}^{(p\mathscr{k} + p\mathscr{k} + p)}$ (row 3e in Table 2). The predictors are then used to estimate $\tau^{(n)}$ (row 3f in Table 2). Eq. (13b) is then solved again using $\tau^{(n)}$ and the predictors to obtain the desired "centered corrector" directions $\left[\Delta\pi_{\mathscr{C}\mathrm{PCIP}}^{(n)}; \Delta\theta_{\mathrm{PCIP}}^{(n)}; \Delta\lambda_{\mathrm{PCIP}}^{(n)}\right] \in \mathbb{R}^{(p\mathscr{k} + p\mathscr{k} + p)}$ (row 3g in Table 2), which are then applied to update $\pi_{\mathscr{C}}^{(n)}$ (row 3h in Table 2). Although the PC steps require solving the linear system of search directions twice for each voxel, this can be efficiently parallelized and is computationally trivial compared to forward and back-projections of $\pi^{(n)}$ involved in obtaining $\nabla \varphi(\pi^{(n)}; \pi^{(n)})$ and $\mathscr{H}_{\varphi}(n)$.

We applied momentum-based acceleration and OS to the CMBMD optimization in Table 2; these modifications have been implemented in the standard manner and are omitted from Table 2 for clarity.

2.5. IDD with Volume Conservation Constraint

To investigate the benefits of beam hardening correction inherent to CMBMD, we performed a comparison to VCP-constrained IDD (Granton *et al* 2008, Ding *et al* 2018). This approach also achieves three-material decomposition, but since it acts on individual FDK reconstructions of LE and HE projection data, it does not account for beam hardening unless additional *ad hoc* pre-processing is applied. The IDD decomposition was performed by minimizing a constrained least-squares problem in each voxel:

$$\hat{\boldsymbol{\pi}}_j = \arg\min_{\boldsymbol{\pi}_j} \frac{1}{2} \left\| \mathbf{U}\boldsymbol{\pi}_j - \begin{bmatrix} z_j^{\{LE\}} \\ z_j^{\{HE\}} \end{bmatrix} \right\|_2^2 \quad \text{s.t.} \quad \begin{cases} \mathbf{1}^T \boldsymbol{\pi}_j - 1 = 0 \\ \boldsymbol{\pi}_j \geq \mathbf{0} \end{cases} \tag{16}$$

where $z_j^{\{LE\}}$ and $z_j^{\{LE\}}$ are the jth voxel of LE and HE FDK reconstructions, respectively, and $\mathbf{U} \in \mathbb{R}^{2 \times 3}$ contains the linear attenuations of the three materials at effective energies of the LE and HE beams. Since the problem in Eq. (16) is convex, it was efficiently optimized using an IP algorithm analogous to Sec. 2.3.

3. Experimental Validation

3.1 Simulation Study

We evaluate CMBMD in simulations emulating quantitative DE CBCT of bone composition in the presence of surgical implants made from Titanium (Ti). Three-material decomposition is performed to discriminate adipose tissue (yellow marrow), bone, and Ti under the assumption that the inclusion of Ti in the CMBMD base material set will partly mitigate metal artifacts.

3.1.1. Digital Phantom and Metal Implants—Fig. 1 shows the multi-material phantom used in the simulations. It was obtained by segmentation of an CBCT scan of a knee. Voxel size was $0.25 \times 0.25 \times 0.25$ mm^3. Volume dimensions were $150 \times 150 \times 30$ mm^3. The relatively narrow longitudinal extent of the volume was consistent with x-ray beam collimation assumed in the simulations and benchtop experiments (see below). The phantom consisted of an outer adipose (fat) layer surrounding a layer of muscle tissue (simulated as water) and two bone regions – tibia and fibula.

The tibia and fibula were modeled as a mixture of cortical bone, fat, and water. First, thresholding was applied to identify voxels containing calcified tissue. The volume concentration of cortical bone in these voxels was assumed proportional to the attenuation coefficient, with fractions of one assigned to the most attenuating voxels in this set. The remaining volume of each calcified voxel was filled with fat and water at a ratio of 9:1. Voxels interior to tibia and fibula that fell below the calcified tissue threshold were simulated as a mixture of water and fat only, also at a ratio of 9:1 by volume, consistent with a dominant contribution of yellow marrow (Zankl *et al* 2007). Linear attenuation coefficients and densities of all phantom materials were obtained from Spektr3 (Punnoose *et al* 2016).

Two Titanium implants (uniform density of 4,510 mg/mL) emulating an intramedullary (IM) nail (14 mm outer diameter, 8 mm inner diameter) and a locking screw (3 mm diameter, 60 mm length) were assembled on the tibia as illustrated in Fig. 1. The Ti construct was surrounded by four target inserts (14 mm diameter each) to assess BMD accuracy in the vicinity of metal. Cortical bone concentrations in the inserts were 100 mg/mL, 150 mg/mL, 200 mg/mL and 250 mg/mL; the remainder of insert volume consisted of fat and water at a 9:1 ratio.

In addition to the decompositions of the phantom with the Ti construct, we also provide reference decompositions of the phantom without the metal, where the Ti regions were filled with bone, fat, and water based on the original segmented CBCT image.

3.1.2. CBCT System Model and DE Protocol—Polyenergetic simulations of DE imaging were performed in a configuration emulating an interventional C-arm CBCT system (Schafer *et al* 2011). The source-axis distance (SAD) was 600 mm; the source-detector distance (SDD) was 1,200 mm. For the purpose of this initial feasibility study, we assumed a kV-switching DE protocol, whereby the x-ray source energy alternated between projection views. In total, 180 LE and 180 HE views were acquired over a 360° rotation at 1° angular steps. The LE beam was 60 kV (+2.0 mm Al, +0.25 mm Cu), and the HE beam was 140 kV (+2.0mm Al, +0.25 mm Cu) with an additional high-Z filtration of + 0.225 mm Ag. X-ray spectra were obtained from Spektr3 (Punnoose *et al* 2016). The area of the simulated x-ray field was 300 mm (trans-axial) × 50 mm (axial) at the detector.

Projection data were modeled following Eq. (1). We assumed a flat-panel detector consisting of 0.388×0.388 mm^2 pixels with a 0.6 mm thick CsI:Tl scintillator with quantum detection efficiency included in the spectral response term of Eq. (1). Incident quanta were assumed to be Poisson noise with a mean of 2×10^5 photons/pixel.

To focus the study on the fundamental performance characteristics of CMBMD, we did not include scatter in the current simulation. Such an assumption is reasonable given a high-fidelity scatter correction, e.g., using Monte Carlo simulations (Sisniega *et al* 2013) or deep-learning models (Erath *et al* 2021) as well as axially collimated, narrow-beam data acquisition, similar to the settings in the experiments described below.

3.1.3. Reconstruction Methods—Prior to reconstruction, the projections were binned to 0.776×0.776 mm^2 pixels. This step guards against the "simulation crime" (using matched discretization in simulation and reconstruction) and approximates the non-linear partial volume effect that is present in real data. In all reconstructions, the voxel size was $0.5 \times 0.5 \times 0.5$ mm^3.

For CMBMD experiments, the air voxels surrounding the knee were identified from an initial CBCT image and assigned to the unconstrained region ($\boldsymbol{\pi}_{\mathscr{U}}$). In reconstructions of the digital phantom containing implants, fat, cortical bone and Ti formed the decomposition basis set for the constrained region ($\boldsymbol{\pi}_{\mathscr{C}}$), and fat and cortical bone were used in the unconstrained region. We used 1/3 to initialize each base in $\boldsymbol{\pi}_{\mathscr{C}}$, whereas $\boldsymbol{\pi}_{\mathscr{U}}$ was initialized with 1 for fat fractions and 0 for cortical bone fractions. Reference CMBMD reconstructions of the phantom without metal were also obtained. Here, the decomposition basis set in $\boldsymbol{\pi}_{\mathscr{C}}$ was fat and cortical bone, each initialized with 1/2. In this configuration, the VCP was only applied on the sum of bone and fat volume fractions.

During CMBMD optimization, the search directions for both PCIP (constrained region) and NR (unconstrained region) were obtained by voxel-wise Gaussian elimination parallelized in the CPU; the Jacobians were first perturbed following Vanderbei and Shanno (1999) whenever needed to achieve positive definiteness. The step length in NR was fixed to 0.9. In

PCIP, we applied the searched MSL (Nocedal *et al* 2009) scaled by 0.995 for the centered corrector step.

Unconstrained MBMDs of the entire volume were obtained for comparison with CMBMD. Two base material sets were considered: i) fat and cortical bone (denoted fat-bone MBMD), and ii) fat and Ti (fat-Ti MBMD). The search directions and step length for the MBMD updates were obtained in the same manner as for the unconstrained region in CMBMD.

Both algorithms incorporated a quadratic regularizer ψ for each base material κ:

$$\psi\left(\boldsymbol{\pi}^{\{\kappa\}}\right) = \frac{1}{4} \sum_x \sum_{w \in \mathbb{N}(x)} \left(\boldsymbol{\pi}_x^{\{\kappa\}} - \boldsymbol{\pi}_w^{\{\kappa\}}\right)^2 \tag{17}$$

where $\mathbb{N}(x)$ represents the indices of six nearest neighbors of the xth voxel. A voxel separable surrogate of Eq. (17) was obtained using additive convexity (De Pierro 1995) to compute the gradient and Jacobian terms associated with ψ. Penalty strengths were as follows: 10^2 for fat in all algorithms, 10^3 for cortical bone in fat-bone MBMD, fat-bone CMBMD and fat-bone-Ti CMBMD, 10^3 for Ti in fat-Ti MBMD, and 10^{-2} for Ti in fat-bone-Ti CMBMD. These values were obtained from a parameter sweep to minimize the normalized root mean squared error (NRMSE) between the true and estimated bone content of the test inserts (see Sec. 4.3). MBMD and CMBMD optimizations involved 500 iterations computed with 15 subsets. All algorithms were implemented in Python 3.7 and executed on a personal workstation equipped with Intel Xeon processor (2.30 GHz) and NVIDIA Quadro RTX 5000 GPU (a custom GPU library was used for forward and back-projection).

In addition to MBMD and CMBMD, VCP-constrained IDD (Sec. 2.5) of the phantom with metal was also performed. The input to IDD consisted of FDK reconstructions of the LE and HE projections obtained using a Hann-apodized ramp filter with cutoff at 0.8x Nyquist frequency and the same $0.5 \times 0.5 \times 0.5$ mm^3 voxels as in the model-based decomposition algorithms.

3.2. Experiments on an Imaging Test Bench

Initial empirical validation involved an imaging test bench configured to replicate the interventional C-arm CBCT geometry of the simulation study. The system employed a 400 $\times$ 300 mm^2 Varex 4030CB (Varex Imaging, Salt Lake City, UT, USA) flat-panel detector (FPD) with 0.194×0.194 mm^2 pixels and a 0.6 mm thick CsI:Tl scintillator. The same kV-switching DE protocol as in the simulations was executed (180 views at 1° step for each energy, 60 kV LE beam, 140 kV + 0.225 mm Ag HE beam). Axial collimation was 15 mm to minimize x-ray scatter.

The x-ray spectra for use in CMBMD and MBMD were estimated by adjusting the added filtration in Spektr3 (Punnoose *et al* 2016) for the LE and HE beams to match exposure measurements obtained at varying thicknesses of added Al attenuation (1 – 6 mm).

The imaging phantom (Fig. 2) consisted of a cylindrical water bath (~100 mm diameter, ~150 mm height) containing three Gammex (Gammex Inc., Middleton, WI, USA) bone-mimicking inserts (28 mm diameter, 60 mm height) and metal surgical hardware. Calcium

concentrations in the inserts were 100 mg/mL, 200 mg/mL, and 400 mg/mL Ca (Fig. 2C). Two metal implant configurations were considered: 1) a Ti nail (12 mm diameter) shown in Fig. 2A, and 2) the same Ti nail accompanied by a Ti plate (16×4 mm^2 in the axial plane) in Fig. 2B. Projections of the phantom without the metal implants were also acquired as a reference.

Glare deconvolution followed by fast Monte Carlo scatter estimation and subtraction (Sisniega *et al* 2013, Liu *et al* 2021b) were performed prior to reconstruction. Projections were binned 2×2 to 0.388×0.388 mm^2 pixel size and reconstructed with $0.5 \times 0.5 \times 0.5$ mm^3 voxels. The constrained region in CMBMD was the interior of the phantom, and the unconstrained region was the surrounding air, identified by thresholding a preliminary FDK HE reconstruction. Other CMBMD and MBMD settings were the same as in the simulation studies.

3.3. Performance Metrics

The decompositions were displayed as a composite monochromatic image $\hat{\boldsymbol{\mu}}_{\mathrm{mono}}(\epsilon) \in \mathbb{R}^{\mathcal{I}}$, equal to a sum of the estimated volume fractions of the base materials scaled by their linear attenuation at energy ϵ, here set to 70 keV:

$$\hat{\boldsymbol{\mu}}_{\mathrm{mono}}(\epsilon) := \sum_{\kappa=1}^{K} \mu^{\{\kappa\}}(\epsilon)\hat{\boldsymbol{\pi}}^{\{\kappa\}} \tag{18}$$

To quantify the uniformity of the monochromatic image, we evaluated the coefficient-of-variation (CV) of soft-tissue regions-of-interest (ROIs) indicated with dashed circles in Fig. 1. CV was computed individually for each axial slice of the reconstruction as the ratio of ROI standard deviation to its mean. In objects without the Ti hardware, CV is mostly affected by noise. Since the same regularization settings were used in fat-bone MBMD and fat-bone CMBMD, any CV differences in the metal-free cases might reflect potential additional noise suppression due to the VCP constraint. In the objects with metal, CV is primarily a metric of image artifacts.

For the simulation studies, we report the distribution of CV (mean and standard deviation) in two subsets of 16 slices each. One subset was vertically centered on the screw (denoted as "Nail + Screw"), whereas the other one was centered on a nail-only region superior to the screw (denoted as "Nail"); the central slices of these subsets are shown in Fig. 1. (The metal-free case was evaluated using the same two axial regions and thus we will use the "Nail" and "Nail + Screw" terminology also in the implant-free setting.) For the test bench experiments, the CV distributions are reported over all water voxels of the monochromatic image (24 slices total). These voxels were identified by thresholding an initial CBCT reconstruction of the metal-free phantom.

Decomposition accuracy was evaluated in terms of the normalized root-mean squared error (NRMSE) of the estimated cortical bone volume fraction in the four target inserts in the tibia (Fig. 1) for the simulation study, and in the three Ca inserts for the physical experiments. For each of the inserts, the NRMSE was computed individually in each axial slice as:

$$\text{NRMSE} = \frac{\sqrt{\text{mean}\left\{\left(\widehat{\boldsymbol{\pi}}^{\{\text{bone}\}} - \widetilde{\boldsymbol{\pi}}^{\{\text{bone}\}}\right)^2\right\}}}{\widetilde{\boldsymbol{\pi}}^{\{\text{bone}\}}} \tag{19}$$

where $\widehat{\boldsymbol{\pi}}^{\{\text{bone}\}}$ and $\widetilde{\pi}^{\{\text{bone}\}}$ are the estimated and true bone fractions in a given axial cross section of the insert (excluding a 2-voxel annulus at its boundary). To investigate the decomposition error fluctuations within the reconstructed volume, we report mean and standard deviation of insert NRMSEs across the same sub-volumes as in the CV analysis: the Nail and Nail + Screw slices in the simulation study, and the water voxels in the benchtop experiment.

To enable this performance assessment procedure for fat-Ti MBMD, the estimated fat and Ti volume fraction maps were first linearly combined to yield a bone map. The coefficients of this combination were found as the weights of a weighted sum of linear attenuations of fat and Ti that best approximated linear attenuation of cortical bone.

4. Results

4.1. Simulation Study

Fig. 3 shows MBMD and CMBMD decompositions of the digital phantom. For reference, Fig. 3A provides the results for the implant-free case. In this setting, there are no noticeable differences between fat-bone MBMD (left column) and fat-bone CMBMD (right column); both decompositions are free of major artifacts and present adequate visualization of the test inserts and sufficient uniformity in soft-tissue and fat regions. In the presence of metal (Fig. 3B), MBMD suffers from strong beam-hardening effects evident as blooming, ringing, and streaks. The artifacts are particularly severe in slices containing both the nail and the screw (bottom row), likely because the long axis of the nail is parallel to the axial plane. The artifacts are somewhat diminished in fat-Ti MBMD (2nd column of Fig. 3B) compared to fat-bone MBMD (1st column of Fig. 3B), suggesting a benefit to including the metal directly in the base material set; however, this precludes discrimination of endogenous tissues when using unconstrained MBMD.

Compared to MBMD, fat-bone-Ti CMBMD (3rd column of Fig. 3B) shows substantially reduced metal artifacts. The visual quality appears less affected by the complexity of the implant; both for the nail-only and nail-and-screw slices, the visualization of the test inserts is similar to MBMD of metal-free object. The remaining slight blooming and streaks are attributed to nonlinear partial volume effects and regularization mismatches between base material volumes. We note the piecewise constant appearance of the fat region, where noise appears suppressed because the constraint favors solutions with fat fractions of 1.0.

Fig. 4 compares CV for the decompositions presented in Fig. 3; the results for the slices in the Nail-only axial region (see Sec. 3.3) are in Fig. 4A, the Nail + Screw slices are in Fig. 4B. In the implant-free case, CVs are small both for MBMD and CMBMD (mean of ~0.012), regardless of the axial location. CMBMD achieves slightly lower CV in the measurement region located closer to the bone (Location 2), likely because some noise

fluctuations are suppressed by the VCP constraint. In the presence of metal, the mean CV of fat-bone and fat-Ti MBMDs reach ~0.025 – 0.04 for the Nail slices (simple hardware) and >0.04 for the Nail + Screw slices (complex hardware), reflecting the artifacts visible in Fig. 3. Using the proposed three-material CMBMD, the CV is reduced to <0.020 for the Nail slices and <0.024 for the Nail + Screw slices, approaching the values measured in the reference metal-free phantom and consistent with substantial artifact reduction observed in Fig. 3.

Fig. 5 summarizes the accuracy of estimated cortical bone concentrations in the test inserts, identified by their nominal bone concentrations on the horizontal axis. For the reference MBMD of the implant-free phantom (dash-dot lines), NRMSE across slices is 6.2% - 14.8%, decreasing with increasing bone volume fraction; the cross-slice variability was 0.4% - 1.1%. The NRMSE of two-material CMBMD of the metal-free object is within 1% difference from the NRMSE of MBMD; hence, it was omitted from Fig. 5 for clarity. In the presence of metal, the accuracy of MBMD is substantially reduced, with average NRMSE for nail-only slices of 37.2% - 81.9% for the fat-bone base material set (Fig. 5A) and 29.5% - 50.0% for the fat-Ti base set (Fig. 5B). The error increases to >60% for nail-and-screw slices (light gray solid lines). The wide cross-slice variability of the NRMSE of MBMD (error bars) reflects the artifacts visible in Fig. 3 and is quantified using CV in Fig. 4.

By incorporating the metal as the third basis, CMBMD yields improved quantification of cortical bone concentrations. As shown in Fig. 5C, CMBMD reduces NRSME to below 20% and 26% for the Nail and Nail + Screw slices, respectively, approaching the accuracy of the implant-free reference.

Fig. 5D investigates the convergence of MBMD and CMBMD in the phantom with metal. The data fidelity term of the PWLS objective, $l(\boldsymbol{\pi}^{(n)})$, is plotted as a function of iterations n. Fat-bone MBMD requires ~500 iterations to achieve a stable solution. However, this solution is of relatively poor fidelity (large values of $l(\boldsymbol{\pi}^{(n)})$) – consistent with the severe artifacts in Fig. 3. Replacing bone with metal in the MBMD base set improves data fidelity compared to fat-bone MBMD, but leads to poorer convergence. CMBMD, which includes both metal and bone in its three-material basis set, achieves the fastest convergence and the lowest error compared to either form of unconstrained two-material MBMD. In terms of computing time, CMBMD is slightly slower than MBMD due to additional variables and the application of the PC framework: the times per iteration using the hardware reported in Sec. 3.1.3 are ~3.25 min (15 sub-iterations) for CMBMD compared to ~1.63 min for MBMD.

4.2. Comparison to Constrained IDD

Fig. 6 shows constrained IDD decompositions of the phantom with the implant. In slices containing only the nail, IDD achieves somewhat better delineation of the Ca inserts than fat-bone MBMD, but appears overall visually inferior to CMBMD. This is confirmed by the NRMSE measurements, which range from 15% to 25% (depending on Ca concentrations) in the Nail slices, which is slightly worse than for CMBMD. However, IDD was unable to overcome metal artifacts in the Nail + Screw slices, resulting in a decomposition that is much less accurate than CMBMD, with NRMSE ranging from 40% to 100%.

4.3. Experiment on an Imaging Test Bench

Fig. 7A shows the reference fat-bone MBMD and CMBMD of the physical phantom without metal implants. Visually, both approaches produce similar decompositions, with the CMBMD image exhibiting more pronounced piecewise constant characteristics. Fig. 7B compares the test bench results for the objects with metal (top row: Ti Nail-only, bottom-row: Ti Nail + Plate) using fat-bone MBMD, fat-Ti MBMD and fat-bone-Ti CMBMD. CMBMD reduces metal artifacts relative to MBMD in both implant configurations, yielding a more uniform appearance of the Ca inserts even in areas adjacent to the metal.

The visual image assessment is confirmed by CV measurements in Fig. 8A and NRMSE of Ca concentrations in Fig. 8B–D. In the object without metal, the two-material MBMD and CMBMD result in CV values (0.0136 and 0.0132, respectively) that are comparable but slightly lower for CMBMD, likely due to noise mitigation provided by the VCP constraint. In the presence of metal, CMBMD achieves ~2-fold reduction in CV compared to MBMD and more accurate Ca volume fraction estimates (e.g., ~2x improvement in NRMSE for 100 mg/mL and 200 mg/mL Ca inserts) regardless of the implant configuration. Interestingly, fat-Ti MBMD does not achieve the same improvement in image quality and accuracy over fat-bone MBMD as in the simulation studies (note also that we did not find improved convergence for fat-Ti MBMD in real data). This discrepancy, as well as the residual artifacts in CMBMD, are attributed to mismatches between the forward model and the experimental setup that might result in local minima in the decomposition – for example, incomplete scatter subtraction and inaccuracies in x-ray spectrum and Ti alloy composition.

5. Discussion and Conclusions

We developed an optimization-based three-material decomposition algorithm for DE CT that incorporates volume conservation as a pair of local linear constraints. Initial validation of the proposed CMBMD was performed in application to DE CBCT measurements of bone densities in the presence of metal. The results of simulations and physical test bench experiments demonstrated improved image quality (e.g., fewer streaks and less blooming) and quantitative accuracy using CMBMD compared to the previously developed unconstrained two-material MBMD (Tilley *et al* 2019). Artifact mitigation was reflected in ~2x decrease of monochromatic image CV measured in soft-tissue regions neighboring the implants both in the simulations and the benchtop study. The reduction of bone volume fraction NRMSE ranged from ~5-fold to ~10-fold in the simulations and ~2-fold to ~2.5-fold in the experiments, depending on the amount of metal and nominal bone concentration. The more pronounced error mitigation in the simulations can be attributed primarily to differences in implant configurations between the two studies. Compared to the physical phantom, where the metal was oriented longitudinally, the digital phantom involved a locking screw placed lengthwise in the axial plane, which resulted in more severe artifacts – and thus larger initial NRMSE - across the entire volume (because of beam divergence, the artifacts propagated into slices that did not contain the screw). The absolute bone volume fraction errors after the application of CMBMD were similar in the simulations and experiments, yielding a ~10% - 20% accuracy that approached the performance of two-material MBMD of the object without metal.

While the primary objective of this work was to demonstrate the application of three-material CMBMD in the presence of metal, the results presented in the manuscript also include constrained two-material decomposition of metal-free cases. In this setting, the performance of MBMD and CMBMD is generally comparable. However, CMBMD appears to provide a degree of additional noise suppression, likely because it prohibits random fluctuations that violate the VCP. Future work will investigate whether VCP might similarly reduce artifacts due to missing data, e.g. in slow kV-switching DE CBCT with sparse view sampling.

In addition to the VCP constraint, two other features of the proposed algorithm merit a brief discussion. First, while MBMD was originally derived using an objective function surrogate that was separable both in voxels and materials, the CMBMD surrogate is only separable in voxels. The new formulation is consistent with the strong correlation between material fractions in each voxel and typically leads to faster convergence than the fully separable surrogate (Tivnan *et al* 2020). The improved convergence tends to offset the increased computational burden associated with the need to solve a linear system of equations at each voxel to compute the image update, which was not required in the original MBMD implementation. The second important feature of the proposed CMBMD is that it permits regional imposition of the constraints, whereby the NR updates are applied to the unconstrained voxels and the IP updates are applied in the constrained region – a partitioned optimization scheme that relies on the voxel-separable surrogate. This approach allowed us to accommodate the presence of air voxels that cannot be represented as a linear combination of the base materials under VCP. In this case, the air region is relatively easy to delineate and can be initialized close to its known solution of zero, such that the convergence of the algorithm is determined by the updates occurring in the constrained voxels. Further investigation is merited to assess the performance of this approach in the more general case of an unconstrained region that contains unknown mixture of tissues, or in a partitioned scheme where there are multiple constrained sub-volumes with different sets of base materials. In these settings, the algorithm might be less robust to mismatched regularization or different convergence rates in the different regions compared to the application considered here.

It needs to be recognized that the volume conservation principle implies fairly strict assumptions on object composition. It not only assumes that all materials are in the span of the base material set as in conventional DE, but also that the decomposition coefficients are always positively add to one. To use an example from our experiments, we cannot reconstruct fat voxels with density lower than nominal adipose density – even if that was more consistent with the measured voxel attenuation due to noise or bias, for example – leading to the piecewise-constant appearance of the decompositions. Our results indicate that the VCP constraints are not overly restrictive for the proposed application in orthopedic extremity imaging, where tissues appear to be reasonably well represented using fat and cortical bone. In other body sites with more complex composition, however, robust three-material decomposition might require new optimization schemes where the constraint is not strictly enforced, such as solving the IP linear system inexactly (Arreckx and Orban 2018, Estrin *et al* 2020) or directly replacing the constraint with inexact smooth penalty

(Nocedal and Wright 2006). Alternatively, it might be possible to replace VCP with another constraint, e.g., total variation of a virtual monochromatic image as in Barber *et al* 2016.

VCP has conventionally been applied in image domain (Granton *et al* 2008, Ding *et al* 2018, Jiang *et al* 2020), where its performance might be hampered by the need for accurate artifact pre-corrections during reconstruction of the LE and HE projections. Our comparison of the proposed CMBMD to VCP-constrained IDD demonstrated the advantages of the direct decomposition approach in CMBMD, particularly in the settings where a complex metal structure is present in the object (e.g., the Nail + Screw slices of the digital phantom). In this configuration, IDD would likely require sophisticated multi-pass water-bone-Ti beam hardening corrections, possibly combined with metal in-painting, whereas CMBMD was able to produce nearly artifact-free reconstructions owing to the use of the spectral forward model. However, despite the lack of advanced pre-corrections in the experiments presented here, the performance of VCP-constrained IDD was not substantially worse than two-material MBMD (fat-bone or fat-Ti). For accurate direct decomposition in the presence of metal, the spectral model needs to capture beam hardening properties of both the metal and the endogenous tissues, which appears to benefit from the three-material capability provided by CMBMD.

As discussed in the previous paragraph, the primary benefit of CMBMD in DE imaging in the presence of surgical hardware is to provide the decomposition with accurate models of beam hardening in both the bone tissue and the metal component. For the Ti implants considered here, this approach yields substantial artifact reduction compared to two material MBMD, where only one of the two highly attenuating materials can be accounted for. However, in configurations where photon starvation is significant – e.g., involving large and/or higher density hardware – the resulting missing projection data may preclude CMBMD from converging to the global minimum to avoid artifacts and quantitative inaccuracies. For such scenarios, we are investigating a combination of CMBMD with a prior implant model (Stayman *et al* 2012, Xu *et al* 2017, Liu *et al* 2020b), with CMBMD applied at the boundary of the known metal region to resolve partial-volume effects and other localized mismatches between the component model and measured data.

The clinical motivation to focus the initial feasibility evaluation of CMBMD on bone composition measurements in the vicinity of metal is to improve monitoring of fracture healing, where a DE technique could be used to quantify callus mineralization. New technologies to evaluate bone fusion are highly desirable because of the substantial healthcare burden of delayed unions and non-unions: for example, ~25% of tibial shaft fractures exhibit protracted healing (Heckman and Sarasohn-Kahn 1997), leading to additional hospitalizations and re-operations. Typically, non-unions are diagnosed only after 6–9 months of observation (Antonova *et al* 2013), suggesting that there is a need for better biomarkers of bone healing. Prior work indicates that DE CT might achieve more accurate BMD measurements than single-energy quantitative CT (Kalender *et al* 1987, Wichmann *et al* 2014, Arentsen *et al* 2017), thus motivating the proposed use of DE imaging for assessment of callus composition. The ability to perform material decomposition in the presence of metal is essential to enable this application, since metal fixation hardware is commonly used in complex fractures. This capability might be particularly valuable

in combination with the recently introduced CBCT devices for weight-bearing imaging (for example, extremity CBCT (Koskinen *et al* 2013, Carrino *et al* 2014, Demehri *et al* 2015, Thawait *et al* 2015) or multi-purpose robotic x-ray systems (Grunz *et al* 2020, 2021, Falkowski *et al* 2021), which could provide quantitative assessment of fracture stability as an additional marker of healing. Related clinical applications include evaluation of bone quality around joint replacement and spine fusion constructs.

Future translation of CMBMD to clinical applications will likely require accelerating the computations. The execution times are primarily affected by forward and back-projections of the volume (here, a total of 3 forward and 3 back-projections implemented per iteration) and the application of the spectral model. To address these bottlenecks, we are currently investigating: i) multiresolution reconstruction that employs a finer voxel grid in the regions of interest (Cao *et al* 2016, Sisniega *et al* 2021), and ii) polynomial approximation of the polyenergetic forward model (Abella *et al* 2019). Another important consideration for clinical translation is ensuring robust convergence of the algorithm. Both the NR update in the unconstrained region and the IP update in the constrained region were originally developed for convex nonlinear programming; however, convexity is not guaranteed for the CMBMD objective, and this might affect algorithm convergence. Although we did not observe such issues in our study, it is likely that applications of CMBMD to problems that are more poorly posed, for example fat-water-bone decomposition in for detection of bone marrow edema (Pache *et al* 2010) will require algorithm modifications to address possible non-convexity. These might include i) better variable initialization (e.g., optimized image-based DE decomposition), ii) Jacobian perturbation (Shanno and Vanderbei 2000), and iii) cubic regularization (Benson and Shanno 2014) with merit- or filter-based line searching (Benson *et al* 2002).

In summary, we presented a new algorithm for one-step three-material decomposition of DE data. The proposed CMBMD incorporates volume conservation as an optimization constraint in the previously developed MBMD algorithm. Initial simulation and test-bench validation studies in DE imaging of bone composition in the vicinity of metal demonstrated improved quantitative accuracy and substantial artifact reduction compared to unconstrained two-material MBMD. Potential applications of the algorithm include DE configurations in which a lack of coinciding LE and HE projection rays precludes conventional projection-domain decomposition, for example orthopedic DE CBCT using the recently introduced extremity CBCT system with a three-source x-ray unit (Gang *et al* 2018, Liu *et al* 2020a, 2020b). The three-material decomposition capability provided by CMBMD is of interest not only for DE imaging in the presence of metal, but also for body DE CT/CBCT with multiple contrast agents (Symons *et al* 2017, Tao *et al* 2019) and detection of bone marrow edema (Pache *et al* 2010, Zbijewski *et al* 2015, Ali *et al* 2018, Gosangi *et al* 2020, Müller *et al* 2020).

Acknowledgements

Appendix

I. Voxel-wise Separable Surrogate

To optimize the PWLS problem in Eq. (5), Tilley *et al* 2019 originally implemented the Separable Parabolic Surrogate (SPS) algorithm that computes coordinate descent steps based on consecutive separable and quadratic approximations of Φ at $\boldsymbol{\pi} = \boldsymbol{\pi}^{(n)}$; the final surrogate is separable per voxel j and per material κ. However, when material bases are strongly inter-correlated (possessing similar mass attenuation coefficients over a wide range of energies, e.g., cortical bone and titanium), the material-wise separability used in SPS could be inefficient, resulting in slow speed of convergence.

We improve the convergence by computing the NR step based on a voxel-wise separable surrogate of $\Phi(\boldsymbol{\pi})$ at $\boldsymbol{\pi} = \boldsymbol{\pi}^{(n)}$, with no quadratic or material-wise separable surrogates. We first re-write $\Phi(\boldsymbol{\pi})$ as:

$$
\begin{aligned}
\Phi(\boldsymbol{\pi}) &= \Phi\left(\sum_{j=1}^{J}\sum_{\kappa=1}^{K} \pi_j^{\{\kappa\}}\tilde{e}_j^{\{\kappa\}}\right) \\
&= \Phi\left(\boldsymbol{\pi}^{(n)} + \sum_{j=1}^{J}\sum_{\kappa=1}^{K}\left(\pi_j^{\{\kappa\}} - \pi_j^{\{\kappa\}^{(n)}}\right)\tilde{e}_j^{\{\kappa\}}\right) \\
&= \Phi\left(\sum_{j=1}^{J}\gamma_j^{(n)}\boldsymbol{\pi}^{(n)} + \sum_{j=1}^{J}\sum_{\kappa=1}^{K}\left(\pi_j^{\{\kappa\}} - \pi_j^{\{\kappa\}^{(n)}}\right)\tilde{e}_j^{\{\kappa\}}\right) \\
&= \Phi\left(\sum_{j=1}^{J}\gamma_j^{(n)}\left(\boldsymbol{\pi}^{(n)} + \frac{\sum_{\kappa=1}^{K}\left(\pi_j^{\{\kappa\}} - \pi_j^{\{\kappa\}^{(n)}}\right)\tilde{e}_j^{\{\kappa\}}}{\gamma_j^{(n)}}\right)\right)
\end{aligned}
\tag{A.1}
$$

where $\pi_j^{\{\kappa\}}$ is an entry of $\boldsymbol{\pi}$ at the κth material of the jth voxel, and $\tilde{e}_j^{\{\kappa\}} \in \mathbb{R}^{JK}$ is the standard basis vector that has an entry of 1 at the κth material of the jth voxel and 0 elsewhere. The coefficients $0 < \gamma_j^{(n)} \in \mathbb{R}$ sum to unity:

$$
\sum_{j=1}^{J}\gamma_j^{(n)} = 1
\tag{A.2}
$$

Assuming the Hessian $\mathcal{H}_{\Phi}^{(n)}$ has been properly perturbed (whenever necessary) to be positive definite, the voxel-wise separable surrogate $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ can be obtained following additive convexity (De Pierro 1995, Erdogan and Fessler 1999):

$$\varphi\left(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)}\right) := \sum_{j=1}^{\mathcal{J}} \gamma_j^{(n)} \Phi\left(\boldsymbol{\pi}^{(n)} + \frac{\sum_{\kappa=1}^{\mathcal{K}} \left(\boldsymbol{\pi}_j^{\{\kappa\}} - \boldsymbol{\pi}_j^{\{\kappa\}(n)}\right) \tilde{\boldsymbol{e}}_j^{\{\kappa\}}}{\gamma_j^{(n)}}\right)$$

$$\geq \Phi\left(\sum_{j=1}^{\mathcal{J}} \gamma_j^{(n)} \left(\boldsymbol{\pi}^{(n)} + \frac{\sum_{\kappa=1}^{\mathcal{K}} \left(\boldsymbol{\pi}_j^{\{\kappa\}} - \boldsymbol{\pi}_j^{\{\kappa\}(n)}\right) \tilde{\boldsymbol{e}}_j^{\{\kappa\}}}{\gamma_j^{(n)}}\right)\right) = \Phi(\boldsymbol{\pi}) \qquad \text{(A.3)}$$

Note that $\sum_{\kappa=1}^{\mathcal{K}} \boldsymbol{\pi}_j^{\{\kappa\}} \tilde{\boldsymbol{e}}_j^{\{\kappa\}} \in \mathbb{R}^{\mathcal{J}\mathcal{K}}$ in the numerator of $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ contains material fractions only at the jth voxel and zeros elsewhere, indicating that an individual update for each voxel can be derived (i.e., voxel-wise separability).

One can easily show that $\varphi(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)})$ in Eq. (A.3) satisfies the sufficiency conditions of optimization transfer defined in Erdogan and Fessler 1999:

$$\begin{cases} \varphi\left(\boldsymbol{\pi}; \boldsymbol{\pi}^{(n)}\right) \geq \Phi(\boldsymbol{\pi}) \\ \varphi\left(\boldsymbol{\pi}^{(n)}; \boldsymbol{\pi}^{(n)}\right) = \Phi\left(\boldsymbol{\pi}^{(n)}\right) \\ \nabla \varphi\left(\boldsymbol{\pi}^{(n)}; \boldsymbol{\pi}^{(n)}\right) = \nabla \Phi\left(\boldsymbol{\pi}^{(n)}\right) \end{cases} \qquad \text{(A.4)}$$

Because of voxel separability, the Hessian of the surrogate $\mathcal{H}_\varphi^{(n)}$ has zero cross-voxel second-order derivatives and can be constructed as (Tivnan *et al* 2020):

$$\left[\mathcal{H}_\varphi^{(n)}\right]_{j_1, j_2} = \begin{cases} \left[\mathcal{H}_\Phi^{(n)}\right]_{j_1, j_2} & \text{if } j_1 = j_2 \\ 0 & \text{otherwise} \end{cases}, \quad \forall j_1, j_2 = 1, \ldots, \mathcal{J} \qquad \text{(A.5)}$$

where $\mathcal{H}_\Phi^{(n)}$ can be computed following Eq. (8b).

Reference

Abella M, Martínez C, Desco M, Vaquero JJ, & Fessler JA (2019). Simplified Statistical Image Reconstruction for X-ray CT with Beam-Hardening Artifact Compensation. IEEE transactions on medical imaging, 39(1), 111–118. [PubMed: 31180844]

Ali IT, Wong WD, Liang T, Khosa F, Mian M, Jalal S, & Nicolaou S (2018). Clinical utility of dual-energy CT analysis of bone marrow edema in acute wrist fractures. American Journal of Roentgenology, 210(4), 842–847. [PubMed: 29470155]

Alvarez RE, & Macovski A (1976). Energy-selective reconstructions in x-ray computerised tomography. Physics in Medicine & Biology, 21(5), 733. [PubMed: 967922]

Antonova E, Le TK, Burge R, & Mershon J (2013). Tibia shaft fractures: costly burden of nonunions. BMC musculoskeletal disorders, 14(1), 1–10. [PubMed: 23281775]

Arentsen L, Hansen KE, Yagi M, Takahashi Y, Shanley R, McArthur A, … & Hui SK (2017). Use of dual-energy computed tomography to measure skeletal-wide marrow composition and cancellous bone mineral density. Journal of bone and mineral metabolism, 35(4), 428–436. [PubMed: 27942979]

Arreckx S, & Orban D (2018). A regularized factorization-free method for equality-constrained optimization. SIAM Journal on Optimization, 28(2), 1613–1639.

Barber RF, Sidky EY, Schmidt TG, & Pan X (2016). An algorithm for constrained one-step inversion of spectral CT data. Physics in Medicine & Biology, 61(10), 3784. [PubMed: 27082489]

Benson HY, & Shanno DF (2014). Interior-point methods for nonconvex nonlinear programming: cubic regularization. Computational optimization and applications, 58(2), 323–346.

Benson HY, Vanderbei RJ, & Shanno DF (2002). Interior-point methods for nonconvex nonlinear programming: Filter methods and merit functions. Computational Optimization and Applications, 23(2), 257–272.

Cao Q, Liu S, Osgood G, Demehri S, Siewerdsen J, Stayman J, & Zbijewski W (2019). Cone-beam CT of load-bearing surgical hardware using a mechanical model of implant deformation. In Medical Physics (Vol. 46, No. 6, pp. E399–E399).

Cao Q, Zbijewski W, Sisniega A, Yorkston J, Siewerdsen JH, & Stayman JW (2016). Multiresolution iterative reconstruction in high-resolution extremity cone-beam CT. Physics in Medicine & Biology, 61(20), 7263. [PubMed: 27694701]

Carrino JA, Al Muhit A, Zbijewski W, Thawait GK, Stayman JW, Packard N, … & Siewerdsen JH (2014). Dedicated cone-beam CT system for extremity imaging. Radiology, 270(3), 816–824. [PubMed: 24475803]

Demehri S, Muhit A, Zbijewski W, Stayman JW, Yorkston J, Packard N, … & Siewerdsen JH (2015). Assessment of image quality in soft tissue and bone visualization tasks for a dedicated extremity cone-beam CT system. European radiology, 25(6), 1742–1751. [PubMed: 25599933]

De Pierro AR (1995). A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. IEEE transactions on medical imaging, 14(1), 132–137. [PubMed: 18215817]

Ding Q, Niu T, Zhang X, & Long Y (2018). Image-domain multimaterial decomposition for dual-energy CT based on prior information of material images. Medical physics, 45(8), 3614–3626.

Erath J, Vöth T, Maier J, Fournié E, Petersilka M, Stierstorfer K, & Kachelrieß M (2021). Deep learning-based forward and cross-scatter correction in dual-source CT. Medical physics.

Erdogan H, & Fessler JA (1999). Ordered subsets algorithms for transmission tomography. Physics in Medicine & Biology, 44(11), 2835. [PubMed: 10588288]

Estrin R, Friedlander MP, Orban D, & Saunders MA (2020). Implementing a smooth exact penalty function for equality-constrained nonlinear optimization. SIAM Journal on Scientific Computing, 42(3), A1809–A1835.

Falkowski AL, Kovacs BK, Benz RM, Tobler P, Schön S, Stieltjes B, & Hirschmann A (2021). In vivo 3D tomography of the lumbar spine using a twin robotic X-ray system: quantitative and qualitative evaluation of the lumbar neural foramina in supine and upright position. European Radiology, 31(5), 3478–3490. [PubMed: 33119812]

Fiacco AV, & McCormick GP (1990). Nonlinear programming: sequential unconstrained minimization techniques. Society for Industrial and Applied Mathematics.

Gang GJ, Zbijewski W, Mahesh M, Thawait G, Packard N, Yorkston J, … & Siewerdsen JH (2018). Image quality and dose for a multisource cone-beam CT extremity scanner. Medical physics, 45(1), 144–155. [PubMed: 29121409]

Goodsitt MM, Rosenthal DI, Reinus WR, & Coumas J (1987). Two postprocessing CT techniques for determining the composition of trabecular bone. Invest Radiol, 22(3), 209–215. [PubMed: 3557896]

Gosangi B, Mandell JC, Weaver MJ, Uyeda JW, Smith SE, Sodickson AD, & Khurana B (2020). Bone marrow edema at dual-energy CT: a game changer in the emergency department. RadioGraphics, 40(3), 859–874. [PubMed: 32364883]

Granton PV, Pollmann SI, Ford NL, Drangova M, & Holdsworth DW (2008). Implementation of dual- and triple-energy cone-beam micro-CT for postreconstruction material decomposition. Medical physics, 35(11), 5030–5042. [PubMed: 19070237]

Grunz JP, Weng AM, Kunz AS, Veyhl-Wichmann M, Schmitt R, Gietzen CH, … & Gassenmaier T (2020). 3D cone-beam CT with a twin robotic x-ray system in elbow imaging: comparison of image quality to high-resolution multidetector CT. European Radiology Experimental, 4(1), 1–9. [PubMed: 31900683]

Grunz JP, Pennig L, Fieber T, Gietzen CH, Heidenreich JF, Huflage H, … & Gassenmaier T (2021). Twin robotic x-ray system in small bone and joint trauma: impact of cone-beam computed tomography on treatment decisions. European Radiology, 31(6), 3600–3609. [PubMed: 33280057]

Heckman JD, & Sarasohn-Kahn J (1997). The economics of treating tibia fractures. Bulletin Hospital for Joint Diseases, 56(1), 63–72.

Hudson HM, & Larkin RS (1994). Accelerated image reconstruction using ordered subsets of projection data. IEEE transactions on medical imaging, 13(4), 601–609. [PubMed: 18218538]

Jiang Y, Xue Y, Lyu Q, Xu L, Luo C, Yang P, … & Niu T (2019). Noise suppression in image-domain multi-material decomposition for dual-energy CT. IEEE Transactions on Biomedical Engineering, 67(2), 523–535. [PubMed: 31095473]

Johnson TR, Krauss B, Sedlmair M, Grasruck M, Bruder H, Morhard D, … & Becker CR (2007). Material differentiation by dual energy CT: initial experience. European radiology, 17(6), 1510–1517. [PubMed: 17151859]

Joseph PM, & Spital RD (1978). A method for correcting bone induced artifacts in computed tomography scanners. Journal of computer assisted tomography, 2(1), 100–108. [PubMed: 670461]

Kalender WA, Klotz E, & Suess C (1987). Vertebral bone mineral analysis: an integrated approach with CT. Radiology, 164(2), 419–423. [PubMed: 3602380]

Kaza RK, Platt JF, Cohan RH, Caoili EM, Al-Hawary MM, & Wasnik A (2012). Dual-energy CT with single-and dual-source scanners: current applications in evaluating the genitourinary tract. Radiographics, 32(2), 353–369. [PubMed: 22411937]

Koskinen SK, Haapamäki VV, Salo J, Lindfors NC, Kortesniemi M, Seppälä L, & Mattila KT (2013). CT arthrography of the wrist using a novel, mobile, dedicated extremity cone-beam CT (CBCT). Skeletal radiology, 42(5), 649–657. [PubMed: 22990597]

Liu SZ, Tilley S II, Cao Q, Siewerdsen JH, Stayman JW, & Zbijewski W (2019). Known-component model-based material decomposition for dual energy imaging of bone compositions in the presence of metal implant. In 15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Vol. 11072, p. 1107213). International Society for Optics and Photonics.

Liu SZ, Cao Q, Osgood GM, Siewerdsen JH, Stayman JW, & Zbijewski W (2020a). Quantitative assessment of weight-bearing fracture biomechanics using extremity cone-beam CT. In Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging (Vol. 11317, p. 113170I). International Society for Optics and Photonics.

Liu SZ, Cao Q, Tivnan M, Tilley S II, Siewerdsen JH, Stayman JW, & Zbijewski W (2020b). Model-based dual-energy tomographic image reconstruction of objects containing known metal components. Physics in Medicine & Biology, 65(24), 245046. [PubMed: 33113519]

Liu SZ, Siewerdsen JH, Stayman JW, & Zbijewski W (2021a). Quantitative dual-energy imaging in the presence of metal implants using locally constrained model-based decomposition. In Medical Imaging 2021: Physics of Medical Imaging (Vol. 11595, p. 115951C). International Society for Optics and Photonics.

Liu S, Shi G, Osgood G, Siewerdsen J, Stayman J, & Zbijewski Z (2021b). Quantitative dual-energy imaging of bone compositions on multisource cone-beam CT with model-based artifact corrections. Medical Physics, 48(6).

Liu SZ, Zhao C, Herbst M, Weber T, Vogt S, Ritschl L, Kappler S, Siewerdsen JH, & Zbijewski W (2022) Feasibility of dual-energy cone-beam CT of bone marrow edema using dual-layer flat panel detectors. In Medical Imaging 2022: Physics of Medical Imaging (Vol. 12031, pp. 379–386). International Society for Optics and Photonics.

Liu X, Yu L, Primak AN, & McCollough CH (2009). Quantitative imaging of element composition and mass fraction using dual-energy CT: Three-material decomposition. Medical physics, 36(5), 1602–1609. [PubMed: 19544776]

Long Y, & Fessler JA (2014). Multi-material decomposition using statistical image reconstruction for spectral CT. IEEE transactions on medical imaging, 33(8), 1614–1626. [PubMed: 24801550]

Maaß C, Baer M, & Kachelrieß M (2009). Image-based dual energy CT using optimized precorrection functions: A practical new approach of material decomposition in image domain. Medical physics, 36(8), 3818–3829. [PubMed: 19746815]

Macari M, Spieler B, Kim D, Graser A, Megibow AJ, Babb J, & Chandarana H (2010). Dual-source dual-energy MDCT of pancreatic adenocarcinoma: initial observations with data generated at 80

kVp and at simulated weighted-average 120 kVp. American Journal of Roentgenology, 194(1), W27–W32. [PubMed: 20028887]

Mory C, Sixou B, Si-Mohamed S, Boussel L, & Rit S (2018). Comparison of five one-step reconstruction algorithms for spectral CT. Physics in Medicine & Biology, 63(23), 235001. [PubMed: 30465541]

Müller FC, Gosvig KK, Børgesen H, Gade JS, Brejnebøl M, Rodell A, … & Boesen M (2020). Dual-energy CT for suspected radiographically negative wrist fractures: a prospective diagnostic test accuracy study. Radiology, 296(3), 596–602. [PubMed: 32662760]

Nesterov Y (2005). Smooth minimization of non-smooth functions. Mathematical programming, 103(1), 127–152.

Nicolaou S, Liang T, Murphy DT, Korzan JR, Ouellette H, & Munk P (2012). Dual-energy CT: a promising new technique for assessment of the musculoskeletal system. American Journal of Roentgenology, 199(5_supplement), S78–S86. [PubMed: 23097171]

Nocedal J, & Wright S (2006). Numerical optimization. Springer Science & Business Media.

Nocedal J, Wächter A, & Waltz RA (2009). Adaptive barrier update strategies for nonlinear interior methods. SIAM Journal on Optimization, 19(4), 1674–1693.

Pache G, Krauss B, Strohm P, Saueressig U, Blanke P, Bulla S, … & Baumann T (2010). Dual-energy CT virtual noncalcium technique: detecting posttraumatic bone marrow lesions – feasibility study. Radiology, 256(2), 617–624. [PubMed: 20551186]

Patino M, Prochowski A, Agrawal MD, Simeone FJ, Gupta R, Hahn PF, & Sahani DV (2016). Material separation using dual-energy CT: current and emerging applications. Radiographics, 36(4), 1087–1105. [PubMed: 27399237]

Punnoose J, Xu J, Sisniega A, Zbijewski W, & Siewerdsen JH (2016). Spektr 3.0—A computational tool for x-ray spectrum modeling and analysis. Medical physics, 43(8Part1), 4711–4717. [PubMed: 27487888]

Rigie D, & La Riviere PJ (2014, March). Fast model-based restoration of noisy and undersampled spectral CT data. In Medical Imaging 2014: Physics of Medical Imaging (Vol. 9033, p. 90333F). International Society for Optics and Photonics.

Schafer S, Nithiananthan S, Mirota DJ, Uneri A, Stayman JW, Zbijewski W, … & Siewerdsen JH (2011). Mobile C-arm cone-beam CT for guidance of spine surgery: Image quality, radiation dose, and integration with interventional guidance. Medical physics, 38(8), 4563–4574. [PubMed: 21928628]

Shanno DF, & Vanderbei RJ (2000). Interior-point methods for nonconvex nonlinear programming: orderings and higher-order methods. Mathematical Programming, 87(2), 303–316.

Sisniega A, Abascal J, Abella M, Chamorro J, Desco M, & Vaquero JJ (2014). Iterative Dual-Energy material decomposition for slow kVp switching: a compressed sensing approach. In XIII Mediterranean Conference on Medical and Biological Engineering and Computing, pp. 491–494.

Sisniega A, Stayman JW, Capostagno S, Weiss CR, Ehtiati T, & Siewerdsen JH (2021). Accelerated 3D image reconstruction with a morphological pyramid and noise-power convergence criterion. Physics in Medicine & Biology, 66(5), 055012. [PubMed: 33477131]

Sisniega A, Zbijewski W, Badal A, Kyprianou IS, Stayman JW, Vaquero JJ, & Siewerdsen JH (2013). Monte Carlo study of the effects of system geometry and antiscatter grids on cone-beam CT scatter distributions. Medical physics, 40(5), 051915. [PubMed: 23635285]

Stayman JW, Otake Y, Prince JL, Khanna AJ, & Siewerdsen JH (2012). Model-based tomographic reconstruction of objects containing known components. IEEE transactions on medical imaging, 31(10), 1837–1848. [PubMed: 22614574]

Symons R, Krauss B, Sahbaee P, Cork TE, Lakshmanan MN, Bluemke DA, & Pourmorteza A (2017). Photon-counting CT for simultaneous imaging of multiple contrast agents in the abdomen: an in vivo study. Medical physics, 44(10), 5120–5127. [PubMed: 28444761]

Szczykutowicz TP, & Chen GH (2010). Dual energy CT using slow kVp switching acquisition and prior image constrained compressed sensing. Physics in Medicine & Biology, 55(21), 6411. [PubMed: 20938070]

Tao S, Rajendran K, McCollough CH, & Leng S (2019). Feasibility of multi-contrast imaging on dual-source photon counting detector (PCD) CT: An initial phantom study. Medical physics, 46(9), 4105–4115. [PubMed: 31215659]

Thawait GK, Demehri S, AlMuhit A, Zbijweski W, Yorkston J, Del Grande F, … & Siewerdsen JH (2015). Extremity cone-beam CT for evaluation of medial tibiofemoral osteoarthritis: initial experience in imaging of the weight-bearing and non-weight-bearing knee. European journal of radiology, 84(12), 2564–2570. [PubMed: 26388464]

Tilley S, Jacobson M, Cao Q, Brehler M, Sisniega A, Zbijewski W, & Stayman JW (2017). Penalized-likelihood reconstruction with high-fidelity measurement models for high-resolution cone-beam imaging. IEEE transactions on medical imaging, 37(4), 988–999.

Tilley S II, Zbijewski W, & Stayman JW (2019). Model-based material decomposition with a penalized nonlinear least-squares CT reconstruction algorithm. Physics in Medicine & Biology, 64(3), 035005. [PubMed: 30561382]

Tivnan M, Wang W, & Stayman JW (2020). A Preconditioned Algorithm for Model-Based Iterative CT Reconstruction and Material Decomposition from Spectral CT Data. arXiv preprint arXiv:2010.01371.

Vanderbei RJ, & Shanno DF (1999). An interior-point algorithm for nonconvex nonlinear programming. Computational Optimization and Applications, 13(1–3), 231–252.

Wang W, Ma Y, Tivnan M, Li J, Gang GJ, Zbijewski W, … & Stayman JW (2021). High-resolution model-based material decomposition in dual-layer flat-panel CBCT. Medical physics.

Wichmann JL, Booz C, Wesarg S, Kafchitsas K, Bauer RW, Kerl JM, … & Khan MF (2014). Dual-energy CT–based phantomless in vivo three-dimensional bone mineral density assessment of the lumbar spine. Radiology, 271(3), 778–784. [PubMed: 24475863]

Wortman JR, Uyeda JW, Fulwadhva UP, & Sodickson AD (2018). Dual-energy CT for abdominal and pelvic trauma. Radiographics, 38(2), 586–602. [PubMed: 29528816]

Xu S, Uneri A, Khanna AJ, Siewerdsen JH, & Stayman JW (2017). Polyenergetic known-component CT reconstruction with unknown material compositions and unknown x-ray spectra. Physics in Medicine & Biology, 62(8), 3352. [PubMed: 28230539]

Yu L, Liu X, & McCollough CH (2009, March). Pre-reconstruction three-material decomposition in dual-energy CT. In Medical Imaging 2009: Physics of Medical Imaging (Vol. 7258, p. 72583V). International Society for Optics and Photonics.

Zanelli A, Domahidi A, Jerez J, & Morari M (2020). FORCES NLP: an efficient implementation of interior-point methods for multistage nonlinear nonconvex programs. International Journal of Control, 93(1), 13–29.

Zankl M, Eckerman KF, & Bolch WE (2007). Voxel-based models representing the male and female ICRP reference adult—the skeleton. Radiation protection dosimetry, 127(1–4), 174–186. [PubMed: 17545663]

Zbijewski W, Gang GJ, Xu J, Wang AS, Stayman JW, Taguchi K, … & Siewerdsen JH (2014). Dual-energy cone-beam CT with a flat-panel detector: Effect of reconstruction algorithm on material classification. Medical physics, 41(2), 021908. [PubMed: 24506629]

Zbijewski W, Sisniega A, Stayman JW, Thawait G, Packard N, Yorkston J, … & Siewerdsen JH (2015). Dual-energy imaging of bone marrow edema on a dedicated multisource cone-beam CT system for the extremities. In Medical Imaging 2015: Physics of Medical Imaging (Vol. 9412, p. 94120V). International Society for Optics and Photonics.

Zhao C, Liu SZ, Wang W, Herbst M, Weber T, Vogt S, … & Zbijewski W (2021). Effects of x-ray scatter in quantitative dual-energy imaging using dual-layer flat panel detectors. In Medical Imaging 2021: Physics of Medical Imaging (Vol. 11595, p. 115952A). International Society for Optics and Photonics.
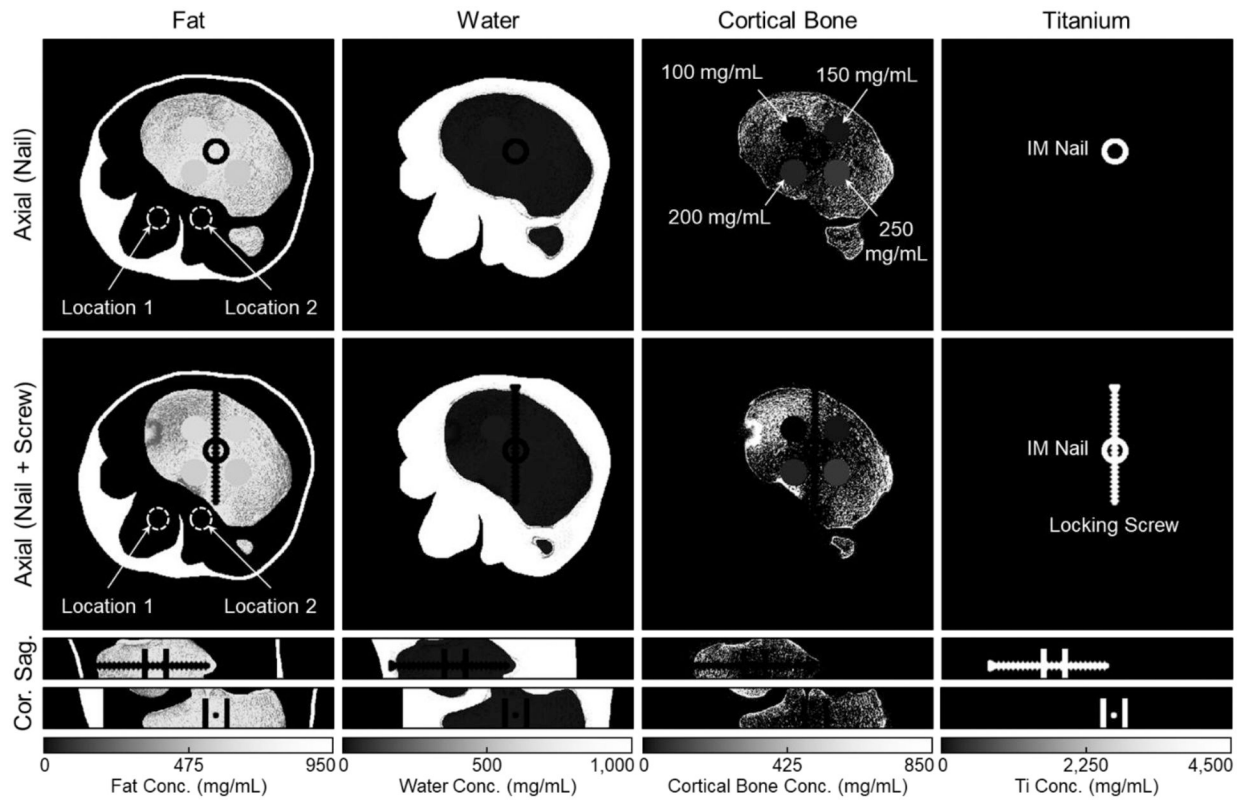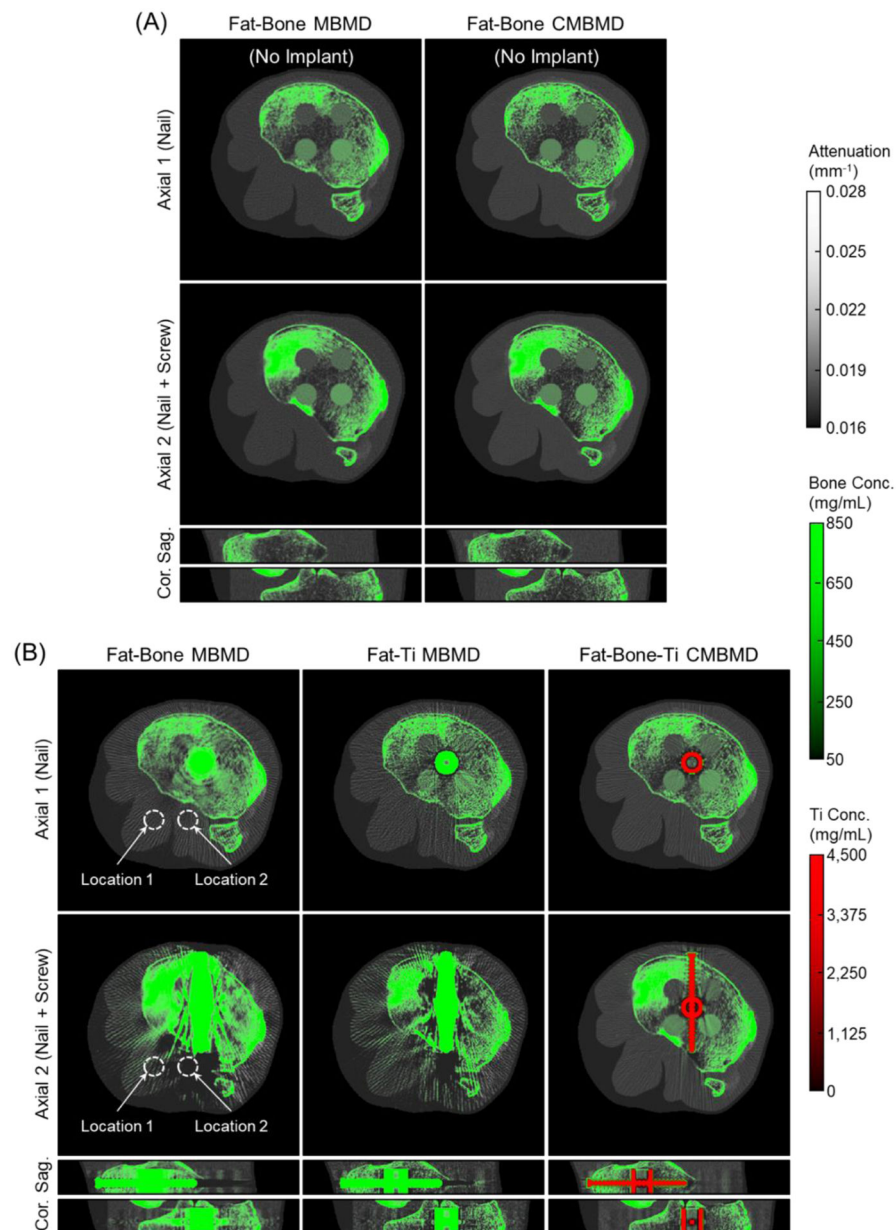
Figure 1.
The knee phantom consisting of fat, water, and cortical bone. Four fat-water-bone target inserts were placed inside the trabecular region, with bone densities indicated on the image. Two Ti implants, including an IM nail and a locking screw, were assembled as illustrated. Dashed circles on the fat axial images denote the regions for CV measurement (see Sec. 3.3).
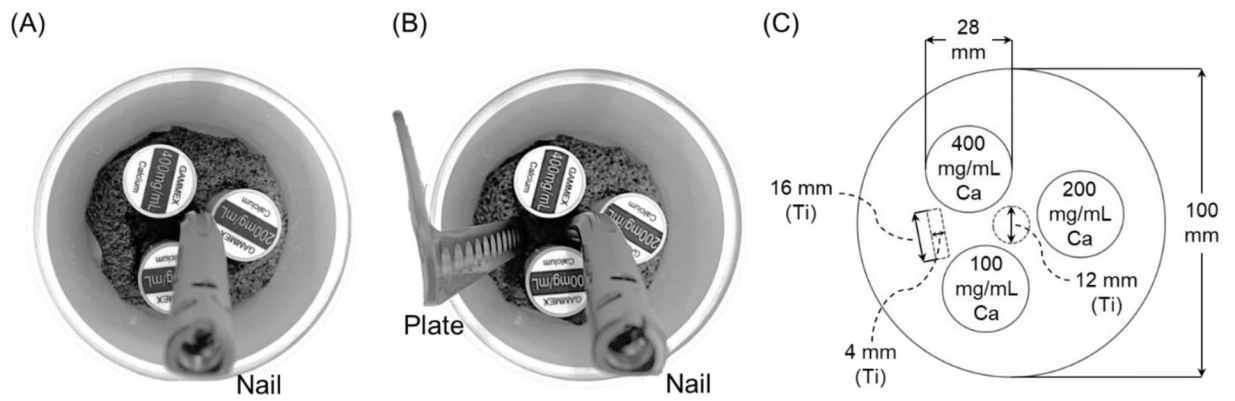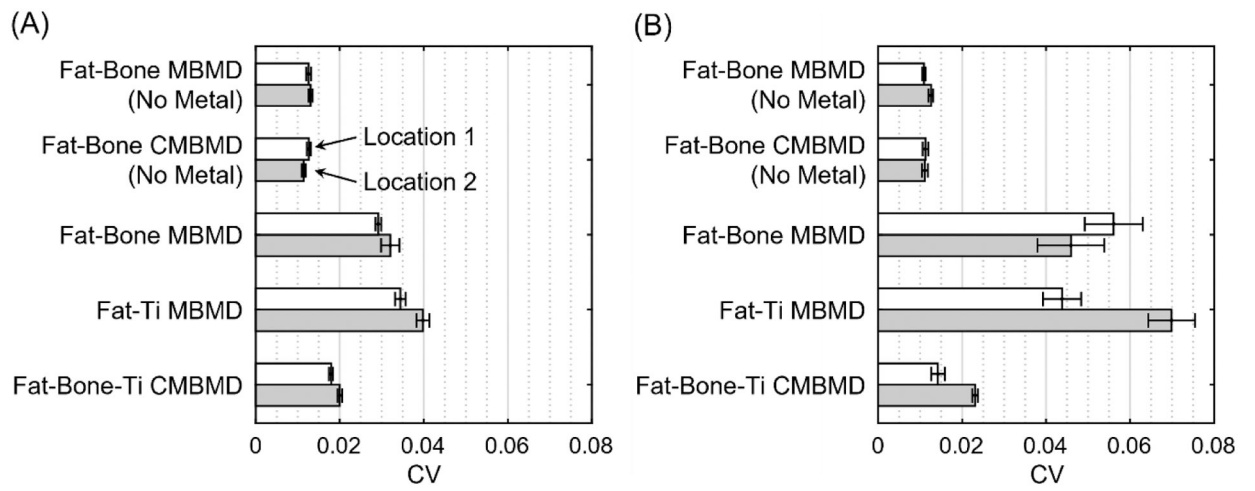
Figure 2.
The physical phantom containing a water bath with three Gammex Ca inserts. Two phantom realizations with different Ti implant configurations were used: (A) tibial nail only, and (B) tibial nail and a plate. Phantom specifications in terms of dimensions and material concentration are shown in (C) for the nail and plate case.

Figure 3.
Composite monochromatic (70 keV) images of decompositions of (A) the reference case with no metal, and (B) the phantom with implants. The images in (A) were obtained using two-material MBMD (left) and two-material CMBMD (right). In (B), the left and central columns are for two-material MBMD, the right column is three-material CMBMD. Two axial views and the central sagittal and coronal views are shown for each decomposition. Estimated concentration maps of cortical bone are overlaid in green; Ti maps (for CMBMD only) are in red. Dashed circles indicate the regions for CV measurement (same as in Fig. 1).
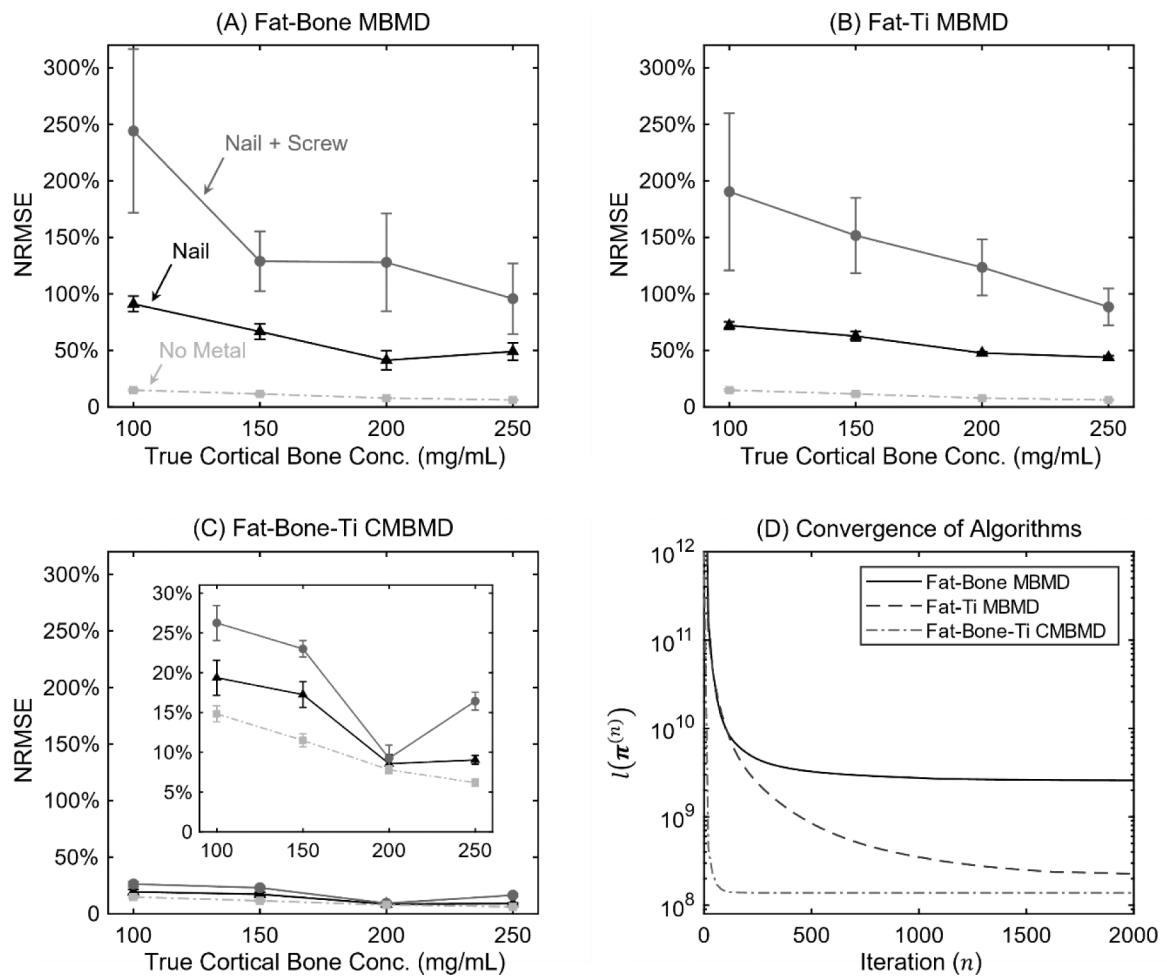
Figure 5.
NRMSE of cortical bone concentration estimates for (A) fat-bone MBMD, (B) fat-Ti MBMD, and (C) fat-bone-Ti CMBMD (note than narrow-range NRMSE view in the inset). Regions with simple (nail only) and complex (nail+screw) metal configurations are assessed separately. Markers and error bars denote the mean and the standard deviation of the NRMSE, respectively. Dotted lines represent reference fat-bone MBMD of the implant-free phantom. (D) Convergence of the data fidelity term for the three decomposition methods.
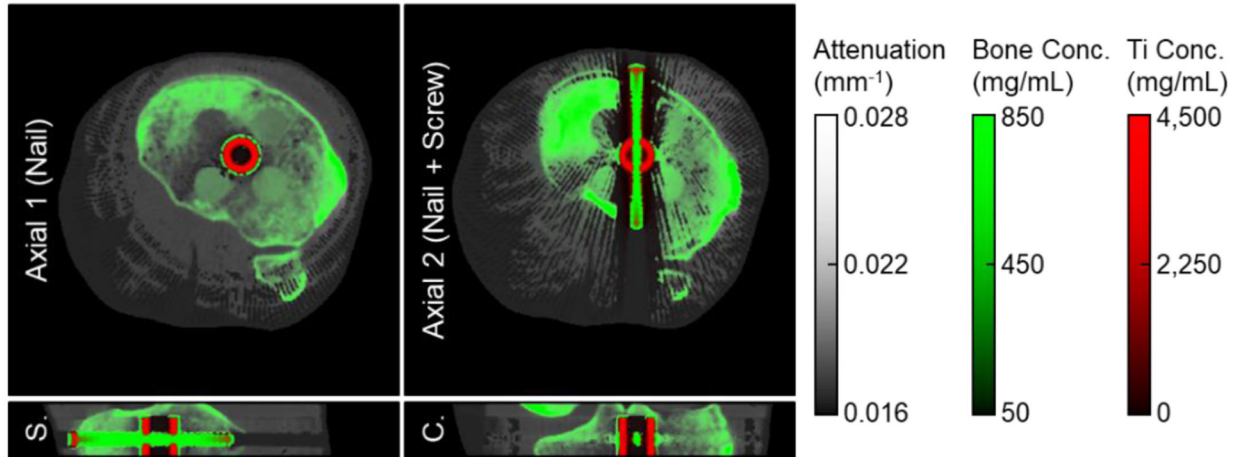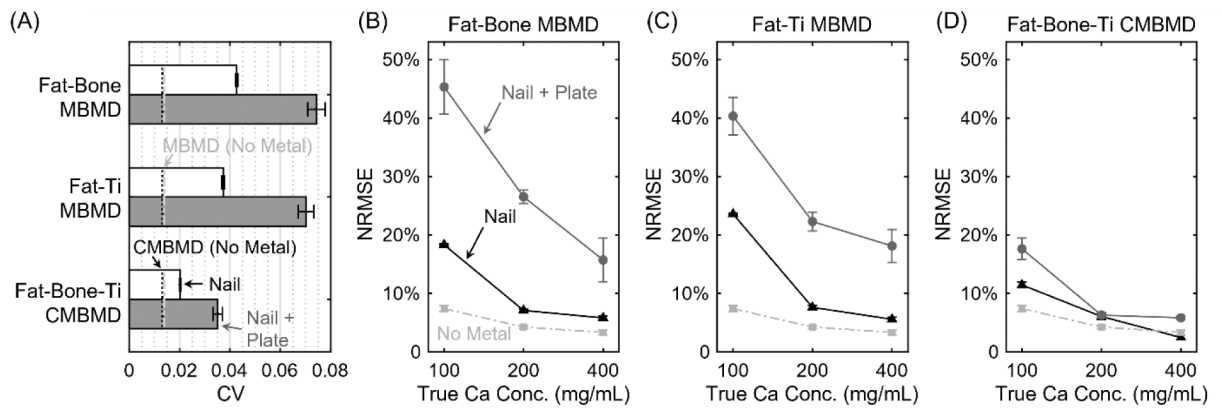
Figure 6.
VCP-constrained three-material IDD decomposition of the phantom with the implant. Two axial views and the central sagittal and coronal views are shown. Estimated concentration maps of cortical bone and Ti are overlaid in green and red, respectively.
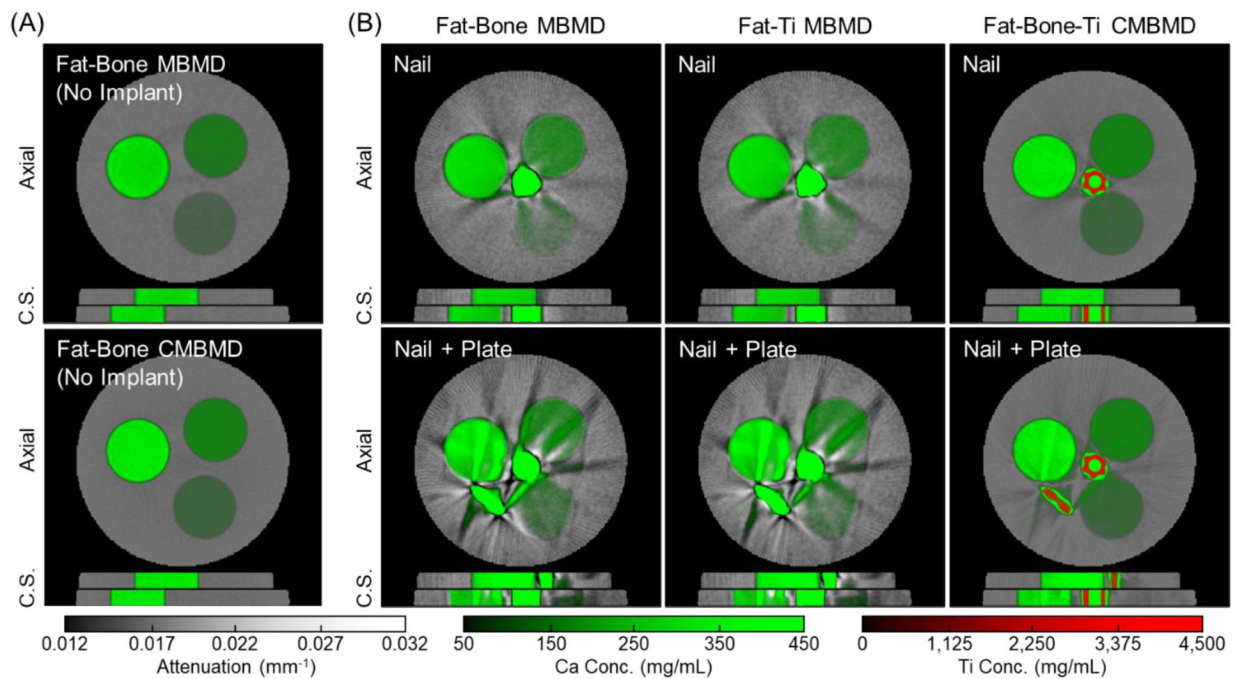
Figure 7.
Two-material MBMD, two-material CMBMD and three-material CMBMD decompositions of experimental CBCT test-bench data. (A) Reference cases with no metal using MBMD (top) and CMBMD (bottom). (B) Objects with metal. The top row in (B) is a phantom with tibial nail only, and the bottom row is a more complex configuration with a nail and a plate. Estimated Ca and Ti (for CMBMD only) concentration maps are overlaid onto a composite monochromatic (70 keV) image in green and red, respectively.

Figure 8.
CV (A) and NRMSE of Ca concentration estimates (B-D) measured in test bench experiments. White bars in (A) represent the phantom with a simple implant configuration (nail-only), gray bars are for a complex configuration (nail and plate). Error bars denote the cross-slice standard deviations. Mean CVs of the phantom without metal are provided as reference: the dash-dot line is for fat-bone MBMD, the dotted line is for fat-bone CMBMD. The NRMSE of fat-bone MBMD is shown in (B), fat-Ti MBMD is in (C), and fat-bone-Ti CMBMD is in (D). Errors for the Ti nail-and-plate phantom (complex implant scenario) are shown using gray continuous lines, black lines are for the nail-only phantom, and dash-dot lines are for the reference MBMD case without metal (CMBMD results for the metal-free case are nearly identical). Markers and error bars denote the mean and the standard deviation in NRMSE for each insert.

Table 1.

Summary of mathematical notation.

Variable	Description	Unit
Forward Model		
i	Number of rays	-
j	Number of image voxels	-
k	Number of base materials	-
$\boldsymbol{\pi}$	Volume fraction vector of reconstructed image	-
$\mu^{\{\kappa\}}(\epsilon)$	Linear attenuation for the κth material at energy ϵ	1/mm
$S_i(\epsilon)$	System spectral response for the ith ray at energy ϵ	photons/kV
e	Number of energy bins in the discretization of $\mu^{\{\kappa\}}$ and s_i	-
$\mathbf{A}_i$	System matrix for the ith ray	mm
$\mathbf{M}_i$	Matrix operator to compute base material line integrals	mm^3/g
$\mathbf{B}_i$	Matrix operator to apply system spectral response	photons
Partitioned Optimization		
p	Number of constrained voxels	-
q	Number of unconstrained voxels	-
$\boldsymbol{\pi}_{\mathcal{U}}$	Volume fraction vector of unconstrained voxels	-
$\boldsymbol{\pi}_{\mathcal{C}}$	Volume fraction vector of constrained voxels	-
$\boldsymbol{\lambda}, \boldsymbol{\theta}$	KKT multiplier vectors for constrained voxels	-
τ	Barrier strength	-
Special Matrices and Operators		
$\exp(\cdot)/\ln(\cdot)$	Element-wise exponentiation/logarithm	-
$\mathbf{0}/\mathbf{1}$	Vector of all zeros/ones	-
$\mathbb{I}$	Identity matrix	-
$\mathbf{D}\{\cdot\}$	Diagonalization operator	-

Table 2.

Pseudocode for the regionally constrained CMBMD framework. Line search, OS, and momentum-based accelerations are omitted. The unconstrained (NR) and constrained (IP) blocks are performed in parallel. $\nabla\varphi^{\pi\mathcal{U}}$, $\left(\nabla\varphi^{\pi\mathcal{C}}\right)$, and $\mathcal{H}_\varphi^{\pi\mathcal{U}(n)}$ $\mathcal{H}_\varphi^{\pi\mathcal{C}(n)}$ represent the unconstrained (constrained) voxels of $\nabla\varphi$ and $\mathcal{H}_\varphi^{(n)}$. In row 3f, $\alpha^* \in \mathbb{R}$ is the maximal step length for the predictor that can be obtained following Nocedal *et al* 2009.

---

1) Initialize $\pi^{(0)}$

2) Set $\xi^{(0)} = 1$ and initialize $\theta^{(0)}$, $\lambda^{(0)}$ using KKT equalities in Eq. (12).

3) **For iteration**
$n = 1, …, N$:

a) Compute $\nabla\varphi\left(\pi^{(n)}; \pi^{(n)}\right)$ and $\mathcal{H}_\varphi^{(n)}$ following Eq. (8).

b) Partition into constrained and unconstrained regions:

$$
\begin{aligned}
\pi^{(n)} &\xrightarrow{\text{Partition}} \left\{\pi_\mathcal{C}^{(n)}, \pi_\mathcal{U}^{(n)}\right\} \\
\nabla\varphi\left(\pi^{(n)}; \pi^{(n)}\right) &\xrightarrow{\text{Partition}} \left\{\nabla\varphi^{\pi\mathcal{C}}\left(\pi_\mathcal{C}^{(n)}; \pi_\mathcal{C}^{(n)}\right), \nabla\varphi^{\pi\mathcal{U}}\left(\pi_\mathcal{U}^{(n)}; \pi_\mathcal{U}^{(n)}\right)\right\} \\
\mathcal{H}_\varphi^{(n)} &\xrightarrow{\text{Partition}} \left\{\mathcal{H}_\varphi^{\pi\mathcal{C}(n)}, \mathcal{H}_\varphi^{\pi\mathcal{U}(n)}\right\}
\end{aligned}
$$

    # Unconstrained NR Update Block:

c) Solve for NR search direction $\Delta\pi_{\mathcal{U}\text{NR}}^{(n)}$: $\mathcal{H}_\varphi^{\pi\mathcal{U}(n)}\Delta\pi_{\mathcal{U}\text{NR}}^{(n)} = \nabla\varphi^{\pi\mathcal{U}}\left(\pi_\mathcal{U}^{(n)}; \pi_\mathcal{U}^{(n)}\right)$

d) Update unconstrained voxels following Eq. (7a).

    # Constrained IP Update Block:

e) Solve for predictors $\Delta\pi_\mathcal{C}^*$, θ^*, λ^* with $\tau = 0$: $\mathcal{J}\nabla\mathcal{L}_0^{(n)}\begin{bmatrix}\Delta\pi_\mathcal{C}^* \\ \Delta\theta^* \\ \Delta\lambda^*\end{bmatrix} = \nabla\mathcal{L}_0\left(\begin{bmatrix}\pi_\mathcal{C}^{(n)} \\ \theta^{(n)} \\ \lambda^{(n)}\end{bmatrix}\right)$

where $\nabla\mathcal{L}_\tau$ and $\mathcal{J}\nabla\mathcal{L}_\tau$ are defined in Eqs. (14) and (15).

f) Set barrier strength $\tau^{(n)}$: $\tau^{(n)} = \sigma\xi^{(n)} = \sigma\dfrac{\pi_\mathcal{C}^{(n)T}\theta^{(n)}}{p\hbar}$

where the centering parameter σ is: $\sigma := \left(\dfrac{\xi^*}{\xi^{(n)}}\right)^3 = \left(\left(\dfrac{\left(\pi_\mathcal{C}^{(n)} + \alpha^*\Delta\pi_\mathcal{C}^*\right)^T\left(\theta^{(n)} + \alpha^*\Delta\theta^*\right)}{p\hbar}\right)/\xi^{(n)}\right)^3$

g) Solve for centered correctors $\Delta\pi_{\mathcal{C}\text{PCIP}}^{(n)}$, $\theta_{\text{PCIP}}^{(n)}$, $\lambda_{\text{PCIP}}^{(n)}$:

$$\mathcal{J}\nabla\mathcal{L}_0^{(n)}\begin{bmatrix}\Delta\pi_{\mathcal{C}\text{PCIP}}^{(n)} \\ \Delta\theta_{\text{PCIP}}^{(n)} \\ \Delta\lambda_{\text{PCIP}}^{(n)}\end{bmatrix} = \nabla\mathcal{L}_0\left(\begin{bmatrix}\pi_\mathcal{C}^{(n)} \\ \theta^{(n)} \\ \lambda^{(n)}\end{bmatrix}\right) + \begin{bmatrix}\mathbf{0} \\ \mathbf{D}\{\Delta\pi_\mathcal{C}^*\}\mathbf{D}\{\Delta\theta^*\}\mathbf{1} + \sigma\xi^{(n)}\mathbf{1} \\ \mathbf{0}\end{bmatrix}$$

h) Update constrained voxels and KKT multipliers following Eq. (13a): $\begin{bmatrix}\pi_\mathcal{C}^{(n+1)} \\ \theta^{(n+1)} \\ \lambda^{(n+1)}\end{bmatrix} = \begin{bmatrix}\pi_\mathcal{C}^{(n)} \\ \theta^{(n)} \\ \lambda^{(n)}\end{bmatrix} - \alpha_{\text{PCIP}}^{(n)}\begin{bmatrix}\Delta\pi_{\mathcal{C}\text{PCIP}}^{(n)} \\ \Delta\theta_{\text{PCIP}}^{(n)} \\ \Delta\lambda_{\text{PCIP}}^{(n)}\end{bmatrix}$

i) Combine the updated $\pi_\mathcal{U}^{(n+1)}$ and $\pi_\mathcal{C}^{(n+1)}$ into the updated volume $\pi^{(n+1)}$.