Calculation of the expected output spectrum for a mid-infrared supercontinuum source based on As_2S_3 chalcogenide photonic crystal fibers

R. J. Weiblen,[1,*] A. Docherty,[1] C. R. Menyuk,[1] L. B. Shaw,[2] J. S. Sanghera,[2] and I. D. Aggarwal[3]

[1]*Department of Computer Science and Electrical Engineering, TRC 203*
University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA
[2]*Naval Research Laboratory, Code 5620, Washington, D.C. 20375, USA*
[3]*Sotera Defense Solutions, Crofton, MD 21114, USA*

[*]*ro2@umbc.edu*

Abstract: We computationally investigate supercontinuum generation in an As_2S_3 solid core photonic crystal fiber (PCF) with a hexagonal cladding of air holes. With a goal of obtaining a supercontinuum output spectrum that can predict what might be seen in an experiment, we investigate the spectral and statistical behavior of a mid-infrared supercontinuum source using a large ensemble average of 10^6 realizations, in which the input pulse duration and energy vary. The output spectrum is sensitive to small changes (0.1%) in these pulse parameters. We show that the spectrum can be divided into three regions with distinct characteristics: a short-wavelength region with high correlation, a middle-wavelength region with minimal correlation, and a long-wavelength region where the behavior is dominated by a few rare large-bandwidth events. We show that statistically significant fluctuations exist in the experimentally expected output spectrum and that we can reproduce an excellent match to that spectrum with a converged shape and bandwidth using 5000 realizations.

References and links

1. H. Kano and H. O. Hamaguchi, "Vibrationally resonant imaging of a single living cell by supercontinuum-based multiplex coherent anti-Stokes Raman scattering microspectroscopy," Opt. Express **13**(4), 1322–1327 (2005).
2. A. Schliesser, N. Picque, and T. Hänsch, "Mid-infrared frequency combs," Nat. Photonics **6**(7), 440–449 (2012).
3. C. S. Colley, J. C. Hebden, D. T. Delpy, A. D. Cambrey, R. A. Brown, E. A. Zibik, W. H. Ng, L. R. Wilson, and J. W. Cockburn, "Mid-infrared optical coherence tomography," Rev. Sci. Instrum. **78**(12), 123108 (2007).
4. J. S. Sanghera, L. B. Shaw, and I. D. Aggarwal, "Chalcogenide glass-fiber-based mid-IR sources and applications," IEEE J. of Sel. Topics Quantum Electron. **15**, 114–119, (2009).
5. D. R. Solli, C. Ropers, P. Koonath, and B. Jalali, "Optical rogue waves," Nature **450**, 1054–1057 (2007).
6. B. Wetzel, K. J. Blow, S. K. Turitsyn, G. Millot, L. Larger, and J. M. Dudley, "Random walks and random numbers from supercontinuum generation," Opt. Express **20**, 11143–11152 (2012).
7. B. Wetzel, A. Stefani, L. Larger, P. A. Lacourt, J. M. Merolla, T. Sylvestre, A. Kudlinski, A. Mussot, G. Genty, F. Dias, J. M. Dudley, "Real-time full bandwidth measurement of spectral noise in supercontinuum generation," Sci. Rep. **2**, 882 (2012).

8. R. J. Weiblen, A. Docherty, J. Hu, and C. R. Menyuk, "Calculation of the expected bandwidth for a mid-infrared supercontinuum source based on As_2S_3 chalcogenide photonic crystal fibers," Opt. Express **18**, 26666–26674 (2010).

9. U. Møller, S. T. Sørensen, C. Jakobsen, J. Johansen, P. M. Moselund, C. L. Thomsen, and O. Bang, "Power dependence of supercontinuum noise in uniform and tapered PCFs," Opt. Express **20**, 2851–2857 (2012).

10. J. M. Dudley, G. Genty, and S. Coen, "Supercontinuum generation in photonic crystal fiber," Rev. Mod. Phys. **78**, 1135–1184 (2006).

11. R. J. Weiblen, J. Hu, C. R. Menyuk, L. B. Shaw, J. S. Sanghera, and I. D. Aggarwal, "Maximizing the super-continuum bandwidth in As_2S_3 chalcogenide photonic crystal fibers," *Conference on Lasers and Electro-Optics* OSA Technical Digest (CD) (Optical Society of America, 2010), paper CTuX7.

12. J. Hu, C. R. Menyuk, L. B. Shaw, J. S. Sanghera, and I. D. Aggarwal, "Maximizing the bandwidth of supercontinuum generation in As_2Se_3 chalcogenide fibers," Opt. Express **18**, 6722–6739 (2010).

13. J. Dudley and S. Coen, "Coherence properties of supercontinuum spectra generated in photonic crystal and tapered optical fibers," Opt. Lett. **27**, 1180–1182 (2002).

1. Introduction

Supercontinuum (SC) generation in the mid-infrared (mid-IR) is an area of great research interest because of its importance for many applications, including biological spectroscopy [1], optical frequency metrology [2], optical tomography [3], optical cutting and welding, and aircraft defense [4]. The statistical behavior of SC generation, including its stability, coherence, and noise properties, has become the focus of recent studies because of interest in both improving SC sources and studying other systems, such as rogue waves in fluids [5–7].

A common principal aim in SC generation, especially in the mid-IR, is to obtain a broad bandwidth. Additionally, it is often desirable to generate a spectrum that is as flat as possible over the wavelength range of interest. For example, in infrared countermeasures, the ideal spectrum would be flat over the atmospheric mid-IR transmission windows from 3–5 μm and 8–13 μm. Several characteristics of single-shot simulation results appear to indicate that SC sources cannot achieve these goals when the input pulse energies become large [8], as long as the energies are not high enough to saturate the loss and other fiber effects [9]. First, variations as large as 20 dB can exist over a single-shot spectrum. Also, the frequency of these variations depends on the bandwidth resolution. Finally, output bandwidths vary over a large range. This behavior when the input pulse energy is large contrasts strongly with the behavior when the input pulse energy is small, and supercontinuum generation is highly reproducible from shot to shot. It has been shown that the number of solitons in the input pulse plays a critical role in determining the sensitivity of the supercontinuum spectrum to small variations in the input parameters [10].

However, although many real-time measurement methods have been developed for analyzing SC, the theoretically-predicted shot-to-shot variations cannot be observed in experiment or applications where the detector is slow, and we will show in this article that they are an artifact of single-shot simulations. That is the case even when broadband (white) noise is negligible, and only the pulse energy and duration vary randomly.

In many experiments, the repetition rate of the input pulse source is much greater than the sampling rate of the optical spectrum analyzer. Therefore, the measured output spectrum of a SC source consists of an ensemble average over many thousands of pulse realizations, which include laser noise and pulse parameter (pulse duration and peak power) differences due to environmental effects. Because SC generation is extremely sensitive to small changes in the input pulse when the input pulse energy is large, a single numerical simulation will not yield an accurate representation of the experimental SC output or its bandwidth.

To obtain an accurate representation, it is necessary to use an ensemble average. That leads to the question of how large the ensemble must be in order to realistically simulate an experimental system. It is desirable for computational efficiency to use the smallest number of realizations

possible. In previous work [8, 11] on mid-IR SC generation in As_2S_3 photonic crystal fibers, we found that changes of 0.01% in either the pulse duration or the peak power lead to a SC spectrum that is almost completely uncorrelated with the initial spectrum in the intermediate wavelength range of 2700–4600 nm. Since the pulse duration and peak power can vary by 10% or more in the experimental systems of interest to us [4], this result suggests that the experiments generate as many as $10^3 \times 10^3 = 10^6$ independent realizations and it may be necessary to simulate an ensemble with as many as 10^6 realizations. We will show, however, that a much smaller ensemble with 5,000 or fewer realizations is sufficient to reproduce the experimentally expected spectral power density and bandwidth.

In this work, we again consider mid-IR SC generation in an As_2S_3 photonic crystal fiber. We completely characterize the output spectrum using a large ensemble with 10^6 realizations. Its statistical properties are not simple.

We find that the spectrum has three distinct spectral regions whose characteristics are the same for single-shot realizations with similar pulse parameters. The first is the short wavelength region due to initial four-wave mixing with a spectral shape determined mostly by material loss. In this region, while there are significant variations as a function of wavelength, there is a large coherence from shot to shot and the variations never average out.

The second spectral region is a middle wavelength region with a complex structure and rapid fluctuations in the output power as a function of wavelength. These fluctuations are governed by complicated soliton dynamics that are only slightly correlated from one realization to another because of the high soliton number of this system ($N \approx 30$). In this region, most, but not all, of the variations can be eliminated in an ensemble average. An ensemble average of about 5000 realizations reduces the fluctuations by two orders of magnitude. The third region is a long wavelength region that is relatively smooth and represents the falloff from the longest-wavelength soliton, which exists further into the infrared, extending the bandwidth of the spectrum. A few rogue events dominate the bandwidth, but an ensemble of approximately 5000 realizations is sufficient for the output bandwidth to converge.

2. Simulation methods

In this study, we aim to reproduce the expected SC output spectrum for a system with a hyperbolic secant input pulse with a duration of 1 ps and a peak power of 1 kW from a laser source with an operating wavelength of $\lambda = 2.8$ μm, using an As_2S_3 chalcogenide photonic crystal fiber with five hexagonal rings of air-holes, with a pitch of $\Lambda = 3.5$ μm and a hole-diameter-to-pitch ratio of $d/\Lambda = 0.4$ to maintain single mode operation. The input pulse duration and energy are both expected vary within a range of 10% from shot to shot. The fiber length is set at 0.5 m to optimize the balance between nonlinear effects and loss. This system was previously studied in [8], and we use the same computational method to simulate SC generation that is described in [8] and in more detail by Hu et. al. [12]. The equation that governs the evolution of a single SC simulation is the generalized nonlinear Schrödinger equation (GNLSE) given by [10]:

$$\frac{\partial A(z,t)}{\partial z} + \frac{a(\omega)}{2}A - i\,\mathrm{IFT}\Big\{[\beta(\omega_0+\Omega)-\beta(\omega_0)-\Omega\beta_1(\omega_0)]\tilde{A}(z,\Omega)\Big\}$$
$$= i\gamma\Big(1+\frac{i}{\omega_0}\frac{\partial}{\partial t}\Big)\Big[A(z,t)\int_{-\infty}^{t}R(t-t')|A(z,t')|^2\,dt'\Big], \quad (1)$$

where $A(z,t)$ is the electric field envelope and $\tilde{A}(z,\omega)$ is its Fourier transform. Here, $a(\omega)$ is the linear loss as a function of frequency, comprised of the material loss and the fiber loss, which

is given in Fig. 2 of [8]. The Kerr coefficient, γ, is defined by $\gamma = n_2 \omega_0 / (cA_{\mathrm{eff}})$, where n_2 is the nonlinear refractive index given by $n_2 = 3.285 \times 10^{-18}$ m^2 W^{-1}, ω_0 is the angular frequency of the optical carrier, c is the speed of light, and A_{eff} is the fiber's effective area, which depends on the air-hole diameter and pitch. The quantity $R(t)$ is the nonlinear response function defined by $R(t) = (1 - f_R)\delta(t) + f_R h_R(t)$, which includes both the instantaneous Kerr contribution, $\delta(t)$, and the delayed Raman response, $h_R(t)$. We use f_R to denote the fraction of the nonlinear response function due to the Raman effect, which in this case is taken as $f_R = 0.2$. We use IFT{} to denote the inverse Fourier transform. We use a time window of 2^{32} points and 10^4 steps in the longitudinal direction. It has been verified that increasing the resolution of these numerical settings does not change the results of the simulation over the parameter range of interest.

The ensemble average is taken from realizations in the two-dimensional parameter space defined by the input pulse duration and the input pulse peak power. The parameter space is centered around the point with a 1 kW peak power and a 1 ps pulse duration and includes a 10% variation in each dimension. Thus, the parameter space extends from 950 fs to 1050 fs in pulse duration and from 950 W to 1050 W in pulse peak power. Realizations are chosen 100 at a time. Each set of 100 realizations forms a uniform gridding of the parameter space, such that realizations in each set are separated as far apart as possible. Successive sets are created by shifting the sampling grid so that no realization is repeated and each new set is as far away from previous sets as possible. We show this gridding scheme schematically in Fig. 1. Each set of 100 is averaged, and we repeated this 100-point sampling until we had 10^6 individual SC realizations, or 10^4 100-realization ensemble averages. Then, these 10^4 averages were averaged cumulatively to investigate the behavior of the ensemble as more realizations are included.
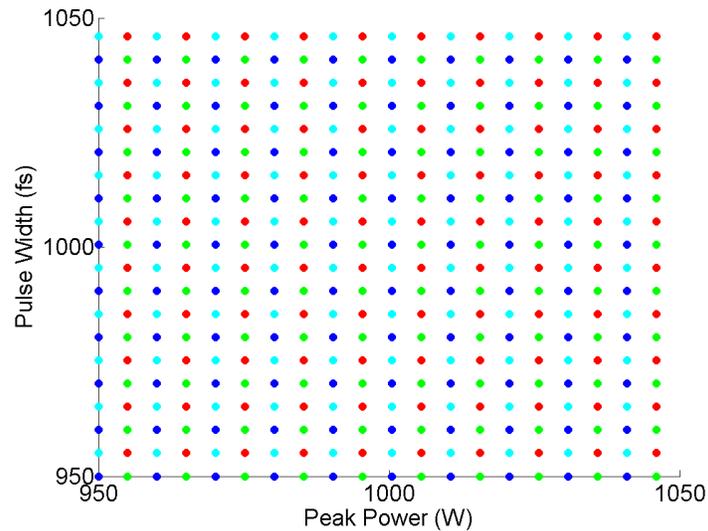

Fig. 1. Uniform gridding scheme. Realizations of the same color are averaged (see Media 1).

We do not include broadband (white) noise in our simulations. While adding white noise does change the SC output spectrum that results from a particular set of parameters when its variance becomes sufficiently large, so does a small change of 0.1% in one or both of the input parameters. In order to verify that adding white noise has little effect on the statistics,

we carried out a simulation with 1000 realizations using white noise and a variance that was sufficiently large to generate data similar to Fig. 3. The results were statistically similar to the results that we obtain by varying the input pulse energy and duration. We note that the statistical properties of the laser noise are not well-characterized in the experiments of interest to us [4] over the spectral bandwidth of the pulse, in contrast to the large, random variations of the pulse duration and energy, and there is no *a priori* reason to assume that the noise is either large or white. Finally, it is not surprising that white noise induces sensitivity in the output spectrum, but it might be considered surprising that slight changes in just the output pulse duration and energy do the same. For some applications — notably those for which it is desirable to obtain a broadband, noisy, and reasonably flat spectrum — this sensitive dependence on the pulse energy and duration is useful, since these parameters naturally vary over a wide range. For all these reasons, we omitted white noise in the simulations that we present here.

3. Spectral correlation in parameter space

Figure 2 shows a typical supercontinuum output spectrum with the approximate delineation of the three wavelength regions. Region 1, the short wavelength region, from around 1.5 μm to 2.7 μm shows the least change as the pulse parameters are varied. The spectrum in Region 1 is primarily due to early four-wave mixing with a spectral shape determined primarily by material loss, which increases rapidly for wavelengths shorter than 2 μm, and by the four-wave mixing resonances. Region 2, the middle wavelength region, from around 2.7 μm to around 4.6 μm shows the most change as the pulse parameters are varied slightly. However, while Region 2 shows a large amount of variation, the effect on the bandwidth is small, since most of the variation is due to a fine structure that varies rapidly with wavelength. Region 3, the long wavelength region, from around 4.6 μm and longer, changes rarely but significantly. The power spectrum in this region is due almost entirely to the longest-wavelength, highest-energy soliton and has the greatest effect on the bandwidth of the SC output.
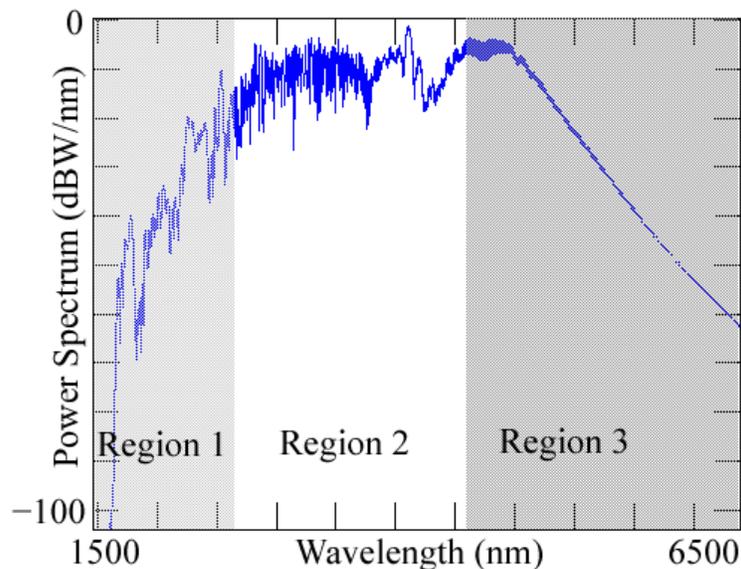

Fig. 2. Example supercontinuum spectrum showing three regions.

In Fig. 3, we show the average Pearson correlation coefficient between realizations with dif-

ferent pulse parameters from the parameter space for the short and middle wavelength regions. The Pearson correlation coefficient is computed as

$$\rho = \frac{\sum_{\lambda=\lambda_1}^{\lambda_n}(X_\lambda - \bar{X})(Y_\lambda - \bar{Y})}{\left[\sum_{\lambda=\lambda_1}^{\lambda_n}(X_\lambda - \bar{X})^2 \sum_{\lambda=\lambda_1}^{\lambda_n}(Y_\lambda - \bar{Y})^2\right]^{1/2}}, \tag{2}$$

where X and Y are two different output spectra or output spectrum regions with corresponding mean values $\bar{X}$ and $\bar{Y}$ and X_λ and Y_λ are the values of those output spectra at a particular wavelength λ. For the short wavelength region, $\lambda_1 = 1000$ nm and $\lambda_n = 2700$ nm. For the intermediate wavelength region, $\lambda_1 = 2700$ nm and $\lambda_n = 4600$ nm. In both cases, the values of λ are separated by 1 nm.

Figure 3(a) shows the average correlation coefficient for the short wavelength region and middle wavelength region between realizations where the input pulse peak power is varied by 0.1 W. Figure 3(b) shows the average correlation coefficient for the short wavelength region and middle wavelength region between realizations where the input pulse duration is varied by 0.1 ps. The average correlation coefficient is taken over 1000 realizations between 950 W and 1050 W for the peak power plots and over 1000 realizations between 950 fs and 1050 fs for the pulse duration plots. These values represent a 10% variation in pulse parameters for a 1 ps, 1 kW input pulse. The long correlation length for the short wavelength region is evident for both dimensions of the two-dimensional pulse parameter space. By contrast, the middle wavelength region has a very short correlation length in both peak power and duration, but we note that its correlation does not go to zero.

These correlation coefficient plots seem to indicate that there are many independent realizations of output spectra present in this parameter space. Because the correlation drops quickly for parameter differences greater than about 0.1 in either dimension (peak power in W or pulse duration in fs) for the middle wavelength region, we might assume that there will be $100/(0.1) = 1000$ independent realizations in either dimension, where the number 100 represents a 10% variation from 1000 fs or 1000 W. In two dimensions, we would thus estimate that there are 10^6 distinct realizations of output spectrum in this parameter space. It is from this space that we will draw our samples.
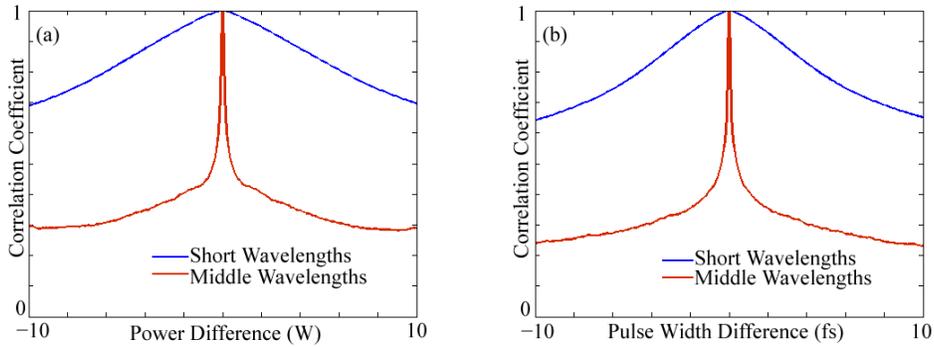

Fig. 3. Average correlation coefficient for short and middle wavelength regions when (a) the peak power is varied and (b) when the pulse duration is varied.

4. Spectral convergence

Figure 4 shows a typical single-shot SC output spectrum for an input pulse with a peak power of 1 kW and a pulse duration of 1 ps, which shows large fluctuations in output power, and the ensemble-averaged output spectrum for the full ensemble including 10^6 samples taken from the parameter space as described in Section 2. In addition to the large-scale fluctuations in output power present in the single-shot result, it is readily apparent that this single-shot result has a slightly larger bandwidth than the full ensemble—about 100 nm wider. This variability in the output spectrum and its bandwidth, as demonstrated here, is the main cause of uncertainty in single-shot results. We note that the short wavelength region still shows large, persistent fluctuations, even in the full ensemble average, which we attribute to periodic phase-matching of the four-wave mixing. Also visible in the ensemble-averaged result are some small-scale fluctuations in the middle wavelength region, particularly on the short wavelength side, which we attribute to residual soliton interactions, as we will discuss later.
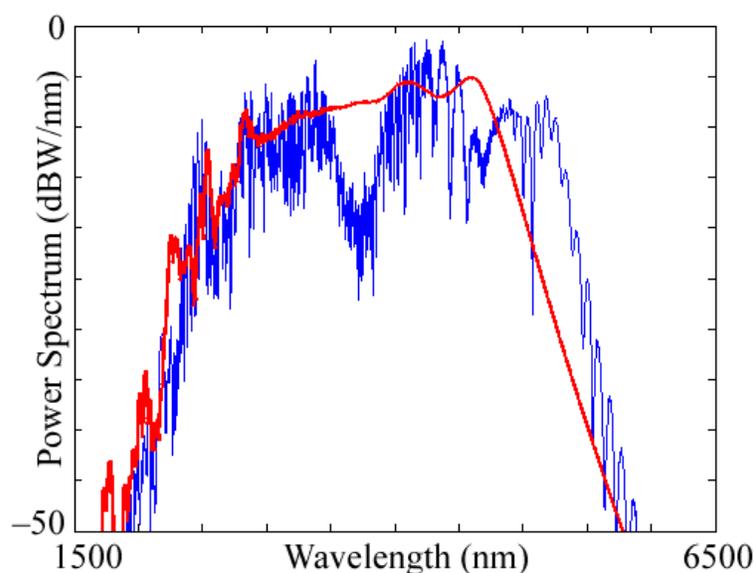

Fig. 4. Single-shot simulation output (blue) and full 10^6 ensemble average output (red).

As mentioned above, the variations in wavelength region 3 are most responsible for the variability in the output bandwidth. Figure 5 shows a comparison of two single-shot output spectra, which are similar for most wavelengths except in the long wavelength region. The difference in bandwidth between these two spectra is about 700 nm. The large bandwidth of the spectrum shown in blue is due to a rogue event. These events occur rarely, but often enough to affect the ensemble average.

In Fig. 6, we show a histogram of the output bandwidth for 10^5 realizations using 50 bins. This distribution is bimodal. Most output spectra in the ensemble have a bandwidth around 2.8 μm, while a smaller group of bandwidths is clustered around 2.5 μm. However, even in this (relatively) limited sample space of noise-free simulations, spectra with bandwidths greater than 3.2 μm, more than 25% greater than the average, exist in the ensemble.

It has been shown [10, 13], that for SC systems with soliton numbers greater than about $N = 10$, the mutual coherence between different realizations of SC generation will decrease rapidly. Hence, the wavelengths of the SC outputs will be largely uncorrelated from shot-to
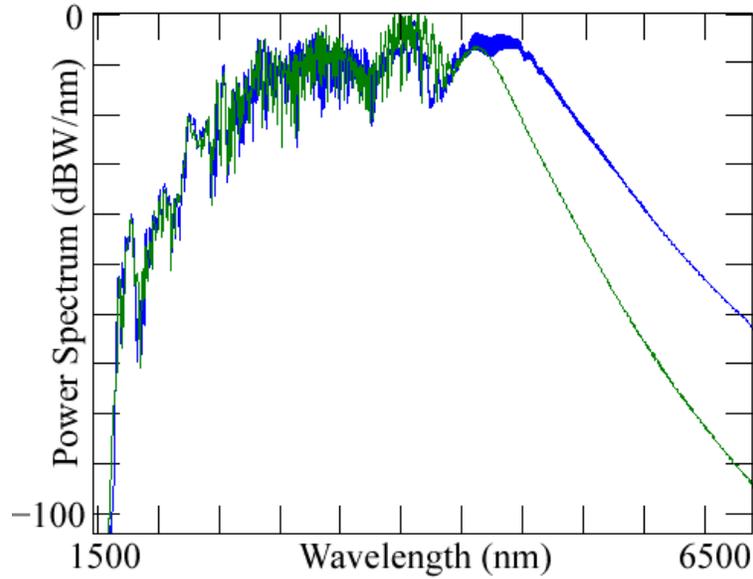
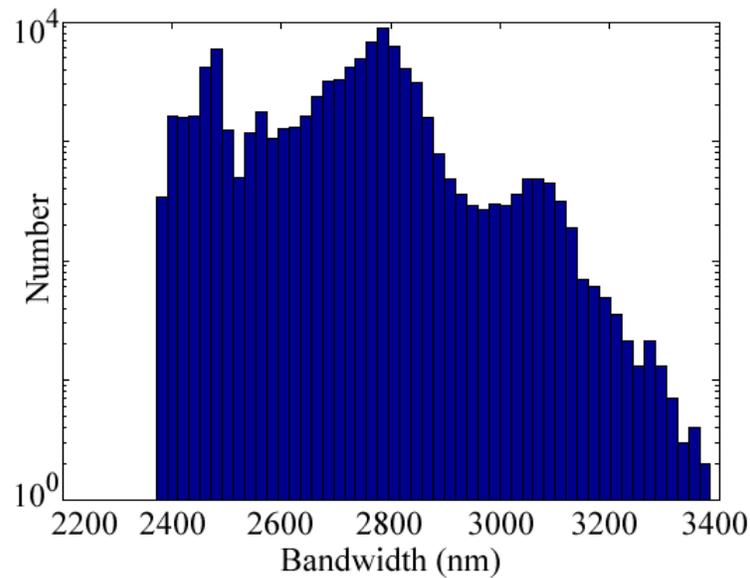Fig. 5. Comparison of output spectra with different bandwidths: 2.5 μm (green) and 3.2 μm (blue).

Fig. 6. Histogram of output bandwidths.

shot. The large soliton number for this system ($N \approx 30$) is consistent with the great sensitivity in the output spectrum that we observe when the input pulse duration and peak power change. At the same time, the number is small enough that the residual correlations remain visible. We attribute the bumps that are visible in the middle range of wavelengths in the 10^6 ensemble-averaged spectrum in Fig. 4 to these residual correlations. In Fig. 7(a), we show the spectrogram

evolution for a single realization, and in Fig. 7(b), we show the average spectrogram evolution for 1000 realizations. The solitons that are visible in a single realization remain perceptible in the average, although greatly diminished in amplitude relative to the background.
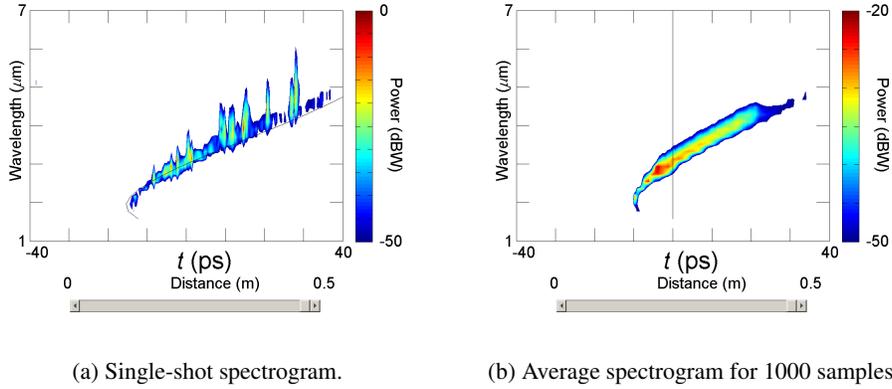

(a) Single-shot spectrogram. (b) Average spectrogram for 1000 samples.

Fig. 7. Spectrogram plots for (a) a single simulation (see Media 2) and (b) a 1000-simulation average (see Media 3).

5. Statistical characterization

The location and amplitude of the large-scale, fast fluctuations in the spectrum appear to be randomly varying from one realization to the next, although, as just noted, appearances are deceiving. If the samples from the parameter space were statistically uncorrelated, we would expect the variance of these fluctuations to reduce with an increasing number of samples in the ensemble like N^{-1}, where N is the number of samples in the ensemble. From the plots of the correlation coefficient in Section 3, we might expect that the highly uncorrelated middle wavelength region would show a reduction in variance that follows this rule. The short wavelength region has a much broader correlation function, and so we would expect the variance of the fluctuations in this region to reduce more slowly as more realizations are included in the ensemble average.

To test this hypothesis, we computed the variance of the fluctuations in both regions, using the final ensemble-averaged spectrum as a basis for comparison. The variance of the fluctuations in a spectrum is computed as

$$\sigma^2 = \frac{1}{n} \sum_{\lambda=\lambda_1}^{\lambda_n} \left[(X_\lambda - \overline{X_\lambda}) - \mu \right]^2, \tag{3}$$

where n is the number of wavelengths, or the size of the region, under consideration, X_λ is the amplitude of the spectrum at wavelength λ, and $\overline{X_\lambda}$ is the amplitude of the full ensemble-averaged spectrum at that wavelength, and μ is the mean value of $X_\lambda - \overline{X_\lambda}$. For the short wavelength region, $\lambda_1 = 1000$ nm and $\lambda_n = 2700$ nm. For the intermediate wavelength region, $\lambda_1 = 2700$ nm and $\lambda_n = 4600$ nm. Again, the values of λ are separated by 1 nm.

The variances of the fluctuations in the short and middle wavelength regions are shown in Fig. 8(a) and Fig. 8(b), respectively. In both cases, the variance decreases at a similar rate, approximately equal to $N^{-5/6}$. The variance decreases rapidly as the number of samples approaches 10^6 because the data are being compared to the ensemble average with 10^6 samples. While both regions have a similar overall rate of decrease, the decrease in variance is relatively

smooth for the middle wavelength region compared to the short wavelength region. The short wavelength region fluctuation variance reduces in jumps, with a seemingly self-similar pattern that repeats approximately two times every decade. We attribute the lack of smoothness in this case to the high correlation in this wavelength range and the residual spectral structure that does not average out, which is due to a combination of material loss and four-wave mixing resonances.
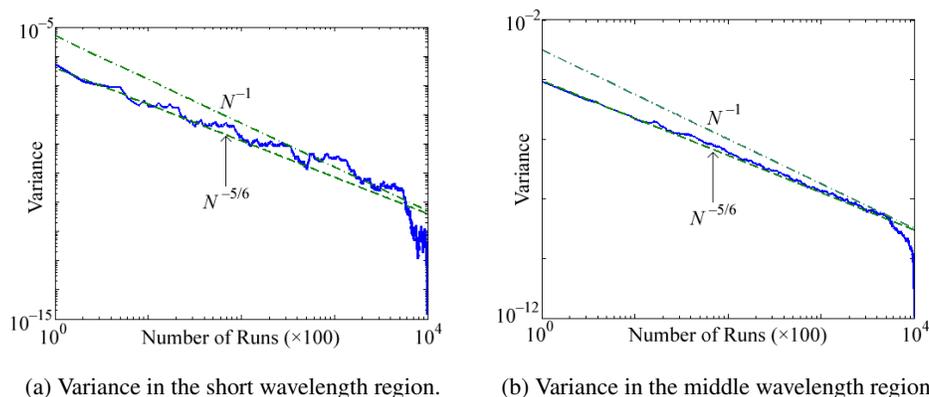

(a) Variance in the short wavelength region.   (b) Variance in the middle wavelength region.

Fig. 8. Variance of the fluctuations in the (a) short and (b) middle wavelength region. Lines corresponding to $N^{-5/6}$ and N^{-1} are also shown for comparison.

The $N^{-5/6}$ rate of reduction for the middle wavelength region means that with 5000 samples, the variance of the fluctuations has reduced by approximately two orders of magnitude from its original value. This reduction is sufficient to yield a good quantitative estimate of the shape, smoothness, and amplitude of the experimentally expected output spectrum. A comparison of the ensemble average including 5000 realizations and 10^6 realizations is shown in Fig. 9. The corresponding spectra match almost exactly. The differences between the output spectra are small scale fluctuations that are almost imperceptible.
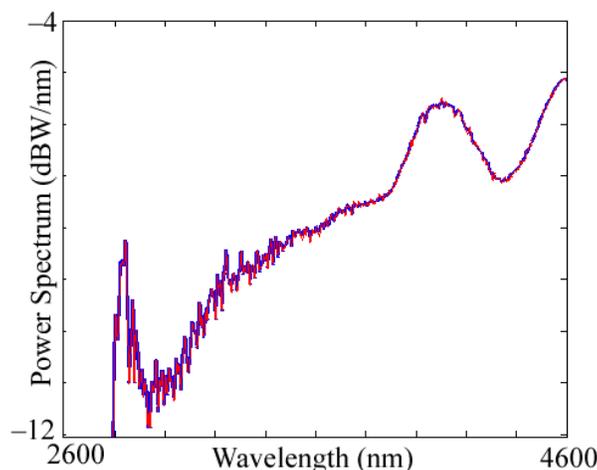

Fig. 9. Ensemble average output for 5000 samples (red) and 10^6 samples (blue).

Because the rate of decrease for the middle wavelength region is slower than N^{-1}, we infer that there are residual correlations between the realizations due to fluctuations as a function of wavelength that never average out. These residual fluctuations are visible in Fig. 9 in both large-scale fluctuations that vary slowly as a function of wavelength and small-scale fluctuations that vary rapidly as a function of wavelength. We attribute these residual correlations to the restricted number of solitons that make up the spectrum and are visible in Figs. 7(a) and 7(b). The presence of these residual correlations is consistent with the Pearson correlation coefficients in Fig. 3. The correlation coefficient does not approach zero and diminishes slowly for separations greater than about $\Delta\tau = \pm 2$ fs in pulse duration or $\Delta P = \pm 2$ W in pulse peak power.

6. Long wavelengths and bandwidth convergence

Figure 10 shows the bandwidth of the ensemble-averaged output spectrum as the the number of samples in the ensemble increases. While the bandwidth comes within 5 nm of its final, converged value after about 1000 realizations, it does not reach its final value until around 5000 realizations. Even though there is still some settling down after 5000 realizations, the fluctuation in the bandwidth is small after that point. This value of 5000 realizations is consistent with the convergence of the middle wavelength region of the spectrum as described in the previous section. We also found similar bandwidths using different sets of 5000 realizations.

In comparing Fig. 6, the histogram of bandwidths, and Fig. 10, we see that the bandwidth of the average spectrum is not the same as the average bandwidth. The average bandwidth is around 2700 nm, while the bandwidth of the average spectrum is close to 2900 nm. This difference indicates that, as expected, the ensemble average behavior in the long-wavelength region is dominated by the large-bandwidth rogue events, even though they occur much less frequently than the average-bandwidth output spectra. It is the bandwidth of the average spectrum that would be observed in an experiment.
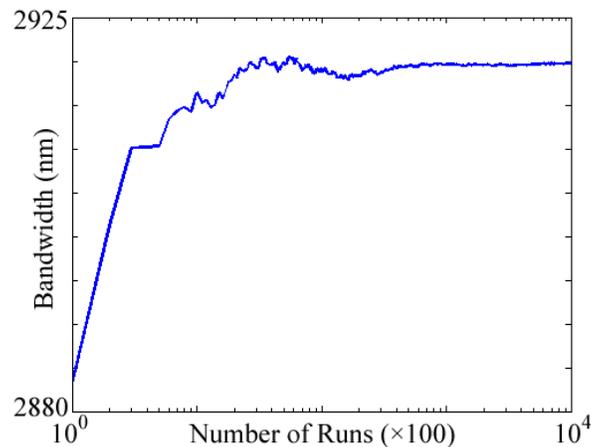

Fig. 10. Average bandwidth versus number of samples in the ensemble average.

7. Conclusion

We have investigated the spectral and statistical characteristics for the output of a supercontinuum source based on an As_2S_3 hexagonal photonic crystal fiber as the pulse duration and energy

vary over a range of 10%. Small variations of just these parameters are sufficient to induce large changes in the output spectrum.. We identified three wavelength regions with different statistical properties and computed a large ensemble average of 10^6 realizations to further investigate the behavior of the output spectrum as the ensemble size increases. We used a gridding scheme in which the realizations are spaced as far apart as possible.

We showed that even with 10^6 realizations, there are still small-scale and large-scale fluctuations present in the ensemble average output spectrum, due to residual correlation of the realizations in the parameter space. We found that the Pearson correlation coefficient between realizations does not become zero when the sample separation is large, which is consistent with this observation. We also found that the variance of the ensemble-averaged spectrum decreases proportional to $N^{-5/6}$, which is once again consistent with the presence of correlations.

In the mid-wavelength region, we used spectrograms to show that the residual fluctuations are due to the limited number of solitons (~30) that create the spectrum. The bandwidth of a single realization is determined by the amplitude of the largest soliton and depends sensitively on the parameters. We found that the ensemble-averaged bandwidth is determined by a small number of large-bandwidth realizations that are correlated with rogue solitons.

We used these studies to determine that 5000 realizations is sufficient to obtain a converged bandwidth and spectrum that is expected to be an excellent predictor of the output spectrum seen in an experiment.

Acknowledgments