

In Regards to Su et al. 2022: "[DIRECT-Net: A unified mutual-domain material decomposition network for quantitative dual-energy CT imaging](#)"

Xiaochuan Pan^{1,2} and Emil Sidky¹

¹Department of Radiology, The University of Chicago, Chicago, USA

²Department of Radiation & Cellular Oncology, The University of Chicago, Chicago, USA

In Response to Pan & Sidky

Ting Su¹, Xindong Sun², Jiecheng Yang¹, Yikun Zhang³, Haodi Wu⁴, Shibo Fang¹, Yang Chen³, Hairong Zheng⁵, Dong Liang^{1,5} and Yongshuai Ge^{1,5}

¹ Research Center for Medical Artificial Intelligence, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

² School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

³ School of Computer Science and Engineering, Southeast University, Nanjing 210096, China.

⁴ Wuhan National Laboratory for Optoelectronics (WNLO), Huazhong University of Science and Technology (HUST), Wuhan 430074, China

⁵ Paul C Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

Reply: Quantitative image reconstruction in dual-energy computed tomography (CT) remains a topic of active research. We read with interest "DIRECT-Net: A unified mutual-domain material decomposition network for quantitative dual-energy CT imaging," which, referred to as the paper hereinafter, appears in the 2022 February Issue of Med Phys.¹ In the paper the author propose a deep-learning (DL) method, referred to as the Direct-Net method, to address the problem of quantitative image reconstruction directly from data in full-scan dual-energy CT (DECT). We would like to comment on the study and conclusion in the paper. (Words and numbers in italic below are quoted from the paper.)

Response: We thank Drs. Pan and Sidky for this opportunity to discuss the published work. For convenience, comments from Drs. Pan and Sidky are referred to as "Pan & Sidky comments" and the published paper is referred to as "the paper".

Reply: No meaningful evidence is provided in support of the paper's conclusion. An extensive body of research work in the literature on theoretical development and practical application of methods exists that can well address the problem. The paper concludes that "[the] DIRECT-Net has promising benefits in improving the DECT image quality." As the conclusion stands, evidence is needed to demonstrate the "benefits" of the Direct-Net method proposed for improving over, i.e., performing better than, relevant existing methods addressing the same problem. As commented below, however, no meaningful evidence can be identified in the paper supporting the conclusion.

Response: Thank you for the comment. We agree that DECT reconstruction is not a new research topic and that there exists an extensive body of prior works. As Pan & Sidky have made clear in the reply opening statement, this is still a "topic of active research." Given the fact that DL offers a new paradigm to study both new and old research topics, the purpose of the published paper is to apply this new methodology to DECT reconstruction problems.

Clearly, different people may have different understandings and definitions to the meaning of "promising benefits." To us, a research method such as the proposed DL method in our work that offers flexibility to incorporate information

hidden in the training samples into DECT reconstruction problems is a benefit to the methodology itself. In the published paper, we did compare the proposed method with other attempts that use DL to tackle similar problems to show qualitative and quantitative improvement, and the presented results did show “promising benefits.” As to whether a conclusion is “meaningful” or not, different scientists may have different standards and definitions. Without knowing Pan & Sidky’s specific standards and definitions, it can be very difficult for us and the scientific community to openly discuss the scientific issues raised in a fair and objective manner.

Nevertheless, the authors do admit that the comparisons performed in the published work may not be thorough enough to include a systematic performance comparison among many other existing methods and the authors did not claim that the performance study is exhaustive either. In this regard, the authors would also appreciate if Pan & Sidky can make it clear what the “relevant existing methods” are in Pan & Sidky’s mind. In the authors’ view, whether a work/method is relevant or not could be subjective. Again, different scientists may have different standards to define “relevant existing methods” for comparison. Without a commonly agreed-upon set of “relevant existing methods” and a well-accepted state-of-the-art reference method in the scientific community, it might be difficult for researchers to conduct comparisons to satisfy everyone.

Finally, the conclusion in the paper was phrased as “promising benefits,” meaning “potential benefits.” In the authors’ view, any benefits/advantages of an algorithm can only be claimed as “potential” or “promising” before it is applied to a well-curated clinical data set to perform comparisons with rigorous statistical analysis.

Reply: The level of relative errors in concentration estimations for the calcium chloride and iodine solution phantoms in the paper, as reported in Tables 2 and 3, appears notably inconsistent with that of relative errors reported in the literature for the methods compared. Much of the relative errors obtained in the paper for the comparison methods (references cited), i.e., ID-EP(Ref. [7]),² Direct-JSI (Ref. [14]),³ Butterfly-Net (Ref. [20]),⁴ and Clark-Net (Ref. [23]),⁵ are over 20% in Table 2 and over 10% in Table 3, with maximum relative errors > 90%. In contrast, in Table II of Ref. [7],² an average relative error of 4.2% with a max of 12.83% has been reported; in Table II of Ref. [14],³ relative errors can be computed as <1% for bone and tissue regions; in Table I of Ref. [20],⁴ relative errors can be estimated to be <1%; and in Ref. [23],⁵ the relative error has been reported as about 3%. While the study conditions in Refs. [7],² [14],³ [20],⁴ and [23]⁵ may not be identical to those in the paper, the difference in estimation errors between the Direct-Net method and the comparison methods appears perplexingly large that it calls for proper explanation.*

Accurate concentration estimates depend primarily on accurate physics modeling of the measurement, and estimation error can occur when the modeled physics such as spectra is not the same as the physics in simulation and/or experiment DECT. It is unclear if there is any mismatch between the physics in the data and the modeled physics in the methods compared.

For the non-DL methods considered in the paper, results are obtained with an image-domain decomposition (IDD) method, referred to as the ID-EP method (Ref. [7]),² and a direct iterative decomposition method, referred to as the Direct-JSI method (Ref. [14]).³ The concentration estimations of the Direct-JSI method are much worse than those of the ID-EP method in the paper. The former appears to use the proper physics modeling, while the latter makes a linear approximation of the logarithm of equation (1) in the paper; one would thus expect that the performance of the former is at least as good as that of the latter. It is unclear as to why the result of the Direct-JSI method appears considerably worse than that of the ID-EP method.

Response: Thank you. The authors appreciate this comment, but the authors want to say that these relative errors are real and can reoccur if repeating the experiments with the same parameters, e.g., the spectra, iterative parameters,

and so on. To demonstrate that the large errors (or gaps) of the DIRECT-JSI algorithm during the implementations and validations are real, the entire algorithm was revisited carefully, and the results are shown below.

I. Validation of the algorithm implementation with simulated noise-free data

The implementation of the DIRECT-JSI algorithm in this study was based on the code provided in [19]. In order to make sure the numerical implementation is not a confounding factor, DECT results were regenerated using a numerical phantom having 7 inserts with different iodine concentrations (2 mg/cm³, 2.5 mg/cm³, 5 mg/cm³, 7.5 mg/cm³, 10 mg/cm³, 15 mg/cm³, 20 mg/cm³). The imaging geometry and energy spectra were identical to those used in the paper. To focus on the accuracy, as emphasized in Pan & Sidky’s comment, noise was not added in numerical simulation studies. After the sanity check was performed, the DIRECT-JSI method was applied to generate results for performance evaluations. Note that the regularization parameter was also set to zero since there was no added noise in this case. Moreover, the number of subsets was set to be 4, and the number of iterations was set to be 500 to achieve empirical convergence. The results presented in Fig. 1 show the change of the cost function and also the root mean square error (RMSE) with the increase of iteration numbers. Figure 2 shows the quantification results of the seven iodine inserts. Note that the relative error ($|\sum_{j \in \text{Insert}} \rho_j / N_{\text{insert}} - \rho_{\text{true}}| / \rho_{\text{true}}$) of these seven iodine inserts is about negligible (0.00001%~0.01%) after 100 iterations. This sanity check clearly indicates that the numerical implementation of the reported DIRECT-JSI is correct.

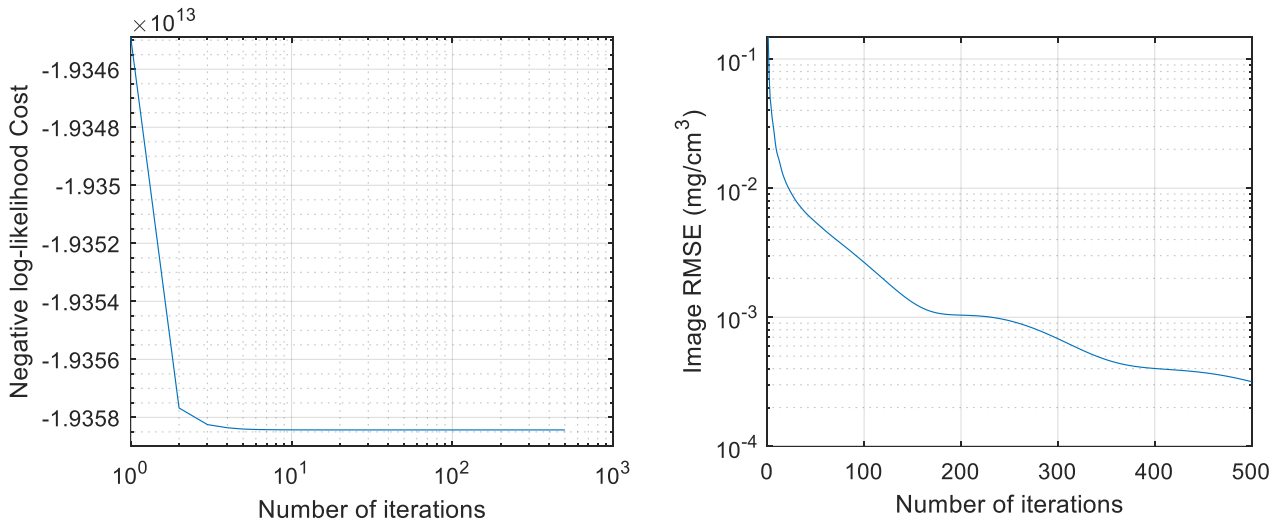


Fig. 1. (left) The negative log-likelihood cost and (right) the image RMSE as a function of the number of iterations for the simulated noise-free data.

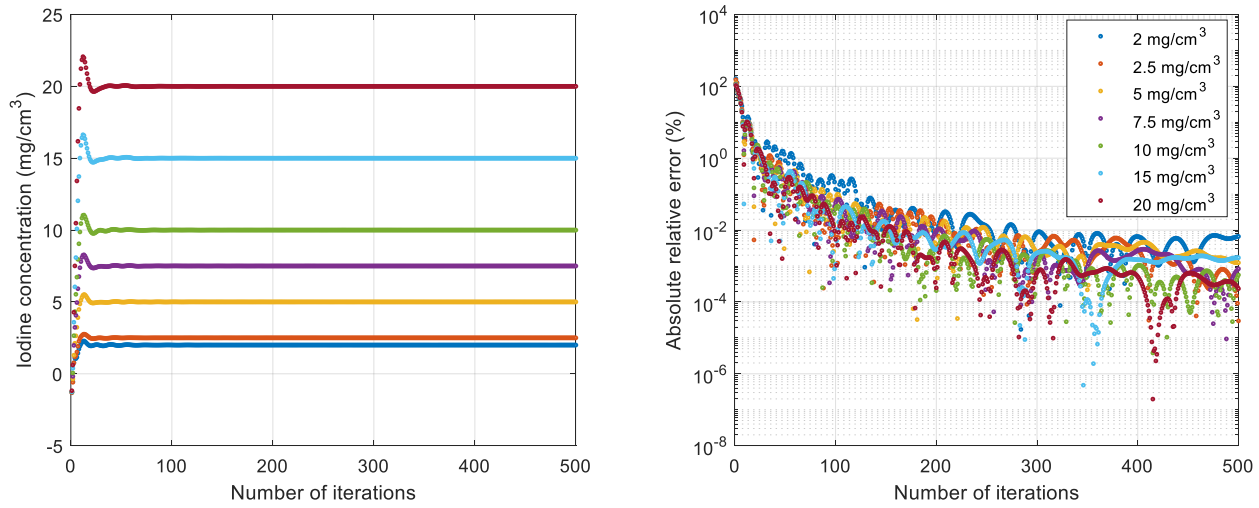


Fig. 2. (left) The quantified iodine concentrations and (right) the absolute relative errors as a function of the number of iterations for the simulated noise-free data.

II. Influence of the iteration parameters on experimental data

After the above validations, the authors re-implemented the DIRECT-JSI algorithm on the experimental iodine phantom data using the same parameters to obtain Fig. 8(c) in the published paper, except that the iteration number increased up to 500. Specifically, the number of subsets was set to 4, the regularization weight λ_b was 10^3 for water basis and 10^7 for iodine basis, the tuning parameter γ in the Huber regularization was set to be 0.1 for water basis and 0.001 for the iodine basis, and the geometrical neighborhood N_j was 3×3 . As one can observe in the results presented in Fig. 3 below, the empirical convergence was observed after 100 iterations. Similar to the results presented in the paper, the absolute relative errors are between 20%~50%.

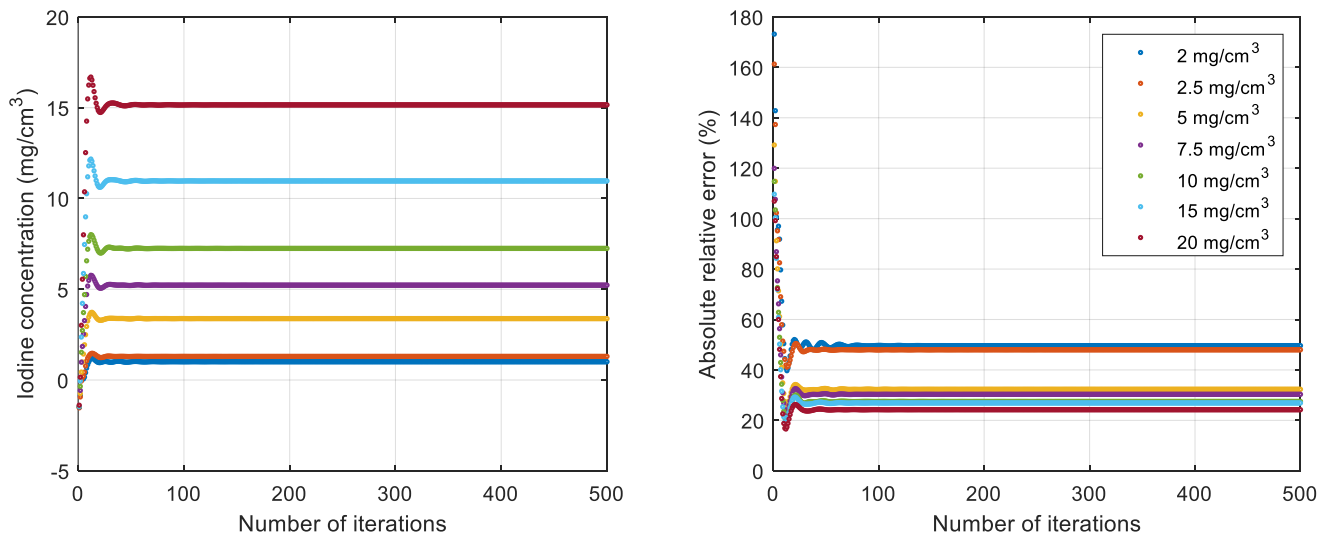


Fig. 3. (left) The quantified iodine concentrations and (right) the absolute relative errors as a function of the number

of iterations for the experimental iodine-phantom data.

To investigate the potential causes of the relatively large relative errors above, several additional experiments were performed to examine the impact of the regularization parameter on the quantification results. Results are presented in Fig. 4.

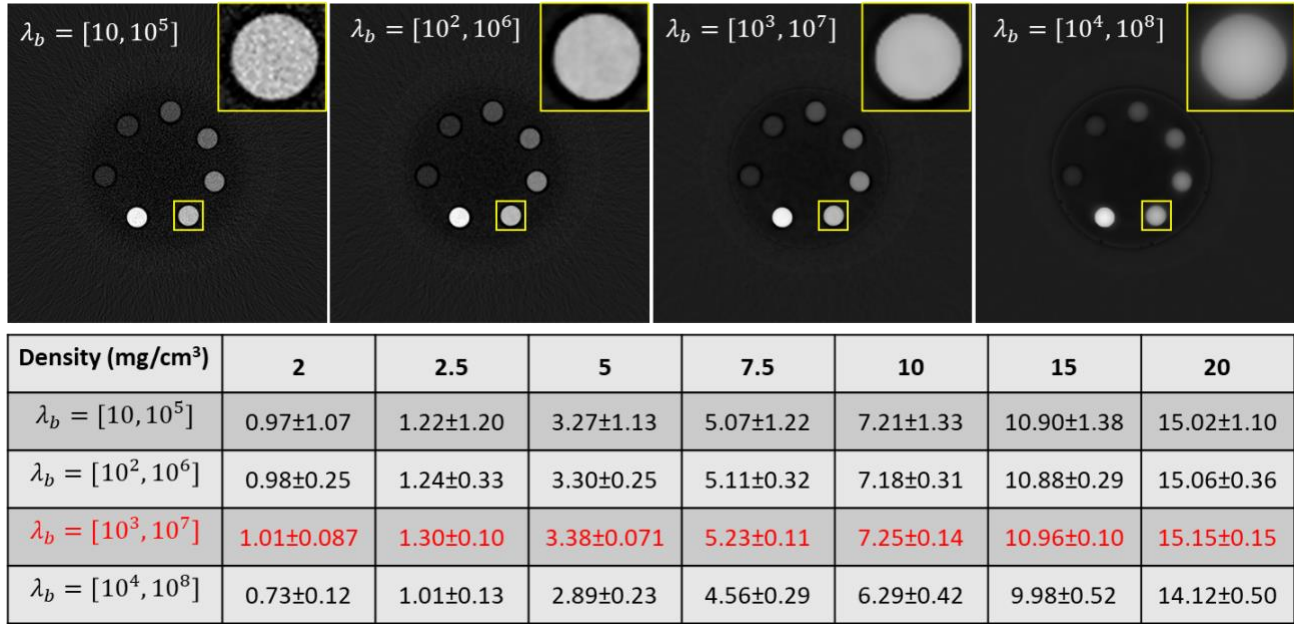


Fig. 4. Top row: the decomposed iodine images with different regularization weights λ_b . The display window is $[-2, 16]$ mg/cm³. Bottom row: the mean value and standard deviation of the calculated iodine concentrations. The highlighted parameter $\lambda_b = [10^3, 10^7]$ was used in the paper.

III. Influence of the spectra on experimental data

To investigate the influence of the dual-energy spectra on the decomposition performance, we simulated three low-energy (LE) spectra (S_{LE}^1 , S_{LE}^2 , and S_{LE}^3) and three high-energy (HE) spectra (S_{HE}^1 , S_{HE}^2 , and S_{HE}^3) by SpekCalc (see Fig. 5). Note that S_{LE}^2 and S_{HE}^2 are the spectra used for material decomposition in the paper, and the other spectra were generated by varying the copper filtration thickness. Next, the DIRECT-JSI algorithm is performed multiple times using different LE and HE spectral combinations. The absolute relative errors shown in Fig. 6 indicated that the material decomposition accuracy strongly depends on the choice of energy spectra: the spectra combinations S_{LE}^1/S_{HE}^2 and S_{LE}^2/S_{HE}^3 generated the lowest overall quantification error; the spectra combination S_{LE}^1/S_{HE}^3 yield less than 4% error for concentrations between 7.5 mg/cm³ and 20 mg/cm³. However, the spectra combinations S_{LE}^2/S_{HE}^1 and S_{LE}^3/S_{HE}^1 show much larger quantification errors. The decomposed basis images of the spectral combination of S_{LE}^1/S_{HE}^3 and S_{LE}^3/S_{HE}^1 are shown in Fig. 7.

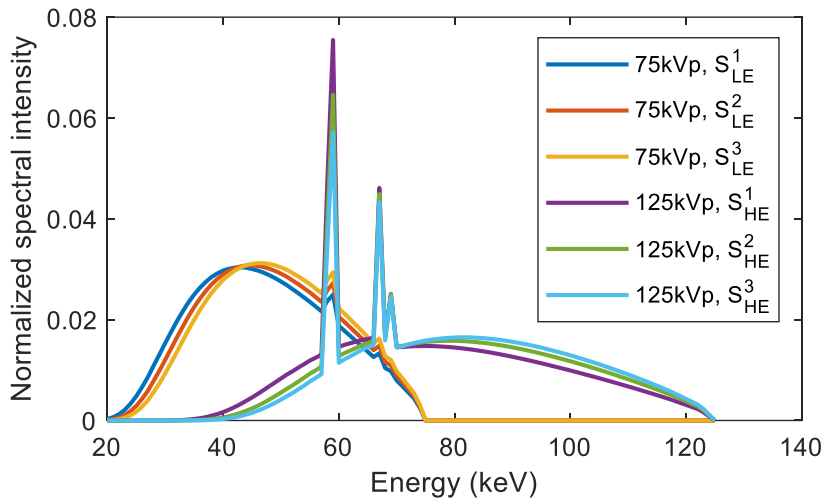


Fig. 5. Different LE and HE spectra.

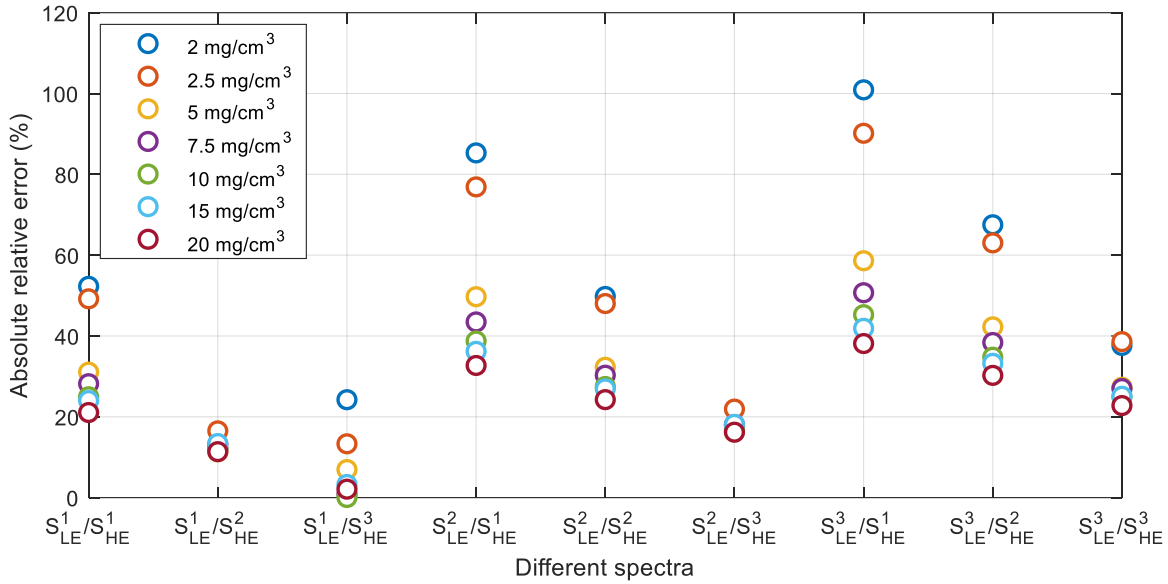


Fig. 6. The absolute relative errors of the quantified iodine concentrations using different LE and HE spectra.

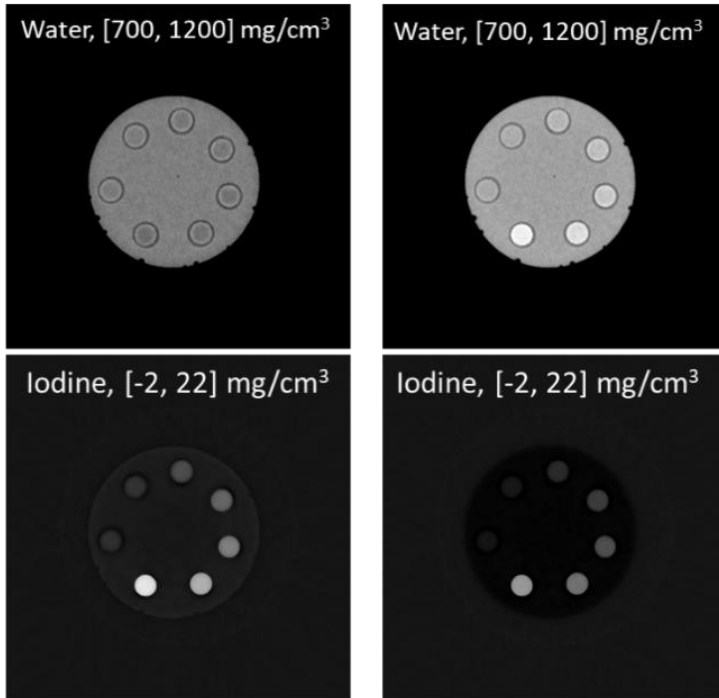


Fig. 7 The DIRECT-JSI algorithm-decomposed water and iodine images with the spectral combinations of (left) S_{LE}^1/S_{HE}^3 and (right) S_{LE}^3/S_{HE}^1 .

The LE and HE beam spectra used in the paper were simulated from SpekCalc with known beam filtrations. In particular, the 1.5 mm Al was estimated from the half-value-layer (HVL) measurements at 80 kVp, and the 0.2/1.2 mm Cu was added. As clarified in the manuscript, the generation of the network training data were also based on these assumed LE and HE beam spectra.

These simulated spectra were used in all the comparison algorithms/networks (except for the ID-EP method in which the LE and HE effective attenuation coefficients were measured directly on the FBP reconstructed CT image) when evaluating the experimental data. As presented in the paper, it was observed that the performance of DIRECT-JSI algorithm similar to the above results in Fig. 7. In contrast, the proposed DIRECT-Net method outperforms the others in generating precise basis images. In summary, these results led the authors to conclude that the “DIRECT-Net has promising benefits in improving the DECT image quality” in the published paper.

Reply: The low- and high-kVp FBP reconstructions observed in Figs. 6(left), 7(left), and 8(left) in the paper display little visible beam-hardening artifacts (possibly due to the small cross-section size and material composition of the phantoms and knuckle specimen used.) As such, accurate concentration estimations of calcium and iodine can be expected, respectively, using the standard IDD method from the FBP reconstructions (with a simple calibration if necessary), the well-established data-domain decomposition (DDD) methods, and the recent one-step algorithms. In other words, the study with insignificant beam-hardening effect demonstrates neither that the Direct-Net method proposed can accurately reconstruct images in full-scan DECT with a realistic level of beam-hardening effect nor that it performs better than do the relevant, existing methods for full-scan DECT. It is indeed unclear why concentration estimations with significant differences were obtained by use of the methods compared for this case only with a weak beam-hardening effect.

The Direct-Net method's performance evaluated under the condition of a realistic level of beam-hardening effect, clearly more prominent than that considered in the paper, would be necessary to support the paper's conclusion.

Response: Thank you. The comments touched two different aspects: 1) When beam hardening effects were relatively weak, you did not comment that the DIRECT-Net method outperformed other "well-established data-domain decomposition methods, and the recent one-step algorithms" and 2) Pan & Sidky commented whether DIRECT-NET can still outperform the mentioned existing methods under the condition of "a realistic level of beam-hardening effect." This is an interesting comment and a healthy question. The authors thank Pan & Sidky for raising it for us to further investigate the performance of the DIRECT-Net.

A group of completely new training and testing data were generated for the iodine-water insert phantom having much larger sizes (32 cm x 32 cm physical dimension with a 512 x 512 image matrix). The results of the monochromatic 50 keV image are shown in Fig. 8. As one can observe, the beam-hardening shading artifacts between iodine inserts in the LE CT image are significantly mitigated in the monochromatic image, indicating that the DIRECT-Net indeed has the potential to reduce strong beam-hardening artifacts. Results also show consistent performance of the DIRECT-Net with the one-step direct DECT decomposition algorithm despite the fact that some blooming artifacts appeared around the inserts and further fine tuning of the neural network parameters is needed. Nevertheless, the results in Fig. 8 indeed show a potential benefit of the DIRECT-Net in generating DECT basis images with mitigated beam hardening shading artifacts for large sized objects (32 cm).

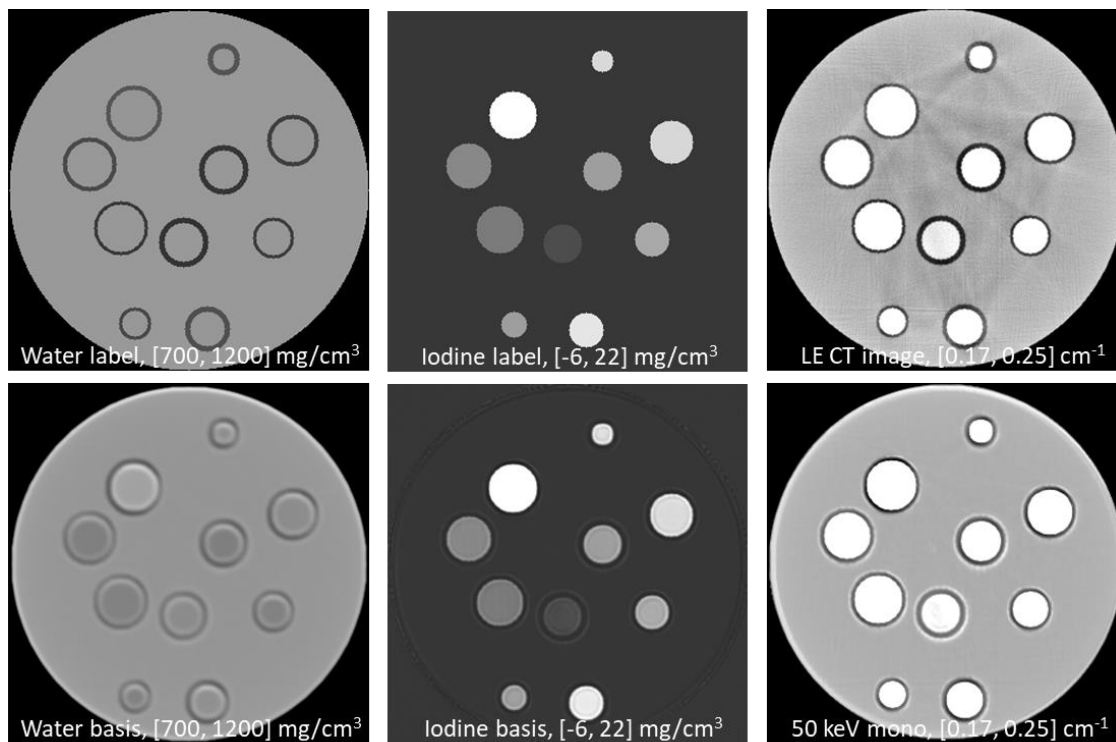


Fig. 8 The decomposition results of the DIRECT-Net on the data with increased beam-hardening effect for large sized objects (32 cm).

Reply: Numerous methods have been reported in the literature for quantitative image reconstruction and accurate concentration estimations in full-scan DECT. To support the paper's conclusion, the Direct-Net method logically needs

to be compared with the relevant, established methods, including the three types of the IDD methods, the DDD methods, and especially the recent one-step algorithms because both Direct-Net method and recent one-step algorithms reconstruct images directly from data in full-scan DECT. The statement in the paper that “Comparisons are performed with different DECT decomposition algorithms” is not supported by sufficient evidence because the Direct-Net method was not compared appropriately with the relevant, existing algorithms or methods in each of the three types.

The standard IDD method, the most widely used in clinical and other applications, directly inverts a 2X2 matrix of equation (1) in Ref. [7] 2 for obtaining basis images from low- and high-kVp FBP reconstructions. It subsequently estimates material concentrations from the basis images obtained. No result can be found in the paper comparing the Direct-Net and standard IDD methods.

The DDD methods are used widely in academic research on, and industrial products of, DECT to minimize effectively beam-hardening artifacts, thus yielding accurate concentration estimation. No result can be found in the paper comparing the Direct-Net and DDD methods.

The recent one-step algorithms have been demonstrated to yield accurate estimations of material concentrations in not only full-scan, but also partial-scan, DECT. While both the Direct-Net method and one-step algorithms reconstruct basis images directly from data in full-scan DECT, no results are presented in the paper comparing the Direct-Net method with the recent one-step algorithms, and appropriate citations of relevant, recent one-step algorithms appear missing in the paper.

Response: The authors agree with Pan & Sidky that “numerous methods have been reported in literature,” but the authors are not aware of any professional guidelines or consensus standard of performance comparison in the research community yet. Nevertheless, the authors are happy to take this opportunity to present readers with the requested performance results for both IDD and DDD methods since these methods are straightforward in implementations. Particularly, the authors admire the recently published one-step algorithms from Pan & Sidky group^{6,7,8,9}, for example.

However, the authors are afraid of not being able to implement these above algorithms and choosing the right parameters to generate fair results for comparison. If Pan & Sidky are able to share a copy of executable code of Pan & Sidky’s preferred numerical implementation of the one-step algorithm in either GitHub or any other public platform for the authors to generate the results, we would be happy to do so.

As requested, the standard IDD method (direct matrix inversion of CT images) was implemented, and the results are presented below in Fig. 9 and Table 1. The results showed similar mean values as those of the ID-EP method in Table 2 and 3 in the published paper, but with much larger standard deviations (std). Such observation is consistent with the results in Ref [7].

The DDD method¹⁰ was implemented and the decomposition results are shown in Fig. 10 and Table 1. One can observe that the quantified mean densities of CaCl_2 and iodine are similar to those from DIRECT-JSI, but the standard deviations are much larger.

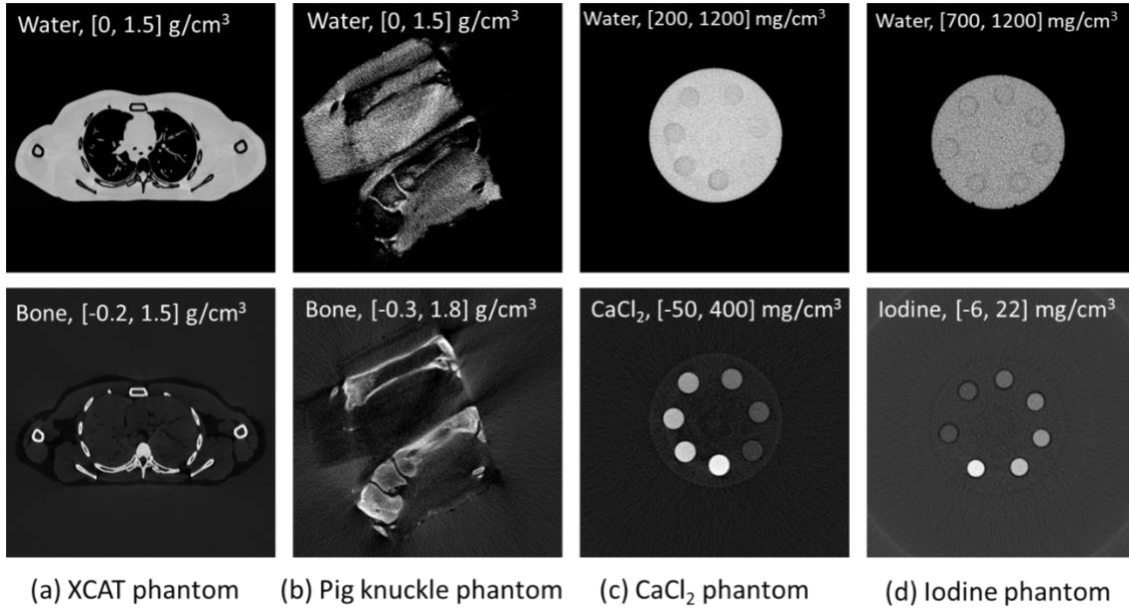


Fig. 9 The decomposition results from the standard IDD method for four different phantoms. Results in the first column are obtained from the numerical simulations, and the other results are obtained from the real experiments.

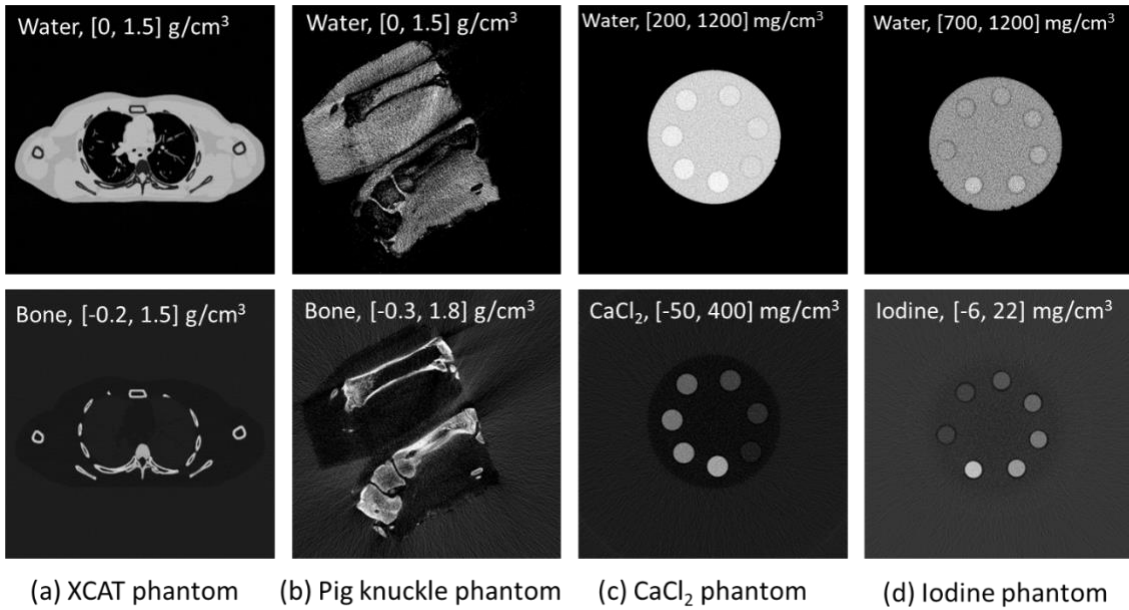


Fig. 10 The decomposition results from the DDD method for four different phantoms. Results in the first column are obtained from the numerical simulations, and the other results are obtained from the real experiments.

Table 1 The quantified iodine and CaCl₂ concentrations (mean±std) from the IDD and the DDD method.

CaCl ₂ concentrations (mg/cm ³)							
True	37.5	75	125	200	250	300	375

IDD	52.16±26.58	90.30±25.85	142.54±25.6 7	219.39±26.4 2	261.61±28.5 4	305.89±29.9 2	369.64±30.6 7
DDD	2.82±28.16	29.40±27.19	64.14±27.37 7	118.39±30.3 7	151.20±31.3 6	185.29±34.3 4	233.61±34.7 7
Iodine concentrations (mg/ cm ³)							
True	2	2.5	5	7.5	10	15	20
IDD	2.25±1.22	2.60±1.32	5.24±1.20	7.59±1.33	10.26±1.34	14.85±1.38	19.95±1.27
DDD	0.95±1.39	1.22±1.47	3.26±1.49	5.08±1.55	7.24±1.58	10.90±1.72	14.98±1.53

Reply: It is unclear that the performance rankings of the methods compared in the paper are conclusive. Because a method such as those considered in the paper involves parameters (please also see the Reply comment below,) its reconstructed image, concentration estimation, and any goodness metrics are thus functions in a multi-dimensional space formed by the parameters involved. The results presented in the paper were obtained only at a single point in each of the respective multi-dimensional parameter spaces of the methods considered.

Therefore, the conclusion about the relative rankings of their performance obtained from these single-parameter-point results is unlikely to be conclusive. This is because, when different points are chosen in these parameter spaces, the results and performance rankings of the methods are likely to vary, possibly leading to observations and conclusion contrary to those made in the paper.

Response: The authors totally agree with Pan & Sidky that the parameters will have to be well-selected for a certain algorithm to generate optimal results, and this was what the authors wrote in the paper, except that the authors did not study the performance dependence on X-ray spectra. The authors liked Pan & Sidky's comment, and the authors will continue to investigate the performance of the DIRECT-Net in the future and would like to share these results with the entire scientific community when they are available. As Pan & Sidky emphasized throughout the comments, without clear scientific evidence, it would be premature to jump to the conclusion that certain parameters would "[lead] to observations and conclusion[s] contrary to those made in the paper."

Reply: In DECT or spectral CT, the relationship between basis functions and model data are characterized often by imaging model equation (1) in the paper. In terms of basic research on reconstruction algorithms, quantitative image reconstruction is referred to often as accurate inversion of imaging model equation (1). Numerical evidence has been reported in the literature that the recent one-step algorithms quantitatively reconstruct images by accurately inverting equation (1) directly from model data in full-scan and partial-scan DECT. No evidence can be found in the paper showing the "benefits" of the Direct-Net method for "improving" over the recent one-step algorithms in accurate image reconstruction in terms of accurate inversion of equation (1) from full-scan or partial-scan data in DECT.

Response: The authors congratulate Pan & Sidky on their great accomplishment of having successfully solved the DECT inverse problem defined by the imaging model in Eq. (1). However, the authors did not claim that we solved the DECT inverse problem in the paper. Rather, the authors only presented a new DL-based approach to generate the DECT basis images which shows improved performance within the defined conditions.

Reply: It was stated in the paper that the one-step algorithms "may have strong spectra dependence and parameter

selection issues, and they may also require long time.” The statement is simply spurious, as discussed below: The development of recent one-step algorithms for DECT was motivated by the design and enabling of novel DECT scan configurations. The research presented in the paper is for image reconstruction in DECT with a standard, full-scan configuration that is well-studied, and its relevance and significance thus diminish in terms of researching on novel algorithms for image reconstruction in full-scan DECT.**

Response: Thank you. The authors should have elaborated a bit more in the paper instead of this single sentence comment about the one-step algorithms. We would like to elaborate a bit further here to avoid further misunderstanding.

When the published paper says that the one-step algorithms “may have strong spectral dependence,” the authors mean that we did not find a systematic study on how the one-step algorithms perform at different spectral conditions. We might have overlooked some works from the Pan & Sidky group in addressing this problem, so if Pan & Sidky can point the authors to the right references, we would greatly appreciate it.

When we say that the one-step algorithms may have “parameter selection issues,” this is only a general comment for all iterative reconstruction algorithms and is NOT specific to just Pan & Sidky algorithms. We did not intend to attack you or Pan & Sidky work in this comment. Retrospectively, we should have acknowledged that the DL method including the DIRECT-Net method also requires fine tuning of parameters for optimal performance.

When we say that “they may also require long time,” this is also a general comment and well-known fact that iterative image reconstruction algorithms have clinically unacceptably long reconstruction times due to the need to iteratively reconstruct images for each clinical case. In contrast, DL methods leave the computational overhead offline and thus people generally believe that it has some edge in this regard. Additionally, the combination of DL method and iterative reconstruction method is also an active research area that holds the promises to further shorten the reconstruction time and optimize the selections of reconstruction parameters.

Reply: It is unclear what “strong spectra dependence” precisely is referred to. If the paper suggests that the robustness of the recent one-step algorithms is highly sensitive to the choice of spectra, no evidence can be identified in the paper to support the suggestion. In contrary, works in the literature reveal that the recent one-step algorithms are not more sensitive to the choice of spectra than any other methods in terms of accurate image reconstruction in full-scan (and partial scan) DECT.

Response: As the authors elaborated in the opening response to Pan & Sidky’s above comment, the authors did not state that Pan & Sidky recent one-step algorithm is “highly sensitive to the choice of spectra”. We simply said it “may have strong dependence”. Be aware that this statement is just a very general comment and is not to any specific algorithm (including the one-step algorithms developed by Pan & Sidky).

Reply: Any methods, including the Direct-Net method proposed and non-DL methods, will involve parameters and their value selections. For example, in the Direct-Net method proposed,¹ the design of the subnetworks and the DT module involves a slew of parameters listed in Table 1 and the specific forms indicated in equations (3) and (4); equally importantly, the design and implementation of the loss function, training schemes, and training/target image curation also involve multiple parameters and choices; there are additional parameter selections as discussed in the paragraph below equation (7); and even the Direct-Net’s architecture design itself constitutes a parameter of choice, as the “room to improve the DIRECT-Net” architecture was discussed in Sec. 5 of the paper. In fact, given data, image reconstruction of any methods involves parameters, including the choices of image representations such as voxels and

blobs and their sizes, imaging model forms, tube spectra and detector response, objective functions, algorithms, and network parameters. Therefore, like any reconstruction methods, the Direct-Net method proposed involves parameters and their value selections.

Response: As elaborated in the opening response to Pan & Sidky comment above, the authors admit that the DIRECT-Net method also encounters the challenges of parameter selection in the training stage. Again, this is not specific to Pan & Sidky recent one-step algorithms.

Reply: The recent one-step algorithms seek to accurately invert imaging model equation (1), i.e., accurate image reconstruction, for various scanning configurations of potential interest in DECT and spectral CT. Therefore, the iteration number (i.e., computation time) is not a leading factor of research concern in terms of accurately inverting the imaging model. The recent one-step algorithms were compared with the standard IDD and DDD methods in studies with simulated noisy data and real data collected in experiments; and the comparison results demonstrate that the performance of the recent one-step algorithms is not inferior to that of the well-established standard IDD and DDD methods in full-scan DECT. Additionally, for partial-scan DECT considered, the recent one-step algorithms have also been demonstrated to invert accurately equation (1), yielding image reconstruction and concentration estimation comparable to those obtained from data in full-scan DECT.***

Response: Once again, the authors congratulate Pan & Sidky on the success in the development of one-step algorithms for various innovative “scanning configurations of potential interest in DECT and spectral CT.” The authors would like to mention that the focus of the study is the DECT method that could be potentially applied to the currently available DECT systems in clinical practices. For example, the DECT imaging system with dual-layer detector configurations.

In summary, the authors thank you for this precious opportunity to discuss Pan & Sidky’s work and our work. This open dialogue also gives the authors an opportunity to share with the medical physics community some new results that Pan & Sidky requested. We hope that this open discussion would help us for more rigorous studies on the performance of different DECT image generation methods in future.

Footnotes:

* It is stated in the paper that “the degraded decomposition performance of DIRECT-JSI method is mainly caused by the inconsistencies between the assumed Poisson noise model in the algorithm and the real Gaussian-like noise distribution in the experiment.” The statement is unlikely to be correct, as it can be demonstrated in a straightforward study described below: Using noiseless data generated from the numerical phantoms with equation (1) in the paper for full-scan DECT, one can reconstruct images with the DIRECT-JSI and Direct-Net methods. Since there is no data noise in this deterministic case, the data-noise-statistics maximum-likelihood argument becomes irrelevant. If there remains a “degraded decomposition performance of DIRECT-JSI method” in the study, it cannot be attributed to the choice of noise models anymore; and it might be instead the result of other issues such as incorrect implementation of the methods. This study can be used indeed specifically for verifying the implementation correctness of the methods because, in this noiseless, ideal case, they should yield, up to the computer precision, numerically identical, accurate reconstructions.

** The problem of image reconstruction from full-scan data in DECT is well-studied, just like the problem of image

reconstruction from the full-scan Radon transform (RT) that is well-studied. Basic research on additional algorithms only for inverting the full-scan RT, i.e., only for image reconstruction from knowledge of the full-scan RT, seems of little relevance and significance.

*** The paper cites that “the ID-EP method spent about 625 s for 25000 iterations” and that “the Direct-JSI method took about 627 s for 40 iterations.”

First, the standard IDD method, which was not compared against, is by far the most widely used in clinical and other applications of DECT. The standard IDD method obtains basis images by directly inverting the 2X2 matrix of equation (1) in Ref. [7] ² from the low- and high-kVp FBP reconstructions, certainly without any iterations. Indeed, the results in Ref. [7] ² appear to show that the standard IDD and ID-EP methods are of a comparable level of performance in terms of image visualization and concentration estimation. Therefore, the statement “the ID-EP method spent about 625 s for 25000 iterations” can be misleading, as it may feed a false impression that any capable IDD methods would require intensive computation for basis image decomposition.

Second, the DIRECT-JSI method seeks to reconstruct images by solving the optimization problem, i.e., equation (8) in Ref. [14] ³. As discussed above, when its convergence is of concern in terms of solving an optimization problem, the number of iterations (and thus computation time) is not a leading factor of relevance.

Computation time of image reconstruction is clearly an extremely important consideration in a real-world application or product. However, it is a misconception to expect that theoretical algorithms reported in research articles in the literature can be applied directly to addressing optimally a real-world, practical reconstruction application, because the algorithms are not developed, optimized, and evaluated under conditions unique to the workflow and task-specific utility metrics of the real-world application. In fact, it is unlikely that the specific conditions and requirements in an application or product, and even the application (or product) itself, are known at the time an algorithm is investigated and developed as a research work published in the literature.

It is a reconstruction procedure, along with its specifically selected parameters, that is adopted in a practical application or product. The reconstruction procedure necessarily is tailored to, and optimized for, the unique workflow settings (such as fast reconstruction time,) subject and data conditions, and task-specific metric requirements in the application or product, even though the reconstruction procedure itself generally does not solve the image model that an algorithm is developed to solve in a research work published in the literature. It may be the case that the design of only some portions of a reconstruction procedure of practical utility is motivated by an algorithm published in the literature. To simply put, an algorithm published as a research work in the literature is generally unlikely to be equivalent to a reconstruction procedure optimized for DECT practical application or product. Evidence is not provided that the proposed Direct-Net method has been optimized and evaluated in any practical application or product.

References:

1. Su T, Sun X, Yang J, Mi D, Zhang Y, Wu H, Fang S, Chen Y, Zheng H, Liang D, Ge Y. DIRECT-Net: A unified mutual-domain material decomposition network for quantitative dual-energy CT imaging. *Med Phys.* 2022;49(2):917-934
2. Niu T, Dong X, Petrongolo M, Zhu L. Iterative image-domain decomposition for dual-energy CT. *Med Phys.* 2014;41(4):041901.

3. Mechlem K, Ehn S, SELLERER T, Braig E, Munzel D, Pfeiffer F, Noel PB. Joint statistical iterative material image reconstruction for spectral computed tomography using a semi-empirical forward model. *IEEE Trans Med Imaging*. 2018;37:68-80.
4. Zhang W, Zhang H, Wang L, Wang X, Hu X, Cai A, Li L, Niu T, Yan B. Image domain dual material decomposition for dual-energy CT using butterfly network. *Med Phys*. 2019;46(5):2037-2051.
5. Clark DP, Holbrook M, Badea CT. Multi-energy CT decomposition using convolutional neural networks. *Medical Imaging 2018: Physics of Medical Imaging*. International Society for Optics and Photonics. 2018;105731:1057310.
6. Chen, B., Zhang, Z., Sidky, E.Y., Xia, D. and Pan, X., 2017. Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT. *Physics in Medicine & Biology*, 62(22), p.8763.
7. Chen, B., Zhang, Z., Xia, D., Sidky, E.Y. and Pan, X., 2018. Algorithm-enabled partial-angular- scan configurations for dual-energy CT. *Medical Physics*, 45(5), pp.1857-1870.
8. Chen, B., Zhang, Z., Xia, D., Sidky, E.Y. and Pan, X., 2021. Non-convex primal-dual algorithm for image reconstruction in spectral CT. *Computerized Medical Imaging and Graphics*, 87, p.101821.
9. Zhang, Z., Han, X., Pearson, E., Pelizzari, C., Sidky, E.Y. and Pan, X., 2016. Artifact reduction in short-scan CBCT by use of optimization-based reconstruction. *Physics in Medicine & Biology*, 61(9), p.3387.
10. Schlomka, et al. "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography." *Physics in Medicine & Biology* 53.15 (2008): 4031.