CT Material Decomposition using Spectral Diffusion Posterior Sampling

Xiao Jiang,

Grace J. Gang,

J. Webster Stayman

Abstract

In this work, we introduce a new deep learning approach based on diffusion posterior sampling (DPS) to perform material decomposition from spectral CT measurements. This approach combines sophisticated prior knowledge from unsupervised training with a rigorous physical model of the measurements. A faster and more stable variant is proposed that uses a "jumpstarted" process to reduce the number of time steps required in the reverse process and a gradient approximation to reduce the computational cost. Performance is investigated for two spectral CT systems: dual-kVp and dual-layer detector CT. On both systems, DPS achieves high Structure Similarity Index Metric Measure(SSIM) with only 10% of iterations as used in the model-based material decomposition(MBMD). Jumpstarted DPS (JSDPS) further reduces computational time by over 85% and achieves the highest accuracy, the lowest uncertainty, and the lowest computational costs compared to classic DPS and MBMD. The results demonstrate the potential of JSDPS for providing relatively fast and accurate material decomposition based on spectral CT data.

I. INTRODUCTION

SPECTRAL CT has enabled a number of applications including density estimation, virtual monochromatic imaging, and contrast agent enhancement. Many of these applications rely on material decomposition, which reconstructs basis material density maps from the spectral projections[1]. Material decomposition is an ill-conditioned nonlinear inverse problem without an explicit solution. Existing material decomposition algorithms can be categorized into three main types: analytical decomposition[2], iterative/model-based decomposition[3], and learning-based decomposition[4].

Deep learning methods have been extensively used in medical image formation including spectral CT[4], [5], [6], [7], [8]. Such approaches leverage prior knowledge learned from large datasets and can generally surpass classic approaches in terms of image quality. However, many deep learning methods do not directly leverage a physical model of the measurements, raising concerns about robustness, network hallucinations, etc. Some methods integrate a physical model as part of network training to improve the data consistency[9], [10]. However, these trained networks are tailored to specific system models, requiring network retraining for system changes in protocol, technique, device, etc.

Diffusion Posterior Sampling (DPS)[11] has established a new framework to integrate a learned prior and physics model. Specifically, DPS starts with unsupervised training of a score-based generative model (SGM) to capture the target domain distribution, then application of a reverse process to estimate image parameters - alternating between SGM reverse sampling to drive the image towards the target distribution and model-based updates to improve the data consistency with measurements. As a result, the final output adheres to both the prior distribution and the actual measurements. A major advantage of DPS is that its network training is not specific to any physical model, allowing for application across different imaging systems. Recently, we have applied DPS to nonlinear CT reconstruction[12], demonstrating its effectiveness for both low-mA and sparse-view single-energy CT reconstruction.

In this work, the original DPS framework is expanded to Spectral DPS (SDPS) specifically for material decomposition in spectral CT. To further enhance this method, we propose a "jumpstarted" sampling strategy with gradient approximation, which significantly stabilizes the sampling process and reduces computational costs. The performance of both SDPS and its jumpstarted variant (JSDPS) are demonstrated in dual-kVp[13] and dual-layer CT[14].

II. METHODOLOGY

A. Spectral Deep Posterior Sampling

1) Score-based Generative Model (SGM) for Inverse Problems: A SGM[15] defines a forward process which continuously perturbs the target-domain sample with time-dependent noise. The corresponding reverse process generates new samples by inverting the perturbation process. Both forward and reverse processes are described by a stochastic differential equation (SDE). Specifically, the SDEs for a Denoising Diffusion Probabilistic Model (DDPM)[16] have the following form at time t:

$$Forward \; : \; \mathrm{d}\mathbf{x} = -\frac{\beta_t}{2}\mathbf{x}\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\mathbf{w}$$

(1a)

$$Reverse \; : \; \mathrm{d}\mathbf{x} = \left[-\frac{\beta_t}{2}\mathbf{x} - \beta_t\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)\right]\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\mathbf{w}$$

(1b)

where β_t is the time-dependent noise variance and $\mathrm{d}\mathbf{w}$ is the standard Wiener process. The unknown score function $\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)$ is approximated by a deep neural network $\mathbf{s}_\theta(\mathbf{x}_t, t)$. A SGM is also capable of posterior sampling from conditional distribution $p(\mathbf{x} \mid \mathbf{y})$:

$$\mathrm{d}\mathbf{x} = \left[-\frac{\beta_t}{2}\mathbf{x} - \beta_t\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t \mid \mathbf{y})\right]\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\mathbf{w}$$

(2)

In the context of CT, $\mathbf{y}$ and $\mathbf{x}$ are the projection measurements and the image volume, respectively. Leveraging Bayes rule $p(\mathbf{x} \mid \mathbf{y}) \propto p(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})$, the reverse sampling (2) may be reformulated as:

$$
\begin{aligned}
d\mathbf{x} = &\left[-\frac{\beta_t}{2}\mathbf{x} - \beta_t \nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t) \right] dt + \sqrt{\beta_t} d\mathbf{w} \\
&- \beta_t \nabla_{\mathbf{x}_t} \log p_t(\mathbf{y} \mid \mathbf{x}_t) dt
\end{aligned}
$$

(3)

The conditional distribution $p_t(\mathbf{y} \mid \mathbf{x}_t)$ can be approximated[11] as

$$
p_t(\mathbf{y} \mid \mathbf{x}_t) \approx p_t(\mathbf{y} \mid \hat{\mathbf{x}}_0), \quad \text{where} \quad \hat{\mathbf{x}}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t + (1 - \bar{\alpha}_t))\mathbf{s}_\theta(\mathbf{x}_t, t)
$$

(4)

Based on this approximation, the deep posterior sampling process could be expressed as[12]:

$$
\begin{aligned}
d\mathbf{x} = &\left[-\frac{\beta_t}{2}\mathbf{x} - \beta_t \nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t) \right] dt + \sqrt{\beta_t} d\mathbf{w} \\
&- \beta_t \nabla_{\hat{\mathbf{x}}_0} \log p_t(\mathbf{y} \mid \hat{\mathbf{x}}_0) \nabla_{\mathbf{x}_t} \hat{\mathbf{x}}_0 dt
\end{aligned}
$$

(5)

Note that the first two term is exactly same as the unconditional DDPM sampling (1b), and $p_t(\mathbf{y} \mid \hat{\mathbf{x}}_0)$ is the likelihood term for the measurements. Therefore, the reverse sampling (5) integrates the prior information captured by DDPM and the physical measurement model provided by the likelihood function.

The forward process is discretized into T time steps [16]:

$$
\begin{aligned}
&\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon \\
&\text{where } \bar{\alpha}_t = \prod_{i=1}^{t}(1 - \beta_t), \epsilon \sim \mathcal{N}(0, \mathbf{I}), t = 1, 2 \ldots T
\end{aligned}
$$

(6)

Since $\epsilon = -\sqrt{1 - \bar{\alpha}_t} \log p_t(\mathbf{x}_t)$, DDPM trains a network $\epsilon_\theta(\mathbf{x}, t)$ to predict the noise ϵ. That is, network parameters, θ, are estimated

$$
\theta^* = \operatorname{argmin} \; \mathbb{E}_{\mathbf{x}_0} \mathbb{E}_{\epsilon, t} \| \epsilon_\theta (\mathbf{X}_t, t) - \epsilon \|_2^2
$$

(7)

Using the trained network and physical model, the image $\mathbf{x}_0$ may be reconstructed by solving the reverse SDE[16], [17].

2)   Diffusion Posterior Sampling for Material Decomposition: This work aims to apply the DPS for the material decomposition problem, which estimates basis material

densities from the spectral CT measurements[1]. We adopt a general spectral CT model[14] where measurements **y** are assumed to follow a multi-variate Gaussian distribution with covariance **K**: $\mathbf{y} \sim \mathcal{N}(\bar{\mathbf{y}}, \mathbf{K})$ and mean:

$$\bar{\mathbf{y}} = \mathbf{BS} \ \exp(-\mathbf{QAx})$$

(8)

where **x** are densities for each basis volume, **Q** are mass attenuation coefficients for each basis, **A** represents (channel-specific) projection matrices, **S** are channel-specific spectral response, and **B** captures overall gain effects. Substituting Eq.(8) into Eq.(5), we have the spectral DPS (SDPS) for material decomposition:

$$d\mathbf{x} = \left[-\frac{\beta_t}{2}\mathbf{x} - \beta_t \nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t) \right] dt + \sqrt{\beta_t} d\mathbf{w}$$
$$- \beta_t \nabla_{\widehat{\mathbf{x}}_0} \|\mathbf{BS} \ \exp(-\mathbf{QA}\widehat{\mathbf{x}}_0) - \mathbf{y}\|_{\mathbf{K}^{-1}}^2 \nabla_{\mathbf{x}_t}\widehat{\mathbf{x}}_0 dt$$

(9)

The SDPS pseudo-code is shown below and illustrated in Fig. 1.

Algorithm 1

Spectral Diffusion Posterior Sampling (SPDS)

1: T: diffusion steps

2: η_t: step size

3: **y**: spectral CT measurement

4: $\mathbf{x}_T \sim \mathcal{N}(0, \boldsymbol{I})$

5: **for** $t = T$ to 1 **do**:

6:    $\mathbf{z} \sim \mathcal{N}(0, \boldsymbol{I})$

7:    $\widehat{\mathbf{x}}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t))$

8:    $\mathbf{x}'_{t-1} = \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t}\widehat{\mathbf{x}}_0 + \sigma_t \mathbf{z}$

9:    $\mathbf{x}_{t-1} = \mathbf{x}'_{t-1} - \eta_t \nabla_{\mathbf{x}_t} \|\mathbf{BS} \ \exp(-\mathbf{QA}\widehat{\mathbf{x}}_0) - \mathbf{y}\|_{\mathbf{K}^{-1}}^2 \ \nabla_{\mathbf{x}_t}\widehat{\mathbf{x}}_0$

10: **end for**

B.   Strategies for fast and stable SPDS

1)   Instability of SDPS: The SDPS framework has the advantage of combining a learned prior and a physics model. However, we have observed that the stability of SDPS is highly dependent on the step size, η_t, and that even with careful tuning, decompositions present large variations over posterior samples and hallucinations. Careful design of a step size scheduler is required to balance the prior information and measurements, particularly for the earlier steps in the reverse process. In early steps, $\mathbf{x}_t$ is far from the solution. While large step sizes may be favorable to enhance data consistency, this risks disrupting diffusion updates - making it challenging to pick optimal step sizes.

2) Jumpstarted Sampling: For many imaging problems, an initial estimate is readily available through fast computation. For example, in spectral CT, projection-domain or image-domain material decomposition [18], [19] can be used as the first-pass estimate, which we denote as $\hat{\mathbf{x}}_0^f$. Forward diffusion can be directly performed on $\hat{\mathbf{x}}_0^f$ as $\hat{\mathbf{x}}_t^f = \sqrt{\overline{\alpha}_t}\hat{\mathbf{x}}_0^f + \left(\sqrt{1 - \overline{\alpha}_t}\right)\epsilon$. The difference between the distribution of $\hat{\mathbf{x}}_t$ and $\hat{\mathbf{x}}_t^f$ can be quantified by the KL divergence:

$$D_{KL}(p(\mathbf{x}_t \mid \mathbf{x}_0) \| p(\hat{\mathbf{x}}_t^f \mid \hat{\mathbf{x}}_0^f)) = \frac{\overline{\alpha}_t}{2(1 - \overline{\alpha}_t)} \|\hat{\mathbf{x}}_0^f - \mathbf{x}_0\|_2^2.$$

(10)

In forward diffusion, $\overline{\alpha}_t/(1 - \overline{\alpha}_t)$ progressively diminishes, ultimately converging towards zero. Consequently, Eq.(10) implies that we can expect for sufficiently large t, the difference between the distribution of $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t^f$ is small enough to be ignored, then the reverse sampling can start from $\hat{\mathbf{x}}_0^f$ instead of pure noise. In effect, we can skip early time steps by using an approximate solution to which an appropriate amount of noise has been added - yielding a more stabilized decomposition by using an initialization closer the solution as well as a faster solution having "jumpstarted" over many early time steps.

3) Gradient Approximation: SDPS computes the data fidelity gradient term via chain rule: $\nabla_{\mathbf{x}_t}\log p_t(\mathbf{y} \mid \hat{\mathbf{x}}_0) = \nabla_{\hat{\mathbf{x}}_0}\log p_t(\mathbf{y} \mid \hat{\mathbf{x}}_0)\nabla_{\mathbf{x}_t}\hat{\mathbf{x}}_0$. The $\nabla_{\mathbf{x}_t}\hat{\mathbf{x}}_0$ term may be expanded as:

$$\nabla_{\mathbf{x}_t}\hat{x}_0 = \frac{1}{\sqrt{\overline{\alpha}_t}}(1 - \sqrt{1 - \alpha_t}\nabla_{\mathbf{x}_t}\epsilon_\theta(\mathbf{x}_t, t))$$

(11)

The computation of the Jacobian $\nabla_{\mathbf{x}_t}\epsilon_\theta(\mathbf{x}_t, t)$ is both time- and memory-intensive, particularly for high dimensional images and deep neural networks. We have experimentally observed that the Jacobian is well-approximated by a diagonal matrix for each time step, and the diagonal elements are of limited range. Therefore, we approximate the Jacobian by a constant (which could be further absorbed into the step size). Incorporating both the jumpstarted sampling and the gradient approximation, we propose the jumpstarted spectral DPS (JSDPS), which is outlined in Algorithm 2 below.

Algorithm 2

Jumpstarted Spectral DPS (JSDPS)

1: T, T': training steps, sampling steps, $T' \ll T$

2: η_t: step size

3: $\mathbf{y}$: spectral CT measurement

4: $\hat{\mathbf{x}}_0^f$: image-domain decomposition

5: $\epsilon \sim \mathcal{N}(0, I)$

6: $\mathbf{x}_{T'} = \hat{\mathbf{x}}_0^{T'} = \sqrt{\overline{\alpha}_{T'}}\mathbf{x}_0 + \sqrt{1 - \overline{\alpha}_{T'}}\epsilon$

7: **for** $t = T'$ to 1 **do**:

8:     $\mathbf{z} \sim \mathcal{N}(0, \boldsymbol{I})$

9:     $\widehat{\mathbf{x}}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t))$

10:     $\mathbf{x}'_{t-1} = \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t}\widehat{\mathbf{x}}_0 + \sigma_t\mathbf{z}$

11:     $\mathbf{x}_{t-1} = \mathbf{x}'_{t-1} - \eta_t \nabla_{\widehat{\mathbf{x}}_0} \|\mathbf{BS} \exp(-\mathbf{QA}\widehat{\mathbf{x}}_0) - \mathbf{y}\|^2_{\mathbf{K}^{-1}}$

12: **end for**

C. Unconditional Multi-Material Generation

1) Dataset Generation: This work focuses on two material decomposition with water and calcium as bases. To build spectral CT and material density dataset, we collected 32000 clinical chest CT slices from the CT Lymph Nodes dataset[20]. Those images are pre-processed to convert Hounsfield Units to attenuation coefficients and remove patient beds. An upper bound value (2000HU) was also applied to remove large attenuation values from metal. Two soft threshold functions were applied to approximate water and calcium densities (unit: $/ml$):

$$\rho_w(\mu) = \begin{cases} k_w\mu, & \text{if } \mu \leq \mu_w \\ k_w\mu_w - k_{wc}(\mu - \mu_w), & \text{if } \mu_w < \mu < \mu_c \\ 0, & otherwise \end{cases}$$

(12a)

$$\rho_c(\mu) = \begin{cases} k_c(\mu - \mu_c) + k_{cw}(\mu_c - \mu_w), & \text{if } \mu \geq \mu_c \\ k_{cw}(\mu - \mu_w), & \text{if } \mu_w < \mu < \mu_c \\ 0, & otherwise \end{cases}$$

(12b)

In this work, k_w, k_{wc}, k_{cw}, k_c were empirically set to $5.18, -8.77, 5.69, 2.12$ g/cm^2, while μ_w and μ_c were 0.22 cm^{-1} and 0.35 cm^{-1}. The water and calcium densities were used as ground truth to simulate spectral measurements according to Eq.8.

2) Network Training: DDPM[16] is employed as the SGM framework, with the Residual Unet[21] as the backbone network. Paired water and calcium images are concatenated to form the 2 channel image $\mathbf{x}_0$. The discretized diffusion process uses $T = 1000$ time steps with a linear variance scheduler from $\beta_1 = 1e^{-4}$ to $\beta_{1000} = 0.02$. Training and validation datasets contain 25000 slices and 5000 slices, respectively. The loss function (7) is minimized by the Adam optimizer with a batch size of 16 and a learning rate of 10^{-4}. Training terminated after 200 epochs.

D. Conditional Posterior Sampling for Different Spectral Systems

The unconditional DDPM training captures the material bases distribution without specification of the spectral CT device. The trained model can be used in material decomposition for arbitrary spectral CT systems. Here we investigate the SDPS/JSDPS performance on a simulated dual-layer CT [14] and dual-kVp CT system [13]. The dual-layer CT uses a detector with $300\mu m$ CsI and $600\mu m$ CsI scintillators in the top and bottom layers, respectively (with a 5mm gap between layers). The x-ray tube operates at 120 kVp. The dual-kVp system uses single-layer detector with $600\mu m$ CsI. The tube voltage alternates between 80 kVp and 120 kVp every other view. For both systems, 800 projections were simulated with Poisson noise equivalent to an exposure of 0.05 mAs/view. Voxel size and detector pixel size were set to 0.8 *mm* and 1.0 *mm*, respectively.

E. Evaluation

We first evaluated the performance of JSDPS with and without gradient approximation. A standard image-domain decomposition from filtered backprojection reconstructions, i.e., the initialization of the JSDPS algorithm, was also shown for comparison. Second, we compared the performance of several material decomposition algorithms including model-based material decomposition (MBMD) [3], SDPS, and JSDPS. MBMD was formulated as the following optimization problem:

$$\hat{\mathbf{x}} = \mathrm{argmin} \|\mathbf{BS} \ \exp(-\mathbf{QAx}) - \mathbf{y}\|_{\mathbf{K}^{-1}}^2 + \lambda_w \mathbf{R}(\mathbf{x}_w) + \lambda_c \mathbf{R}(\mathbf{x}_c)$$

(13)

where $\mathbf{R}$ is a quadratic gradient roughness penalty. Regularization strengths, λ_w, λ_c, for water and calcium were set to 10^{-4} and 4×10^{-4}, respectively. We used 10000 iterations of separable paraboloidal surrogates updates, and final change in loss is $\sim 0.01\%$ from the previous iteration. SDPS was implemented according to Algorithm 1, and the step size scheduler follows [11]: $\eta_t = \eta / \|\mathbf{BS} \ \exp(-\mathbf{QAx}) - \mathbf{y}\|_{\mathbf{K}^{-1}}$. We swept η from 0.1 to 10, and the optimal η was determined as the one with minimal MSE from ground truth. JSDPS used a constant step scheduler with $T' = 150, \eta_t = 0.02$. Performance was evaluated for computation time (calculated for eight parallel decompositions), and image quality - quantified by Structural Similarity Index Measure(SSIM) and Peak Signal-to-Noise Ratio (PSNR). Additionally, to compare variability and biases in the posterior samples provided by SDPS and JSDPS, we compute sample bias and standard deviation over an ensemble of 16 outputs:

$$\mathrm{Bias} = \|\mathbb{E}\{\hat{\mathbf{x}}\} - \mathbf{x}\|_2 \quad \mathrm{Std} = \|\mathbb{E}\{(\hat{\mathbf{x}} - \mathbb{E}\{\hat{\mathbf{x}}\})^2\}\|_2$$

(14)

III. RESULTS

A. Illustration of JSDPS variants

Fig.2 displays JSDPS results as compared with standard image-domain decomposition (initialization) and the ground truth. Image-domain decomposition shows significant image noise and streaking. JSDPS without gradient approximation effectively captures most of the high-contrast structures, including smaller pulmonary vessels. However, as seen in the zoomed-in ROIs, low-contrast soft tissue hallucinations are present. In the contrast enhanced areas of the calcium image, such as the aorta, deformations are also observed. By using the gradient approximation, the reverse sampling achieves 20.2% and 29.3% computation time decrease for the dual-layer and dual-kVp system, respectively. Gradient approximation also effectively mitigates the hallucinations and enhances the consistency with ground truth for vascular structures.

B. Comparison of SDPS, JSDPS, and MBMD

1) Image Quality: Figure 3 compares the performance of SDPS, JSDPS, and MBMD in water and calcium bases for the dual layer and dual kV systems. Four zoomed in areas are shown in the (1) lung, (2) heart (iodine-filled aorta in the calcium basis), (3) low-contrast soft tissue, and (4) spine. Comparing across all three algorithms, MBMD produces images with the highest noise (evident in ROIs 2 and 3) at the lowest spatial resolution (evident in ROI 1). The SDPS results have low noise and high spatial resolution, but are prone to hallucinations observed as erroneous anatomy in all four ROIs. While SDPS produces results that are more closely aligned with the prior distribution, the instability of the algorithm is problematic for the low fidelity (low dose) data simulated in this work. The JSDPS algorithm outperforms both MBMD and SDPS from both visual observations of anatomical structures faithful to the ground truth and according to SSIM and PSNR. Good decomposition results were obtained for both the dual layer and dual kV system, demonstrating generalizability to different spectral systems.

Figure 4 shows the bias and standard deviation maps computed from an ensemble of 16 individual output samples. SDPS exhibits large variability and bias around edge features, suggesting substantial image uncertainty. In contrast, JSDPS displays markedly reduced bias and standard deviation, demonstrating the capability to effectively stabilize the sampling process.

2) Computational Cost: Table I summarizes the computational cost of different material decomposition algorithms for eight output estimates. MBMD is the slowest due to the large number of iterations (10000) required for convergence. SDPS substantially reduces the number of iterations (1000), thereby shortening the computation time. Leveraging the jumpstarted sampling strategy and the gradient approximation, JSDPS requires only 150 iterations, which results in an additional 86.83% and 88.07% reduction in computation time for dual-layer and dual kVp CT, respectively. Furthermore, a 23.16% and 42.54% memory saving was also achieved compared to SDPS for the two spectral systems due to the absence of Jacobian computation.

IV.   CONCLUSION AND DISCUSSION

Spectral CT material decomposition is a challenging and ill-conditioned problem, often exhibiting excessive noise and slow convergence. The novel SDPS framework is designed to tackle these issues - combining a sophisticated learned SGM prior and a physical model, to achieve low image noise while enhancing image quality. We additionally adopted jumpstarted sampling and gradient approximation strategies which were found to be effective in stabilizing performance as well as reducing computation time and memory consumption. in evaluations on both dual-layer and dual-kVp systems, JSDPS achieves the highest accuracy, the lowest uncertainty, and the lowest computational costs compared to SDPS and MBMD. This work demonstrated that JSDPS is a promising approach for spectral CT decomposition.

Acknowledgments

REFERENCES

[1]. Long Y and Fessler JA, IEEE transactions on medical imaging, vol. 33, no. 8, pp. 1614–1626, 2014. [PubMed: 24801550]

[2]. Jiang X et al. Medical physics, vol. 48, no. 9, pp. 4843–4856, 2021. [PubMed: 34289129]

[3]. Tilley S et al. Physics in Medicine & Biology, vol. 64, no. 3, p. 035005, 2019. [PubMed: 30561382]

[4]. Gong H et al. Medical physics, vol. 47, no. 12, pp. 6294–6309, 2020. [PubMed: 33020942]

[5]. Wu W et al. Neural Networks, vol. 144, pp. 342–358, 2021. [PubMed: 34560584]

[6]. Abascal JF et al. IEEE Access, vol. 9, pp. 25 632–25 647, 2021.

[7]. Clark DP et al. in Medical Imaging 2018: Physics of Medical Imaging, vol. 10573. SPIE, 2018, pp. 415–423.

[8]. Zhu J et al. Physics in Medicine & Biology, vol. 67, no. 14, p. 145012, 2022.

[9]. Fang W et al. Physics in Medicine & Biology, vol. 66, no. 15, p. 155013, 2021.

[10]. Eguizabal A et al. arXiv preprint arXiv:2208.03360, 2022.

[11]. Chung H et al. arXiv preprint arXiv:2209.14687, 2022.

[12]. Li S et al. arXiv preprint arXiv:2312.01464, 2023.

[13]. Cassetta R et al. Physics in Medicine & Biology, vol. 65, no. 1, p. 015013, 2020. [PubMed: 31775131]

[14]. Wang W et al. Medical physics, vol. 48, no. 10, pp. 6375–6387, 2021. [PubMed: 34272890]

[15]. Song Y et al. arXiv preprint arXiv:2011.13456, 2020.

[16]. Ho J et al. Advances in neural information processing systems, vol. 33, pp. 6840–6851, 2020.

[17]. Song J et al. arXiv preprint arXiv:2010.02502, 2020.

[18]. Stenner P et al. Medical physics, vol. 34, no. 9, pp. 3630–3641, 2007. [PubMed: 17926967]

[19]. Mendonça PR et al. IEEE transactions on medical imaging, vol. 33, no. 1, pp. 99–116, 2013. [PubMed: 24058018]

[20]. Roth HR et al. in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part I 17. Springer, 2014, pp. 520–527.

[21]. Jiang X et al. in 7th International Conference on Image Formation in X-Ray Computed Tomography, vol. 12304. SPIE, 2022, pp. 488–492.

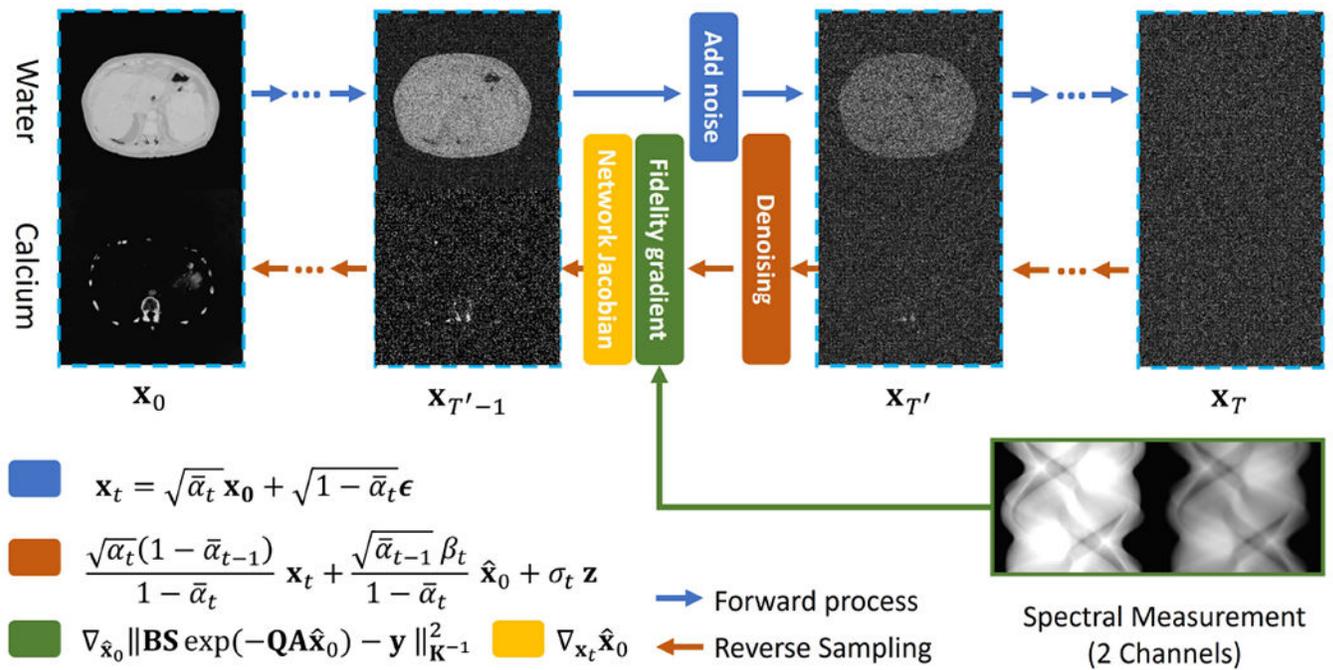**Fig. 1.**
SDPS Workflow: Basis material images are concatenated to form $\mathbf{x}_0$. The forward process progressively perturbs images with noise, and a noise prediction network is trained. The reverse process uses DDPM sampling to progressively denoise the image and uses a likelihood-based update to enforce data consistency.
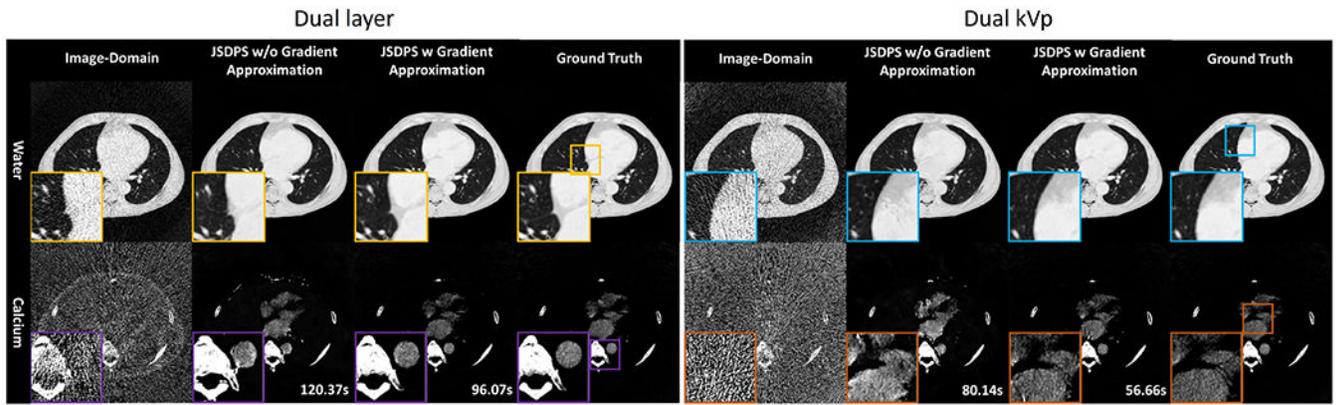
Fig. 2.

Decomposed water (top, W/L: 1.2/0.6 g/ml) and Calcium (bottom, W/L: 0.05/0.1 g/ml) images. Left to right: Image-domain decomposition, JSDPS without gradient approximation, JSDPS with gradient approximation, ground truth. Computational time for eight outputs is provided in the bottom left corner.
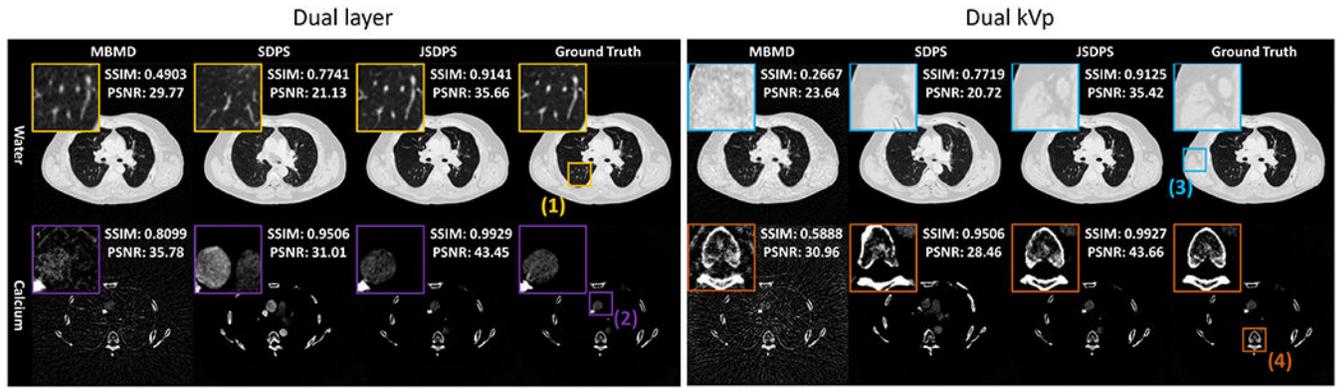
Fig. 3.

Decomposed water (top, W/L: 1.2/0.6 g/ml) and Calcium (bottom, W/L: 0.05/0.1 g/ml) for (left to right): MBMD, SDPS, JSDPS, ground truth.
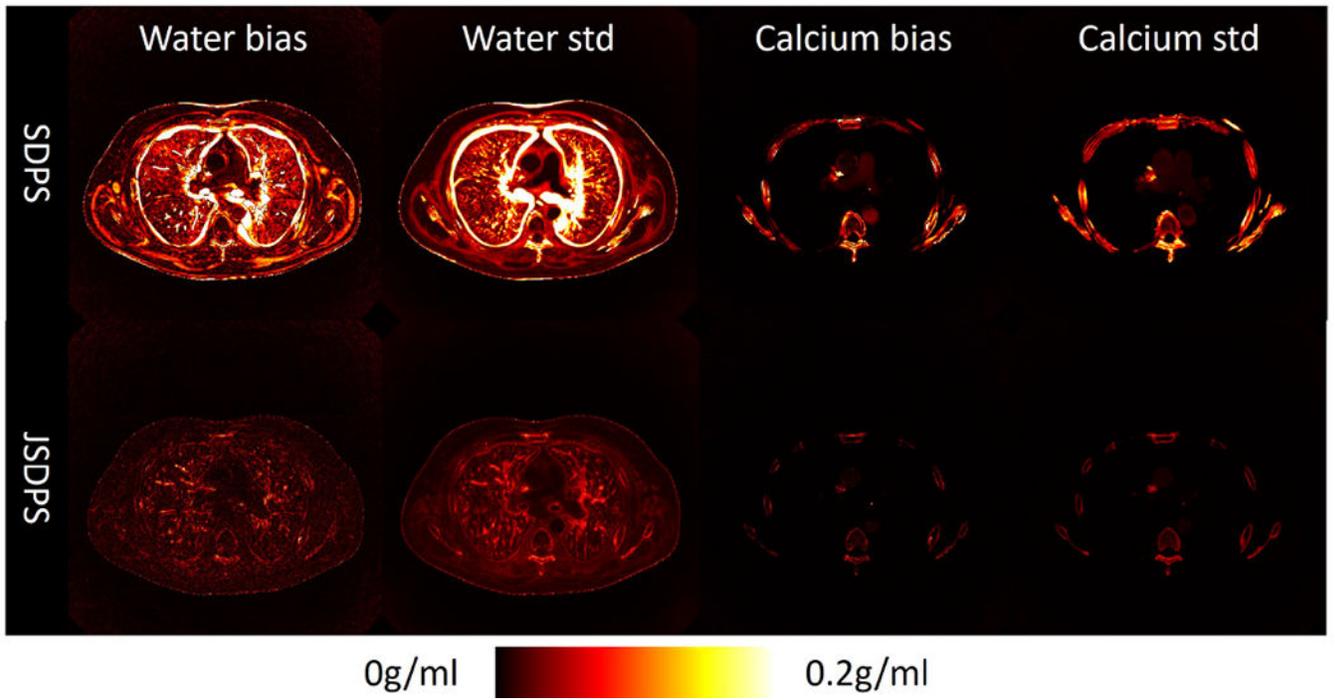
Fig. 4.
Bias and standard deviation map for SDPS and JSDPS. Here we only show the results of dual-kVp system, which is similar to that of the dual-layer system.

TABLE I

C OMPUTATIONAL COST OF MATERIAL DECOMPOSITION ALGORITHMS

	Dual layer		Dual kVp	
	Time(s)	Memory(GB)	Time(s)	Memory(GB)
MBMD	6406.3	11.3	2802.5	4.8
SDPS	728.2	17.7	478.6	13.4
JSDPS	95.9	13.6	57.1	7.7