Official Journal of CIOMP
elight.springeropen.com

RESEARCH ARTICLE

Open Access

Image encryption using spatial nonlinear optics

Junfeng Hou[1] and Guohai Situ[1,2*]

Abstract

Optical technologies have been widely used in information security owing to its parallel and high-speed processing capability. However, the most critical problem with current optical encryption techniques is that the cyphertext is linearly related with the plaintext, leading to the possibility that one can crack the system by solving a set of linear equations with only two cyphertext from the same encryption machine. Many efforts have been taken in the last decade to resolve the linearity issue, but none of these offers a true nonlinear solution. Inspired by the recent advance in spatial nonlinear optics, here we demonstrate a true nonlinear optical encryption technique. We show that, owing to the self-phase modulation effect of the photorefractive crystal, the proposed nonlinear optical image encryption technique is robust against the known plaintext attack based on phase retrieval. This opens up a new avenue for optical encryption in the spatial nonlinear domain.

Keywords: Optical encryption, Spatial nonlinear optics, Self-defocusing

1 Introduction

Light, as a carrier of information, poccesses a number of unique features that can be processed to secure information. For example, the diffraction and temporal spectrum of light can be incooperated with optical variable devices, providing important security features for the anti-counterfeit of valuable documents and credit cards [1]. The scattering of light by a volumetric random material can form a unique fingerprint that may be used as a physical unclonable function [2, 3]. On another level, one can engineer the phase of the light field in a random manner by using a random phase mask (or, key), scrambling the information it carries in both the spatial and the spatial frequency domains by using a coherent optical information system. With a proper way to compensate the scrambling introduced to the phase, one can recover the information carried by the light. Based on this principle and taking the advantage of the ultra-high bandwidth and capability of coherent optical systems [4], researchers have developed various optical systems for security verification authentication [5, 6] and image encryption [7–9]. Owing to the ultra-large key space, it is unlikely to find the random phase mask using brute force attack within a reasonable time, making it promising for secure optical storage [10–13] and many others [14, 15].

Under the framework of the classical double random-phase encoding, the cyphertext $g(x, y)$ is related to the plaintext $f(x, y)$ as $g(x,y) = [f(x,y)R_1(x,y)] \otimes h(x,y)$, where $h(x, y)$ is the impulse response in the form of $h(x,y) = \mathcal{F}\{R_2(\mu,\nu)\}(x,y)$, where $R_1(x,y)$ and $R_2(\mu,\nu)$ stand for the random phase masks located at the input plane and the Fourier plane of the coherent 4f system, respectively, and $\mathcal{F}$ denotes the Fourier transform. It has been strictly proved [7] that the cyphertext $g(x, y)$ is stationary white noise provided that $R_1(x,y)$ and $R_2(\mu,\nu)$ are statistically independent uniform distributions in $[0, 2\pi]$. Therefore it is robust against blind deconvolution. However, recent studies have shown that such phase-encoding modality has certain security issues [16–21]. In particular, when someone manages to obtain information about the cyphertext and the corresponding plaintext,

*Correspondence: ghsitu@siom.ac.cn
[1] Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, 201800 Shanghai, China
Full list of author information is available at the end of the article

the cryptanalysis is reduced to nothing more than a phase retrieval problem, which can be solved by using, for example, projection-based optimization algorithms [22]. Although the presence of phase singularities prohibits the retrieval of the random phase key precisely, feasible solutions to the problem are sufficient to reveal information about the plaintext encrypted by the same set of keys [17–21]. It has been pointed out that this security issue stems from the linearity of the encryption modelity [18]. Thus, many efforts have been taken recently to resolve the linearity issue. Examples include the introduction of polarization encoding [23], photon-counting technique [24], computational ghost imaging (CGI) [25], coherence effect [26], and the bilinearity of phase-space distribution functions [27]. These methods rely on the increment of computation complexity, yet fail to offer a true nonlinear solution. As a consequence, they are also vulnerable to cryptanalysis. Indeed, recent studies have demonstrated the crypanalysis to the polarization [28] and CGI-based [29] encoding.

Thus, an essential revolution that is based on nonlinearity should be called for in order to address the security issue from the ground. Indeed, nonlinear optical effects have great potential in information security applications. For example, in the temporal domain, it has been demonstrated that the intensity fluctuation from a chaotic semiconductor laser can be adopted to generate random numbers in ultra-high speed [30, 31]. Unfortunately, this nonlinear technique cannot apply to image encryption in the spatial domain directly, although chaotic maps have been used for random phase encryption, mostly as keys to permuting [32] the plaintext. Thus it is both intuitively and practically important to develop spatial-nonlinear-optics-based image encryption techniques. Here we demonstrate one of such schemes. In the proposed scheme, the spatial nonlinearity is provided by using a Kerr-like crystal. Unlike security storage [10–13] where the nonlinear crystal is solely used to record the interference patterns, what it matters here is the screened photorefractive effect that offers a mechanism to mix the modes of the plaintext image when they propagate through the crystal. This is the most distinguishing feature of the proposed scheme in comparison to all the linear ones, in which the modes of the image propagate independently from the input to the output. As a consequence, the linear relation between the cyphertext and the plaintext is broken, making the proposed scheme resistant to all the existing optical cryptanalysis techniques in principle.

2 Results and discussion

2.1 Experimental setup

The basic experimental set-up is schematically shown in Fig. 1a. The plaintext image, $f(x_0, y_0)$, where (x_0, y_0) denotes the coordinates of the input plane, was displayed on an amplitude-only spatial light modulator (SLM-A). A 4-f imaging system was then used to project the plaintext image displaced on SLM-A to the input of the proposed encryption engine, which is depicted in Fig. 1b. The engine has a cascaded structure, each of which is composed of a phase-only SLM (SLM-P) and a photorefractive crystal [which was a $Sr_{0.61}Ba_{0.39}Nb_2O_6$ (SBN:61) in this study]. In our experiments, the first phase-only SLM ($SLM\text{-}P_1$) was placed on the conjugation plane of SLM-A so as to introduce a random phase, $R_0 = \exp[j\phi(x_0, y_0)]$, to the plaintext image, resulting in a random-phase-encoded image $\psi_0(x_0, y_0) = f(x_0, y_0)\exp[j\phi(x_0, y_0)]$, where the subscript 0 in ψ stands for the axial position

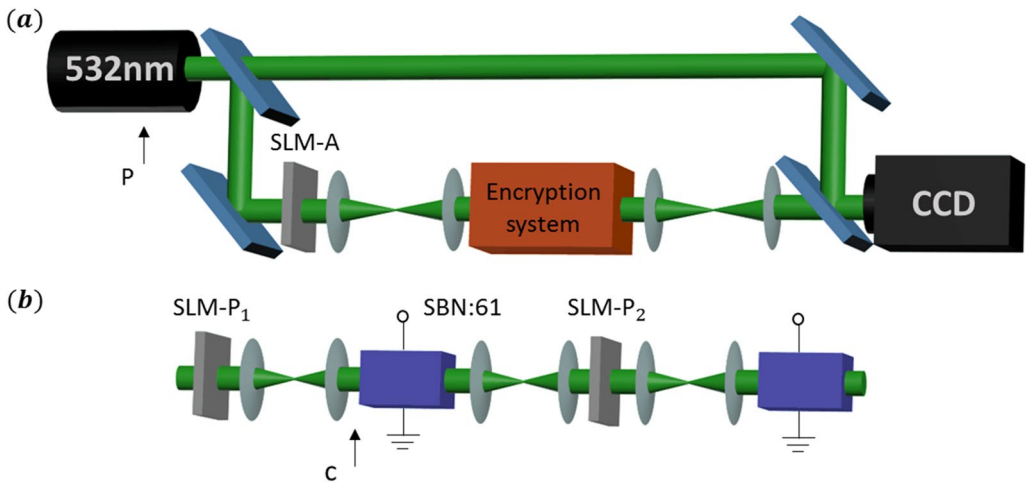

Fig. 1 **a** Schematic illustration of the optical experimental set-up. **b** Detailed illustration of the proposed nonlinear encryption engine

$z = 0$. This random-phase-encoded image $\psi_0(x_0, y_0)$ was then projected to the front surface of a SBN:61 crystal whose crystalline c-axis was perpendicular to the beam propagation direction. The complex wave field at the back surface of the SBN:61 crystal was then projected to SLM-P_2, the other phase-only SLM that was used to random-phase encode the incoming light field by $R_1 = \exp[j\varphi(x_1, y_1)]$ displaced on it. The resulting complex image is called the cyphertext image, written as $g(x, y)$ for convenience. This complex cyphertext image was recorded holographically by interfering with an additional reference beam, usually a plane wave with a known carrier frequency, as shown in Fig. 1a. For the SBN crystal, we used the self-defocusing nonlinearity, which is evoked by applying an external negative electric field E along the c-axis. Technically, this responses to the change of refractive index of $\delta n \propto r_{33} E\bar{I}/(1+\bar{I})$, where $\bar{I}$ is the input intensity $|\psi_0(x,y)|^2$ measured relative to a background (dark current) intensity [33], and $r_{33} = 255$ pm/V is the electro-optic coefficient relative to the applied field E and the c-axis [34].

2.2 Theory and experimental results

In the experimental demonstration, we used a binary image shown in Fig. 2a as the plaintext for simplicity. The direct image of it through our experimental set-up is shown in Fig. 2b. It was taken by the camera when displaying the binary image shown in Fig. 2a on the SLM-A while the other two SLMs-P and the external electric field E switching off. The distortion exhibits in the image was mainly due to the imperfection of the crystal and the aberration of the imaging optic. More careful alignment of the optic did not make significant improvement in our experiments. Nevertheless, we take it as the ground truth plaintext image in our proof-of-principle demonstration. To encrypt the plaintext image, we displayed two statistically independent random phases on the two phase-only SLMs, and turned on the nonlinearity. Mathematically, this nonlinear encryption process can be written as

$$\begin{aligned} g(x,y) &= T\{\exp[j\varphi(x_1,y_1)]\psi_{z_1}(x_1,y_1); z_2\} \\ &= T\{\exp[j\varphi(x_1,y_1)]T\{\psi_0(x_0,y_0); z_1\}; z_2\}, \end{aligned} \tag{1}$$

where the transform $T\{\cdot\}$ is defined as the nonlinear Schrödinger transform whose integral form is given by [35]

$$\psi_z(x,y) = \mathrm{FST}\{\psi_0(x_0,y_0); z\} - j\int_0^z U(z-z')\delta n(|\psi_{z'}(x',y')|^2)\psi_{z'}(x',y')\mathrm{d}z', \tag{2}$$

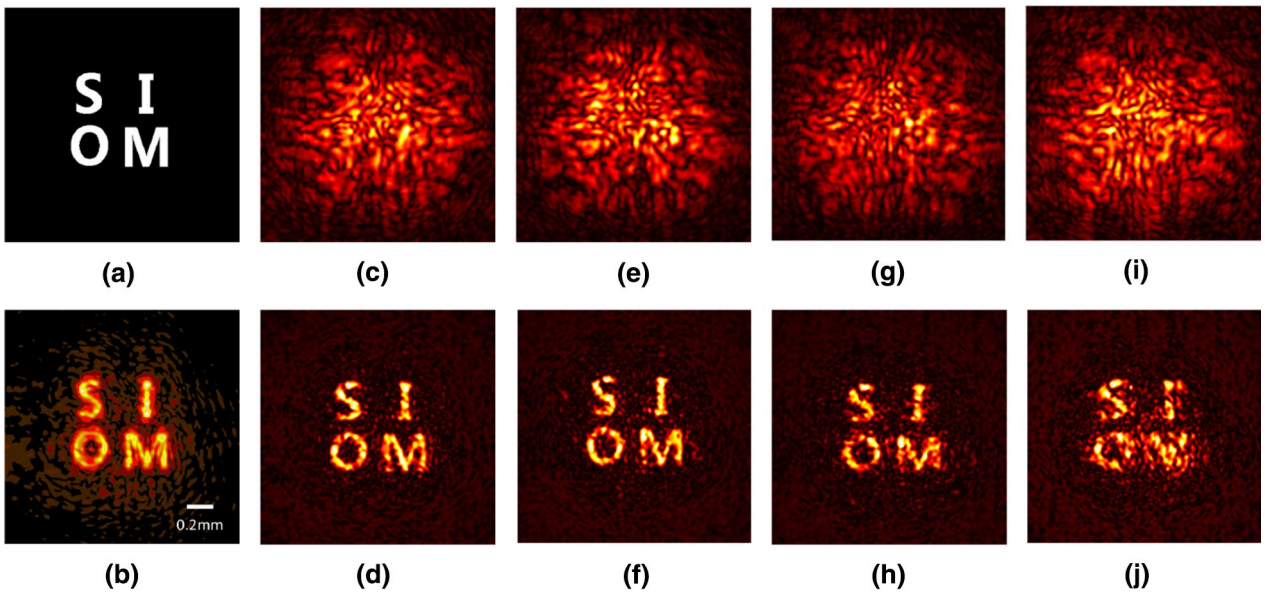

Fig. 2 Main experimental results. **a** the plaintext and **b** its direct image, which serves as the ground truth. The cyphertext images at $E = 0\,\mathrm{Vcm}^{-1}$ (c), $-500\,\mathrm{Vcm}^{-1}$ (**e**), $-1000\,\mathrm{Vcm}^{-1}$ (**g**) and $-2000\,\mathrm{Vcm}^{-1}$ (**i**), and **d**, **f**, **h**, **j** the corresponding decrypted image with the correct keys. All the images are normalized and plotted with the colormap hot

where FST$\{\psi_0(x_0, y_0); z\}$ denotes the linear propagation of $\psi_0(x_0, y_0)$ within the crystal with the length of z, $U(z)$ is the free Schrödinger operator given by $U(z) \propto \exp[ik\Delta/z]$, where k is the wave vector and Δ denotes the transverse Laplacian, and $\delta n(|\psi_{z'}(x', y')|^2)$ is the index of refraction induced by the nonlinearity of the crystal at the plane z'. The nonlinear term in Eq. (2) suggests that the original changes to the beam will be accumulatively augmented upon propagation. As a consequence, the spatial modes of the beam evolve in a coupled manner even with the generation of new ones owing to the wave mixing process [36], rather than propagating independently as in a linear system [7, 8] that all the current techniques for optical image encryption are operating on. It is in this way that the proposed scheme can break the linearity.

As mentioned above, the cyphertext obtained in this way is a complex-valued image [the intensity of which is shown in Fig. 2c]. It should be recorded using interferometry-based techniques like digital holography [37]. This allows the encryption as the process described by Eq. (1) to be reversible provided that the nonlinear medium is fully characterized and the amplitude and phase of the cyphertext image $g(x, y)$ are known [36, 38]. Thus the plaintext can be reconstructed from the digital hologram of the cyphertext numerically, with the conjugations of the two random phase keys presented in the first places, respectively, to demodulate the random phase

$$f(x_0, y_0) = \exp[-j\phi(x_0, y_0)]T\{\exp[-j\varphi(x_1, y_1)]T\{g(x, y); -z_2\}; -z_1\} \tag{3}$$

The decrypted image with the correct keys is shown in Fig. 2d. This demonstrates that the numerical decryption can reverse the wave-mixing process and demodulate the random phase. Here the external voltage that applied across the c-axis of the SBN crystal was $E = -500\ \mathrm{Vcm}^{-1}$, and the geometric parameters $z_1 = 9.7$ mm and $z_2 = 8$ mm. We need to mention that only the first nonlinear transform was performed optically since only one SBN:61 crystal (with the size of $4.4 \times 4.4 \times 9.7\ \mathrm{mm}^3$) was used in this experiment because there is only one such crystal at hand. The second nonlinear transform was performed numerically. Note that various algorithms have been proposed for the numerical solution of the nonlinear Schrödinger equation [39]. Here we simply employed the split-step Fourier propagation method, which has been intensively used in the studies in nonlinear optics [34, 36, 40].

In the extreme case that the nonlinearity reduces to zero (no applied voltage across the crystal), i.e., the second term in Eq. (2) is absent, the system becomes a Fresnel-based system [8], except that it propagates in the crystal other than in free space. The intensities of the cyphertext and the decrypted image are plotted in Fig. 2c and d, respectively. One can see that the plaintext image can be recovered in the linear case is comparable to that in the nonlinear case shown in Fig. 2f. Both these two images are lightly distorted in comparison to the ground truth as the optics were not perfectly aligned in our proof-of-principle experiments, or the numerical reconstruction algorithm did not take the imperfection of the crystal into account. Further calibration of the algorithmic with respect to the experimental setting will help improve the reconstructed results [36]. Comparing to the linear counterpart, the nonlinearly encrypted cyphertext image is more obliterated by virtue to the nonlinear self-defocusing and light-induced scattering that arises from the augment of the beam scattered by the imperfection of the crystal [41]. Such difference in the intensity patterns has been observed in the case even without random phase modulation [36], and thus has the potential to add additional physical security features.

The plaintext image can be recovered even when the nonlinearity was further increased. However, it can be more seriously distorted because the light-induced scattering effect is stronger in this case. In Fig. 2h, we plot the reconstructed plaintext image when the external voltage was tuned to $E = -1000\ \mathrm{Vcm}^{-1}$. It is clearly seen that the noise is augmented as the nonlinearity increases, and thus the recovered image is distorted. It becomes severely when the external voltage goes up to $-2000\ \mathrm{Vcm}^{-1}$ (Fig. 2j), even all the keys are correctly presented. It is quite challenging to get rid of the light-induced scattering effect in the numerical decryption algorithm because it can be invoked by the imperfection anywhere inside the crystal [42] or even on its surface [43]. In addition, the actual nonlinear effects can be even more intriguing [44]. For example, wave-mixing in the self-defocusing crystal can induce focusing as well [45, 46]. But the numerical decryption algorithm at the current stage does not take them into account. Fortunately, these effects can be ignored when the nonliearity strength is not too strong as in our study. The experimental results confirms that the proposed encryption system works well at small nonlinearity as the light-induced scattering can be very weak in this case [47]. Indeed, in numerical simulation, the recovered plaintext images are perfectly identical with the ground truth regardless of the nonlinearity strength. Details can be found in the Appendix.

2.3 Toleration analysis

For an optical encryption system, it is important to analyze how the misalignment of the keys affects the

performance of decryption since it explicitly relies on the reversibility of the system. It is expectable that the decryption is sensitive to the alignment as otherwise the modulation cannot be feasibly undone. However, a certain level of toleration against misalignment is desirable for the sake of practice.

As described in Eq. (1), there are several keys to the proposed system: the random phases R_0 and R_1, their geometric positions in the system, and the nonlinearity. To perform the toleration analysis, we should make an assumption that the correct random phases R_0 and R_1 should be presented. Otherwise it is not possible to recover any meaningful image. This has been well studied in the linear counterpart [7]. One can expect that it will not become better in the nonlinear case. Thus we focus on the toleration to the misalignment of the random phases along the transverse and longitudinal directions and to the change of nonlinearity strength for decryption with respect to that for encryption. And we will examine these factors independently.

First, we analyze the toleration to the displacement of the random phase key along a transverse direction. To perform the analysis, one can first calculate the complex conjugation of a cyphertext image $g^*(x,y)$, and then numerically reverse the second nonlinear transform in Eq. 1a. The resulting complex disturbance can be written as $\exp[-j\varphi(x_1,y_1)]T\{\psi_0^*(x_0,y_0);-z_1\}$. If R_1 is not placed at its original position, but transversely translated over a distance Δx_1 along the x-axis, the demodulated image can be written as $\exp[j\varphi(x_1-\Delta x_1,y_1)]\exp[-j\varphi(x_1,y_1)]T\{\psi_0^*(x_0,y_0);-z_1\}$. This means that the random phase cannot be demodulated completely in this case. The residual phase distortion $\exp\{j[\varphi(x_1-\delta x_1,y_1)-\varphi(x_1,y_1)]\}$ invokes speckle noise, which is accumulatively augmented upon the nonlinear propagation through the crystal [42, 47]. As a result, the recovered plaintext image is corrupted by noise, the quaility of which can be evaluated by using some standard criterion indicator such as the normalized mean-squared error (NMSE). One can expect that the NMSE value monotonously increases along with Δx_1 from 0 to l_x, the correlation length of R_1. Indeed, we observed a linear relation between them as shown in Fig. 3a. In comparison with the linear counterpart [48], the proposed nonlinear encryption engine is more sensitive to the transverse translation of R_2, as one can see from the inset in Fig. 3a that the position mismatch of $l_x/2$ is sufficient to make the decrypted image totally corrupted.

The response to the axial translation of R_1 can be clearly seen by writing the decrypted image $\hat{f}(x_0,y_0)=\exp[-j\phi(x_0,y_0)]T\{\exp[-j\varphi(x_1,y_1)]T\{g(x,y);-z_2-\Delta z\};-z_1\}$ according to Eq. (3). The noise comes from the mismatch Δz, and is further augmented by the second nonlinear transform in the decryption process. Thus it is expected that the level of noise increases as Δz increases either in the + or − direction, as evidenced by the experimental results shown in Fig. 3b. But the NMSE value increases quickly from 0 to about 0.5 as $|\Delta z|$ increases from 0 to 4 mm, and then become steady as $|\Delta z|$ increases further. One can clearly see that the proposed engine is more tolerant to the misalignment of R_1 in the longitudinal direction in comparison to the transverse one. This is reasonable because the latter one is due to an effectively wrong random phase key.

The analysis of the toleration to the misalignment of R_0 is straightforward. The transverse misalignment of R_0 does not have any effect to the decrypted image when the

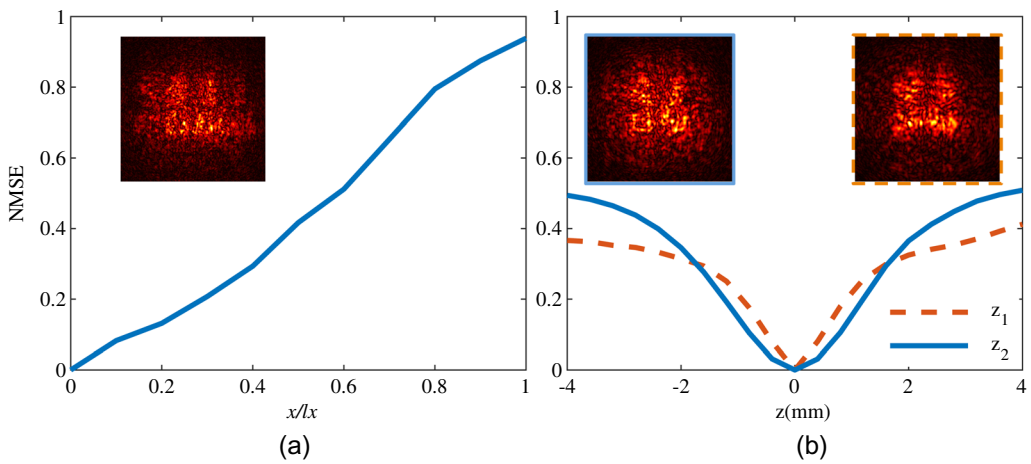

Fig. 3 Toleration to the misalignment of the random phase keys in the transverse (**a**) and longitudinal (**b**) directions

plaintext image $f(x_0, y_0)$ is real, as in our study. However, if R_0 is misaligned in the longitudinal direction, there will be residual random phase and give rise to noise effect. Because of the absence of further amplification, the decryption is more tolerant to R_0 than R_1, as depicted by Fig. 3b.

Next we examine the robustness to additive noise of the decrypted image. This is done by adding zero-mean Gaussian white noise to the cyphertext image so that $g'(x,y) = g^*(x,y) + \alpha n(x,y)$, where α is a weighting factor that specifies the strength of the noise with respect to the signal, and $n(x,y) \sim \mathcal{N}(0,\sigma)$, where σ is the standard deviation. Owing to the nonlinearity, the additive noise $n(x, y)$ is coupled with the signal term $g^*(x,y)$ on the way that $g'(x,y)$ is propagating back to the original input plane. The resulting noise on the recovered plaintext $\hat{f}'$ then is not additive anymore. The other immediate consequence of the nonlinear coupling is that the strength of the noise on $\hat{f}'$ is not linearly proportional to $n(x, y)$. Indeed, it has been reported that a portion of the noise power can be transferred to the signal [49]. As a result, the proposed nonlinear encryption technique should be more robust to noise, although the PSNR of $\hat{f}'$ should be a nonlinear function of $\mathrm{SNR} = -10\log_{10}|g^*|^2/(\alpha|n|)^2$ of g'. Indeed, we observed such a nonlinear dependence in our experiment (Fig. 4). The NMSE value decreases nonlinearly as the strength of $n(x, y)$ linearly increases. As an example, we plot in Fig. 4a–d the recovered plaintext when the SNR of $g'(x,y)$ is 10, 0, −10, and −20 dB, respectively. It is clearly seen that the detail of the plaintext retains even the SNR of $g'(x,y)$ is 0 dB. In contrast, linear dependence is expected in its linear counterpart [50] as additive noise on g^* is transformed to additive noise on $\hat{f}'$, and the power of noise conserved due to the canonical nature of this linear encryption system [51].

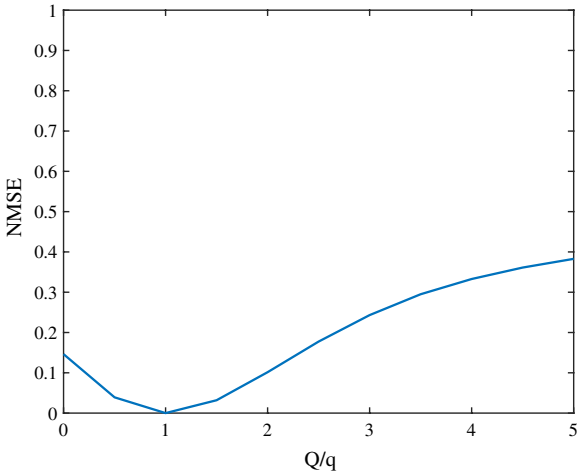

Fig. 5 Toleration to the mismatch of the nonlinearity strength in the decryption with respect to the encryption

We also examined how the decrypted image is affected by the deviation of the strength of nonlinearity (denoted by Q) alone for decryption with respect to that for encryption (denoted by q). Specifically, this can be seen by the change of the NMSE value with respect to Q/q. The result is plotted in Fig. 5. It suggests that the decryption is quite robust to the change of nonlinearity. The NMSE value is less than 0.1 when $Q/q = 2$, and is about 0.4 even when $Q/q = 5$. Even when the decryption is carried out with $Q = 0$, one can obtain a plaintext with acceptable quality (NMSE ≈ 0.15) if the two random phase masks and the length of the crystals are known. This is reasonable because of the fact that the nonlinear refractive index δn is four orders of magnitude smaller than the linear one [34, 36]. However, this does not mean the introduction of spatial nonlinearity is trivial. In fact, the nonlinearity does not mean to use in this way. It is

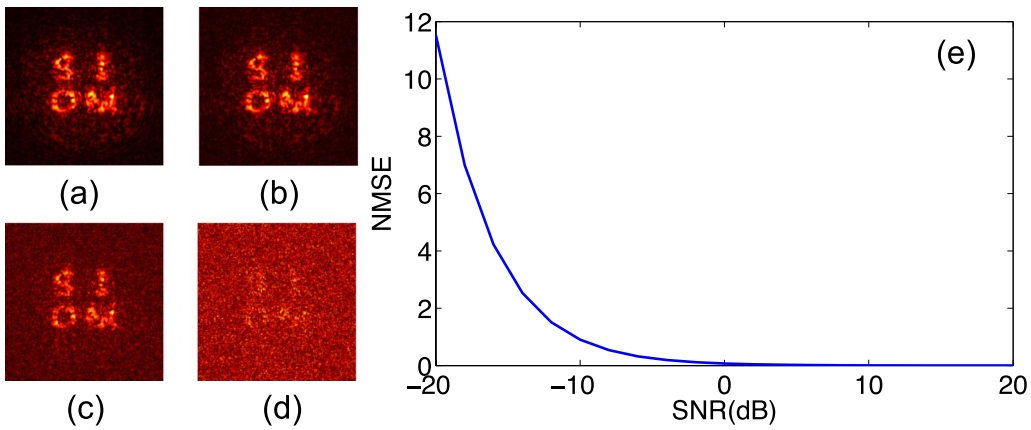

Fig. 4 Robustness to additive noise

used to protect the system from cryptalaysis when the random phase keys are unknown. We will show in Sect. that it has a significant impact to the enhancement of the security.

2.4 Security analysis

Most of the cryptanalysis techniques [17–20] rely on Kerckhoffs's principle [52] that an intruder is assumed to have full access to the cyphertext image $g(x, y)$ and/or the corresponding plaintext image $f(x_0, y_0)$. Thus, one more transform of $g(x, y)$ does not add significant intrinsic security. This is in particular true for a linear system, in which one can easily calculate the Fourier spectrum of the cyphertext $g(x, y)$ and subsequently recover the plaintext image by using phase retrieval algorithms owing to the memory effect [20]. Here we show that the proposed nonlinear encryption technique is immune to such phase-retrieval-based known-plaintext attack (KPA).

According to Kerckhoffs's principle [52], we are assumed to know M pairs of cyphertext–plaintext images, i.e., $[g_m(x,y), f_m(x_0,y_0)]$, where $m = 1, \ldots, M$. To examine the KPA, we also assume that the strength of nonlinearity and the length of the crystal z_1 and z_2 are known as well. It is straightforward to calculate $\psi_{z_1,m}(x_1,y_1)\exp[j\varphi(x_1,y_1)]$ from $g_m(x,y)$ using nonlinear digital holography [36]. Since the random phase $\mathrm{R}_1 = \exp[j\varphi(x_1,y_1)]$ is unknown, it is not possible to directly use digital holography to reconstruct $f_m(x_0,y_0)$ from $\psi_{z_1,m}(x_1,y_1)\exp[j\varphi(x_1,y_1)]$. Note that the random phase R_1 does not change the magnitude $|\psi_{z_1,m}(x_1,y_1)|$. Thus an alternative approach is to retrieve the unknown phases $\varphi(x_1,y_1)$ and therefore, $\phi(x_0,y_0)$, from $f_m(x_0,y_0)$ and $|\psi_{z_1,m}(x_1,y_1)|$. In contrast to the linear counterpart, a nonlinear phase retrieval algorithm is needed in this case [53, 54]. If such KPA successes, the retrieved phase, denoted as $\hat{\varphi}(x_1,y_1)$, should be used to decrypt any other cyphertext image, $g_t(x,y)$, encrypted by the same system and the same set of keys. For the cryptanalysis of a linear double random-phase encoding [7, 8], the multiple-phase retrieval algorithm [19] has been demonstrated to be implicitly feasible. Here we adopt the routine of this algorithm but replacing the linear canonical transform in [19] with the nonlinear Schrödinger transform to perform the attack.

Apparently, if nothing but a noise-like pattern is recovered, we can conclude that the proposed nonlinear encryption method is immune to the phase-retrieval-based KPA. This can be verified on experimental data. However, one may argue that this may attribute to the defect of the crystal or noise in the system as this may break the reversibility of the system [55]. Thus, we endeavor to examine the security via numerical experiments, which can be regarded as a fundamental baseline.

In the numerical study, we used $M = 4$ pairs of cyphertext-plaintext images to perform the aforementioned KPA. The 4 plaintext images are shown in Fig. 6a–d, and the corresponding cyphertext images are shown in Fig. 6e–h, respectively. The KPA algorithm attempts to recover the random phase $\varphi(x_1,y_1)$ and uses it to decode the cyphertext of an unknown plaintext image shown in Fig. 6i. The recovered random phase key $\hat{\varphi}(x_1,y_1)$ is shown in Fig. 6j. The difference between it and the original phase $\varphi(x_1,y_1)$ is shown in Fig. 6k. Its random-like distribution implies that the KPA algorithm [19], although has been demonstrated to be very efficient to attack a linear encryption system, is not able to retrieve the phase key of the proposed nonlinear encryption system. Indeed, nothing about the image to be analyzed (Fig. 6i) is revealed in the recovered image (Fig. 6l). Instead, it is some information about the known plaintext images that is revealed. In the specific case shown in Fig. 6l, it is a clear 'S' together with a dimmed 'M' against a noisy background that is recovered with $\hat{\varphi}(x_1,y_1)$. With a close look at the positions of S and M, it is not difficult to see that they appear at their original positions as in Fig. 6a, d as if they were memorized by the retrieved phase key $\hat{\varphi}(x_1,y_1)$. This exotic phenomenon is mainly due to the fact that phase evolution in a nonlinear optical system is significantly dependent on intensity-induced refractive index changes [36]. Although nonlinear refractive index is small comparing to the base one, it does have a significant impact to the enhancement of the security level, protecting it from the powerful KPA analysis. The appearance of which known plaintext image in the recovered image is determined by where the KPA algorithm stops. In the case of Fig. 6l, the KPA algorithm stops after the use of the known plaintext image 'S' (Fig. 6a) iteratively compute the phase key $\hat{\varphi}(x_1,y_1)$. And thus $\hat{\varphi}(x_1,y_1)$ *memorizes* clearly the information of the image 'S'. One iteration previous to this is the use of the image 'M'. And thus $\hat{\varphi}(x_1,y_1)$ still has a little memory of it. This phenomenon has never been observed in the linear counterparts [16–20] or in phase retrieval using nonlinear diversity [53, 54].

3 Conclusion

In conclusion, we have experimentally demonstrated for the first time a true nonlinear optical encryption method. In comparison with the conventional linear counterpart [7, 8] that relates the cyphertext and plaintext by using linear canonical transforms, the unique feature of the proposed method is the employment of the nonlinear Schrödinger transform. Thus it offers a way to resolve the security vulnerability mainly owing to the linearity [16–20]. Indeed, we have demonstrated that the proposed method is resistant to a phase-retrieval-based KPA, a

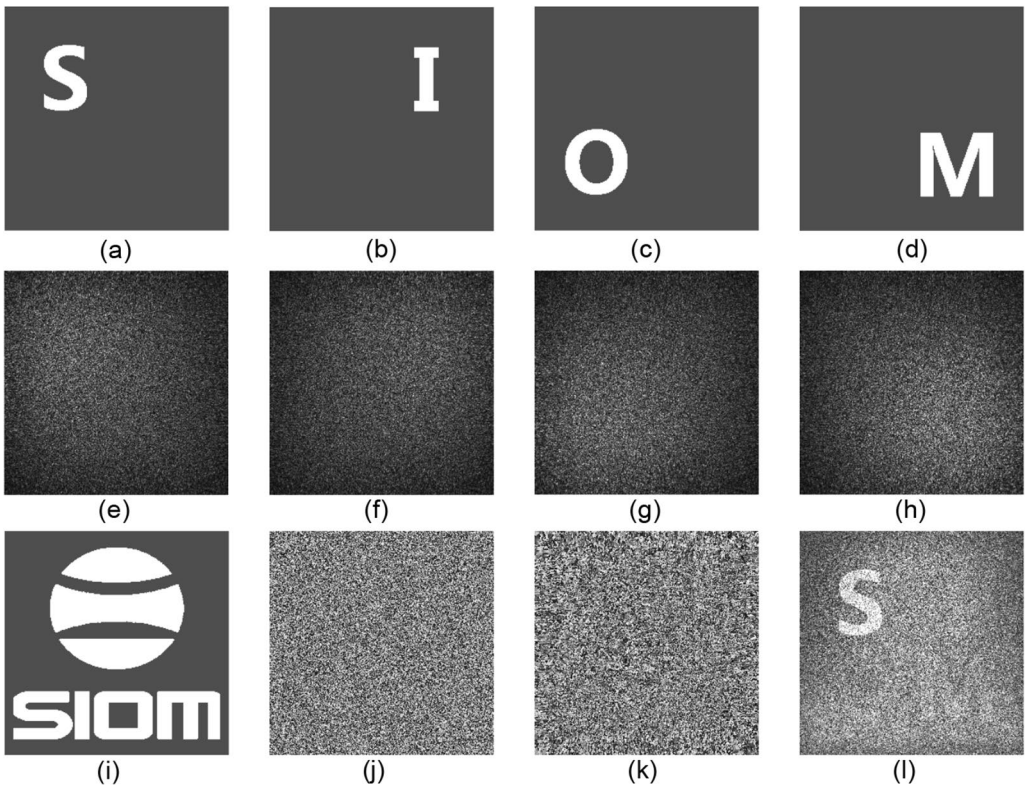

Fig. 6 Robustness against KPA. **a**–**d** The four known plaintext images and **e**–**h** the corresponding cyphertext images, **i** the plaintext image the KPA attempts to reveal, **j** the retrieved R_2, **k** the difference between the retrieved R_2 and the original one, and **l** the result of the KPA. One can clearly seen that the recovered image does not resemble the original plaintext at all

class of cryptanalysis methods that has been shown to be powerful to crack the conventional linear optical encryption engines [16–20]. This opens up a new avenue for optical image encryption in the spatial nonlinear domain.

One may argue that with the growing power of machine learning, the proposed nonlinear encryption method should be vulnerable to it. Indeed, we have demonstrated that the conventional double random-phase encryption is vulnerable to a deep-learning-based attack [21]. However, we believe that the proposed nonlinear encryption method is robust to it. A deep neural net requires a large set of plaintext–ciphertext pairs from the same system to train. But owing to the self-defocusing effect, the potential induced inside the crystal is plaintext-dependent (as manifested in Fig. 6). As a result, even though the plaintext-ciphertext pairs in the training set might have been obtained through the same crystal, the induced potential was effectively different for different plaintext in the training set. Thus one can imagine that different pairs of plaintext-ciphertext are associated with a different setting of the crystal so that it is infeasible to learn a common mapping function among them.

Appendix

In the simulation, the virtual nonlinear image encryption system is composed of two SBN:75 crystals with the length of $z_1 = 9.7$ cm and $z_2 = 8$ cm, respectively, as in our experimental counterpart. A plaintext image (as the one shown in Fig. 2a) together with a random phase key (implemented as a 2D array, each element of which is a pure phase that obeys a uniformed distribution between 0 and 2π) is virtually placed at the front surface of the first crystal. At the back surface, we virtually place the other random phase key that obey the same distribution function, but is statistically independent with respect to the first one. Because it is a virtual system, the second crystal can be placed immediately behind the second random phase key. We assume that the system is illuminated by a coherent plane wave with the wavelength of $\lambda = 532$ nm. The size of the images is 256×256 pixels. The nonlinear propagation of the laser beam is implemented by the split-step Fourier propagation method [36, 40]. The defects of the crystal and the noise of the image sensor are neglected in our simulation. The results are plotted in Fig. 7.

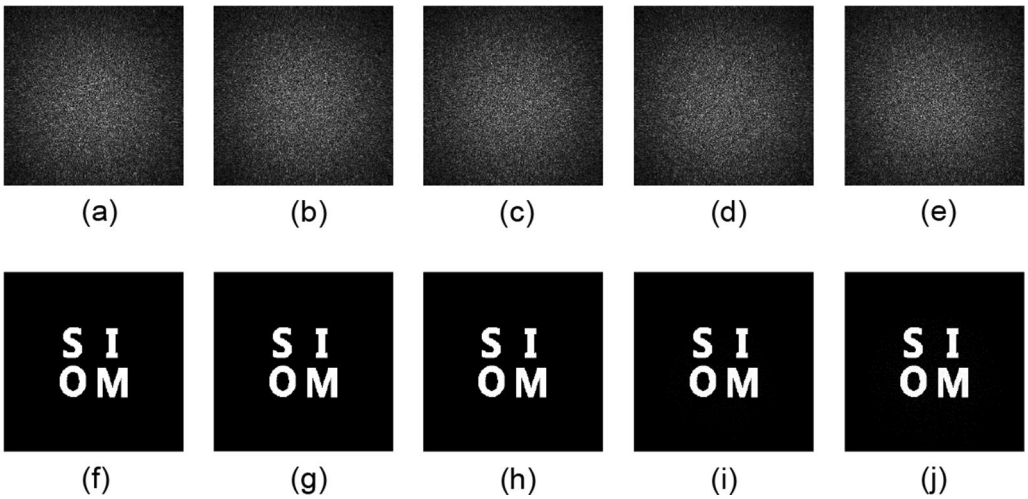

Fig. 7 Simulation results. The cyphertext image at **a** $E = 0$, **b** $-500\,\mathrm{V\,cm^{-1}}$, **c** $-1000\,\mathrm{Vcm^{-1}}$, **d** $-1500\,\mathrm{Vcm^{-1}}$ and **e** $-2000\,\mathrm{Vcm^{-1}}$, and **f**–**j**, the corresponding decrypted image when all the keys are correct

Acknowledgements
We'd like to thank Prof. Xinzheng Zhang with Nankai University for helpful discussion and the use of the SBN crystal.

Authors' contributions
GS conceives the concept; JH performed the experiments. JH and GS analyzed the data and contributed to the writing of the manuscript. GS supervised the project. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (61991452, 62061136005) and the Sino-German Center (GZ1391).

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare no conflict of interest.

Author details
[1]Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, 201800 Shanghai, China. [2]Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, 310024 Hangzhou, China.

Received: 7 April 2021 Revised: 4 December 2021 Accepted: 7 December 2021
Published online: 07 February 2022

References

1. R.L. Renesse, *Optical Document Security* (Artech House, New York, 2004)
2. R. Pappu, B. Recht, J. Taylor, N. Gershenfeld, Physical one-way functions. Science **297**, 2026–2030 (2002)
3. S.A. Goorden, M. Horstmann, A.P. Mosk, B. Škorić, P.W.H. Pinks, Quantum-secure authentication of a physical unclonable key. Optica **1**, 421–424 (2014)
4. D. Psaltis, Coherent optical information systems. Science **298**, 1359–1363 (2002)
5. B. Javidi, J.L. Horner, Optical pattern recognition for validation and security verification. Opt. Eng. **33**, 1752–1756 (1994)
6. B.L. Volodin, B. Kippelen, K. Meerholz, B. Javidi, N. Peyghambarian, A polymeric optical pattern-recognition system for security verification. Nature **383**, 58–60 (1996)
7. P. Refregier, B. Javidi, Optical image encryption based on input plane and Fourier plane random encoding. Optics Lett. **20**, 767–769 (1995)
8. G. Situ, J. Zhang, Double random-phase encoding in the Fresnel domain. Opt. Lett. **29**, 1584–1586 (2004)
9. X. Li, T.-H. Lan, C.-H. Tien, M. Gu, Three-dimensional orientation-unlimited polarization encryption by a single optically configured vectorial beam. Nat. Commun. **3**, 988–993 (2012)
10. O. Matoba, B. Javidi, Encrypted optical memory system using three dimensional keys in the Fresnel domain. Opt. Lett. **24**, 762–764 (1999)
11. O. Matoba, B. Javidi, Encrypted optical storage with wavelength key and random codes. Appl. Optics **38**, 6785–6790 (1999)
12. O. Matoba, B. Javidi, Encrypted optical storage with angular multiplexing. Appl. Optics **38**, 7288–7293 (1999)
13. X. Tan, O. Matoba, T. Shimura, K. Kuroda, B. Javidi, Secure optical storage using fully phase encryption. Appl. Optics **39**, 6689–6694 (2000)
14. B. Javidi, A. Carnicer, M. Yamaguchi, T. Nomura, E. Pérez-Cabré, M.S. Millán, N.K. Nishchal, R. Torroba, J.F. Barrera, W. He, X. Peng, A. Stern, Y. Rivenson, A. Alfalou, C. Brosseau, C. Guo, J.T. Sheridan, G. Situ, M. Naruse, T. Matsumoto, I. Juvells, E. Tajahuerce, J. Lancis, W. Chen, X. Chen, P.W.H. Pinkse, A.P. Mosk, A. Markman, Roadmap on optical security. J. Opt. **18**, 083001 (2016)
15. O. Matoba, T. Nomura, E. Perez-Cabre, M.S. Millan, B. Javidi, Optical techniques for information security. Proc. IEEE **97**, 1128–1148 (2009)
16. A. Carnicer, M. Montes-Usategui, S. Arcos, I. Juvells, Vulnerability to chosen-cyphertext attacks of optical encryption schemes based on double random phase keys. Optics Lett. **30**, 1644–1646 (2005)
17. X. Peng, P. Zhang, H. Wei, B. Yu, Known-plaintext attack on optical encryption based on double random phase keys. Optics Lett. **31**, 1044–1046 (2006)
18. Y. Frauel, A. Castro, T.J. Naughton, B. Javidi, Resistance of the double random phase encryption against various attacks. Optics Express **15**, 10253–10265 (2007)
19. G. Situ, U. Gopinathan, D.S. Monaghan, J.T. Sheridan, Cryptanalysis of optical security systems with significant output images. Appl. Opt. **46**, 5257–5262 (2007)
20. G. Li, W. Yang, D. Li, G. Situ, Cyphertext-only attack on the double random-phase encryption: Experimental demonstration. Opt. Express **25**(8), 8690–8697 (2017). https://doi.org/10.1364/oe.25.008690

21. M. Liao, S. Zheng, S. Pan, D. Lu, W. He, G. Situ, X. Peng, Deep-learning-based ciphertext-only attack on optical double random phase encryption. Opto-Electr. Adv. **4**, 200016 (2021)
22. Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, M. Segev, Phase retrieval with application to optical imaging: A contemporary overview. IEEE Signal Processing Magazine **32**(3), 87–109 (2015). https://doi.org/10.1109/msp.2014.2352673
23. A. Alfalou, C. Brosseau, Dual encryption scheme of images using polarized light. Optics Lett. **35**, 2185–2187 (2010)
24. M. Cho, B. Javidi, Three-dimensional photon counting double-random-phase encryption. Optics Lett. **38**, 3198–3201 (2013)
25. W. Chen, X. Chen, Ghost imaging for three-dimensional optical security. Appl. Phys. Lett. **103**, 221106 (2013)
26. D. Peng, Z. Huang, Y. Liu, Y. Chen, F. Wang, S.A. Ponomarenko, Y. Cai, Optical coherence encryption with structured random light. PhotoniX **2**, 6 (2021)
27. J. Liu, X. Xu, Q. Wu, J.T. Sheridan, G. Situ, Information encryption in phase space. Opt. Lett. **40**, 859–862 (2015)
28. L. Wang, Q. Wu, G. Situ, Chosen-plaintext attack on the double random polarization encryption. Opt. Express **27**, 32158–32167 (2019)
29. S. Yuan, L. Wang, X. Liu, X. Zhou, Forgery attack on optical encryption based on computational ghost imaging. Optics Lett. **45**, 3917–3920 (2020)
30. A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, P. Davis, Fast physical random bit generation with chaotic semiconductor lasers. Nat. Photonics **2**, 728–732 (2008)
31. I. Kanter, Y. Aviad, I. Reidler, E. Cohen, M. Rosenbluh, An optical ultrafast random bit generator. Nat. Photonics **4**, 58–61 (2010)
32. A.M. Elshamy, A.N.Z. Rashed, A.E.A. Mohamed, O.S. Faragalla, L. Mu, S.A. Alshebeili, F..E.. Abd El-Samie, Optical image encryption based on chaotic baker map and double random phase encoding. J. Lightwave Technol. **31**, 2533–2539 (2013)
33. M. Segev, B. Crosignani, P.D. Porto, A. Yariv, G. Duree, G. Salamo, E. Sharp, Stability of photorefractive spatial solitons. Optics Lett. **19**, 1296–1298 (1994)
34. G. Situ, J.W. Fleischer, Dynamics of the Berezinskii-Kosterlitz-Thouless transition in a photon fluid. Nat. Photon. **14**, 517–522 (2020)
35. J. Ginibre, G. Velo, On a class of nonlinear Schrödinger equations. I. the Cauchy problem, general case. J Funct. Anal. **32**, 1–32 (1979)
36. C. Barsi, W. Wan, J.W. Fleischer, Imaging through nonlinear media using digital holography. Nat. Photonics **3**, 211–215 (2009)
37. U. Schnars, W. Jüptner, *Digital Holography* (Springer, Heidelberg, 2005)
38. M. Tsang, D. Psaltis, F.G. Omenetto, Reverse propagation of femtosecond pulses in optical fibers. Optics Lett. **28**, 1873–1875 (2003)
39. G. Arora, V. Joshi, R.C. Mittal, Numerical simulation of nonlinear Schrödinger equation in one and two dimensions. Math. Models Computer Simulat. **11**, 634–648 (2019)
40. E. Figueiras, D. Olivieri, A. Paredes, H. Michinel, An open source virtual laboratory for the Schrödinger equation. Eur. J. Phys. **39**, 055802 (2018)
41. V.V. Voronov, Photo-induced light scattering in cesium doped variant strotnium niobate crystals. Soviet J. Quant. Electr. **10**, 1346 (1980)
42. Q.W. Song, C.-P. Zhang, P.J. Talbot, Self-defocusing, self-focusing, and speckle in $LiNbO_3$ and $LiNbO_3$: Fe crystals. Appl. Optics **32**, 7266–7271 (1993)
43. G. Zhang, Q.X. Li, P.P. Ho, S. Liu, Z.K. Wu, R.B. Alfano, Dependence of specklon size on the laser beam size via photo-induced light scattering in $LiNbO_3$:Fe. Appl. Optics **25**, 2955–2959 (1986)
44. C. Denz, M. Schwab, C. Weilnau, *Transfers-Pattern Formation in Photorefractive Optics* (Springer, New York, 2003)
45. G.P. Agrawal, Induced focusing of optical beams in self-defocusing nonlinear media. Phys. Rev. Lett. **64**, 2487–2490 (1990)
46. J.M. Hickmann, A.S.L. Gomes, C.B. de Araújo, Observation of spatial cross-phase modulation effects in a self-defocusing nonlinear medium. Phys. Rev. Lett. **68**, 3547–3550 (1992)
47. G. Zhang, G. Tian, S. Liu, J. Xu, G. Zhang, Q. Sun, Noise amplification mechanism in $LiNbO_3$: Fe crystal sheets. J. Opt. Soc. Am. Opt. Phys. **14**, 2823–2830 (1997)
48. B. Wang, C.C. Sun, W.S. Su, A.E.T. Chiou, Shift-tolerance property of an optical double-random phase-encoding encryption system. Appl. Optics **39**, 4788–4793 (2000)
49. D.V. Dylov, J.W. Fleischer, Nonlinear self-filtering of noisy images via dynamical stochastic resonance. Nat. Photon. **4**, 323–328 (2010)
50. B. Javidi, A. Sergent, G. Zhang, L. Guibert, Fault tolerance properties of a double phase encoding encryption technique. Opt. Eng. **32**, 992–998 (1997)
51. J.J. Healy, M.A. Kutay, J.T. Sheridan, *Linear Canonical Transforms: Theory and Applications* (Springer, New York, 2016)
52. W. Stallings, *Cryptography and Network Security* (Prentice Hall, Englewood Cliffs, NJ, 2004)
53. M. Puida, F. Ivanauskas, Light beam phase retrieval in nonlinear media: a computer simulation. Liet. Matem. Rink. **45**, 504 (2005)
54. C.-H. Lu, C. Barsi, M.O. Williams, J.N. Kutz, J.W. Fleischer, Phase retrieval using nonlinear diversity. Appl. Optics **52**, 92–96 (2013)
55. A. Sagiv, A. Ditkowski, R.H. Goodman, G. Fibich, Loss of physical reversibility in reversible systems. Physica D **410**, 132515 (2020)