modes, we can still reconstruct to a fair degree the higher intensity modes. Lastly, in comparing the uniformly weighted (RUN_UP) and properly weighted (RUN_WP) runs on the simulated Poisson shot noise data set $\mathbf{y_p}$, it is evident that using only uniform weighting results in higher noise in the modes.

3.3. Experiment

The experimental arrangement consists of two parts – the Schell-model partially coherent light source and the modified phase space tomography capture system. A Thorlabs LEDC13 530nm collimated LED light source was used as the initial light source, and its light was filtered by a 532nm bandpass filter with a FWHM of 10nm in order to enhance the monochromaticity of the light. This light was then passed through a 2-f optical system consisting of a 100mm focal length plano-convex lens with a 100nm slit at the front focal plane and a 500nm slit at the back focal plane. The flat side of the plano-convex lens was facing the 100nm slit. This arrangement generates a Schell-model partially coherent field immediately after the 500nm slit.

This partially coherent field was then propagated 150mm onwards until another plano-convex lens, this time having a focal length of 50mm and its flat side facing away. A uEye UI-1460SE-C USB color camera mounted on a Newport XMS50 motorized translation stage was placed behind this lens. The translation stage with its 50mm travel distance allowed the sensor plane on the camera to travel between locations 45mm and 95mm behind the 50mm lens.

A custom Python script was used to control both the translation stage and the camera to capture focal stacks of 201 images each, at 0.25mm step intervals. Two focal stacks were captured using the setup, one with the LED turned on (LED_ON) and one with the LED turned off (LED_OFF). The latter was captured for background subtraction.

Although the camera used was a Bayer-based color camera, the capture process captured only the raw pixel values from the sensor without demosaicing. For each stage position, the following procedure was performed:

1. Retrieve a 256×256 sub-image `img_on` from the entire 2048×1536 LED_ON image.
2. Retrieve a sub-image `img_off` covering the same exact pixels from the LED_OFF image.
3. Set `line_on` to be a set of values where each value is the mean value of only the green pixels in each column of `img_on`.
4. Set `line_off` to be a set of values where each value is the mean value of only the green pixels in each column of `img_off`.
5. Set `line_on_1` to be the set of pixel values for all the green pixels in the center 256×2 portion of `line_on`.
6. Set `line_on_std` to be the sample standard deviation of green pixel values in each column of `img_on`.
7. Subtract `line_off` from `line_on` and append to intensity measurement vector $\mathbf{y_{exp}}$.
8. Subtract `line_off` from `line_on_1` and append to intensity measurement vector $\mathbf{y_{exp1}}$.
9. Append `line_on_std` to the noise standard deviation estimate vector $\boldsymbol{\sigma}_{\mathbf{exp}}$.

The collected data is shown as focal stacks in Fig. 11.

Four separate runs of the factored form descent algorithm were then performed, as shown in Table 2. The _U runs were with uniform weighting, i.e. $\sigma_m = 1$ for all m, whereas the _W

#180864 - $15.00 USD Received 29 Nov 2012; revised 3 Feb 2013; accepted 4 Feb 2013; published 1 Mar 2013

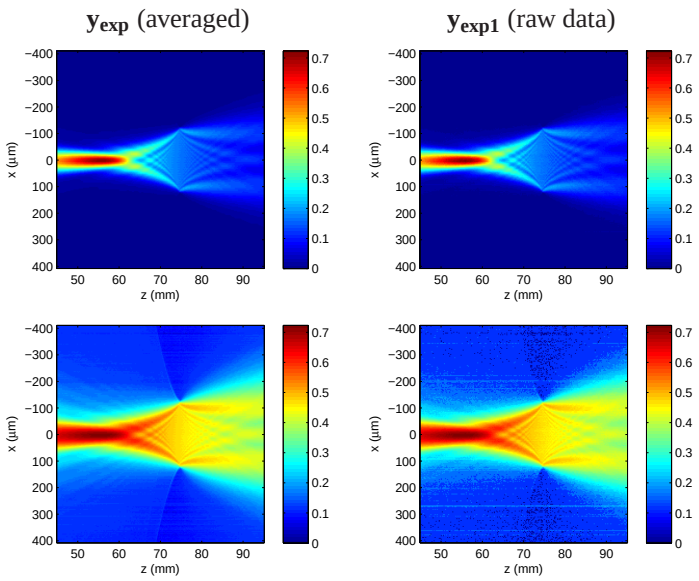

Fig. 11. The data collected during the experiment, visualized as focal stacks (top row) and gamma-boosted focal stacks (bottom row).

Table 2. Algorithm Runs on Experimental Data

Name	*Input*	*Iterations*	*Weighting*
RUN_EXP_U	$\mathbf{y_{exp}}$	500	uniform
RUN_EXP_W	$\mathbf{y_{exp}}$	500	$\sigma_{\mathbf{exp}}$
RUN_EXP1_U	$\mathbf{y_{exp1}}$	500	uniform
RUN_EXP1_W	$\mathbf{y_{exp1}}$	500	$\sigma_{\mathbf{exp}}$

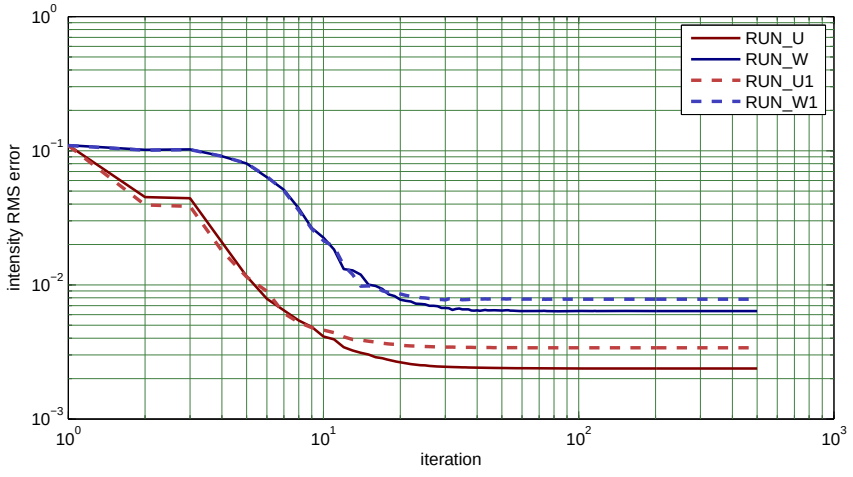

Fig. 12. Convergence of RMS error between the input experimental data and the intensity computed from the current iterate of the mutual intensity in the factored form descent algorithm. Both the error axis and the iteration (time) axis are shown in log scale.

#180864 - $15.00 USD Received 29 Nov 2012; revised 3 Feb 2013; accepted 4 Feb 2013; published 1 Mar 2013

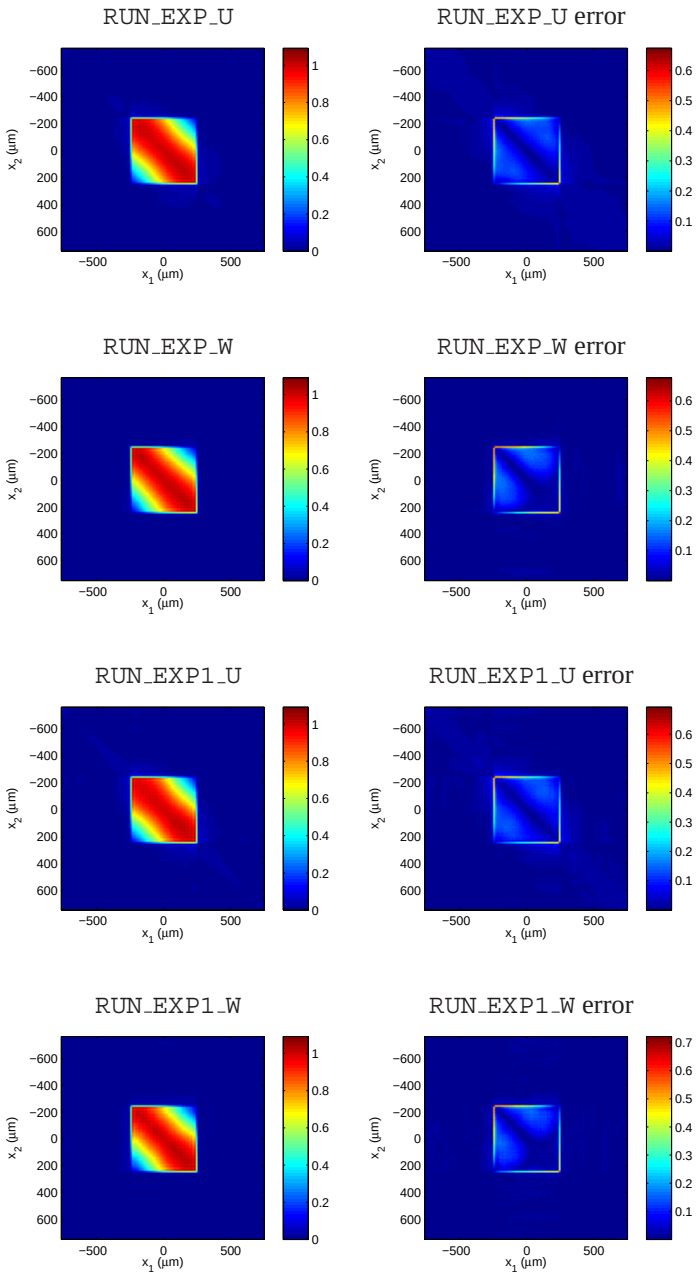

Fig. 13. Images corresponding to the resulting mutual intensity computed by runs on the experimental data sets. The left column contains images of the magnitude of the mutual intensity and the right column contains images of the magnitude of the difference between the attained mutual intensity and the theoretical mutual intensity.

#180864 - $15.00 USD Received 29 Nov 2012; revised 3 Feb 2013; accepted 4 Feb 2013; published 1 Mar 2013

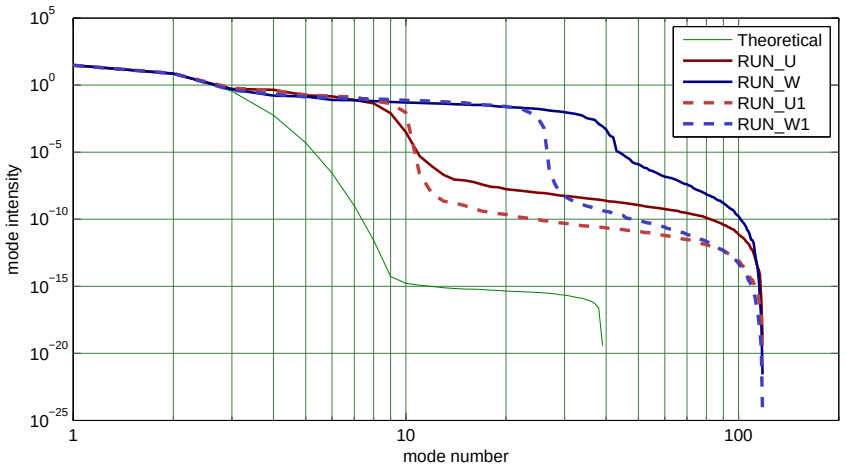

Fig. 14. A plot of the energy contained in each mode after performing a coherence mode decomposition on the theoretical mutual intensity as well as the computed mutual intensity from each experimental run. The horizontal axis gives the mode number, with modes sorted by decreasing energy, and the vertical axis gives the energy on a log scale.

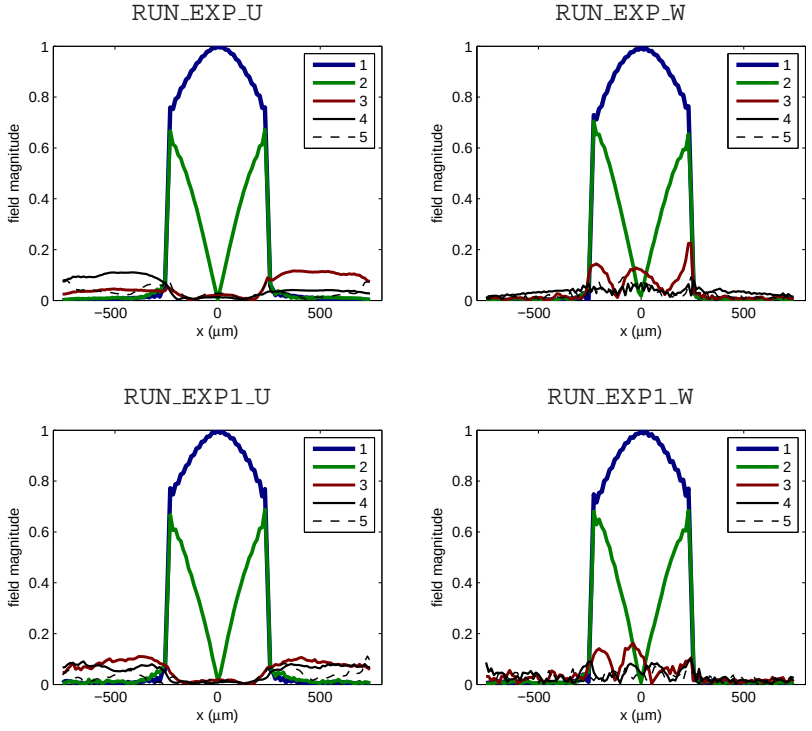

Fig. 15. A plot of the magnitude of the field for the first five coherence modes of the theoretical mutual intensity as well as the computed mutual intensity from each run.

#180864 - $15.00 USD Received 29 Nov 2012; revised 3 Feb 2013; accepted 4 Feb 2013; published 1 Mar 2013

runs were with estimates of the standard deviation. Though technically, the weighting should be $\sigma_{\mathbf{exp}}/\sqrt{128}$ for the RUN_EXP_W case, constant scale factors on the weighting vector have no actual effect on the algorithm; it only rescales the merit function by a constant scale factor. The RUN_EXP1 runs were done to see a more realistic use-case, since it is often not feasible to perform averages to boost SNR in real-world situations. Each run converged by the end of the 500 iterations, as can be seen in Fig. 12.

As expected, unweighted RMS intensity error was reduced in all the runs, and the weighted runs resulted in higher unweighted RMS intensity error. Furthermore, the runs on unaveraged data generally resulted in slightly higher error.

From the mutual intensity images in Fig. 13, it can be seen that the experimental data matches fairly closely the theoretical data, with most of the error in the border regions. It appears that the recovered mutual intensity wasn't as sharp as the theoretical, which is probably due to aberrations and limitations of the imaging system. Furthermore, use of noise standard deviation estimates seemed to have removed the excess energy along the diagonal in regions outside of the central square region, although the resulting mutual intensity no longer looked as spatially symmetric.

The mode-wise energy distribution graph in Fig. 14 shows that the uniformly weighted runs deviate from the theoretical mode drop-off more quickly, although they have smaller "first plateaus". The non-uniformly weighted runs follow the theoretical mode drop-off more closely, but they have larger "first plateaus" after they deviate. The higher fidelity of the first few modes can be seen in the mode field magnitudes plotted in Fig. 15, where RUN_EXP_W managed to reconstruct the first three modes, with the third mode in RUN_EXP1_W still showing some semblance of the theoretical third mode. Runs RUN_EXP_U and RUN_EXP1_U failed to reconstruct the third mode properly and exhibited more noise outside of the region of the 500nm slit.

4. Conclusions

We have constructed a generalized formulation for coherence retrieval and demonstrated a novel optimization algorithm to solve for the global minimum. A verifiable test using both simulated and experimental measurement data of a Schell-model source demonstrates that the algorithm functions correctly, leading to reasonable reconstructions of the original mutual intensity. The ability to weight the fidelity of each measurement by providing noise standard deviations for each point enables the option of boosting reconstruction fidelity by matching noise statistics. In all runs, the theoretical possibility of a pathological saddle point was never encountered.

However, there are two problems with the algorithm that could be the focus of further research. The first problem, as can be seen in Fig. 5 is that although error reduction is initially fast, final convergence is slow. Even when near convergence, there is extraneous energy present in modes that should have zero energy, as shown in Fig. 9. The noiseless case error image in Fig. 7 may provide some clues for a way to precondition the algorithm to solve these problems.

The second problem is that while the algorithm compensates for noise in the measurements, it does not compensate for errors in the specification of the propagation matrix A. In a real-world situation, it is impossible to determine the exact system function of an optical system. Error will either result from aberrations in the physical system or simply from measurement error if one chooses to characterize A empirically. To make the algorithm more robust, we need to improve the algorithm to allow it to tolerate a certain degree of error in A.

In addition to these problems, there are many other potential new avenues of research, including expanding the experimental setup to capture fully two-dimensional fields and applying the algorithm to non-tomographic capture protocols as well as illumination synthesis problems.

Lastly, we would like to discuss the similarities and differences between this work and compressive phase space tomography [6]. While both techniques aim at recovering the mutual inten-

#180864 - $15.00 USD Received 29 Nov 2012; revised 3 Feb 2013; accepted 4 Feb 2013; published 1 Mar 2013

sity function of a partially coherent field and rely on the positivity of mutual intensity matrices, compressive phase space tomography introduces an additional prior that the field has a low number of modes and intentionally undersamples the measurements. This work relies on no such assumption and thus requires many more measurements. Perhaps it would be possible to unify these two approaches in the future.

A. Proof of Theorem 1

Proof. To prove Theorem 1, we will first show that the algorithm is monotonically decreasing and will eventually converge to some value. The monotonic behavior of merit function values is a direct consequence of the use of exact global line searches; for each iteration, the merit function value for the next iteration is bounded above by the current value and bounded below by the global minimum. Therefore, the sequence of merit function values must decrease monotonically and eventually converge to some value. Now let us investigate the conditions required for the algorithm to stop making progress and thus terminate.

Since the iterate $X(i)$ changes by $\alpha(i)S(i)$ each iteration, the factored form descent algorithm would stop making progress if and only if $\alpha(i)S(i)$ became zero for all future iterations. This is only possible if $S(i) = 0$ or if $\alpha(i) = 0$ in the case that $S(i) \neq 0$. We will now consider the first case.

We can split this case into two situations, depending on whether $G(i) = 0$. If $G(i) = 0$, then $\beta(i)$ necessarily has to be zero because the numerator is a dot product with $G(i)$. Therefore, $S(i) = 0$ would be zero as well. If $G(i) \neq 0$, then we can show that $S(i) \neq 0$ as well. The only way $S(i)$ could have been zero were if $G(i) = -\beta(i)S(i-1)$. For $i = 1$, $\beta(i) = 0$ and thus this statement would be false. For $i > 1$, the line search in the previous iteration guarantees that $\langle G(i), S(i-1)\rangle = 0$, and hence it would be impossible for a scalar multiple of $S(i-1)$ to be equal to $G(i)$. Therefore, $S(i) = 0$ can only be true if $G(i) = 0$.

Now let us consider the case when $S(i) \neq 0$ and $\alpha(i) = 0$. A necessary condition for $\alpha(i) = 0$ is for $G(i)$ and $S(i)$ to have an inner product of zero, necessarily implying:

$$0 = \langle S(i), G(i)\rangle = \langle G(i) + \beta(i)S(i-1), G(i)\rangle = \langle G(i), G(i)\rangle \tag{16}$$

Therefore, this case also requires $G(i) = 0$ to be true. Combining the two cases results in the conclusion that when the algorithm stops making progress and terminates, $G(i) = 0$.

Hence, the factored form descent algorithm must produce monotonically decreasing merit function values and converges when $G(i) = 0$. □

Acknowledgments

The authors would like to thank Lei Tian, Laura Waller and Jon Petruccelli for their helpful discussions. This research was supported by the National Research Foundation Singapore through the Singapore-MIT Alliance for Research and Technology's BioSyM and CENSAM research programmes.

#180864 - $15.00 USD Received 29 Nov 2012; revised 3 Feb 2013; accepted 4 Feb 2013; published 1 Mar 2013