Homomorphic-Encrypted Volume Rendering

Sebastian Mazza, Daniel Patel, and Ivan Viola

Abstract—Computationally demanding tasks are typically calculated in dedicated data centers, and real-time visualizations also follow this trend. Some rendering tasks, however, require the highest level of confidentiality so that no other party, besides the owner, can read or see the sensitive data. Here we present a direct volume rendering approach that performs volume rendering directly on encrypted volume data by using the homomorphic Paillier encryption algorithm. This approach ensures that the volume data and rendered image are uninterpretable to the rendering server. Our volume rendering pipeline introduces novel approaches for encrypted-data compositing, interpolation, and opacity modulation, as well as simple transfer function design, where each of these routines maintains the highest level of privacy. We present performance and memory overhead analysis that is associated with our privacy-preserving scheme. Our approach is open and secure by design, as opposed to secure through obscurity. Owners of the data only have to keep their secure key confidential to guarantee the privacy of their volume data and the rendered images. Our work is, to our knowledge, the first privacy-preserving remote volume-rendering approach that does not require that any server involved be trustworthy; even in cases when the server is compromised, no sensitive data will be leaked to a foreign party.

Index Terms—Volume Rendering, Transfer Function, Homomorphic-Encryption, Paillier

1 INTRODUCTION

Volume rendering is extensively used in domains where the underlying data is considered highly confidential. One example includes the field of medicine, where CT, MRI, or PET data are used for diagnostic or treatment-planning purposes. Another such example is hydrocarbon and mineral exploration in energy industries for inspecting the subsurface using seismic scans.

For volume rendering, privacy can currently only be achieved by storing and processing the datasets locally. Volume rendering requires computers with large memory and powerful processing power. Such hardware must be frequently maintained and upgraded. Therefore, for many organizations, it would be advantageous to outsource the rendering to cloud services. As cloud services remove the need to be in close proximity to the rendering hardware, users can now also view volume rendering on thin clients that do not have the required memory or processing power, such as tablets and smart phones. However, hospitals must protect sensitive personal data and energy companies must protect their valuable data assets. Thus, it is essential that their data is not visible to the cloud services, as these either cannot be trusted, or their security might be compromised. Therefore, we want to make it possible to perform direct volume rendering on untrusted hardware while preserving the same level of privacy for the datasets as the privacy achieved with a classical local rendering approach.

The basic concept of our privacy-preserving approach is shown in Figure 2. First, the data is acquired and immediately encrypted by, for example, a machine that is directly connected to a medical scanner. Then the encrypted volume is uploaded to the honest-but-curious public server. This is done only once per volume. When the clients that hold the secure key request rendering, the server performs ray-casting directly on the encrypted volume data. This computation results in an image containing encrypted values, which is then sent to the client. When the client receives the requested image, it is decrypted and displayed to the user. As the server that computed the rendered image will only see encrypted data, our approach maintains privacy.

Our design is constrained by three requirements. The first require-

---

1 The server will perform the algorithm it is requested to compute honestly (correctly); however, there is the potential for possible access by the curious eyes of administrators or hackers.
**Fig. 2.** Our approach consists of a computer that produces, encrypts, and sends volume data to a server, which then renders the data and sends the result to a client. The client decrypts and visualizes the result. The text that belongs to encrypted data or processing is stated in red.

**ment** is that the privacy of the user data is protected by the design of the algorithm and does not depend on hiding implementation details, keeping any part of the system secret, or any other obscure technique that cannot be secure, at least not in the long run (Kerckhoffs’s principle [17]). Such obscure techniques only make it difficult to know the actual security of the system. Therefore, the security of our volume rendering approach solely depends on the security of a well-established cryptographic algorithm, continuously being scrutinized in cryptographic research. We have chosen the well-established Paillier cryptographic algorithm, which is partly homomorphic [24]. The key property of homomorphic encryption (HE) is that arithmetic operations on encrypted data are dual to arithmetic operations on plaintext (original, unencrypted) data. This enables an algorithm to perform a correct 3D volume rendering image synthesis directly on the encrypted data, without being able to ever access the plaintext data. As a consequence, the result of the rendering on the server is an encrypted image.

We are currently not able to show interactive frame rates with the proof-of-concept implementation of our approach. However, one future research goal should attempt to make a remote rendering system fast enough to achieve this. This leads to our second requirement, which is to only use techniques that will not prevent the system from scaling the performance with the computational power available on the server and will not prohibit interactive frame rates. The third requirement is to support thin clients without much memory and computational power. As a result, we consider the client to be a low powered device, which is connected to a mobile or another medium-bandwidth network, while we assume that the server is a powerful machine (e.g., with multiple professional GPUs) or even a compute cluster.

By using encryption schemes like AES [7], it is currently possible to store volume datasets securely in the cloud. However, for rendering images from the datasets, the entire volume needs to be downloaded and decrypted first, and then rendered on the client. A privacy-preserving remote volume rendering can also make the cloud more attractive as a storage space for volume data, because with our proposed technique, it is no longer necessary to download the whole dataset before images can be synthesized from it.

2 RELATED WORK

We have only found two works that address the topic of privacy-preserving rendering of volumetric data. The most similar work to ours is that of Mohanty et al. [22]. They present a cryptosystem for privacy-preserving volume rendering in the cloud. Unlike our approach, they achieve correct alpha compositing. However, to attain this goal, they end up with a solution that cannot be considered secure, that has a fixed transfer function, and that requires that the volume is sent from one server to another server for each rendered frame.

Their approach requires two servers for rendering: a Public Cloud Server and a Private Cloud Server. The first step of their rendering approach is to apply a color and opacity to each voxel before encrypting the volume. This means that the transfer function is pre-calculated and cannot be changed by a user without performing a time-consuming reencryption and uploading of the volume. In the next step, the encrypted data is uploaded to the Public Cloud Server, which stores the volume data. When the Public Cloud Server receives an authorized rendering request from a client, the server calculates all sample positions for the requested ray casting and interpolates the encrypted color and opacity values for each sample position. All interpolated sample values then need to be individually sent to the Private Cloud Server, which decrypts the opacity value of each sample in order to perform the alpha blending along the viewing rays. For alpha compositing, the opacity values of samples represent object structures in the volume; therefore, anyone who can gain access to the Private Cloud Server, such as an administrator or a hacker, will be able to observe these structures in the volume dataset. If an unauthorized person has access to this server, the whole approach collapses. For the task of encrypting and decrypting parts of the volume data on the servers, their approach requires a central Key Management Authority (KMA). While this brings the advantage that an organization can centrally control which users have access to a specific volume, it enforces the attack surface of their system considerably, because the KMA has all keys required for decrypting all volume data. Therefore, the confidentiality of the KMA is constitutional for the privacy of all datasets, no matter who they belong to.

Another weakness of their approach is the required network bandwidth between the Public and the Private Cloud Server because all sample values of a ray casting frame need to be transferred from the Public and the Private Cloud Server (more than 1GB). With our approach, the privacy of the volume data and rendered image depends only on a single secure key. Also, our approach should scale linearly with the computing power of the hardware it is running on.

Chou and Yang [3] present a volume rendering approach that attempts to make it difficult for an unintended observer to make sense of the volume dataset that resides on a server. This is done by, on the client’s side, subdividing the original data into equally sized blocks. The blocks are rearranged in a random order and then sent to the server as a volume. The server then performs volume rendering on each block and sends the result back to the client, which will reorder the individual block renderings and composite them to create a correct rendering. To obfuscate the data further, on the client’s side, the data values in each block are changed using one out of three possible monotonic operations: flipping, scaling, and translating. Monotonic operations are used as they are invertible and associative under the volume rendering integration. Therefore, doing the inverse operators on the resulting rendering gives the same result as doing them on the data values before performing the rendering. This algorithm cannot be considered safe, and the authors acknowledge this as they state that the goal is only to not trivially reveal the volume to unauthorized viewers. A possible attack would be to consider the gradient magnitude of the obfuscated volume. This should reveal the block borders. The gradient magnitude can further be used inside each block to reveal structures in the data that can be used for aligning the blocks correctly.

To attain our goal of developing an approach that is open and secure by design, we use the Paillier cryptosystem developed by Paillier in 1999 [24]. This cryptosystem is an asymmetric encryption scheme, where the secure key contains two large prime numbers p and
q, and the public key contains the product N (modulus) of p and q. The cryptosystem supports an additive homomorphic operation (⊗). If this operation is applied to two encrypted values \([m_1, m_2]\) (\(\{m\}\) means encrypted \(m\)), the decrypted result is the sum of the \(m_1\) and \(m_2\) (\(\text{Dec}(m_1 \otimes m_2) = (m_1 + m_2) \mod N\)). Furthermore, a homomorphic multiplication (\(\odot\)) between an encrypted value and a plaintext value is done by the Paillier encryption scheme. For example, \(\text{Dec}(m_1 \odot d) = (m_1 \times d) \mod N\). Since Paillier’s cryptosystem does not carry over multiplication of two encrypted values to plain text, it is classified as a partially homomorphic encryption (PHE) scheme. Paillier can securely encrypt many values (e.g., \(512^3\) voxels of a volume) from a small number space (e.g., \(2^{10}\) possible density values), because it is probabilistic, which means that during the encryption, the obfuscation can map a single plaintext value randomly to a large number of possible encrypted values. This makes a simple “probing” for finding out the number correspondence impossible. Further details about Paillier’s cryptosystem such as the encryption and decryption algorithm is provided in the Supplementary material document. We are limited to the arithmetic operations supported by Paillier for creating a volume rendering that captures as much structure as possible from the data. This forces us to think unconventionally and creatively in designing the volume renderer.

For homomorphic image processing, the work by Ziad et al. makes use of the additive homomorphic property of Paillier’s cryptosystem. They demonstrate that they are able to implement many image processing filters using the limited operations allowed with Paillier. They implement filters for negation, brightness adjustment, low pass filtering, Sobel filter, sharpening, erosion, dilation and equalization. While most of these filters are computed entirely on the server side, erosion, dilation, and equalization require the client for parts of the computation. There are various works that make use of such a trusted client protocol approach to overcome the limitation of a PHE scheme and enable operations such as addition, multiplication, and comparisons on the encrypted data. A trusted client knows the secure key and can, therefore, perform any computation on the data or convert/re-encrypt it from one encryption scheme to another (e.g., from an additive to a multiplicative homomorphic encryption). These client-side computations introduce latency because the data needs to be transferred back and forth between the server and the client. Furthermore, the client needs to have enough computational power to avoid becoming the bottleneck of the system. To mitigate this problem, automated code conversions can be used that minimize the required client side re-encryptions. While a trusted client approach could theoretically solve many of the problems we face with our untrusted server-only approach, it is not practical for volume rendering. The most demanding problems of volume rendering, such as transferring a voxel value and advanced compositing (alpha blending, maximum intensity projection, ...), need to be done per voxel. Hence, every voxel that could contribute to the image synthesis processes must be transferred to the trusted client and processed there for every rendered frame. The encryption and decryption on the client side are more expensive than the operations required for a classical sample compositing due to the size of encrypted values (e.g., 1000 bit per voxel). If an amount of data in the range of the volume itself needs to be transferred from the server to the client, where the data would need to be encrypted and decrypted, it is pointless to perform any calculations on the server, because the client then has more work to do than in a classical volume rendering on the client. Moreover, it does not save any network traffic as compared to a simple download, decrypt, and process use case. Therefore, we argue that trusted client approaches are not suitable for our work. Furthermore, a trusted client approach will not work with thin clients, which contradicts our third requirement. Our second requirement is also contradicted because in real-world use cases, the network bandwidth between a client like a tablet computer and a cloud server will not have enough bandwidth (e.g., more than 1Gbit/s) to support interactive frame rates.

3 Encrypted Rendering Overview

The first step of the introduced privacy preserving rendering system is the encryption of the volume dataset (Figure 2 Acquisition Device). During the encryption stage, every single scalar voxel value of a volume dataset needs to be encrypted with Paillier’s approach (see Algorithm 5 in the Supplementary Material document). Meta data of the volume such as width, height, depth and the storage order of voxels will not be encrypted. The next step is to upload the encrypted volume dataset to a server (Figure 2 arrow from Acquisition Device to Cloud Server). For our approach, the device that encrypts the volume and uploads it to a server does not even need the secure key, because for encryption, only the public key is required.

When a rendered image is requested to be shown on a client, the client sends a rendering request to the server, which has the encrypted volume dataset (Figure 2 arrow from Client to Cloud Server). The rendering request contains further information about the settings of the rendering pipeline, such as the camera position, view projection, and (depending on the selected rendering type) also information about the transfer function that should be used. After the server receives such a rendering request, it uses the included pipeline settings and the already stored encrypted volume dataset to render the requested image (Figure 2 the rendering pipeline stages of the Cloud Server). To preserve privacy, the server does not have the secure key and can not, therefore, decrypt the volume data. The operations that are used for rendering the image from the encrypted volume dataset are limited to the homomorphic operations addition (\(\oplus\)) and multiply with plaintext (\(\odot\)), which are defined for Paillier’s encryption scheme. When the rendering is finished, the server will send the calculated image data to the client (Figure 2 arrow from Cloud Server to Client). The resulting image that the client receives is still encrypted. Decrypting such an image is only possible for a client that knows the correct secure key. For everyone else, the image will be random noise (shown in Supplementary Video Material). Since every single pixel value is an encrypted number, every single pixel can be decrypted independently of the other pixels. For a gray-scale image, that means one number per pixel. An RGB colored image requires three values that need to be decrypted per pixel.

In Section 4, we explain how the homomorphic operations of Paillier’s HE can be used for X-ray sample integration. Furthermore, we will show how to use Paillier’s cryptosystem with floating-point numbers, which allows us to perform trilinear interpolation. Section 5 explains a more advanced approach that allows the emphasizing of different density ranges in the rendered images.

4 Encrypted X-Ray Rendering

Ray-casting is the most frequently used approach for volume rendering. Furthermore, ray-casting based algorithms can be easily and efficiently parallelized and can be implemented with fewer memory reads than slicing-based algorithms. Memory access is time-consuming, especially if every number that needs to be read is thousands of bits long. Therefore, we implement our privacy-preserving volume rendering approach with ray-casting. However, other direct volume rendering approaches developed for unencrypted data, such as slicing, can be used as well. Slicing on the server can be built by the same encrypted rendering pipeline components (sampling / interpolation, color mapping, compositing), which we will explain anon. Slicing could also be used to just perform the sampling on the server, transfer the slices to the client, and perform the compositing there. However, this would not fulfill our requirements because of the required network bandwidth and the computation time on the client.

The ray casting algorithm first calculates a viewing ray for every pixel of the final image (Figure 2 Ray Traversal - stage of the Server). These viewing rays will be calculated based on the camera position, up vector, opening angle, image resolution, and pixel index. At discrete and equidistant steps along the ray, the data of the volume is sampled (Figure 2 Sampling - stage of the Server). The last step is the compositing, where the final pixel value is calculated based on the sample values of a viewing ray (Figure 2 Compositing - stage of the Server). X-ray rendering is a volume rendering approach where the sample value is mapped to a white color with monotonically increasing opacity, and the compositing is a summation followed by a normalization at the end of the ray traversal. If the sampling of the voxel values is done by nearest-neighbor filtering, the sum along a viewing ray can be
calculated by only using the homomorphic add operation $(\oplus)$ which is already defined for Paillier’s cryptosystem. The final normalization of all samples along a view ray cannot be done directly by the homomorphic operations of Paillier’s encryption scheme because this requires a division that can result in a non-integer value that is not supported. However, the server could send the encrypted sum together with the sample count to the client, which can perform the division after decrypting the sum.

To improve the nearest-neighbor sampling with trilinear interpolation, a mechanism that allows the summing and normalization of encrypted values $([m_1], [m_2])$, which are scaled by some plaintext weights $(\alpha_1, \alpha_2)$, is required. For plaintext integers, the interpolation could be implemented between the integer arithmetic operations add, multiply, and divide (1D example: $([m_1 \cdot \alpha_1 + m_2 \cdot \alpha_2])/(\alpha_1 + \alpha_2)$). Since an arbitrary division is not supported by Paillier’s cryptosystem, this is not directly feasible on encrypted data. A possible solution could be to use fraction types, which has an encrypted denominator and a plaintext numerator for storage and calculations. After the image is rendered, which contains such fractions as pixel values, the client can download it, decrypt the denominators, and perform the deferred division.

However, we decided to use a floating-point encoding, which is easier to implement and allows a shader code development as is usual for hardware accelerated rendering. With a floating-point representation of encrypted values, it is possible to multiply the eight neighboring voxels of a sample position with the distances between the samples and voxel position. These distances, which have a sum of 1, are the weights of the interpolation (1D example: $m_1 \cdot \alpha_1 + m_2 \cdot \alpha_2$). A floating-point encoding will also make the final division of the sample sum for X-ray rendering on the server side possible. While a floating-point encoding does not directly enable divisions in the encrypted domain, it can be used to approximate a division by a multiplication with the reciprocal of the divisor, as shown in Equation 1.

\[
\sum \approx \text{Dec} \left( \left( \sum \right) \cdot \frac{1}{\tilde{n}} \cdot 10^{-\gamma} \right) \cdot 10^{-\gamma}
\]

The sum of samples along a viewing ray is denoted as $\sum$, and $n$ is the count of samples. The precision of the approximation is defined by the count of decimal digits $\gamma$ (e.g., $\gamma = 3$ for thousandth). Before the reciprocal of $n$ is multiplied with the sum of the comma is moved $\gamma$ digits to the right ($10^\gamma$) and then rounded ($\lceil \cdot \rceil$). The multiplication with $10^{-\gamma}$, which moves the comma back to the correct position, can be achieved, by subtracting $\gamma$ from the exponent of the floating-point encoded result. Since the Paillier cryptosystem is defined over $\mathbb{Z}_N$, the result is only correct if no intermediate result is greater than $N - 1$.

We will discuss the used floating-point encoding in Section 4.1. It is called the mantissa. The exponent $e$ defines the position of the comma in the final number. The base $b$ is a constant that is defined upfront (e.g., during the compilation of the application). We used a decimal system for convenience; therefore, our prototype uses $b = 10$. However, $b$ can be any positive integer that is greater or equal to 2.

4.1 Encrypted Floating-Point Numbers

A floating-point number is defined as $m \cdot b^e$, where $m$ is called the mantissa. The exponent $e$ defines the position of the comma in the final number. The base $b$ is a constant that is defined upfront (e.g., during the compilation of the application). We used a decimal system for convenience; therefore, our prototype uses $b = 10$. However, $b$ can be any positive integer that is greater or equal to 2.

To calculate with floating-point arithmetic in the encrypted domain, we have chosen to use the approach developed for Google’s Encrypted BigQuery Client [11]. The idea is to store the mantissa $m$ and the exponent $e$ of a floating-point number in two different integer variables. During the encryption of the floating-point number $(m, e)$, only the mantissa $m$ is encrypted using Paillier’s cryptosystem. The exponent $e$ remains unencrypted, which results in the floating-point number $([m], e)$. This floating-point number representation is also used by the python-paillier library [28], the Java library javallier [29] and in the work by Ziad et al. [30].

For an addition of two such encrypted floating-point numbers, both need to have the same exponent. Therefore, the exponents of both numbers must be made equal before the actual addition, if they are not already equal. Hence, it is not possible to increase the exponent if the mantissa is encrypted because that would require a homomorphic division of the encrypted mantissa, which is not possible. Therefore, the floating-point number with the greater exponent needs to be changed. On the other hand, decreasing the exponent of a floating-point number is not a problem because it requires a homomorphic multiplication of the encrypted mantissa with a plaintext number, which is possible with Paillier. Equation 2 shows how to calculate the new mantissa $[m]$ that is required for decreasing the exponent of the floating-point number $([m], e)$ to the lower exponent $e_n$. The new floating-point number is defined as $([m], e_n)$, which represents exactly the same number as $([m], e)$. It is just another way to store it.

\[
[m] = [m] \otimes b^{e - e_n}
\] (2)

When both floating-point numbers $([m_1], e_1)$ and $([m_2], e_2)$ have the same exponent $e_1 = e_2 = e_n$, the homomorphic sum $[m]$ of both mantissas can be calculated by the add operation defined for Paillier, which results in the final floating-point number $([m], e_n)$. The Algorithm 1 shows this approach for summing two floating-point numbers with encrypted mantissas. The lines from 2 to 10 bring the exponents of both floating-point numbers to the same value ($e_n$), and line number 11 contains the addition of the encrypted mantissas. A multiplication with a floating-point number that contains an encrypted mantissa $([m], e_1)$ and a floating-point number with a plaintext mantissa $([m_2], e_2)$ can be achieved by multiplying the mantissas with the multiplication operation defined for Paillier $([m_1] = [m] \otimes m_2)$ and a plaintext addition of the exponents ($e = e_1 + e_2$). This is also stated in line 10 and 11 of the Algorithm 2 which is sufficient for a correct result. The lines from 2 to 9 contain a performance optimization, which prevents the intermediate result of $[m_1] \cdot m_2$, which is computed before mod $N^2$ is applied in line 10, from being unnecessarily large. This optimization is also used by the python library python-paillier [28] in paillier.py and the java library javallier [15] in PaillierContext.java.

Signed numbers can be represented by using a two’s complement representation for the mantissa $m$. The exponent $e$ does not change. If $v$ is a negative integer, the two’s complement in the integer modulo $N$ can be calculated by: $m = v + N$. In the encrypted domain, the additive inverse $-m$ of $m$ is defined by the multiplicative inverse $[m]^{-1} = [1]$ of $[m]$ in the integers, modulo $N^2$ ($[1] = [1]$ mod $N^2$ and can be computed from $[m]$ and $N^2$ by the extended Euclidean algorithm [18]). This complement representation for encrypted floating-point encoded dataset. For the rendering of the left image, a nearest-neighbor sampling was used, and for the right image, a trilinear interpolation was used. The used dataset contains three objects with different densities: a solid cube in the center wrapped inside a sphere and another sphere at the top left front corner. The same dataset is also used for renderings shown in Figure 4 and Figure 6.
The evaluation of such a transferred value allows for a comparison of two equal voxel values, because of the obfuscation during the encryption. That means an encrypted voxel dataset will probably not contain many voxel values, because it would be at least as big as the volume itself.

Another approach for non-encrypted data is to store just some supporting points that contain the density and color. The evaluation for this transfer function approach is achieved by interpolating the color between the value of the next lower and next greater supporting point. To find the neighboring supporting points of the voxel value that should be transferred, comparison operators such as less than (<) or greater than (>) are required. However, comparison operators cannot exist for probabilistic PHE schemes like Paillier because that would break its security (see Section 7.3). Therefore, the question is how to implement comparison operators such as lower than (<) or greater than (>) for encrypted numbers. Wamser et al. [36] suggest to use a Vandermonde-Matrix as an approximation of those operators. The result of this function is again an encrypted number. A promising approach that can achieve this was presented by Wamser et al. [36] in their work on "oblivious lookup-tables".

5.1 Oblivious Lookup Tables

Let $X = \{x_1, x_2, \ldots, x_n\}$ be an enumeration of values that should be mapped to $Y = \{y_1, y_2, \ldots, y_m\}$ by the lookup function $f(x_i) = y_1$. The idea is to create a vector $\vec{v}_i$ for every $x_i \in X$ with the same cardinality as $X$ ($(\vec{v}_i)_j = |X|$) and define the evaluation of a lookup by the dot product $\vec{v}_i \cdot \vec{l} = y_i$. The scalar value $y_i$ is the result of the lookup. For a transfer function, this would be the value of one color channel. The vector $\vec{l}$ can be calculated form the linear equation $\vec{V} \cdot \vec{l} = \vec{y}$, $\vec{V}$ is a square matrix of full rank with $n = |X|$, that uses all vectors $\vec{v}_i$ as rows. However, this linear equation needs to be solved only once. Therefore, the client can calculate $\vec{l}$ upfront based on unencrypted numbers. The equation $\vec{V} \cdot \vec{l} = \vec{y}$ has a unique solution, if all vectors $\vec{v}_i$ are linearly independent. Hence, the crucial part is to find an approach to extrapolate every vector $\vec{v}_i$ only from one single vector $\vec{x}_i$ so that the $\vec{v}_i$ are linearly independent from each other. Wamser et al. [36] suggest to use a Vandermonde-Matrix as $V$ \begin{equation}
V = \begin{pmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\
1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_m & x_m^2 & \cdots & x_m^{n-1}
\end{pmatrix}
\end{equation}
From the creation rule of the Vandermonde-Matrix, it follows that a $\vec{v}_i$, which is equal to the $i$-th row of the matrix $V$, is defined as $\vec{v}_i = (1, x_i, x_i^2, \ldots, x_i^{n-1})$. The lookup function $f(x_i)$ can, therefore, be stated as: $f(x_i) = (1, x_i, x_i^2, \ldots, x_i^{n-1}) \cdot \vec{l} = y_i$ \begin{equation}
f(x_i) = (1, x_i, x_i^2, \ldots, x_i^{n-1}) \cdot \vec{l} = y_i
\end{equation}
The dot product in \begin{equation} \text{Equation 4} \end{equation} can be calculated even if $\vec{v}_i = (1, x_i, x_i^2, \ldots, x_i^{n-1})$ is encrypted because only the operations $\oplus$ and $\odot$ that are defined for the Paillier HE are required for calculating a dot product. However, it is not possible to calculate the vector $\vec{v}_i$ from an encrypted $[\vec{x}_i]$, because this would involve multiplications of two encrypted numbers, which is not possible with Paillier.

A theoretical solution for this could be to store the vector $\vec{v}_i$ instead of scalar $x_i$ as the value of a voxel. For a volume dataset, where the voxel values have only a resolution of 8 bits, this would lead to a vector length of $n = 2^8$ = 256. Therefore, the required storage size for the volume will increase 256 times.

A volume with $512 \times 512 \times 512$ voxels and a resolution of 8 bits per voxel requires $512^3 \cdot 8$ bits/8 bits = 134,217,728 Bytes = 128 MB. The same volume encrypted with Paillier HE with a pub-

---

**Algorithm 1: Paillier Floating Point Add**

**Parameters:** Encrypted mantissa $[m_1]$, $[m_2]$ and plaintext exponents $e_1$, $e_2$ of the two floating point numbers that should be summed.

**Result:** Encrypted mantissa $[m_3]$ and plaintext exponent $e_n$.

```plaintext
1 procedure fpAdd($[m_1]$, $e_1$, $[m_2]$, $e_2$)
2 if $e_1 > e_2$ then
3    $m_3 = [m_1] \odot b^{e_1-e_2}$
4    $e_n = e_2$
5 else if $e_1 < e_2$ then
6    $m_3 = [m_2] \odot b^{e_2-e_1}$
7    $e_n = e_1$
8 else
9    $e_n = e_1$
10 end
11 $[m_3] = m_3 \odot 2^{e_n}$
12 return ($[m_3]$, $e_n$)
```

**Algorithm 2: Paillier Floating Point Multiply**

**Parameters:** Encrypted mantissa $[m_1]$, plaintext mantissa $m_2$ and the plaintext exponents ($e_1$, $e_2$) of the two floating point numbers that should be multiplied.

**Result:** Encrypted mantissa $[m_3]$ and plaintext exponent $e_n$.

```plaintext
1 procedure fpMultiply($[m_1]$, $e_1$, $m_2$, $e_2$)
2 if $m_2 = N - m_2$, $\|m_2\|$ the negative of $m_2$
3 if $m_2 \leq \text{max. value that can be encrypted by current } N$ then
4    $m_3 = [m_1]^{-1} \mod N^2$
5 else
6    $m_3 = [m_1]$
7    $m_d = m_2$
8 end
9 $e_n = e_1 + e_2$
10 return ($[m_3]$, $e_n$)
```
5.2 Density Range Emphasizing

Our simplified transfer function approach is based on the observation that it is possible to compute the dot product of a vector with encrypted values and a vector with plaintext values. Furthermore, the dot product can be used to calculate an encrypted scalar value indicating the similarity of an encrypted vector and a plaintext vector. This will work if both vectors have length 1. Therefore, our approach is to encode the density values of each voxel as a vector and encrypt each component of this vector by the Paillier encryption algorithm (Supplementary Material Algorithm 5). In order to highlight a user-defined density range, the density value at the center of this range needs to be encoded as a vector. Note that this vector is not encrypted. The encrypted volume rendering engine can now compute the dot product between this vector and the encrypted vector of a sample position. Then the ray-casting algorithm needs to sum up the results of the dot products along a ray instead of the density values. This approach allows a user to emphasize a selectable density range in the rendered image. Figure 4 contains images that were created using this approach. The top left subfigure shows a result of an X-ray rendering for comparison. All other subfigures show results for different density ranges that are emphasized. The density that is encoded as a vector that was used for the dot-product calculation is specified in the caption of each subfigure.

The density-to-vector encoding scheme we used is based on an HSV-to-RGB color conversion. The exact encoding scheme is stated in Algorithm 3 (Figure 5) illustrates the magnitude of the vector components for all possible density values. Furthermore, the response intensities for user-defined emphasizing densities at 0.45 and 0.85 are shown. At the last line of Algorithm 3 the calculated vector is normalized. This is important to make sure that the result of the dot product is always between 0 and 1 and to ensure that the highest possible dot product result (1) is at the user-defined emphasizing density.

There are other and possibly better density-to-vector encoding schemes. However, the HSV-based encoding leads to results that feel natural, especially while smoothly increasing or decreasing the emphasizing density. The encoding scheme should in any case be chosen in such a way that the curve created by the dot product is steep and narrow (see dashed lines in Figure 5), so that the density selected by the user can be seen as clearly as possible in the resulting image. The Algorithm 3 takes not only the density that should be encoded as parameter, but also the count of dimensions of the returned vector. Increasing the count of dimension not only makes the dot product response curve more steep (See Figure 5 and compare the dashed lines in the left and right plot.), but also increases the required storage size of the encoded and encrypted volume dataset. Note that the count of dimensions must be the same during the encryption of the volume and for the encoding of the user-defined emphasizing density. This also means that the amount of computations required for the volume rendering depends on the number of dimensions used for encoding the volume.

5.3 Simplified Transfer Function

It is possible to add RGB colors to the rendered images based on the density range emphasizing described in the last section. This is useful because RGB colors allow a user to emphasize different densities in the same image while keeping the densities distinguishable (see Figure 6). Since the dot product between an encoded and encrypted voxel value and a user-defined encoded density is an encrypted scalar value, a multiplication with another plaintext number is possible. For our simplified transfer function approach, the dot product result needs to be multiplied with a user-defined RGB color vector. As the dot product expresses the similarity between the voxel value and the user-defined density, the intensity of the resulting RGB color will be high if the densities are similar, and low otherwise. Since the RGB color vector is not encrypted, the multiplication between the encrypted dot product result and the RGB color vector can be archived by three separate homomorphic multiplications (⊗) of one encrypted and one plaintext vector by the Paillier encryption algorithm (Supplementary Material). However, the HSV-based encoding leads to results that are created by our encrypted density emphasizing approach. The volume density values are encoded with 4-dimensional vectors.

There are other and possibly better density-to-vector encoding schemes. However, the HSV-based encoding leads to results that feel natural, especially while smoothly increasing or decreasing the emphasizing density. The encoding scheme should in any case be chosen in such a way that the curve created by the dot product is steep and narrow (see dashed lines in Figure 5), so that the density selected by the user can be seen as clearly as possible in the resulting image. The Algorithm 3 takes not only the density that should be encoded as parameter, but also the count of dimensions of the returned vector. Increasing the count of dimension not only makes the dot product response curve more steep (See Figure 5 and compare the dashed lines in the left and right plot.), but also increases the required storage size of the encoded and encrypted volume dataset. Note that the count of dimensions must be the same during the encryption of the volume and for the encoding of the user-defined emphasizing density. This also means that the amount of computations required for the volume rendering depends on the number of dimensions used for encoding the volume.

### Algorithm 3: Encode Density

**Parameters**: The normalized density that should be encoded as a vector with dim dimensions.

**Result**: Vector \( v \)

```python
1 procedure encodeDensity(density, dim)
2     initialize vector \( v \) with length \( dim \) and set all indices to 0
3     \( s = \text{density} \cdot 2 \cdot (\text{dim} - 1) \)
4     \( f = ([s] + 1)/2 \)
5     \( d = [f] \)
6     \( v[d] = 1 \)
7     if \( d > 0 \) and \( f = d \) then
8         \( v[d - 1] = 1 - ((s + f) - 1) \)
9     else if \( d + 1 < \text{dim} \) and \( d < f \) then
10        \( v[d + 1] = s - [x] \)
11     return \( \text{normalize}(v) \)
```

Fig. 5. Visualization of density encoded as 3 dimensional (left) and 6 dimensional (right) vectors. The scalar value (density) of the voxel is represented on the x-axes. The magnitude of each vector component at a specific density is represented by the curves. The first component is drawn in red, the second in green, red, purple, olive and light blue. The dashed curves show the result of the dot product between the encoded voxel value and a TF-Node vector for a density of 0.45 in cyan and a density of 0.85 in orange.
The vectors $\vec{v}$ and $\vec{i}$ are the encoded density and RGB color of the TF-Node with index $i$. The symbol $\odot$ is used as operator for a dot product between one encrypted vector and one plaintext vector. $\vec{c}$ denotes the count of user defined function node (TF-Node) pair, but also for multiple such pairs. For a better understanding, we show the pair consisting of a density and an RGB color a transfer function node (TF-Node).

$\vec{c}_{i} = \bigoplus_{i=0}^{n} (\vec{v} \odot \vec{d}_{i}) \odot \vec{c}_{i}$  \hspace{2cm} (5)

To obtain the final encrypted RGB color of a pixel, the sum of all encrypted RGB sample values $[\vec{c}_{i}]$ along a viewing ray needs to be calculated. The total RGB vector needs to be divided by the sample count as usual for averaging and, furthermore, by the count of TF-Nodes. The method to approximate a division of an encrypted number is stated in Algorithm 5. After calculating this for every image pixel, the entire encrypted image is sent to the client. A client that knows the right secure key can now decrypt each RGB component of each pixel and display the colored image. Example images rendered with this approach are shown in Figure 1 and Figure 6.

6 Results

All performance tests are executed on a Mac Book Pro (15-inch, 2016) with a 2.9 GHz Intel Core i7. All algorithms are implemented in Java and are only single-threaded. The purpose of the implementation is to prove the concept and, in its current form, is not performance-optimized. All runtimes shown in Table 1 and Table 3 are measured with volume size of $100 \times 100 \times 100$ voxels. The rendered image always has a size of $150 \times 150$ pixels.

Table 1 shows the runtime performance required for encrypting a volume with scalar voxel values, X-ray rendering, and image decryption with different public key modulus lengths. The table is divided into four groups of rows. The first two groups show the required time for rendering with nearest-neighbor sampling. Group three and four show the result for the number generation ($r$) that is required for the obfuscation and the calculation of $r^{x}$ (see Supplementary Material). Algorithm 5 has a substantial impact on the time required for encrypting the volume dataset. We use the java.security.SecureRandom class from the java standard runtime framework as random number generator for the obfuscation.

Table 2 shows the required memory size for this volume with a single scalar value per voxel and also for encodings in multiple dimensions at different modulus lengths. Table 3 shows the runtime required for encrypting a volume with different voxel encodings (two, three, and four dimensions), rendering with our simplified transfer function approach at different counts of TF-Nodes (one, two, ... colors) and image decryption. The resulting performance for all these operations is provided for different public key modulus lengths.

The rendering results of Figure 1 show what can be done with our simplified transfer function. The right image demonstrates the utilization in nuclear medicine. During the diagnosis, these datasets are usually investigated either by showing single slices or by X-ray renderings, where the depth cues are provided through rotating the dataset around an axis. This is possible with our homomorphic-encrypted volume rendering with the added privacy, which is useful for diagnosing from such a highly sensitive type of modality and associated pathologies.

7 Discussion

First, we discuss possible performance improvements of our prototype and how the approach could scale to interactive frame rates for larger real-world datasets. Later, starting with general noteworthy considerations, we discuss security-related aspects of our volume rendering approach. Finally, we follow with an explanation for the invisibility of comparisons. In Section 7.4, we will show that the used floating-point encoding with an encrypted mantissa and a plaintext exponent does not weaken the privacy of the encrypted volume data.
Table 3. Simplified transfer function: required time (in seconds) for encryption, rendering and decryption with different modulus lengths.

|                | plain | 128bit | 256bit | 512bit | 1024bit | 2048bit |
|----------------|-------|--------|--------|--------|---------|---------|
| encrypt        | 29.87 | 115.03 | 622.77 | 4405.69| 31796.00|
| decrypt        | 0.05  | 21.85  | 48.59  | 142.36 | 544.35  | 1999.51 |
| render         | 0.80  | 1.37   | 4.95   | 29.7   | 180.40  |
| encrypt        | 0.06  | 23.38  | 58.20  | 131.10 | 659.48  | 2716.51 |
| decrypt        | 0.33  | 1.18   | 3.23   | 27.09  | 187.51  |
| render         | 0.06  | 54.97  | 213.21 | 1199.68| 8512.23 | 62959.05|
| encrypt        | 40.30 | 162.22 | 897.28 | 6271.48| 47030.98|
| decrypt        | 0.59  | 1.21   | 4.83   | 25.46  | 181.86  |
| render         | 0.53  | 0.83   | 4.47   | 25.98  | 127.84  |
| encrypt        | 0.05  | 4.16   | 8.73   | 257.51 | 919.41  | 3685.33 |
| decrypt        | 0.59  | 1.21   | 4.83   | 25.46  | 181.86  |
| render         | 0.06  | 23.38  | 58.20  | 131.10 | 659.48  | 2716.51 |
| encrypt        | 54.97 | 213.21 | 1199.68| 8512.23| 62959.05|
| decrypt        | 0.59  | 1.21   | 4.58   | 25.46  | 181.86  |
| render         | 0.06  | 41.67  | 89.73  | 257.51 | 919.41  | 3685.33 |
| encrypt        | 19.46 | 42.93  | 98.08  | 429.30 | 2292.48 |
| decrypt        | 0.28  | 0.77   | 5.4   | 15.80  | 60.03   |
| render         | 0.07  | 35.41  | 61.60  | 179.45 | 729.11  | 2850.69 |
| encrypt        | 0.46  | 0.76   | 3.10   | 18.36  | 121.76  |
| decrypt        | 0.07  | 46.41  | 93.63  | 268.59 | 1045.56 | 4251.69 |
| render         | 0.45  | 1.11   | 4.52   | 26.15  | 190.63  |
| encrypt        | 0.08  | 63.57  | 123.32 | 343.20 | 1320.86 | 5217.38 |
| decrypt        | 0.67  | 1.51   | 4.68   | 26.82  | 190.49  |
| render         | 0.06  | 23.38  | 58.20  | 131.10 | 659.48  | 2716.51 |

7.3 Encrypted Comparison Operators

It is not possible to compare encrypted numbers with each other. During the encryption of a number, the obfuscation is performed, which randomly distributes the encrypted values between 0 and $2^N - 1$. Therefore, the order of the encrypted values $[M]$ has nothing to do with the order of the underlying numbers $M$ that were encrypted. Consequently, operators such as smaller than ($<$) or greater than ($>$) cannot provide a result that is meaningful for the numbers $M$, if they are applied to encrypted values $[M]$.

We can also argue that comparison operators cannot exist if the Paillier cryptosystem is secure, since the existence of a comparison operator would break the security of the cryptosystem. Consider a less than comparison for example: if such a comparator could be implemented, every value could be decrypted within $\log_2(N)$ comparisons by a binary search. For a modulus $N$ with a length of 2048 bit, an attacker would need to encrypt and then compare only $\log_2(2^{2048}) = 2048$ numbers with the encrypted value $[m]$ in order to find the decrypted number $m$. This would effectively break the security of the encryption scheme.

7.4 Plaintext Exponent Does Not Leak Private Data

At first glance, it may look like the floating-point representation (encrypted mantissa, plaintext exponent) we used will allow an attacker to obtain more important information than within an encoding where all number components are encrypted. However, if it is implemented correctly, an attacker cannot take any advantage from this number representation. First, we will discuss this for the data in the server memory and, in the last paragraph, we will show how the exponent can be protected during the data transfer from the server to the client.

For the following, we will suppose a secure system with at least 2048-bit long modulus $N$ and, therefore, a mantissa $[m]$ with at least 600 decimal digits usable in the plaintext domain. Voxel values that are stored as 10 bit values are probably precise enough for most volume-rendering use cases. To store numbers between 0 and $2^{10} = 1024$, the exponent $e$ is not required at all, because the voxel information can be stored only in the mantissa $m$. Therefore, the exponent $e$ can be 1 for all voxels. This means that the exponent does not even have to be transferred to the server, because the server can implicitly assume that the exponents of all numbers is 1. An addition of any of these numbers that have an exponent of 1 does not change the exponent, because for an addition, the exponent needs to be taken into account only if the summands have different exponents (see Algorithm 1). Therefore, only a multiplication (e.g., an interpolation between voxel values) can change the exponent to anything other than 1. However, the Paillier cryptosystem only supports the multiplication of an encrypted number with an unencrypted number. Consequently, the number $d$ that changes an exponent has to be unencrypted. Furthermore, this number $d$ can only depend on unencrypted data, because Paillier does not support comparison operators (see Section 7.3), which are required for flow control statements like if or for-loops, and arithmetic operations with an encrypted number will result in useless random noise, except those add ($+$) and multiply ($\otimes$) that are defined for the Paillier cryptosystem. Therefore, the number $d$ can only be the result of some computation with other unencrypted variables. This implies that $d$ does not need to be encrypted, because everyone can calculate $d$ itself. In other words, if the variable $d$ can be computed from some variables that need to be considered as publicly available, because they are unencrypted, it is pointless to encrypt $d$. If $d$, which is unencrypted and can only depend
on unencrypted data, influences an exponent, the exponent exposes only the information that is already publicly available.

The important observation here is that an unencrypted value (e.g., an exponent) can influence an encrypted value (e.g., a mantissa), but an encrypted value (e.g., a mantissa) cannot influence an unencrypted value (e.g., an exponent). This means that no information that is only available as encrypted data can ever be exposed in unencrypted values like the exponent.

In our rendering system, a number \( d \) that changes an exponent can either be the result of a computation with a constant or with an unencrypted number that is provided in unencrypted form to the rendering system, such as the camera properties (position of eye point, opening angle, view direction...). Therefore, an attacker could possibly learn the constants used in our program code and data, such as the camera properties that are provided in the unencrypted form, from the exponents of the rendering result (the image). However, we want to develop an approach that is open and semantically secure by design and not secure through obscurity (compare: \([13,26,32,33]\)). Therefore, we have to treat the source code of the application as publicly available, which means that a constant cannot be considered to be private. Furthermore, for our approach, the camera properties need to be provided in an unencrypted form to the rendering system. Therefore, we cannot consider it as private anyway.

It should be noted that the camera properties could possibly provide interesting information to an attacker, because it could be possible to learn something about the volume data by tracking the camera properties over time. For instance, if a user rotates the camera around a specific region for a considerable amount of time, an attacker could guess that the region contains some interesting data. During the transfer of the camera properties from the client to the server over the network, the camera properties could be secured by using an encrypted tunnel, such as IPSec \([16]\) or TLS \([29]\). However, our basic assumption is that we cannot trust the server that hosts our rendering program. This means that an attacker has access to the entire memory of the server and, therefore, can read the camera properties directly from the memory of the server, regardless of the used network transfer method. While the unencrypted camera properties could indirectly expose some information, we will not discuss this further because it is beyond the scope of this work.

Based on the arguments stated in this section, we can conclude that using plaintext exponents for the rendering process on an untrusted computer system does not provide more information to a third party than using encryption for all components of a floating-point number.

The only remaining part that needs to be considered is the transfer of the final image from the server back to the client across a network. Operations like trilinear interpolation will change the exponents during the rendering. Therefore, the final image will contain floating-point numbers with exponents unequal to 1 and, because the interpolation weights that change the exponents depend on the camera properties, the exponents of the final image will provide some information about the camera properties. The privacy of the information that is stored in the exponents is only important if it can be assumed that the server is trustworthy, which contradicts the basic assumption of this work. Therefore, this is somewhat beyond the scope of this work, but we nonetheless discuss it here for the sake of completeness. In order to encrypt as much information as possible during the image transfer from the server to the client, ideally all information should be stored in the encrypted mantissa. While it is not possible to divide an encrypted number, it is possible to multiply an encrypted number. Furthermore, the encrypted mantissa can store numbers in the range from 0 to \(2^{2047}\). Therefore, it is possible to bring all exponents to the value of the smallest exponent of any pixel of the final image. This can be achieved by the calculation shown in Equation 2. For the new exponent \(e_n\), the value of the smallest exponent of any pixel must be used. If this exponent-decrease operation is applied to all image values on the server before transferring the image to the client, the exponent should not contain any important information during the transfer, because all exponents then contain the same value. However, if there is concern that even this might contain something useful, it is possible to encrypt this exponent with the public key because the client that has the secure key can decrypt it anyway. Since it is the same value for every number that is sent back to the client, this exponent needs to be sent and decrypted only once.

8 Conclusions

While the expressiveness of our renderings is far from what is possible with state-of-the-art algorithms for non-encrypted data, we have presented a highly parallelizable direct volume rendering approach that allows not only the outsourcing of the storage of the volume data, but also the outsourcing of the whole rendering pipeline, without compromising the privacy of the data. The approach we propose does not leak any voxel values or any information computed from a voxel value after the volume encryption. Since we encrypt every single bit of voxel data with Paillier’s cryptosystem, which is provably semantically secure (see: \([25,37]\)), it is rather obvious that with our approach, the confidentiality of the volume data (densities, shapes, structures...) and the colors of the rendered image only depends on the privacy of the secure key. If we trust all devices that have seen the volume data before encryption (e.g., MRI-/CT-scanner, the computer that performs the encryption) to safely delete the data after encryption, only the owner of the secure key is able to obtain any useful information of the encrypted volume or rendered images. This is a significant advantage compared to all previous works to date.

This security naturally comes with associated costs. The storage overhead costs for computation are between four and five orders of magnitude compared to plaintext data. Compared to our prototype, an optimized implementation of our approach can reduce the computational complexity by an order of magnitude.

While we hope that further improvements of our approach would lead to rendering results with better expressiveness, it will be a non-trivial task because the security aspect needs to be considered for even the smallest change. Many of the ideas we considered in the algorithmic design eventually led to a leak of sensitive information, which, in our opinion, intolerable, no matter how small it may be. Future work definatively needs to improve the rendering performance. We see that the performance can be tremendously accelerated, as ray-casting is an embarrassingly parallel workload. For practical utilization of our privacy-preserving volume rendering, an efficient GPU-based implementation would be necessary. A single server full with GPUs should be able to provide five orders of magnitude more computational power than a single CPU core can. Based on the measured performance with a non-optimized single threaded implementation, such a server could be able to achieve interactive frame rates for datasets that are small enough to fit into the memory of the graphics cards. Therefore, we see, as a next step, to port the rendering onto GPUs, where the necessary technological piece will be to design efficient big-integer arithmetic. Another possible improvement within the scope of Paillier HE will be the visual quality of compositing. This can be done with gradient-magnitude opacity modulation, where the gradient magnitude will be pre-calculated and encrypted along with the data values. Such representation can already lead to substantial visual quality improvement, although it will still not reach the outcome of compositing using Porter/Duff’s over operator \([27]\). For the Paillier HE scheme, we do not see a way to implement the over operator compositing, as it requires a multiplication of encrypted numbers. To support alpha blending, new research should be oriented on investigating other homomorphic encryption schemes or a combination of those that, unlike Paillier, would support desired secure alpha blending functionality.

Acknowledgments

The authors wish to thank Michal Hojsk for his fruitful discussions on cryptography. The authors would like to thank Michael Cusack from Publication Services at KAUST for proofreading. The research was supported by King Abdullah University of Science and Technology (KAUST) under award number BAS/1/1680-01-01.
Homomorphic Encryption schemes allow computations on encrypted data such that the decrypted results are equal to the result of a mathematical or logical operation applied to the corresponding plaintext data. Therefore, calculations on encrypted data can be performed without decrypting the data first. This property makes it possible to outsource not only the storage of encrypted data, but also the computation on sensitive data to untrusted third parties (e.g., the cloud). The computation on cloud servers has two major advantages. If only the result and not the whole dataset needs to be transferred to the client, a lower network bandwidth is required. Furthermore, the client can be a thin client like a tablet without much computing and storage resources because the computational expensive rendering is done on the server. Homomorphic encryption schemes are classified into three categories:

- partially homomorphic encryption (PHE): are homomorphic with regard to only one type of operation (addition or multiplication)
- somewhat homomorphic encryption (SHE): Can perform more general calculations than PHE, but only a limited number of them.
- fully homomorphic encryption (FHE): Can perform any computation on encrypted data.

Rivest et al. [30] invented the idea of Homomorphic Encryption in 1978. They showed the demand of a secure Homomorphic Encryption scheme that supports a large set of operations. However, it took more than 30 years until the first proposal of such a FHE was made by C. Gentry in 2009 [9]. While the first FHE schemes were just concepts, due to an ongoing development and optimization process, they are currently at least efficient enough for functional implementations.

However, FHE schemes are still not practical for real applications, because the storage and computational costs are too high [1],[4],[21]. Therefore, in our approach, we propose to take advantage of the Paillier PHE scheme, whose homomorphic properties that are relevant for our encrypted volume rendering will be introduced next.

### A.1 The Paillier Cryptosystem

Paillier’s cryptosystem [24] is an additive PHE scheme. The important property of this scheme is that a multiplication of two encrypted numbers is equivalent to an addition in the plaintext domain. This means that it is possible to calculate the sum of plaintext numbers that have been encrypted by multiplying the encrypted numbers. This relation is stated in Equation 6. For Equation 6 and Equation 7 we adopt the notation by Ziad et al. [38]. The ⊕ symbol is used for the operation on encrypted numbers that is equivalent to an addition on plaintext numbers. The encrypted version of the value $m$ is denoted as $[m]$ and $\text{Dec}([m]) = m$ means decrypting $[m]$ to $m$ by the decryption function of Paillier’s cryptosystem (see: Algorithm 6).

Paillier works on finite fields and thus performs a modulo ($\mod N$) operation after each multiplication ensures to stay in the field and does not change the decrypted result because the corresponding plaintext numbers need to be less than the modulus $N$ for a correct decryption anyway. However, the modulo operation makes further calculations more efficient because it prevents unnecessary large numbers ($n$ multiplications will blow up the number length by about $n$ times).

\[
\text{Dec}([m_1] \oplus [m_2]) = \text{Dec}([m_1] \times [m_2]) \mod N^2 \\
= (m_1 + m_2) \mod N
\]  

(6)

Since it is possible to add encrypted numbers to each other, in a for-100p, an encrypted number can be added to itself $d$ times to simulate a multiplication with $d$. However, note that the value $d$ is in plaintext. Since addition is performed by doing multiplication on the encrypted numbers, we can, instead of a for-100p, take the encrypted value to the power of $d$ to get this result, which can be implemented more efficiently than a for-100p. Furthermore, it has the advantage that it also works for $d < 0$. The case $d = -1$ is of special interest, because this makes subtraction of two encrypted numbers possible

$$\text{Dec}([m_1] - [m_2]) = \text{Dec}([m_1] \times [m_2]^{-1}) \mod N^2)$$. The symbol $\otimes$ is used for such an operation on one plaintext number $d$ and one encrypted number $[m_1]$. Equation 7 shows how to calculate this multiplication with one encrypted number.

\[
\text{Dec}([m_1] \otimes d) = \text{Dec}([m_1]^d \mod N^2) \\
= (m_1 \times d) \mod N
\]  

(7)

While the Paillier PHE supports an efficient method to multiply an encrypted and a plaintext number, it does not support the multiplication of two encrypted numbers and is, therefore, not a fully homomorphic encryption scheme.

The Paillier HE is a probabilistic asymmetric encryption scheme like the well-known RSA scheme [31]. That means that an encryption can be performed by using the public key, which is derived from the secure (or private) key. However, for the task of decryption, the secure key is required, which cannot (efficiently) be calculated from the public key because this would require factoring a product of two large prime numbers. It is essential for the security of Paillier’s cryptosystem (and also for RSA) that there is no known fast method for integer factorization of a product of two large prime numbers. The Algorithm 4 shows an example of a key generation function for the Paillier cryptosystem, which sets the generator $g$ always to $N + 1$, as this allows a more efficient encryption function (see: [8]). Furthermore, the algorithm clearly shows that the secure key, which needs to be kept secret, contains the two large prime numbers $p$ and $q$ (e.g., 1024 bit long) and the public key contains the product $N$ (e.g., 2048 bit long) of these two prime numbers.

### A.2 Key Generation Function

Algorithm 5 shows the pseudocode for the encryption and decrypt routines for the Paillier HE. The encryption algorithm also contains the obfuscating of an encrypted number with a random number $r$, which qualifies the Paillier HE as a probabilistic encryption scheme.

This means that a specific plaintext message $m$ can be represented by many possible ciphertexts $[m_1], [m_2], [m_3], \ldots, [m_r]$. The decryption with the right secure key will return the original message $m$ for all the possible ciphertext representations. While this is not required for correct homomorphic calculations with Paillier (imagine $r = 1$), it is important for the semantic security against chosen-plaintext attacks (IND-CPA) that Paillier’s cryptosystem provides [24][37]. Without this obfuscating, it would be possible to decrypt datasets without knowing the secure key because an attacker would only need to encrypt all possible plaintext values with the public key, store the plaintext and the corresponding ciphertext values in pairs, and compare the values of the encrypted dataset with the self-encrypted values for which the correct
Algorithm 6: Paillier Decrypt

Parameters: The ciphertext $c$ that should be decrypted with the secure key $(p, q)$ and the public key $(N, g)$.

Result: Plaintext $m$

```
1 procedure decrypt(c, p, q, N, g)
   // L- and H-function as defined by Paillier [24] at section 7
2    procedure L(x, y)
3       return $\frac{x}{y}$
4    procedure H(x)
5       return $L(g^{x-1} \mod x^2, x)^{-1} \mod x$
6    procedure C(m_p, m_q)
7       // Chinese Remainder Theorem
8       return $m_p + ((m_q - m_p) \cdot (p^{-1} \mod q) \mod q) \cdot p$
9       $h_p = H(p)$ // can be pre computed
10      $h_q = H(q)$ // can be pre computed
11      $m_p = L(c^{p-1} \mod p^2, p) \cdot h_p \mod p$
12      $m_q = L(c^{q-1} \mod q^2, q) \cdot h_q \mod q$
13      $m = C(m_p, m_q)$
14
```

decryption is known. For datasets with a limited number of possible values like the voxel values of a volume, which usually contains no more than $2^{10} = 1024$ different values, this would be a trivial task.