



Comparison of three undersampling approaches in computed tomography reconstruction

Chenyang Shen¹, Yifei Lou², Liyuan Chen¹, Tieyong Zeng³, Michael K. Ng⁴, Lei Zhu⁵, Xun Jia¹

¹Innovative Technology of Radiotherapy Computations and Hardware (iTORCH) Laboratory, Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX, USA; ²Department of Mathematical Sciences, University of Texas Dallas, Richardson, TX, USA; ³Department of Mathematics, Chinese University of Hong Kong, Hong Kong, China; ⁴Department of Mathematics, Hong Kong Baptist University, Hong Kong, China; ⁵Department of Modern Physics, School of Physical Sciences, University of Science and Technology of China, Hefei 230026, China

Correspondence to: Xun Jia; Chenyang Shen. Department of Radiation Oncology, University of Texas Southwestern Medical Center, 2280 Inwood Rd., MC9303, Dallas, TX 75235, USA. Email: xun.jia@utsouthwestern.edu; chenyang.shen@utsouthwestern.edu.

Background: Projection data undersampling is an effective approach to reduce X-ray radiation dose in computed tomography (CT). In modern CT technologies, undersampling is also a favorable method to reduce projection data size to facilitate rapid CT scan and imaging. It is an intriguing question that given an undersampling ratio, what is the optimal undersampling approach that enables the best CT image reconstruction. While this is in general a challenging mathematical question, it is the motivation of this paper to compare three types of undersampling operations, which we hope to shed some light to this question.

Methods: We considered regular view undersampling that acquires X-ray projections at equiangular projection angles, regular ray undersampling that acquires projections at all angles but with X-ray lines blocked within each projection under a periodic pattern, and random ray undersampling that acquires each X-ray line with a certain probability. By representing the undersampling projection operators under the basis of singular vectors of full projection operator, we generated matrix representations of these undersampling operators and numerically perform singular value decomposition (SVD). Singular value spectra and singular vectors were compared.

Results: For a given undersampling ratio, the random ray undersampling approach preserves the properties of the full projection operator better than the other two approaches. This translates to advantages of reconstructing a CT image at a lower error, which has also been demonstrated in the numerical experiments.

Conclusions: We compared three undersampling strategies and found that random undersampling preserves the most information and outperforms the other two in terms of reconstruction quality.

Keywords: Computed tomography (CT) reconstruction; dose reduction; iterative reconstruction; singular value decomposition (SVD); undersampling

Submitted Apr 25, 2019. Accepted for publication Jun 30, 2019.

doi: 10.21037/qims.2019.07.07

View this article at: <http://dx.doi.org/10.21037/qims.2019.07.07>

Introduction

Computed tomography (CT) or cone beam CT (CBCT) are volumetric image tools that have been broadly used for diagnosing diseases or guiding treatments. X-ray radiation exposure associated with them is a clinical concern, especially when they are used in a large population

of healthy patients, e.g., in lung cancer screening, or repeatedly used on a patient, e.g., during a long course of image-guided radiotherapy treatment. Over the years, several dose reduction approaches such as tube-current modulation (1-4) and data undersampling (5-10) have been developed. In the context of data undersampling,

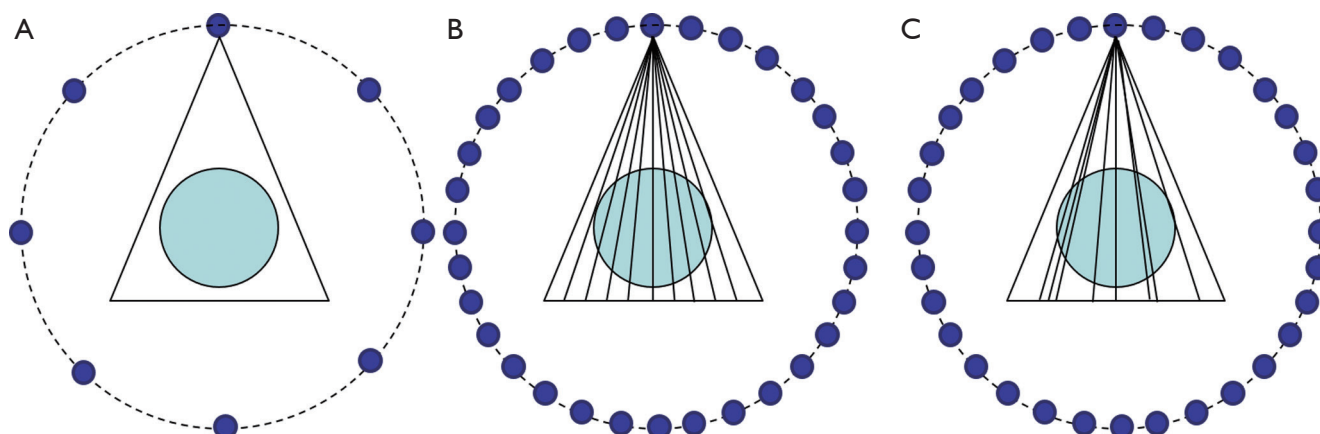


Figure 1 Illustration of three undersampling approaches. Blue circles indicate projection angles at which data are acquired.

conventional analytical reconstruction approaches are not fully capable of addressing the missing data issue, yielding artifacts in reconstructed images. In contrast, iterative reconstruction techniques motivated by the recent booming research of compressive sensing (CS) (11,12) are more robust to undersampling and are able to retrieve images with satisfactory quality from undersampled data. A number of regularization methods in CS have been introduced to the CT reconstruction problem, such as total variation (TV) (5-9), tight frame (10), and nonlocal means (13-16), etc.

Another motivation for undersampling is to facilitate rapid data acquisition and imaging. Along with recent advancements in CT technology, for instance energy resolved photon-counting detector (17-19), data size in a CT projection becomes larger and larger. This leads to the challenge to quickly read out acquired data before moving to the next projection measurement. Addressing this issue is especially important for those clinical applications requiring a high temporal resolution, e.g., cardiac imaging. While advanced electronics have been designed to improve readout speed, it is also highly desirable to explore novel directions to reduce projection data size without sacrificing quality of reconstructed CT images.

Within the framework of projection data undersampling, it is a natural question regarding the most effective way of sampling. Let us first consider data acquisition without undersampling. This can be expressed in a simple linear system $Pf = g$, where P is the projection operator that maps an image f to a number of densely placed projection angles. The undersampling operation corresponds to an operator S that has zero or one in its diagonal elements, yielding a data acquisition system $P_s f = SPf = Sg = g_s$. Here

P_s denotes the undersampling projection operator and g_s the corresponding measurements, a subset of the full data vector g . The mathematical properties of P_s depend on the specific way of sampling, namely the operator S . This in turn governs the difficulty of solving the linear system $P_s f = g_s$, namely CT reconstruction. Hence, it is interesting to ask which operator S permits the best image restoration. Generally speaking, this is a difficult question. In this paper, it is our motivation to compare three types of undersampling operations, which we hope to shed some light to this question of interest.

There are two typical undersampling approaches that have been widely discussed in the literature. The first one is view undersampling, as illustrated in Figure 1A. Since CT data acquisition is typically performed in a view-by-view fashion, this view undersampling is straightforward in practice: simply tuning on X-ray exposure in those views of interest. Probably due to this reason, view-undersampling approach is the most widely studied undersampling method in low-dose CT problems (5-10). The second approach is regular ray undersampling, as shown in Figure 1B. While X-ray projections are acquired at all the densely placed angles, the X-ray source is blocked by a blocker that has a periodic pattern. This blocking method was originally proposed to address X-ray scatter contamination problem in CBCT (20,21). The signals measured in the purposely-created blocked areas are assumed to be scatter. By smoothly interpolating them to the open areas, scatter signal can be accurately estimated and then corrected. As missing primary data is an apparent problem in this setup, iterative reconstruction approach was employed to improve reconstructed image quality (22).

There are infinitely many ways of undersampling. While it is a difficult problem to theoretically find the best one, we will perform a comparison study of three different undersampling approaches in this paper. In addition to the aforementioned two methods, the third one that will be investigated is random ray undersampling, i.e., each X-ray line is randomly selected with certain probability, as illustrated in *Figure 1C*. This is motivated by the advancements in CS, where randomness introduced in a data acquisition process typically helps signal restoration. In the rest of this paper, we will compare the properties of the projection matrices in these three undersampling approaches.

Methods

Singular value decomposition (SVD) of a full projection operator

To present theoretical insights of the CT projection operator, we consider CT projection under parallel projection geometry (*Figure 2*). The reason for this choice is due to the existence of a closed form SVD of the full projection operator under this geometry. Current CT imaging systems employ fan-beam geometry, as illustrated in *Figure 1*. While this is different from the parallel-beam geometry here captures some insights in the fan-beam setup. The validity of our analysis is demonstrated to some extent by the agreement with numerical experiments performed under the fan-beam geometry.

Consider a 2D function $f(x)$ defined in a unit circle $\{x \in \mathbb{R}^2: |x| < 1\}$, a parallel X-ray projection operation to all the directions $\theta \in [0, 2\pi]$ maps this function into $\{s, \theta\} \in [-1, 1] \times [0, 2\pi)$ through a linear system $Pf = g$, where P is the full projection operator respect to variables s, θ and x . We use the term “full” to denote the situation in which measurements at the entire projection domain are made. When it comes to undersampling, let us consider a general undersampled projection operator $P_s = SP$, where $S = (s, \theta)$ is the sampling operator. The function value of S at (s, θ) is either 1 or 0 depending on whether the data at the coordinate (s, θ) is sampled or not.

In this paper, one objective is to study SVD forms of different undersampling operators. The motivation to pursue this SVD form comes from its importance in reconstruction. Consider a typical reconstruction problem via the least square approach, i.e., solving $f = \arg \min \|Pf - g\|^2$, where P is

a general projection operator, either full or undersampled. If P has a SVD form $[v_i(x), u_i(s, \theta); \sigma_i]$, the solution to the least square problem is simply $f = \sum_i v_i \frac{\langle g, u_i \rangle}{\sigma_i}$ where $\langle \cdot, \cdot \rangle$ denotes a dot product in the Euclidean space. This naturally implies a reconstruction strategy: (I) decomposing the measured projection data g to the basis functions in the projection domain by computing the coefficients $\langle g, u_i \rangle$; (II) dividing the coefficients by the corresponding singular values σ_i ; and (III) superposing the basic functions $v_i(x)$ in the image domain with the modified coefficients. Note that the least square term is a typical choice to ensure data fidelity in many iterative reconstruction approaches.

One particular property of the full projection operator P is that it has an SVD representation $[v_{m,l}(x), u_{m,l}(s, \theta); \sigma_{m,l}]$ (23) with

$$v_{m,l}(x) = \pi^{-\frac{1}{2}} (m+1)^{\frac{1}{2}} |x|^l P_{m-l}^{(0,l)} \left(2|x|^2 - 1 \right) e^{il \arg x} \quad [1]$$

$$u_{m,l}(s, \theta) = \pi^{-1} w(s) U_m(s) e^{-il\theta} \quad [2]$$

$$\sigma_{m,l} = 2\pi^{\frac{1}{2}} (m+1)^{-\frac{1}{2}} \quad [3]$$

$$Pv_{m,l} = \sigma_{m,l} u_{m,l} \quad [4]$$

where $\arg x$ gives angle of the specific points within the unit circle ($\{x \in \mathbb{R}^2: |x| < 1\}$), $w(s) = (1-s^2)^{\frac{1}{2}}$, $P_n^{(a,b)}(x)$ is the Jacobi polynomial, and $U_m(s)$ is Chebyshev polynomial of second kind with $m > 0$, $-m \leq l \leq m$, and $m+l$ being even. l is the index for the projection angle direction and m is the index within a projection. Note that $v_{m,l}(x)$ indicates spherical harmonics which enforces the orthogonality of v as the singular vector in image domain. This decomposition allows reconstructing the image in the full projection case via the aforementioned strategy (24).

With this basis, we can express the full projection operator P in its matrix representation with matrix elements $\langle u_{m,l}, Pv_{n,k} \rangle$, $m, n > 0$, $|l| \leq m$, and $|k| \leq n$. Denote the matrix as R . Its matrix element is (23)

$$\begin{aligned} R_{m,l,n,k} &= \langle u_{m,l}, Pv_{n,k} \rangle \\ &= \frac{2}{\pi^{3/2} \sqrt{n+1}} \int_0^{2\pi} d\theta e^{i(l-k)\theta} \int_{-1}^1 ds w(s) U_m(s) U_n(s) \\ &= \frac{2}{\pi^{3/2} \sqrt{n+1}} 2\pi \delta_{l,k} \frac{\pi}{2} \delta_{m,n} = \frac{2\pi^{1/2}}{\sqrt{n+1}} \delta_{l,k} \delta_{m,n} \\ &= \sigma_{m,l} \delta_{l,k} \delta_{m,n} \end{aligned} \quad [5]$$

δ is the Kronecker's delta function. The property of Chebyshev polynomials $\int_{-1}^1 ds w(s) U_m(s) U_n(s) = \frac{\pi}{2} \delta_{m,n}$ is used here. This implies that R is diagonal (according to its indices given by order of Jacobi polynomials and Chebyshev polynomials) with diagonal elements of $\sigma_{m,l}$. The matrix R has a block structure in the sense that l and k are block indices corresponding to the projection angle direction and m and n are indices within each block associated with the direction within a projection angle.

The upper limits of these indices are determined by the highest sampling frequency in the corresponding directions. In the continuous setting, the upper limits are infinity. When it comes to a realistic scenario with the number of projections M in the full sampling case, $|l|, |k| \leq M_{ang} = M/2$. Similarly, the finite detector size limits the sampling frequency within a projection. This sets an upper limit of $m \leq M_{det}$ with M_{det} being set to the number of detector elements per projection. Supplementary A presents details on the construction of the matrix R .

SVD of an undersampled projection operator

When it comes to an undersampling problem, presence of the binary undersampling operator S changes the SVD form. For a projection operator P_s , we can write its representation R_s under the basis of $v_{m,l}(x)$ and $u_{m,l}(s, \theta)$ corresponding to the full projection operator P in the continuous setting. This is analogous to define $R_s = U^T P_s V$ in the matrix case. This is feasible, since $\{v_{m,l}(x)\}$ and $\{u_{m,l}(x, \theta)\}$ form complete orthogonal basis sets in the image and the projection domains, respectively. Each element of R_s can be expressed as

$$R_{sm,l,n,k} = \langle u_{m,l}, P_s v_{n,k} \rangle$$

$$= \frac{2}{\pi^{3/2} \sqrt{n+1}} \int_0^{2\pi} d\theta e^{i(l-k)\theta} \int_{-1}^1 ds w(s) S(s, \theta) U_m(s) U_n(s) \quad [6]$$

R_s is no longer diagonal. In the following sections of this paper, we focus on the analysis of three commonly used undersampling techniques: regular view undersampling, regular ray undersampling, and random ray undersampling.

Regular view undersampling

In the case of regular view undersampling, $S(s, \theta)$ is unity for $\theta_j = \frac{2\pi j}{N}$, $j = 1, 2, \dots, N$. The number of projections is

$N = rM$, where $r < 1$ is the ratio of undersampling. Hence, we have

$$R_{sm,l,n,k} = \frac{2}{\pi^{3/2} \sqrt{n+1}} \frac{2\pi}{M} \sum_{j=1}^N e^{i(l-k)\theta_j} \int_{-1}^1 ds w(s) U_m(s) U_n(s) \quad [7]$$

$$= \frac{2}{\pi^{3/2} \sqrt{n+1}} \frac{2\pi}{M} \sum_{j=1}^N e^{i(l-k)\theta_j} \frac{\pi}{2} \delta_{m,n}$$

This implies that the matrix elements are nonzero only when the block indices satisfy $l-k = tN$, $t \in \mathbb{Z}$, such that

$$\frac{2\pi}{M} \sum_{j=1}^N e^{i(l-k)\theta_j} = \frac{2\pi}{M} \sum_{j=1}^N e^{itN \frac{2\pi j}{N}} = \frac{2\pi}{M} N = 2\pi r \quad [8]$$

Since $-m \leq l \leq m$ and $m+l$ even, we can change the notation $m \rightarrow 2m+l$ and $n \rightarrow 2n+k$. Thus, in the matrix block satisfying $l-k = tN$, the only nonzero elements should satisfy $2m+l = 2n+k$, i.e., $m-n = -tN/2$, and the matrix element can be obtained by

$$R_{sm,l,n,k} = \frac{2}{\pi^{3/2} \sqrt{2n+k+1}} 2\pi r \frac{\pi}{2} = \frac{2\pi^{1/2} r}{\sqrt{2n+k+1}} \quad [9]$$

Apparently, R_s is not a diagonal matrix anymore. Compared to R , the matrix R_s is different due to the existence of off-diagonal blocks at locations satisfying $l-k = tN$, $t \in \mathbb{Z}$. These off-diagonal blocks appear at locations that are at distances tN from the diagonal blocks. An immediate conclusion is that, the smaller N is, the more off-diagonal elements there are, and hence larger difference between R_s and R . This is expected, as a more aggressive undersampling approach changes the system more significantly. On the other hand, when $N \rightarrow \infty$, the effect of off-diagonal blocks can be ignored, as those blocks are far away from the diagonal. This corresponds to $r \rightarrow 1$. Comparing Eq. [9] and Eq. [6], we observe that $R_s \rightarrow R$ and hence $P_s \rightarrow P$ as $r \rightarrow 1$.

Regular ray undersampling

In the case of regular ray undersampling, projections at all angles are acquired, but a periodic undersampling pattern is used within each projection angle. Therefore, $S(s, \theta)$ is only a function of s . We consider an example case of $S(s) = 1$, if $s \in [na, na+b]$, $n \in \mathbb{Z}$, $b < a$. Here a can be regarded as the period of the regular undersampling pattern and b is the width of sampled data within each period. The undersampling ratio $r = b/a$. We express the matrix of R_s as follows,

$$\begin{aligned}
 R_{sm,l,n,k} &= \frac{2}{\pi^{3/2}\sqrt{n+1}} \int_0^{2\pi} d\theta e^{i(l-k)\theta} \\
 &\int_{-1}^1 dsw(s) S(s) U_m(s) U_n(s) \quad [10] \\
 &= \frac{4\delta_{l,k}}{\pi^{1/2}\sqrt{n+1}} \int_{-1}^1 dsw(s) S(s) U_m(s) U_n(s)
 \end{aligned}$$

This implies that the matrix is of a block diagonal form. For the diagonal block, with a change of notation $m \rightarrow 2m+l$ and $n \rightarrow 2n+k$,

$$\begin{aligned}
 R_{sm,l,n,k} &= \frac{4}{\pi^{1/2}\sqrt{2n+k+1}} \int_{-1}^1 dsw(s) \\
 &S(s) U_{2m+l}(s) U_{2n+k}(s) \quad [11]
 \end{aligned}$$

Compared to the previous view undersampling case that generates a lot of non-zero elements at the off-diagonal blocks of the matrix, regular view undersampling approach only generates off-diagonal elements in the diagonal blocks with $l = k$.

Random ray undersampling

Random ray under-sampling is a strategy that each X-ray line is sampled with probability r . The resulting undersampling pattern gives a sampling function $S(s, \theta)$. Again, let us start from the general expression in Eq. [5] but with an exchanged order of the two integrals,

$$\begin{aligned}
 R_{sm,l,n,k} &= \frac{2}{\pi^{3/2}\sqrt{n+1}} \int_{-1}^1 dsw(s) U_m(s) U_n(s) \\
 &\int_0^{2\pi} d\theta S(s, \theta) e^{i(l-k)\theta} \quad [12]
 \end{aligned}$$

For each fixed s , $S(s, \theta)$ is a random variable with binary values depending on θ . For the diagonal blocks $l = k$,

$$R_{sm,l,n,l} = \frac{2}{\pi^{3/2}\sqrt{n+1}} \int_{-1}^1 dsw(s) U_m(s) U_n(s) f(s) \quad [13]$$

where $f(s) \equiv \int_0^{2\pi} d\theta S(s, \theta) \rightarrow \frac{2\pi}{M} \sum_{j=1}^M S_j$ is a random valued function with $S_j = S(s, \theta_j)$. The summation follows binomial distribution with mean rM and variance $Mr(1-r)$. It hence follows that $f(s)$ has a mean value of $2\pi r$ and variance of $\frac{4\pi^2 r(1-r)}{M}$. For simplicity, we approximate $f(s)$ with a Gaussian random number following the corresponding mean and variance in our numerical evaluation later. The integral over s has to be evaluated numerically.

For the off-diagonal blocks,

$$R_{sm,l,n,k} = \frac{2}{\pi^{3/2}\sqrt{n+1}} \int_{-1}^1 dsw(s) U_m(s) U_n(s) g(s) \quad [14]$$

where $g(s) \equiv \frac{2\pi}{M} \sum_j S_j e^{i(l-k)\theta_j}$ is a complex random variable. Its real and imaginary parts are random variables, both with mean zero and variance $\frac{2\pi^2 r(1-r)}{M}$ (Supplementary B).

Again, we generate $g(s)$ with its real and imaginary parts sampled according to Gaussian distribution. Then we numerically compute the matrix elements where numerical integration over s is needed. Note that $E[f(s)g(s)] = E[f(s)]E[g(s)]$ (Supplementary C), which implies that $f(s)$ and $g(s)$ are uncorrelated random variables. Hence, we do not consider this correlation when generating off-diagonal block matrix elements.

The above deviation shows that the matrix R_s is dense, although the off-diagonal elements are small, random variables of $O\left(\frac{1}{M}\right)$, compared to the diagonal elements that are of order 1. The impact of randomness is reflected by the existence of these small off-diagonal matrix elements. There are two special cases to consider: (I) In the limit $M \rightarrow \infty$, all the off-diagonal elements $R_{sm,l,n,k} \xrightarrow{M \rightarrow \infty} 0$ and the diagonal elements are $R_{sm,l,m,l} \xrightarrow{M \rightarrow \infty} \frac{2\pi^{1/2}r}{\sqrt{n+1}}$. Comparing with Eq. [5], it follows $R_s \xrightarrow{M \rightarrow \infty} rR$. For a real case with a finite but large number of projections, $R_s \sim rR$; (II) in the case of $r \rightarrow 1$, the variance of all the matrix elements approaches 0, and $R_s \rightarrow R$ as expected.

Numerical SVD

To perform numerical SVD, we consider the discrete case with the projection operator represented using matrix. For each undersampling case, we generate the corresponding matrix representation R_s . All the matrices have the same structure, as shown in Supplementary A. We set $M = 2M_{ang} = 200$ and $M_{det} = 250$ in our numerical experiments. These parameters are smaller than the actual number of projections and detectors per projection in a real case. However, further increase these values will lead to matrices that are too large to handle. We think the parameter values are large enough to capture properties of a real case. For each matrix, we perform numerical SVD to find matrices W and X , such that $R_s = X\Lambda_s W^T$, where Λ_s is a diagonal matrix

with diagonal entries λ_{si} being the singular values of R_s . We would like to remark on the following issues. In the case of regular ray undersampling, the matrix R_s is block diagonalized. Hence, we can perform SVD on each block individually, which is computationally less challenging compared to other cases.

After performing SVD, we first compare singular value spectra in different cases. Particularly, we examine $\mu = \frac{|R_s|_*}{|R|_*}$,

where $|\cdot|_*$ denotes the matrix nuclear norm. At the continuum limit, $|P|_* = |R|_*$ and $|P_s|_* = |R_s|_*$. Hence,

$$|R_s|_* = |P_s|_* = |SP|_* \leq |S|_{\max} |P|_* = |P|_* = |R|_* \quad [15]$$

where $|X|_{\max} = \max_i \sigma_i$ gives the largest singular value of X . This implies that $\mu \leq 1$. For a real discrete case, there is no guarantee that $\mu \leq 1$. But we expect this still holds in the regime of large M_{ang} and M_{det} . We use μ to represent perturbation of the undersampling operation on the full projection case. The smaller μ is, the more dramatically the matrix R_s is altered from R and hence so is P_s compared to P . From reconstruction point of view, it is desired to have a projection operator P_s close to the matrix P to maintain its properties, which means R_s should be also close to R .

The second quantity we examine is matrix rank ratio defined as $\delta = \frac{\text{rank}(R_s)}{\text{rank}(R)} = \frac{\text{rank}(P_s)}{\text{rank}(P)}$. As the undersampling is taken, it is expected that matrix rank is reduced. A large δ value indicates less rank reduction caused by undersampling, which is a favorable property for solving a linear equation, e.g., in the reconstruction problem.

The third quantity of interest is $\beta = \frac{|R_s^\dagger - R^\dagger|_F}{|R^\dagger|_F}$, where X^\dagger is Moore-Penrose pseudoinverse of a matrix X . When we consider least square reconstruction problem for $Pf = g$, the solution is $f = P^\dagger g$. Since $|P_s^\dagger - P^\dagger|_F^2 = |U^T(R_s^\dagger - R^\dagger)V|_F^2 = |R_s^\dagger - R^\dagger|_F^2$, the quantity β characterizes the deviation of the least square solution in the undersampled case from the full projection case. In terms of computation,

$$\begin{aligned} |R_s^\dagger - R^\dagger|_F^2 &= \text{tr}[R_s^{\dagger T} R_s^\dagger] + \text{tr}[R^{\dagger T} R^\dagger] - 2 \text{tr}[R_s^{\dagger T} R^\dagger] \\ &= \sum_i \lambda_{si}^{-2} + \sigma_i^{-2} - 2 \text{tr}[R_s^{\dagger T} R^\dagger] \end{aligned} \quad [16]$$

where tr is matrix trace. For the last term, with numerical SVD of the matrix $R_s = X\Lambda_s W^T$, it is easily to show that $R_s^{\dagger T} R^\dagger = W\Lambda_s^{-1} X^T \Sigma^{-1}$, where Σ is a diagonal matrix formed

by the singular values of R . Hence, β can be evaluated numerically.

In addition, we compare left singular vectors of P_s that serve as the basis functions in the image domain for the reconstructed CT image. Specifically, with the matrix W in the numerical SVD form $R_s = X\Lambda_s W^T$, we compute $v_{si}(x) = \sum_j W_{ij} v_j(x)$. Here $v_j(x)$ is the left singular vector of the full projection operator P , whose expression is given in Eqs. [1–4].

CT reconstruction experiment

We perform CT reconstruction experiments aiming at comparing reconstructed image quality among different undersampling approaches at the same undersampling ratio. The Forbild head phantom with a resolution of 256×256 pixels is used. Forward fan-beam projection of the phantom is computed at $M=360$ equiangular projection angles covering the 2π range. Within each projection, 384 equiangular rays are sampled.

For the regular view undersampling case with a given undersampling ratio $r \leq 1$, we compute the number of projections rM and selected this number of equally spaced projection views. For the regular ray undersampling case, within each projection for each consecutive a rays, we selected $b=2$ rays, such that the ratio $\frac{b}{a} \sim r$. For the random ray undersampling case, each ray was selected in a random fashion with a probability of r . Once the rays for an undersampling case are selected, projection data g_s corresponding to these selected rays are extracted from the already-computed full projection data. Correspondingly, the rows in the system matrix A are selected to form the system matrix A_s for the undersampled case.

We study two reconstruction models. The first model is a simple least-square model which can be formulated as a quadratic optimization problem

$$f = \arg \min_f \frac{1}{2} |P_s f - g_s|_2^2 \quad [17]$$

The global optimal solution of this model can be obtained by seeking for f^* satisfying $P_s^T P_s f - P_s^T g_s = 0$, which can be solved by an iterative linear solver such as conjugate gradient (CG) method. The solution to this problem is directly connected to the SVD form of the system matrix.

The second reconstruction model is based on TV

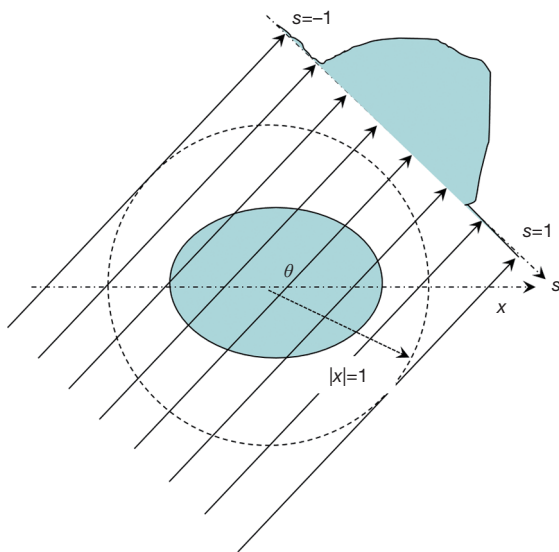


Figure 2 Parallel projection geometry considered in our analysis.

regularization, which solves an optimization problem

$$f = \arg \min_f \frac{1}{2} \|P_s f - g_s\|_2^2 + \tau \|\nabla f\|_1 \quad [18]$$

where τ is a parameter that controls the relative contributions of the two terms in the objective function. This parameter is manually selected in each case for the best image quality.

The optimization problem is solved using the Split-Bregman method (25,26) which has been proved to be equivalent to alternating direction method with multiplier (ADMM) (27,28). We study this model for two reasons. First, in the TV method, we expect the least square term produces the underlying structure of an image, and the TV term regularizes the image and removes image artifacts. Hence, the image quality is largely determined by the projection operator properties. Second, the regularized method can produce images with much improved image quality compared to the non-regularized least square method. It is a meaningful study on the performance of the state-of-the-art TV-method in the three undersampling contexts.

Results

Comparison of projection matrices

μ , δ , and β calculated at different undersampling ratios are shown in Figure 3. μ of the random ray undersampling case is the highest among all the three cases. This implies

that for a given undersampling ratio, the random ray undersampling preserves the most information. The random ray undersampling approach maintains the same rank as the representation of full projection, whereas the other two operators have substantially reduced rank in the regime with a low undersampling ratio. Moreover, the random undersampling also achieved the lowest β value, implying its advantages in pseudo-inverse calculations.

Finally, the comparison of left singular vectors of P_s in the image domain is depicted in Figure 4. For a fair comparison, we select singular vectors in different cases such that they have approximately similar singular values. The singular vector for the full matrix appeared as a ring pattern. As can be seen from $v_{m,l}(x)$ in Eqs. [1–4], the function amplitude only depends on the radius. The phase angle variation leads to changes of image intensities in real and imaginary parts on a ring. For the regular view undersampling, the off-diagonal elements couple singular modes along the angular directions, breaking this ring pattern. In the regular view undersampling case, the matrix elements couple between modes with in each projection view, leading to strong oscillations along the radial direction. This is reflected as more obvious ring patterns in the singular vector. Finally, although there are off-diagonal elements coupling both the angular and the radial directions in the random ray undersampling case, since these elements are small, they almost have no impact on the singular vector image.

CT reconstruction

After presenting properties of the three different projection operators, we examine how they influence the result of CT reconstruction. All the experiments are performed at a given undersampling ratio of $r=0.15$. The result of the least-square model in Eq. [17] is given in Figure 5. Without any spatial regularization in the reconstruction model, artifacts are quite obvious in all the three cases, but appearing with different patterns. It is easy to see the severe ring artifacts in the regular ray undersampling case, while serious streak artifacts can be observed in the image under the regular view undersampling. The artifacts in the random ray undersampling case are more noise-like and do not contain a very strong pattern.

Reconstruction results for the TV-based model are presented in Figure 6. The regularization parameter τ for each undersampling scheme was manually tuned for the best reconstruction quality. The ground truth phantom is

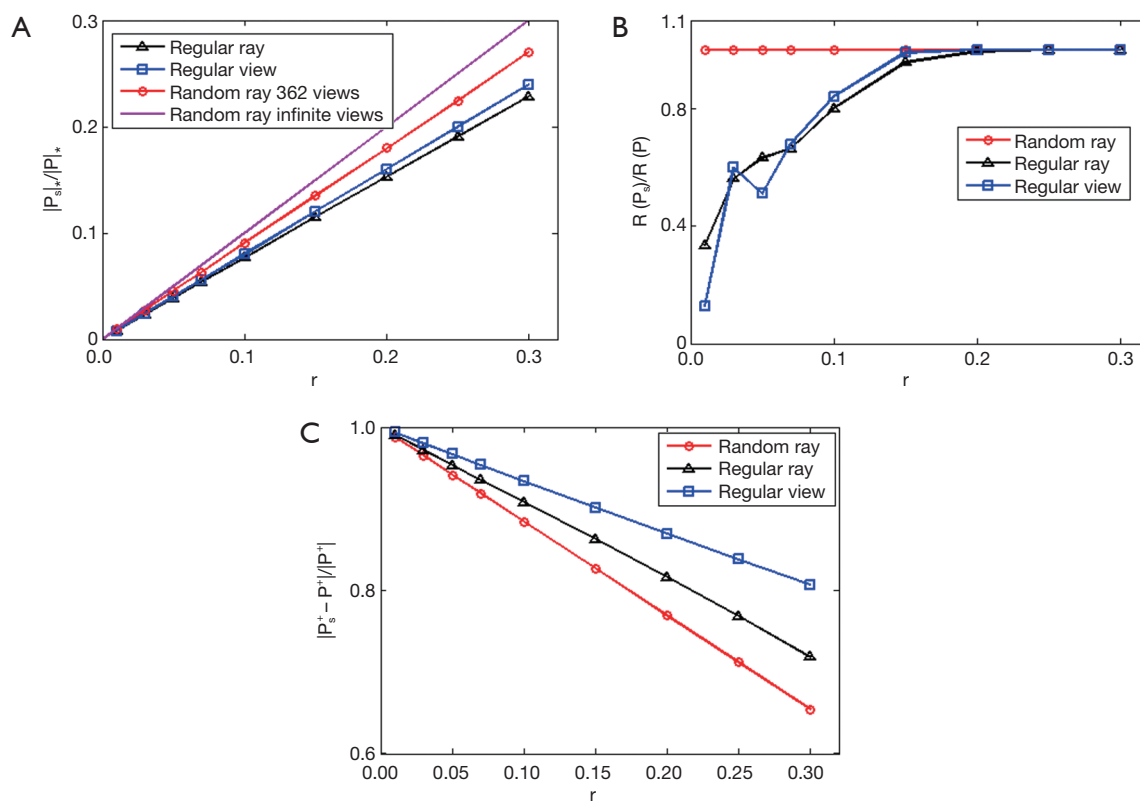


Figure 3 μ (A), δ (B), and β (C) as functions of undersampling ratio for different undersampling operators.

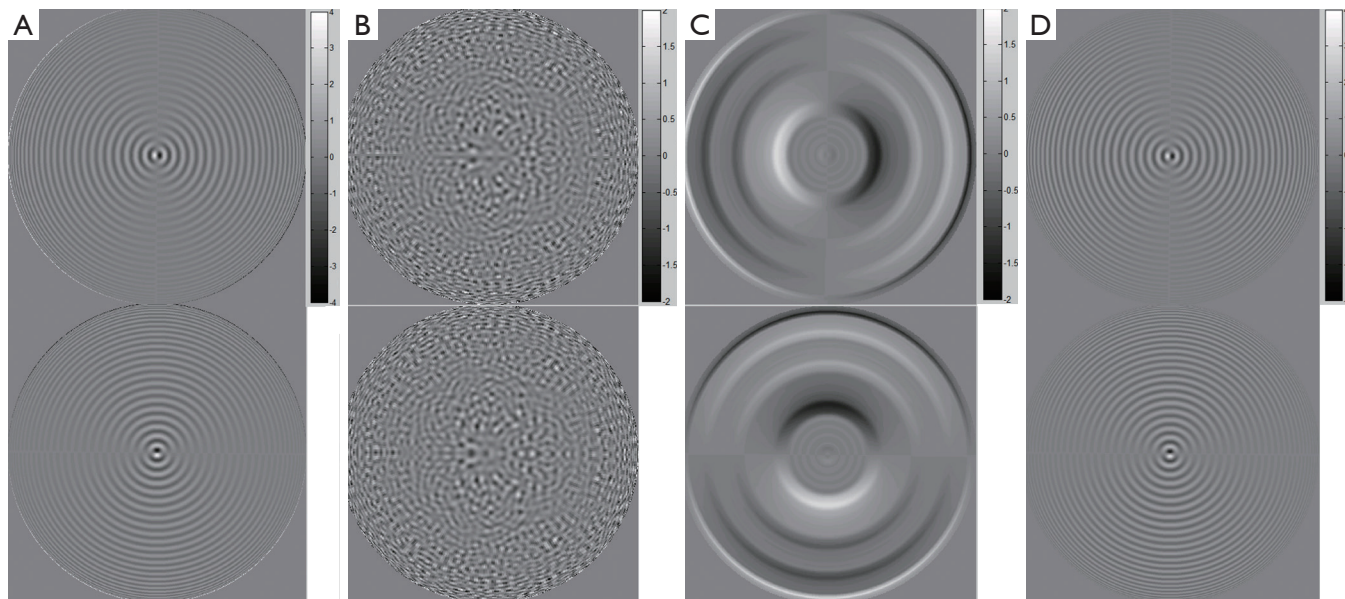


Figure 4 Comparison of left singular functions for (A) full projection, (B) regular view undersample, (C) regular ray undersample, and (D) random ray undersample matrices. Top and bottom rows are real and imaginary part of the functions.

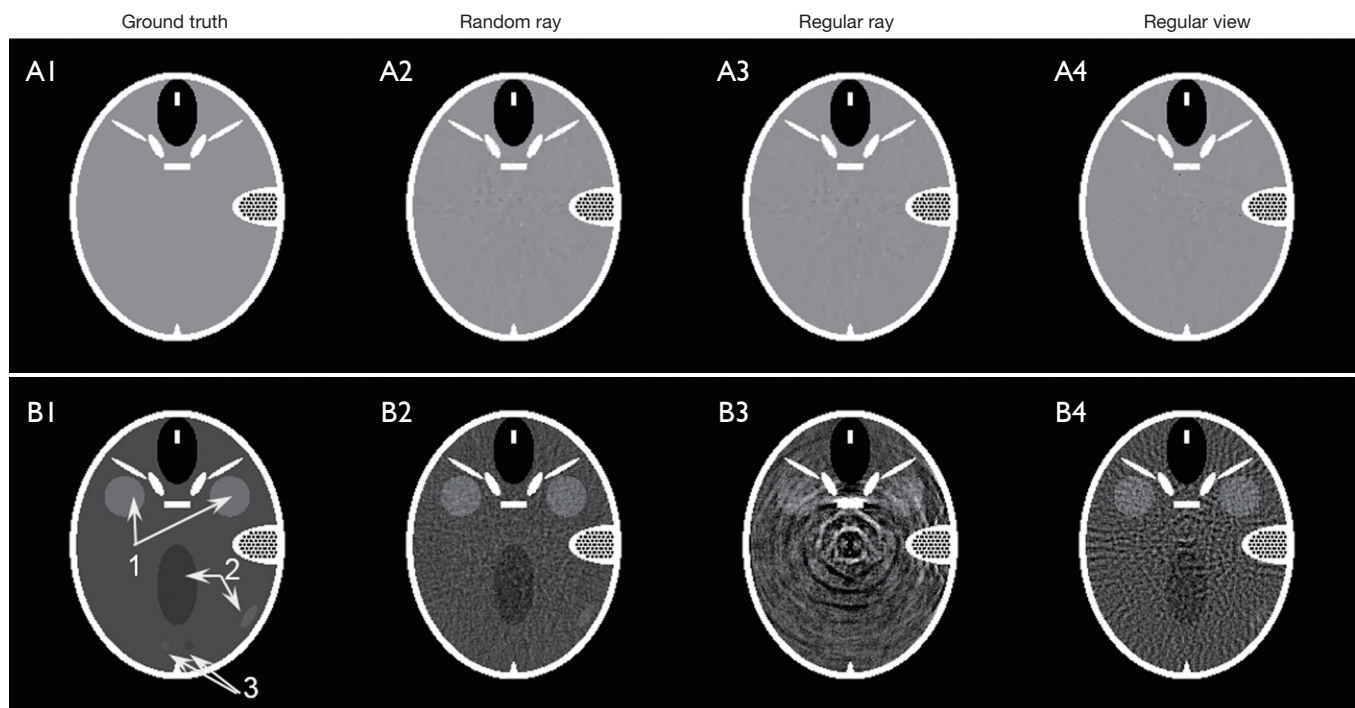


Figure 5 A1 and B1 is the Forbild phantom displayed in a wide window [0, 1.5]; 2–4 are reconstruction results displayed in a narrow window [1.03, 1.1] for different undersampling methods with the same undersampling ratio $r=0.15$.

displayed in *Figure 6A1, B1* in two display windows. The use of a narrow window allows visualization of several low contrast objects that are challenging for the reconstruction problem. While all the reconstruction results are visually close to the ground truth in the wide display window, they are dramatically different when being examined in a narrow display window. Specifically, the residual artifacts behaved differently among undersampling cases. As having been commonly observed in many other previous studies, the regular view method generates streak artifacts. The regular ray undersampling gives rise to ring artifacts, which can be ascribed to the strong ring pattern in the singular vectors, e.g., in *Figure 4C*. In contrast, the random ray approach gives an artifact that appeared as noise. In addition, *Figure 6C* provides a comparison of line profiles along the line indicated by arrow 4 in *Figure 6B1*. The zoom-in part demonstrates that the line profile in the random ray undersampling case is closest to that of the ground truth image. Quantitatively speaking, random ray undersampling approach achieved the highest peak signal-to-noise ratio (PSNR) value among all the cases. *Figure 6D* plots the

residual error $e = \frac{|f - f_0|_2}{|f_0|_2}$ as a function of undersampling

ratio r . The vertical axis is drawn in the logarithmic scale. Among the three cases, the curve for the random ray undersampling approach stayed at the bottom, indicating the better operator properties in this approach for the CT reconstruction problem.

Discussion

In this paper, we have studied and compared properties of three undersampling operators in CT reconstruction: regular view undersampling, regular ray undersampling, and random ray undersampling. By representing the operators under the basis of singular vectors of the full projection operator, we were able to numerically generate matrices for each case and perform SVD to investigate their properties. It was found that, for a given undersampling ratio, the random ray undersampling approach preserves the properties of the full projection operator better than the other two approaches. This translates to the advantages of reconstructing a CT image at a lower error. Numerical experiments performed on a Forbild phantom demonstrated this fact.

Given the theoretical advantages of random undersampling

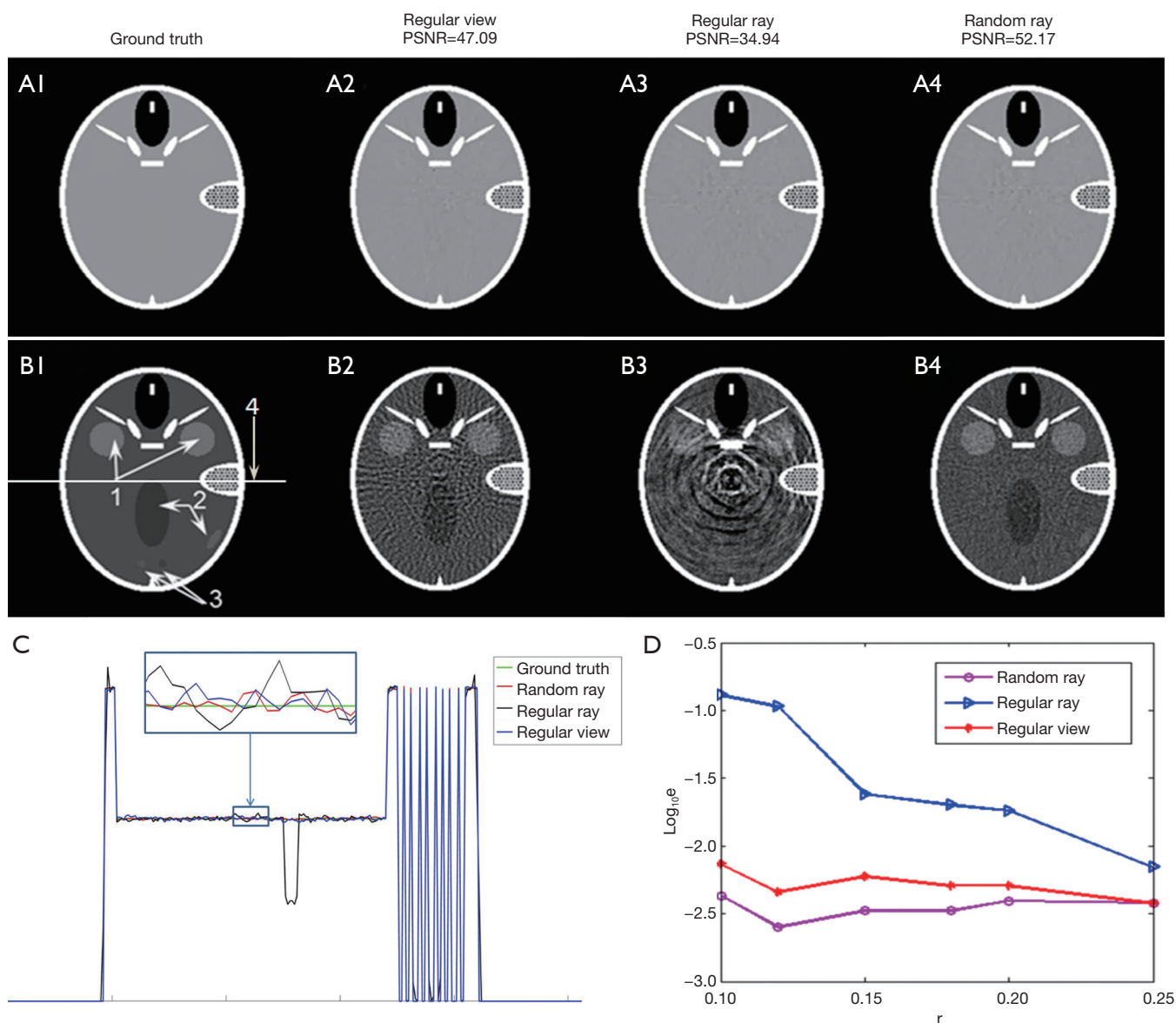


Figure 6 Reconstruction results using TV-based model. A1 and B1 are Forbild phantom displayed in a wide window [0, 1.5] and a narrow window [1.03, 1.1], respectively; A2,A3,A4 and B2,B3,B4 are reconstruction results for different undersampling methods with the same undersampling ratio $r=0.15$ and different display windows; (C) line profiles (see arrow 4 in B1) generated on ground truth phantom image and the reconstruction results; (D) Reconstruction error as a function of undersampling ratio for different undersampling operators.

projection operator, an immediate question is how to realize such a method in a CT or CBCT system. Here we propose two methods. In a CT system, a blocking device similar to the binary multi-leaf collimator used in a Tomotherapy machine could be employed. With a control signal that randomly moves in or out each leaf, random projection data could be acquired. In fact, fluence field modulation has been proposed and investigated in CT reconstruction regime

(29,30). Such an idea has been realized on the Tomotherapy machine recently (31). The proposed random sampling approach generally belongs to the fluence field modulation regime, and is hence expected to be achievable in a similar fashion. In CBCT system, modulating fluence of individual pixel is more challenging. Motivated by the beam blocking idea for CBCT scatter estimation, we propose here a rotating block design. Specifically, a blocker with a random

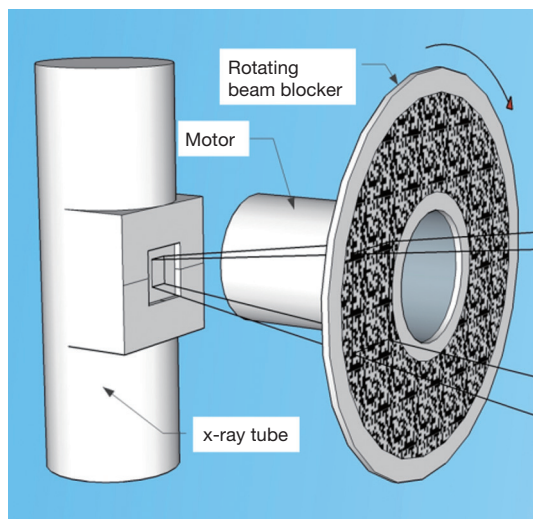


Figure 7 A proposed rotating disk design that could realize the random undersampling on CBCT. CBCT, cone beam computed tomography.

blocking pattern (*Figure 7*) is inserted between the X-ray source and the patient (not drawn). It rotates during CBCT data acquisition under the control of a motor, yielding randomly blocked projections with the blocking pattern varying among projections. One step further, it would be interesting to apply the random projection technique to alternative CT platforms, such as inverse geometry CT system (32-34) which allows reducing of data storage and imaging dose while the time efficiency of CT imaging is boosted.

Main motivation of this study is to introduce randomness in CT data acquisition, in the hope that it would lead to a system matrix that has better numerical properties than the widely used view undersampling or regular ray undersampling methods. Another advantage of it in the CBCT problem is the feasibility for measurement-based scatter estimation and removal. X-ray scatter is the main data contamination factor in CBCT, which degrades accuracy of reconstructed CBCT intensity as well as image contrast. Among many existing methods for scatter estimation and removal, beam blocker-based methods have demonstrated their great potential (20,21). Yet one challenge encountered here was the missing of primary data due to beam blocking. The random undersampling proposed in this paper is compatible with the beam blocker-based scatter removal. With randomly blocked X-ray, the shadow area allows scatter measurement as in the existing

approaches. The randomness in data sampling preserves the properties of the projection matrix, permitting high quality CBCT reconstruction. Feasibility for this idea is currently an ongoing study in our group.

One limitation of the current study is that we used parallel-beam geometry for theoretical analysis but fan-beam geometry for numerical studies. Fan-beam geometry is the most widely used CT geometry. But the advantage of parallel-beam system is its nice mathematical property, especially the explicit formula for SVD, which permits the analysis in this paper. Under the condition of full data acquisition, the two geometries are equivalent, as the projection data can be converted between the two cases via re-binning. However, when it comes to undersampling, the equivalence does not exist anymore. In our study, we hope the analysis performed in the parallel-beam geometry captures some insights in the fan-beam setup. This has been demonstrated to a certain extent by the agreement between results in the numerical studies and the theoretical analysis.

Acknowledgments

Funding: This work is supported in part by NIH grants R01CA227289, R37CA214639, R21EB021545, and NSF grant DMS-1522786.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Jakobs TF, Becker CR, Ohnesorge B, Flohr T, Suess C, Schoepf UJ, Reiser MF. Multislice helical CT of the heart with retrospective ECG gating: reduction of radiation exposure by ECG-controlled tube current modulation. *Eur Radiol* 2002;12:1081-6.
2. Kalra MK, Maher MM, Toth TL, Schmidt B, Westerman BL, Morgan HT, Saini S. Techniques and applications of automatic tube current modulation for CT 1. *Radiology* 2004;233:649-57.
3. Gies M, Kalender WA, Wolf H, Suess C. Dose reduction in CT by anatomically adapted tube current modulation. I. Simulation studies. *Med Phys* 1999;26:2235-47.
4. Kalender WA, Wolf H, Suess C. Dose reduction in CT by anatomically adapted tube current modulation. II. Phantom measurements. *Med Phys* 1999;26:2248-53.

5. Sidky EY, Kao CM, Pan X. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *J X-ray Sci Tech* 2006;14:119-39.
6. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys Med Biol* 2008;53:4777-807.
7. Bian J, Siewerdsen JH, Han X, Sidky EY, Prince JL, Pelizzari CA, Pan X. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. *Phys Med Biol* 2010;55:6575-99.
8. Jia X, Lou Y, Li R, Song WY, Jiang SB. GPU-based Fast Cone Beam CT Reconstruction from Undersampled and Noisy Projection Data via Total Variation. *Med Phys* 2010;37:1757-60.
9. Tian Z, Jia X, Yuan K, Pan T, Jiang SB. Low-dose CT Reconstruction via Edge-preserving Total Variation Regularization. *Phys Med Biol* 2011;56:5949-67.
10. Jia X, Dong B, Lou Y, Jiang SB. GPU-based iterative cone beam CT reconstruction using tight frame regularization. *Phys Med Biol* 2011;56:3787-807.
11. Candes EJ, Tao T. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans Inf Theory* 2006;52:5406-25.
12. Donoho DL. Compressed sensing. *IEEE Trans Inf Theory* 2006;52:1289-306.
13. Lou YF, Zhang XQ, Osher S, Bertozzi A. Image Recovery via Nonlocal Operators. *J Sci Comput* 2010;42:185-97.
14. Tian Z, Jia X, Dong B, Lou Y, Jiang SB. Low-dose 4DCT reconstruction via temporal nonlocal means. *Med Phys* 2011;38:1359-65.
15. Jia X, Tian Z, Lou Y, Sonke JJ, Jiang SB. Four-dimensional cone beam CT reconstruction and enhancement using a temporal nonlocal means method. *Med Phys* 2012;39:5592-602.
16. Li B, Shen C, Chi Y, Yang M, Lou Y, Zhou L, Jia X. Multienergy Cone-Beam Computed Tomography Reconstruction with a Spatial Spectral Nonlocal Means Algorithm. *SIAM J Imaging Sci* 2018;11:1205-29.
17. Pelc NJ. Recent and Future Directions in CT Imaging. *Ann Biomed Eng* 2014;42:260-8.
18. Bennett JR, Opie AM, Xu Q, Yu H, Walsh M, Butler A, Butler P, Cao G, Mohs A, Wang G. Hybrid Spectral Micro-CT: System Design, Implementation, and Preliminary Results. *IEEE Trans Biomed Eng* 2014;61:246-53.
19. Cho HM, Barber WC, Ding H, Iwanczyk JS, Molloy S. Characteristic performance evaluation of a photon counting Si strip detector for low dose spectral breast CT imaging. *Med Phys* 2014;41:091903.
20. Zhu L, Xie YQ, Wang J, Xing L. Scatter correction for cone-beam CT in radiation therapy. *Med Phys* 2009;36:2258-68.
21. Zhu L, Bennett NR, Fahrig R. Scatter Correction Method for X-ray CT Using Primary Modulation: Theory and Preliminary Results. *IEEE Trans Med Imaging* 2006;25:1573-87.
22. Dong X, Petrongolo M, Niu T, Zhu L. Low-dose and scatter-free cone-beam CT imaging using a stationary beam blocker in a single scan: phantom studies. *Comput Math Methods Med* 2013;2013:637614.
23. Natterer F. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, 2001.
24. Bortfeld T, Oelfke U. Fast and exact 2D image reconstruction by means of Chebyshev decomposition and backprojection. *Phys Med Biol* 1999;44:1105-20.
25. Cai JF, Osher S, Shen Z. Split Bregman methods and frame based image restoration. *SIAM J Multiscale Model Simul* 2009;8:337-69.
26. Goldstein T, Osher S. The Split Bregman Method for L1-Regularized Problems. *SIAM J Imaging Sci* 2009;2:323-43.
27. Glowinski R, Le Tallec P. *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*. SIAM, 1989.
28. Glowinski R, Marroco A. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *Revue française d'automatique, informatique, recherche opérationnelle Analyse numérique* 1975;9:41-76.
29. Bartolac S, Graham S, Siewerdsen J, Jaffray D. Fluence field optimization for noise and dose objectives in CT. *Med Phys* 2011;38:S2-17.
30. Bartolac S, Jaffray D. Compensator models for fluence field modulated computed tomography. *Med Phys* 2013;40:121909.
31. Szczykutowicz TP, Hermus J, Geurts M, Smilowitz J. Realization of fluence field modulated CT on a clinical TomoTherapy megavoltage CT system. *Phys Med Biol* 2015;60:7245-57.
32. Schmidt TG, Star-Lack J, Bennett NR, Mazin SR, Solomon EG, Fahrig R, Pelc NJ. A prototype table-top inverse-geometry volumetric CT system. *Med Phys* 2006;33:1867-78.
33. Speidel MA, Wilfley BP, Star-Lack JM, Heanue JA, Van Lysel MS. Scanning-beam digital X-ray (SBDX) technology for interventional and diagnostic cardiac

- angiography. *Med Phys* 2006;33:2714-27.
34. Speidel MA, Tomkowiak MT, Raval AN, Dunkerley DA, Slagowski JM, Kahn P, Ku J, Funk T. Detector, collimator

and real-time reconstructor for a new scanning-beam digital X-ray (SBDX) prototype. *Proc SPIE Int Soc Opt Eng* 2015. doi: 10.1117/12.2081716.

Cite this article as: Shen C, Lou Y, Chen L, Zeng T, Ng MK, Zhu L, Jia X. Comparison of three undersampling approaches in computed tomography reconstruction. *Quant Imaging Med Surg* 2019;9(7):1229-1241. doi: 10.21037/qims.2019.07.07

A. Construction of R with finite size

Following the notations in *Regular view undersampling*, k and l are indices for the angular direction and m and n are those for the direction within a projection. The matrix R has a block structure with block indices l and k , and indices inside each block of m and n (Figure S1). In a realistic scenario, m , n , l and k are all finite with $m, n \leq M_{det}$ and $|l|, |k| \leq M_{ang}$ where M_{det} and M_{ang} are upper bounds determined by the maximal sampling frequency of detector and projection angles respectively. The following constraints still remain: a) $m, n > 0$; b) $-m \leq l \leq m$ and $m+l$ even; and c) $-n \leq k \leq n$ and $n+k$ even. For simplicity we change the notation $m \rightarrow 2m+l$ and $n \rightarrow 2n+k$ and rewrite the constraints as a) $2m+l > 0$, $2n+k > 0$; b) $-2m \leq l \leq 2m+l$; c) $-2n-k \leq k \leq 2n+k$. Hence, combine the finite-size limit, we have $k, l \in [-M_{ang}, \dots, -1, 0, 1, \dots, M_{ang}]$ and

$$\left\{ \begin{array}{l} m \geq 0 \\ m \geq -l \\ m > -\frac{l}{2} \\ m \leq \frac{M_{det}-l}{2} \end{array} \right\}, \dots, \left\{ \begin{array}{l} n \geq 0 \\ n \geq -l \\ n > -\frac{k}{2} \\ n \leq \frac{M_{det}-k}{2} \end{array} \right\} \quad [19]$$

More precisely, $-l \leq m \leq \frac{M_{det}-l}{2}$ for $l < 0$, $0 < m \leq \frac{M_{det}-l}{2}$ for $l=0$ and $0 \leq m \leq \frac{M_{det}-l}{2}$ for $l > 0$. With these constraints ready, the matrix R has the structure shown in Figure S1 and matrix elements in each block are calculated based on Eq. [5]. Similarly, the matrices for regular view, regular ray and random ray undersampling cases are structured in the same way, with the matrix elements calculated according to Eqs. [7], [10], [12], respectively.

B. Mean value and variance of real and imaginary parts of $\frac{2\pi}{M} \sum_j S_j e^{iq\theta_j}$

For the real part, $\frac{2\pi}{M} \sum_j S_j \cos q\theta_j$, its mean value

$$E \left[\frac{2\pi}{M} \sum_j S_j \cos q\theta_j \right] = \frac{2\pi}{M} \sum_j E[S_j] \cos q\theta_j = \frac{2\pi r}{M} \sum_j \cos q\theta_j = 0 \quad [20]$$

Here $\sum_j S_j \cos q\theta_j = 0$ was used. Hence, the variance is

$$\begin{aligned} E \left[\frac{4\pi^2}{M^2} \sum_j S_j \cos q\theta_j \sum_j S_j \cos q\theta_j \right] &= \frac{4\pi^2}{M^2} E \left[\sum_{j,t} S_j S_t \cos q\theta_j \cos q\theta_t \right] \\ &= \frac{4\pi^2}{M^2} E \left[\sum_{j=t} S_j^2 \cos^2 q\theta_j + \sum_{j \neq t} S_j S_t \cos q\theta_j \cos q\theta_t \right] \\ &= \frac{4\pi^2}{M^2} \left[\sum_{j=t} E(S_j) \cos^2 q\theta_j + \sum_{j \neq t} E(S_j) E(S_t) \cos q\theta_j \cos q\theta_t \right] \\ &= \frac{4\pi^2}{M^2} \left[r \sum_{j=t} \cos^2 q\theta_j + r^2 \sum_{j \neq t} \cos q\theta_j \cos q\theta_t \right] \\ &= \frac{4\pi^2}{M^2} \left[r \sum_{j=t} \cos^2 q\theta_j + r^2 \sum_{j,t} \cos q\theta_j \cos q\theta_t - r^2 \sum_{j=t} \cos^2 q\theta_j \right] \\ &= \frac{4\pi^2}{M^2} r(1-r) \sum_j \cos^2 q\theta_j \\ &= \frac{4\pi^2}{M^2} r(1-r) \sum_j \frac{1}{2} [1 + \cos 2q\theta_j] \\ &= \frac{2\pi^2}{M} r(1-r) \end{aligned} \quad [21]$$

C. Proof of $E[f(s)g(s)] = E[f(s)]E[g(s)]$

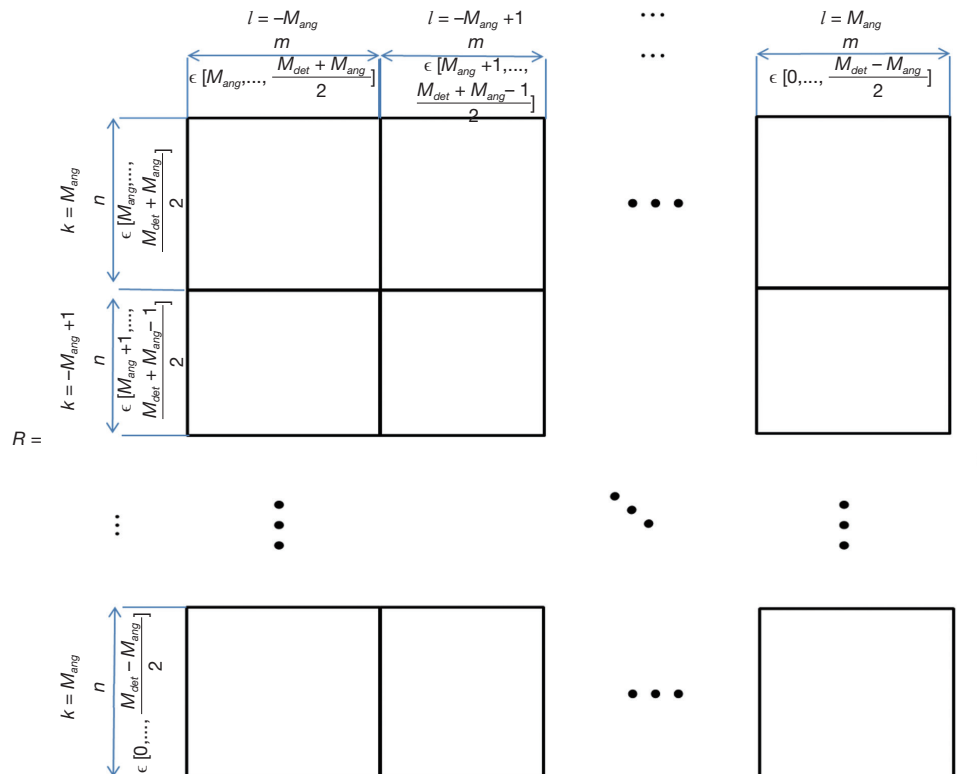
$$\begin{aligned}
E[f(s)g(s)] &= \frac{4\pi^2}{M^2} E\left[\sum_{j,t} S_j S_t e^{iq\theta_j}\right] \\
&= \frac{4\pi^2}{M^2} E\left[\sum_{j=t} S_j^2 e^{iq\theta_j} + \sum_{j \neq t} S_j S_t e^{iq\theta_j}\right] \\
&= \frac{4\pi^2}{M^2} \left[\sum_{j=t} E(S_j) e^{iq\theta_j} + \sum_{j \neq t} E(S_j) E(S_t) e^{iq\theta_j}\right] \\
&= \frac{4\pi^2}{M^2} \left[r \sum_j e^{iq\theta_j} + r^2 \sum_{j \neq t} e^{iq\theta_j}\right] \\
&= \frac{4\pi^2}{M^2} \left[r \sum_j e^{iq\theta_j} + r^2 \sum_{j,t} e^{iq\theta_j} - r \sum_{j=t} e^{iq\theta_j}\right] \\
&= \frac{4\pi^2}{M^2} r^2 \sum_{j,t} e^{iq\theta_j} \\
&= \frac{4\pi^2}{M^2} \sum_j E(S_j) \sum_t E(S_t) e^{iq\theta_j} \\
&= \frac{4\pi^2}{M^2} E\left[\sum_j S_j\right] E\left[\sum_t S_t e^{iq\theta_j}\right] \\
&= E[f(s)] E[g(s)]
\end{aligned} \tag{22}$$


Figure S1 Matrix structure of R.