

# A comparative study of noisy signal evolution in 2R all-optical regenerators with normal and anomalous average dispersions using an accelerated Multicanonical Monte Carlo method

Taras I. Lakoba<sup>1\*</sup> and Michael Vasilyev<sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, University of Vermont, Burlington, VT 05401

<sup>2</sup>Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019

\*Corresponding author: [lakobati@cems.uvm.edu](mailto:lakobati@cems.uvm.edu)

**Abstract:** In [Opt. Express **15**, 10061 (2007)] we proposed a new regime of multichannel all-optical regeneration that required anomalous average dispersion. This regime is superior to the previously studied normal-dispersion regime when signal distortions are deterministic in their temporal shape. However, there was a concern that the regenerator with anomalous average dispersion may be prone to noise amplification via modulational instability. Here, we show that this, in general, is not the case. Moreover, in the range of input powers that is of interest for multichannel regeneration, the device with anomalous average dispersion may even provide less noise amplification than the one with normal dispersion. These results are obtained with an improved version of the parallelized modification of the Multicanonical Monte Carlo method proposed in [IEEE J. Sel. Topics Quantum Electron. **14**, 599 (2008)].

© 2008 Optical Society of America

**OCIS codes:** (060.2330) Pulse propagation and solitons; (060.4510) Optical communications; (060.5530) Fiber optics communications; (070.4340) Nonlinear optical signal processing; (230.1150) All-optical devices; (230.4320) Nonlinear optical devices.

---

## References and links

1. P. V. Mamyshev, "All-optical regeneration based on self-phase modulation effect," in *Proceedings of the 24th European Conference on Optical Communications (ECOC, Madrid, Spain, 1998)*, Vol. **1**, pp. 475–476.
2. L. Provost, C. Finot, P. Petropoulos, K. Mukasa, and D. J. Richardson, "Design scaling rules for 2R-optical self-phase modulation-based regenerators," *Opt. Express* **15**, 5100–5113 (2007).
3. M. Vasilyev and T. I. Lakoba, "All-optical multichannel 2R regeneration in a fiber-based device," *Opt. Lett.* **30**, 1458–1460 (2005).
4. T. I. Lakoba and M. Vasilyev, "A new robust regime for a dispersion-managed multichannel 2R regenerator," *Opt. Express* **15**, 10061–10074 (2007).
5. P. G. Patki, V. Stelmakh, M. Annamalai, T. I. Lakoba, and M. Vasilyev, "Recirculating-loop study of dispersion-managed 2R regeneration," in *Proceedings of the Conference on Lasers and Electro-Optics (CLEO, Baltimore, MD, 2007)*, paper CMZ3.
6. M. Vasilyev, T. I. Lakoba, and P. Patki, "Multiwavelength all-optical regeneration," in *Proceedings of the Optical Fiber Communications Conference (OFC, San Diego, CA, 2008)*, paper OWK3.
7. M. Vasilyev, P. G. Patki, and T. I. Lakoba, "All-optical regeneration of multi-wavelength signals," in *Proceedings of the IEEE LEOS European Winter Topical on Nonlinear Processing in Optical Fibres (Innsbruck, Austria, 2009)*.

8. C. Kouloumentas, L. Provost, F. Parmigiani, S. Tsolakidis, P. Petropoulos, I. Tomkos, and D. J. Richardson, "Four-channel all-fiber dispersion-managed 2R regenerator," *IEEE Photon. Technol. Lett.* **20**, 1169–1171 (2008).
9. N. J. Smith and N. J. Doran, "Modulational instabilities in fibers with periodic dispersion management," *Opt. Lett.* **21**, 570–572 (1996).
10. T. I. Lakoba, "BER degradation by a signal-reshaping processor with non-instantaneous response," *J. Lightwave Technol.* (submitted).
11. F. Öhman, S. Bischoff, B. Tromborg, and J. Mørk, "Noise and regeneration in semiconductor waveguides with saturable gain and absorption," *IEEE J. Quantum Electron.* **40**, 245–255 (2004).
12. T. I. Lakoba, "Multicanonical Monte Carlo study of the BER of an all-optically 2R regenerated signal," *IEEE J. Sel. Top. Quantum Electron.* **14**, 599–609 (2008).
13. B. A. Berg, "Algorithmic aspects of multicanonical simulations," *Nucl. Phys. B (Proc. Suppl.)* **63A–C**, 982–984 (1998); also at <http://www.arxiv.org/paper/hep-lat/9708003>.
14. D. Yevick, "Multicanonical communication system modeling — Application to PMD statistics," *IEEE Photon. Technol. Lett.* **14**, 1512–1514 (2002).
15. R. Holzlöhner and C. R. Menyuk, "Use of multicanonical Monte Carlo simulations to obtain accurate bit error rates in optical communication systems," *Opt. Lett.* **28**, 1894–1896 (2003).
16. It may be noted that the MMC was used in [I. Nasieva, A. Kaliazin, and S.K. Turitsyn, "Multicanonical Monte Carlo modeling of BER penalty in transmission systems with optical regeneration," *Opt. Commun.* **262**, 246–249 (2006)] to study signal transmission through a chain of regenerators. However, these regenerators were modeled by their static transfer functions, which required much less computing time than did solving a partial differential equation for pulse propagation, as we did in [12] and here.
17. B. Charbonnier, N. El Dahdah, and M. Joindot, "OSNR margin brought about by nonlinear regenerators in optical communication links," *IEEE Photon. Technol. Lett.* **18**, 475–477 (2006).
18. T. N. Nguyen, M. Gay, L. Bramerie, T. Chartier, and J.-C. Simon, "Noise reduction in 2R-regeneration technique utilizing self-phase modulation and filtering," *Opt. Express* **14**, 1737–1747 (2006).
19. M. Rochette, J. N. Kutz, J. L. Blows, D. Moss, T. J. Mok, and B. J. Eggleton, "Bit-error-ratio improvement with 2R optical regenerators," *IEEE Photon. Technol. Lett.* **17**, 908–910 (2005).
20. A. O. Lima, I. T. Lima, and C. R. Menyuk, "Error estimation in Multicanonical Monte Carlo simulations with applications to polarization-mode-dispersion emulators," *J. Lightwave Technol.* **23**, 3781–3789 (2005).
21. T.-H. Her, G. Raybon, and C. Headley, "Optimization of pulse regeneration at 40 Gb/s based on spectral filtering of self-phase modulation in fiber," *IEEE Photon. Technol. Lett.* **16**, 200–202 (2004).
22. C. Finot, T. N. Nguyen, J. Fatome, T. Chartier, S. Pitois, L. Bramerie, M. Gay, and J.-C. Simon, "Numerical study of an optical regenerator exploiting self-phase modulation and spectral offset filtering at 40 Gbit/s," *Opt. Commun.* **281**, 2252–2264 (2008).
23. T. Lu and D. Yevick, "Biased multicanonical sampling," *IEEE Photon. Technol. Lett.* **17**, 1420–1422 (2005).
24. D. Yevick and T. Lu, "Improved multicanonical algorithms," *J. Opt. Soc. Am. A* **23**, 2912–2918 (2006).
25. T. Lu and D. Yevick, "Efficient multicanonical algorithms," *IEEE Photon. Technol. Lett.* **17**, 861–863 (2005).

---

## 1. Introduction

All-optical regeneration has the recognized potential to increase the reach of fiber-optical transmission systems without expensive optical-to-electrical signal conversion. In this paper, we will focus on a particular regeneration scheme proposed by P. Mamyshev in [1], which has 2R (reamplification and reshaping) capabilities for the on-off-keying signal format. The Mamyshev regenerator operates by broadening the spectrum of an amplified pulse in a highly nonlinear fiber (HNLF), which is then followed by off-central filtering of a part of the signals spectrum. This device has been demonstrated [1] to improve the quality of both logical ONEs and ZEROs, and has been extensively studied (see, e.g., [2] and references therein) due to its simplicity and robustness. However, in order to become a *practical* alternative to optical-to-electrical signal converters, an all-optical regenerator must be capable of simultaneous processing of multiple wavelength-division multiplexing (WDM) channels. In [3], we proposed such a multichannel version of the Mamyshev regenerator. The key ingredient of our scheme was the dispersion map where the dispersion of each HNLF section was compensated by a periodic (in wavelength) group-delay device (PGDD), so as to reduce detrimental inter-channel nonlinear interaction. As most other implementations of the Mamyshev regenerator, our scheme used a normal (average) dispersion. However, in a later study [4], we pointed out that the filtering amplitude characteristic of a commercially available PGDD dictated the necessity of a different regime

where the average dispersion must be anomalous. This regeneration regime with anomalous average dispersion exhibited greater eye opening improvement and more robust performance than the original, normal-dispersion, regime found in [3]. We experimentally demonstrated that a dispersion-managed implementation of the Mamyshev regenerator can improve the signal degraded by a deterministic distortion (e.g., the linear or nonlinear cross-talk) for a single channel [5, 6] and for eight WDM channels [6, 7]. A four-channel operation of a different dispersion-managed version of the Mamyshev regenerator with anomalous average dispersion was also recently reported in [8].

Despite this demonstrated ability of the regenerator with anomalous average dispersion to reduce deterministic distortions, there was a concern that the amplified spontaneous emission (ASE), which was not considered in [4] and [6]–[8], could be strongly amplified in the anomalous-dispersion regime by modulational instability. An answer to that question is obscured by the fact that our regenerator uses a dispersion map rather than a uniform anomalous-dispersion fiber, and dispersion management may suppress modulational instability [9]. On the other hand, the results of [9] are obtained for a weakly nonlinear regime, while in a regenerator, the propagation is strongly nonlinear. The above question and observations motivate us to perform a numerical study to compare the noise amplification in the dispersion-managed Mamyshev regenerator with normal and anomalous average dispersions.

Let us now explain the approach that we used for this comparison. We distinguish between two “kinds” of noise: (i) the rare large-amplitude bursts that may cause errors if the receiver is placed immediately after the regenerator, and (ii) small-amplitude noise realizations that occur frequently but may cause errors only after accumulation over several regenerated sections of the transmission line. Different “metrics” are needed to assess amplification (or absence of such) of these two “kinds” of noise by a regenerator. Since the bursts (case (i)) are so large that they may cause detection errors, an adequate metric for them is the bit-error rate (BER) of the regenerated signal. Recently, we proved analytically [10] that any 2R regenerator with parameters considered for practical use would cause BER degradation if the receiver is placed immediately after the regenerator (for prior numerical evidence, see [11, 12]). Thus, to compare the amplification of noise bursts, we will numerically estimate which regenerator, that with normal or with anomalous dispersion, degrades the BER stronger.

For case (ii), i.e. for small-amplitude noise realizations that occur most frequently and whose amplitude is within one or two standard deviations, the adequate assessment metric is the  $Q$ -factor, defined, as usual, as

$$Q = \frac{V_{(1)} - V_{(0)}}{\sigma_{(1)} + \sigma_{(0)}}, \quad (1)$$

where  $V_{(0,1)}$  and  $\sigma_{(0,1)}$  are the averages and standard deviations of the received powers of ZEROS and ONES, respectively. That is, computing the  $Q$ -factors of the same signals processed by the regenerators with the two types of average dispersion will show which regenerator can better suppress small signal distortions. We note that, of course, this  $Q$ -factor is not straightforwardly related to the BER of the regenerated signal since the latter’s statistics is far from being Gaussian.

The numerical technique that we employ is a slightly improved version of the modification of the Multicanonical Monte Carlo method (MMC), which (the modification) was proposed in [12]. The original MMC [13, 14, 15] allows one to numerically compute the BER as low as  $10^{-10}$  (and lower) with only  $\sim 10^5$  simulations, without making any assumptions about the statistics of the signal. A hallmark feature of the MMC is that, in stark contrast to the standard Monte Carlo method, it produces noise realizations of *any* amplitude with approximately the same probability. This is ideally suited for our purpose of considering regeneration of both large- and small-amplitude noise. In [12], we proposed a modification of the MMC that paral-

lelizes this algorithm and hence allows one to simulate the required  $\sim 10^5$  noise realizations in a few hours rather than in a few days, as the original MMC would.

In this paper, we apply the method of [12] to theoretically study the effect of ASE on the performance of a dispersion-managed Mamyshev regenerator. More specifically, here we consider only the basic component of this complex problem, namely, the interaction of noise with a single pulse in a single channel. Even though such a reduced model cannot accurately predict the BER or the  $Q$ -factor of the regenerated signal that would come from a transmission system, it can still be used to answer the main question of this work: Is the ASE amplified stronger in the anomalous-dispersion regeneration regime than in the normal-dispersion one? We also note that the simplifications made in the above reduced model are not expected to qualitatively change an answer to this question when inter-channel and adjacent-pulse interactions are “turned back on”. Indeed, in [3] and [4], we already found regenerator parameters that minimize the adjacent-pulse interactions in a 128-bit-long pseudo-random sequence. In the present paper, we use similar parameters, for which these interactions will also be weak. As far as the interaction of the signal with the noise from other channels, it will be negligible due to severe phase-mismatch provided by the PGDD-enabled dispersion map of the regenerator.

The main part of this paper is organized as follows. In Section 2, we will briefly review the idea of the parallelized MMC, proposed in [12], and also discuss the interpretation of the results that will be obtained with it. In particular, we will explain that while a single regenerator placed immediately before the receiver is likely to degrade the BER, a chain of such regenerators distributed along the transmission line can still improve it. In Section 3, we will present the main results of this paper. Namely, we will show and discuss plots of the BER and  $Q$ -factor after the regenerator for different values of its parameters. In Section 4, we will summarize the results, and in the Appendix, will describe a slightly improved version of the parallelized MMC of [12], which we employed in this paper.

## 2. Calculation of the BER of the regenerated signal

We will first review the idea behind the numerical approach which was used to compute both the BER and the  $Q$ -factor of the regenerated signal. While the latter could be obtained with  $\sim 10^3$  standard Monte Carlo simulations, the former requires much more simulations. The original MMC algorithm, mentioned in the Introduction, allows one to compute very low BER values with only  $\sim 10^5$  simulations [14, 15]. However, that algorithm is intrinsically sequential, and it would take a few days to perform the required  $\sim 10^5$  simulations one after the other [16].

The key idea proposed in [12] was to split the procedure into two stages. In the first stage, one runs the MMC *without* the regenerator. Instead, here the noise-degraded signals are processed by the receiver, consisting of an optical bandpass filter (OBPF), photodetector, and electrical filter. Two separate MMC simulations are required, one for ONEs and one for ZEROs. Since one numerical run through the receiver is about two orders of magnitude faster than a run through the regenerator, this stage takes only a few hours (see Appendix). The outcome of this stage contains two pieces of information: (i) probability density functions (PDFs) of non-regenerated ONEs and ZEROs, from which the *input* BER can be computed, and (ii) a large set of *distinct* samples of noisy signal whose received powers *uniformly* fill up the range  $[P_{\min}, P_{\max}]$  where the corresponding PDF is computed. We emphasize that by the nature of the MMC algorithm [13], a significant number of such samples can be collected even in those regions where the PDF is very small. In the Appendix, we describe a modification to the sample collection procedure that improves the one we originally proposed in [12]. Next, in the second stage of our parallelized version of the MMC, the collected noisy samples are processed by the regenerator and then detected by the receiver. The resulting distribution of power values is recorded and used, along with the PDFs of the non-regenerated signal computed in the first

stage, to find the PDFs and hence the BER of the regenerated signal. We will refer to it as the *output* BER. The  $Q$ -factor can be computed using these same data and Eq. (1). Since each signal sample can be run through the regenerator independently from all other samples, the second stage of our procedure can be performed by several computers simultaneously. This results in the overall reduction of computing time from several days to just several hours.

In [12], we used the above numerical technique to demonstrate that the Mamyshev regenerator with a particular set of parameters degrades the BER if the receiver is placed immediately after it. Later, we showed analytically [10] that such BER degradation is a generic phenomenon. To avoid possible confusion, let us note that while a single regenerator degrades the BER, a chain of regenerators distributed along the transmission line can still improve it. Quantitatively, a related result was shown, e.g., in [17], where a single regenerator was placed at some distance before the receiver. To explain this in qualitative terms, let us also consider only one regenerator and, as an additional simplification, suppose that it is placed in the middle of the transmission line where the signal degradation occurs only from the accumulated ASE. First, consider the case with no regenerator present. Given the Gaussian statistics of the noise in this non-regenerated system, it can be straightforwardly shown that the noise bursts that cause errors at the receiver are most likely to occur as a superposition of two smaller bursts of approximately the same amplitude, where one of these “partial bursts” accumulates before, and the other after, the location of the regenerator. Now consider the same system with the regenerator. The regenerator redistributes the noise in such a way that small deviations of the signal from the undistorted ZERO or ONE become even smaller (thus, the eye opening improves). In terms of the PDF, this means that small deviations become less likely. Then, since the area under the PDF must always remain unchanged, large signal deviations must become more likely. A boundary between these “small” and “large” deviations occurs near the decision threshold; see, e.g., Fig. 3 in [17]. Therefore, the aforementioned “partial bursts” of noise qualify as “small” deviations. Then, the first “partial burst” is reduced by the regenerator, and the sum of this reduced “partial burst” and the second, unchanged, “partial burst” is smaller than such a sum in the absence of the regenerator. Consequently, the probability of a detection error is reduced, and the BER is improved.

Thus, in our comparison of regenerators with anomalous and normal average dispersions, we will look at which of the regenerators degrades the BER less and which one improves the  $Q$ -factor more. The “winners” in these two categories will then be declared as providing less amplification to the large- and small-amplitude “kinds” of noise, respectively.

### 3. Results

We considered four regenerators whose output-vs-input power transfer curves are shown in Fig. 1. The parameters of these regenerators are similar to those reported in [4]. Specifically, the regenerators’ dispersion maps contain sections of HNLF (commercial dispersion-compensating fiber) with  $D = -120$  ps/nm/km and  $\gamma = 5$  (W · km)<sup>-1</sup> followed by PGDDs. Two of the regenerators contain 16 map periods with 0.5-km HNLF sections, while the other two have 5 map periods with 1.25-km sections. The corresponding average dispersions are  $-5$ ,  $-2$ ,  $+12$ , and  $+30$  ps/nm/km. We selected these parameters in such a way that the transfer curve of a regenerator with anomalous average dispersion would be similar (at least partially) to the transfer curve of another regenerator with normal average dispersion. Thus, the pairs of similar curves occur for  $D_{av} = -5$  and  $+12$  ps/nm/km and for  $D_{av} = -2$  and  $+30$  ps/nm/km. As we will show below, the similarity of transfer curves does not guarantee a similarity of the small-amplitude noise suppression by the regenerators. Yet, the deterministic amplitude jitter (arising, e.g., from a linear or nonlinear cross-talk) will be suppressed similarly by the regenerators with similar transfer curves. Therefore, it is reasonable to compare how such regenerators suppress other

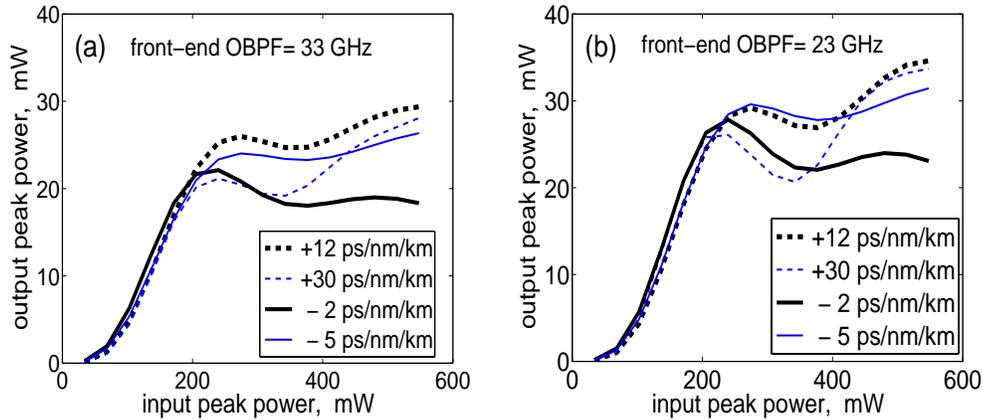


Fig. 1. Static power transfer curves for the four regenerators described in the text. As usual, a static transfer curve is obtained by propagating through the regenerator pulses with the same shape (determined by the shape of the pulse before the front-end OBPF and this OBPF's bandwidth) and variable amplitude. The front-end OBPF's bandwidth is referenced at full-width at half-maximum (FWHM).

kinds of signal distortions.

Continuing with the description of the regenerator parameters, the PGDDs are modeled to have constant (i.e., non-filtering) amplitude characteristics, because a filtering amplitude characteristic would destroy regenerative properties of the devices with normal average dispersion [4]. Immediately in front of the regenerator, we place an OBPF with a 3rd-order super-Gaussian profile whose role is to limit the amount of ASE into the regenerator. (The effect of such a filter's bandwidth on the  $Q$ -factor and BER was previously investigated in [18, 12].) We will refer to this filter as the front-end OBPF, to avoid confusion with the OBPF that carves out the signal's spectrum at the output of the regenerator. The latter filter is offset by 30 GHz from the channel's center and is 13.3-GHz-wide Gaussian, so as to reproduce the width (33 ps, or 33% of the bit slot) of the input pulse measured *before* the front-end OBPF. The input power referenced in Fig. 1 and later on is the peak power measured *after* the front-end OBPF. The receiver placed after the regenerator consists of a photodetector followed by a 8-GHz wide 4th-order Bessel electrical filter. In that part of our simulations where the regenerator is not used (see Section 2), the aforementioned front-end OBPF is placed in front of the receiver. In accordance with our approach whereby we study the noise interaction only with a single pulse, the simulated bit sequence is 01000000, with infinite extinction ratio between ZEROs and ONE. The pulse propagation through the regenerator is modeled by the nonlinear Schrödinger equation without any higher-order effects and for a single polarization of the signal, as in [3] and [4]. Unlike those papers, however, here we model the ASE, with the optical signal-to-noise ratio (OSNR) in one polarization being 17 dB per 0.1 nm. The losses of HNLf and PGDDs were neglected in our simulations, since accounting for such losses does not qualitatively change the dynamics of the signal in the regenerator [4].

In Fig. 2(a), we plot the BERs at the output of all four regenerators versus the FWHM of the front-end OBPF. The input peak power to the regenerator is 350 mW for all data points in this figure. In Fig. 2(b), we plot corresponding data versus the input peak power, while the front-end OBPF is fixed to be 23-GHz wide. As can be seen, the curves in both panels of Fig. 2 coincide within the statistical error (see Appendix to this paper and Appendix B to [12]). Thus, we conclude that large-amplitude noise bursts are not amplified any stronger in the regenerators with anomalous average dispersion than they are in such devices with normal average dispersion having similar power transfer functions. A partial explanation for this will be presented at the end of this Section.

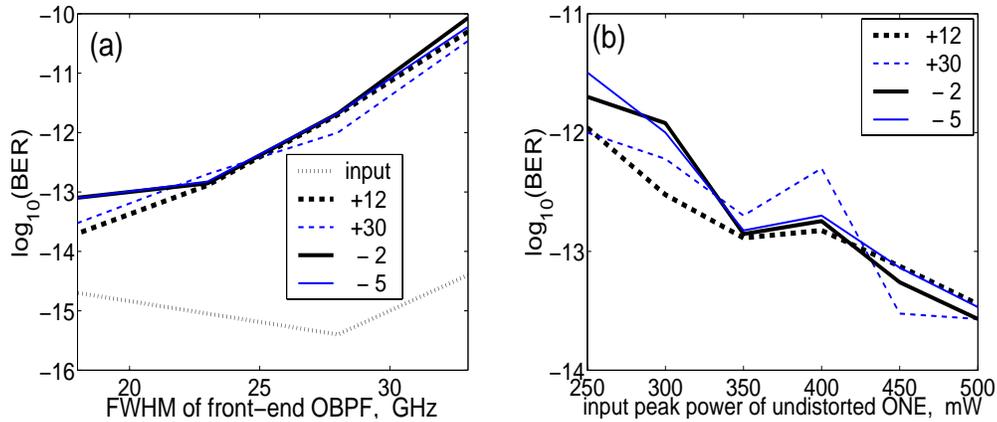


Fig. 2. BER of the regenerated signal plotted versus the front-end OBPF's bandwidth (a) and the input power (b). Other parameters are described in the text above, and the values of the average dispersion (in ps/nm/km), distinguishing the regenerators, are briefly stated in the legends. In (a), the dotted line shows the BER of the signal when the regenerator is absent. Since for the fixed OSNR, this value of BER does not change with the input power, the corresponding curve (i.e., a horizontal line at  $\approx 10^{-15}$ ) is not shown in panel (b).

In Fig. 3, we plot the  $Q$ -factor for the same four regenerators versus the FWHM of the front-end OBPF (panels (a), (b)) and versus the input peak power (panels (c), (d)). Similarly to Fig. 2, the input power is fixed to 350 mW for Figs. 3(a,b) and the OBPF's bandwidth is fixed to 23 GHz for Figs. 3(c,d). These  $Q$ -factors are obtained from Eq. (1) and the PDFs of the regenerated signal. Note that we distinguish between the PDFs computed with the MMC and those inferred from the static transfer curves (for an example of both types of PDFs, see, e.g., Fig. 5 of [12]); the corresponding  $Q$ -factors are shown in Figs. 3(a,c) and 3(b,d), respectively. (For the BER values reported in Fig. 2, counterparts of Figs. 3(b,d) are irrelevant, because the BER computed with inferred PDFs equals the input BER [19].) Let us note that the statistical error (due to different seeds of the random number generator) of the  $Q$ -factor is about 0.1 dB, i.e. much smaller in the relative sense than the statistical error of the BER values reported in Fig. 2 (see Fig. 5 in Appendix). This occurs because the statistical error of the MMC is much greater at the tails of the PDF (which are used to compute the BER) than at the central part of the PDF (which is used to compute the  $Q$ -factor) [20].

Although there is much less obvious structure to the data shown in Fig. 3 than to the data shown in Fig. 2, we can still make a few conclusions. First, addressing our main concern about possible amplification of small-amplitude noise, we can see that there is no evidence that such amplification is consistently stronger in a regenerator with anomalous average dispersion than it

is in a similar device with normal average dispersion. In fact, the  $Q$ -factor of regenerators with  $D_{\text{av}} = +12$  and  $+30$  ps/nm/km for *smaller* input powers, which would be preferred in multichannel operation to minimize inter-channel impairments, is higher than the corresponding values for the regenerators with  $D_{\text{av}} = -2$  and  $-5$  ps/nm/km. Second, we note that while there is some correlation between the curves obtained with computed and inferred PDFs and plotted *versus input power*, there appears to be no such correlation for the corresponding two sets of curves plotted *versus the OBPF's bandwidth*. Moreover, except for the device with  $D_{\text{av}} = +30$  ps/nm/km, which has very low inferred  $Q$ -factor (probably because its transfer curve is the least flat), the inferred  $Q$ -factor values are considerably higher than the actual ones. Finally, the maximum improvement of the computed  $Q$ -factor — seen to be about 2 dB from Figs. 3(a,c) — is consistent with the experimental [21] and numerical [18] values reported earlier for different parameters of the regenerator.

As we noted above, for smaller input powers located near the “knee” of the static transfer curve, the  $Q$ -factor is higher for regenerators with anomalous than with normal average dispersion, given that the two devices have similar transfer curves. Since it is this range of powers that is optimal for multichannel operation [4], we can conclude that the regime with anomalous dispersion that we found in [4] is not prone to noise enhancement via modulational instability. Furthermore, to gain some intuition as to why the  $Q$ -factor is higher in regenerators with anomalous dispersion, we inspected the transfer “bands” of the regenerators with  $D_{\text{av}} = -5$  ps/nm/km and  $+12$  ps/nm/km (see Fig. 4). Such transfer bands are obtained in the second stage of our modified MMC method described in Section 2. Namely, recall that in the first stage, for each small interval  $[P_{\text{in}}, P_{\text{in}} + \Delta P]$  in the considered range of input powers, we have collected distinct noisy samples whose powers (in the absence of the regenerator) fall into that interval. In the second stage, we propagate these samples through the regenerator. Since the ASE causes the shapes of individual signal samples to be (slightly) different, then, for each interval  $[P_{\text{in}}, P_{\text{in}} + \Delta P]$ , one obtains a *range* of output powers, shown in Fig. 4 with red (for ONEs) and green (for ZEROs) contours. Comparison of the transfer bands for ONEs near the input power of 300 mW in panels (a) and (b) immediately reveals a reason behind the lower values of the  $Q$ -factor for the regenerator with  $D_{\text{av}} = -5$  ps/nm/km: The “bulge” seen in panel (a) implies that the variance  $\sigma_{(1)}$  of regenerated ONEs is considerably greater for this regenerator. We verified that a similar bulge also exists in the regenerator with  $D_{\text{av}} = -2$  ps/nm/km, while it is absent in the one with  $D_{\text{av}} = +30$  ps/nm/km. An understanding as to why such a bulge occurs in regenerators with normal, but not with anomalous, average dispersion requires further study.

Parenthetically, let us note that a glance at the transfer band in Fig. 4(a) also reveals that the  $Q$ -factor of the corresponding regenerator should be maximum when the input peak power of a noiseless ONE is about 450–500 mW. Indeed, there, the transfer band is the narrowest. Figure 3(c) confirms this prediction about the  $Q$ -factor. (We verified that beyond the input power of 500 mW, the  $Q$ -factor for this regenerator decreases). Note that for a regenerator with normal average dispersion, the  $Q$ -factor is maximum for input powers that are higher than the values corresponding to the flattest part of the plateau of the static transfer curve (i.e., near 350 mW in Fig. 4(a)). This fact was observed and discussed earlier in [18, 22].

Concluding with the interpretation of Fig. 4, we note that it also explains why the BERs were almost the same for all four regenerators considered (see Fig. 2). First, the input BER was the same for all regenerators. In [10], it was shown that the BER degradation immediately after the regenerator is determined by the width of the transfer band near the decision threshold, i.e. at the steeply rising part of the band. These widths appear to be very similar in the left and right panels of Fig. 4, whence the BER values must also be similar.

In addition to comparing the four regenerators described above, where the PGDD was assumed to have no amplitude filtering, we also computed the BER and  $Q$ -factor of a regenerator

Computed by Monte Carlo

Inferred from transfer curves

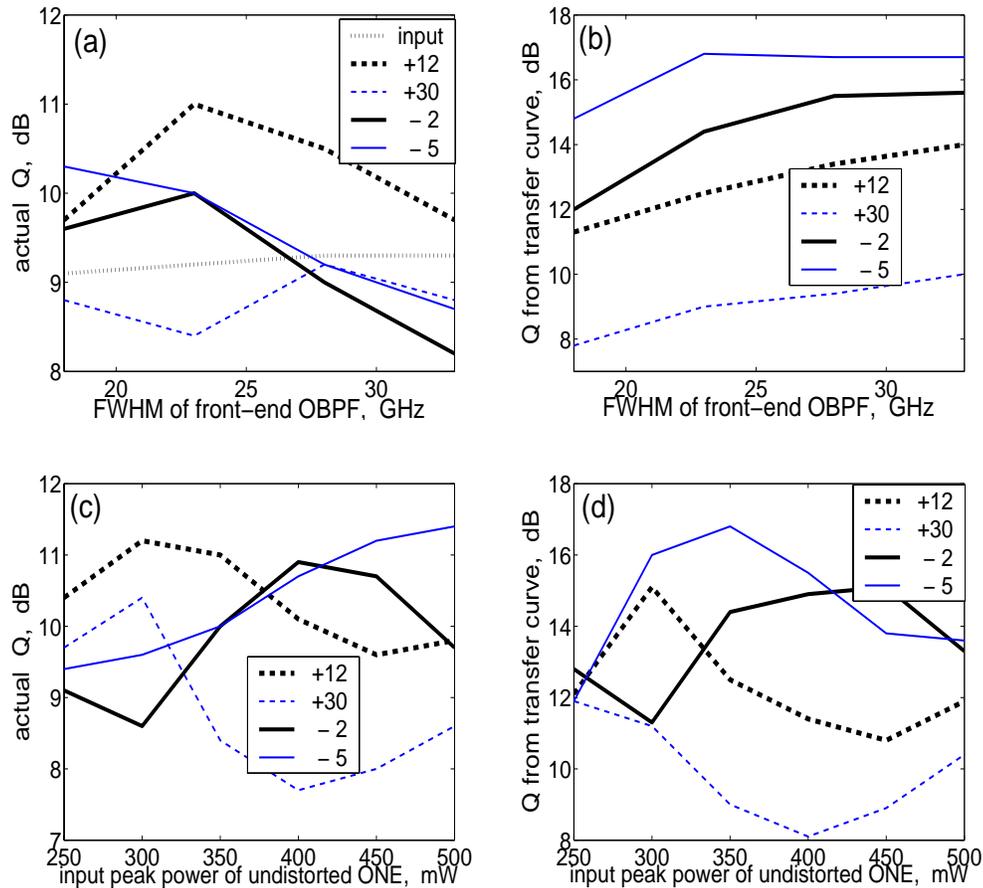


Fig. 3.  $Q$ -factor of the regenerated signal plotted versus the front-end OBPF's bandwidth (a,b) and the input power (c,d). Panels (a,c) and (b,d) are obtained with PDFs computed by the MMC and inferred from static transfer curves (see Fig. 1), respectively. Other relevant parameters are described in the text, and the values of the average dispersion (in ps/nm/km), distinguishing the regenerators, are briefly stated in the legends. In (a), the dotted line shows the  $Q$ -factor of the signal when the regenerator is absent. In (b), this input  $Q$ -factor curve is not shown since it is the same as in (a). Moreover, since for the fixed OSNR, the  $Q$ -factor does not change with the input power, the corresponding horizontal line at 9.2 dB is not shown in panels (c,d).

whose PGDD did have a filtering characteristic with parameters similar to those that were found to be optimal for multichannel operation in [4]. Namely, this regenerator had  $D_{av} = +15$  ps/nm/km and a PGDD with a 90-GHz 3rd-order Gaussian amplitude characteristic shifted from the channel center by 20 GHz. The OBPF at the output of the regenerator was shifted by 25 GHz (to the same side as the PGDD's characteristic), with all the other parameters being the same as listed at the beginning of this section. The results that we found are very similar to those for the regenerator with  $D_{av} = +12$  ps/nm/km described above. Thus, while the filtering

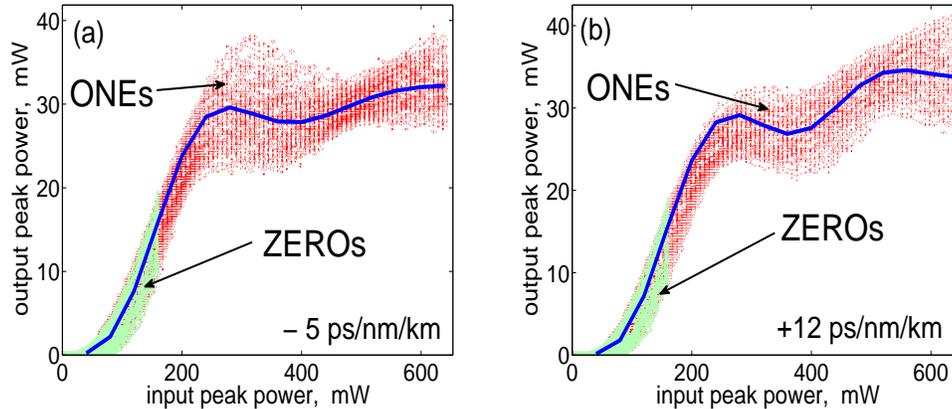


Fig. 4. Transfer bands of the regenerators with  $D_{av} = -5$  ps/nm/km (a) and  $+12$  ps/nm/km (b). The FWHM of the front-end OBPF is 23 GHz, and a ONE with no noise has the input peak power of 400 mW. Static transfer curves are shown by thick blue lines.

by the PGDD is sufficiently strong to destroy regenerative properties of a device with normal average dispersion [4], it is not strong enough to affect the evolution of a noisy signal in the regenerator.

#### 4. Conclusions

The main conclusion of this paper is that the regeneration regime with anomalous average dispersion proposed in [4] does not exhibit any stronger noise amplification (e.g., via modulational instability) than the regime with normal average dispersion. Specifically, by numerically simulating the BER in both regimes, we inferred that the evolution of large-amplitude noise bursts in these regimes is similar; see Fig. 2. To compare the evolutions of small-amplitude noise, we computed the corresponding  $Q$ -factors. Here the results depend on both the bandwidth of the front-end OBPF and the input power of the undistorted ONE; see Fig. 3(a,c). However, for the multichannel regeneration, smaller values of the input power are preferable, so as to minimize inter-channel nonlinear interaction [4]. For the regenerators that we compared in this study with such smaller input powers, the noise on the background of a ONE is amplified more in regenerators with normal than with anomalous average dispersion; see Figs. 3(c) and 4. Thus, we conclude that the particular regeneration regime found in [4] shows better multichannel performance than the regime found earlier in [3] not only for deterministic, but also for random, signal distortions.

In Appendix, we present an improved version of the parallelized modification of the MMC proposed in [12]. In that modification, one first collects noisy samples by the standard MMC and then propagates them through the regenerator to obtain the PDFs. The time taken by the sample collection stage was highly sensitive to the initial guess of the PDF in the standard MMC. In the improved version presented in this work, we employ biased multicanonical sampling to

accelerate the sample collection stage and to greatly reduce its sensitivity to the choice of the initial PDF.

### Appendix: Improved collection of noisy samples via biased multicanonical sampling

Here we describe an improved procedure of collecting *distinct* noisy samples in the first stage of the modified MMC method proposed in [12]. As mentioned in that paper, the speed of such collection was quite sensitive to the initial PDF, which one assumes at the first iteration of the MMC. While our improved algorithm presented herein addresses two different issues, its main thrust is to make the collection speed essentially independent of the choice of the initial PDF.

For the reader's convenience and future reference, we begin by briefly summarizing the steps of the first stage of the modified MMC method described in Section 2. These steps are very similar to those of the original MMC [13, 14, 15]. Let  $\hat{\mathbf{z}} = (\hat{z}_1, \dots, \hat{z}_L)$  be the discrete Fourier spectrum of the numerically generated noise. The  $\hat{z}_i$ 's are independent random quantities having identical Gaussian PDFs. The noise electric field  $\mathbf{z}(t)$  is added to the field  $\mathbf{s}(t)$  of the data and the resulting signal  $\mathbf{s} + \mathbf{z}$  is sent through the receiver described in Section 3. The signal power at the output of the receiver is denoted as  $P(\mathbf{z})$ . Let us subdivide the considered range of powers,  $[P_{\min}, P_{\max}]$ , into small bins  $[P^{(k)}, P^{(k+1)}]$ ,  $k = 1, \dots, K$ . The MMC involves several iterations. At the beginning of each iteration, reset all entries  $H(k)$  of the histogram to zero. Then  $H_n(k)$  at the end of the  $n$ th iteration will indicate how many times the received power  $P(\mathbf{z})$  has been accepted (see below) into bin  $k$  during this iteration. For the first sample in each iteration, generate a random noise realization  $\mathbf{z}$  and compute  $P(\mathbf{z})$ . If that value falls into bin  $k$ , accept it into that bin and hence let  $H_n(k) = 1$ . Denote the Fourier spectrum of this noise realization by  $\hat{\mathbf{z}}_a$ . To obtain each subsequent realization of noise within the  $(n + 1)$ th iteration, do the following.

- 1) Generate a random vector

$$\hat{\mathbf{z}}^{\text{new}} = \hat{\mathbf{z}}^{\text{old}} + \Delta\hat{\mathbf{z}}, \quad (2)$$

where each component of  $\Delta\hat{\mathbf{z}}$  is a symmetrically distributed random variable independent of all the other components.

- 2) For the  $\mathbf{z}^{\text{new}}$  constructed above, compute the received power  $P(\mathbf{z}^{\text{new}})$ . Suppose that it falls into bin  $m$ , and recall that  $P(\mathbf{z}^{\text{old}})$  belongs to bin  $k$ . Accept this random realization with the probability

$$\min(\text{pdf}_n(k)/\text{pdf}_n(m), 1), \quad (3)$$

where  $\text{pdf}_n(k)$  is the approximation to the sought PDF in the  $k$ th bin computed at the  $n$ th iteration (see Eqs. (4), (5) below). At the first iteration, some initial  $\text{pdf}_0$  needs to be assumed. If  $P(\mathbf{z}^{\text{new}})$  is accepted, increment  $H_{n+1}(m)$  by 1 and then rename  $\mathbf{z}^{\text{old}} = \mathbf{z}^{\text{new}}$  and  $k = m$ . If  $P(\mathbf{z}^{\text{new}})$  is rejected, increment  $H_{n+1}(k)$  by 1 and keep  $\mathbf{z}^{\text{old}}$  and  $k$  the same as in the previous realization.

- 3) In addition, if  $P(\mathbf{z}^{\text{new}})$  is accepted, save the signal realization  $\mathbf{s} + \mathbf{z}^{\text{new}}$  so as to use it in the second stage of our procedure (see Section 2).
- 4) Repeat the previous steps a number of times specified for each iteration. At the end of the  $(n + 1)$ th iteration, use the accumulated histogram  $H_{n+1}$  to update the PDF according to the formula [13]:

$$\widetilde{\text{pdf}}_{n+1}(1) = 1, \quad \frac{\widetilde{\text{pdf}}_{n+1}(k+1)}{\widetilde{\text{pdf}}_{n+1}(k)} = \frac{\text{pdf}_n(k+1)}{\text{pdf}_n(k)} \cdot \left( \frac{H_{n+1}(k+1)}{H_{n+1}(k)} \right)^{g_{n+1,0}/g_{n+1}} \quad (4)$$

for  $k = 1, \dots, K - 1$ , and then use

$$\text{pdf}_{n+1}(k) = \widetilde{\text{pdf}}_{n+1}(k) / \sum_{m=1}^K \widetilde{\text{pdf}}_{n+1}(m). \quad (5)$$

The notations  $g_{n+1,0}$  and  $g_{n+1}$  will be explained later.

- 5) Terminate the iterations when the number of collected distinct random realizations of the signal in *all* bins reaches a specified number,  $N_{\text{collected}}$ .

As we noted in the first paragraph of this Appendix, here we present two modifications to the above steps. The first modification concerns which noisy samples we consider “distinct” and subsequently record (see step 3), so as to propagate them through the regenerator later on. Recall that any  $\mathbf{z}^{\text{new}}$  (see (2)) that was accepted by the MMC algorithm was also recorded, as it is distinct from  $\mathbf{z}^{\text{old}}$ . However, a concern here is that the increment  $\Delta \hat{\mathbf{z}}$  of the noise could happen to be “too small” to make  $\hat{\mathbf{z}}^{\text{new}}$  “sufficiently different” from  $\hat{\mathbf{z}}^{\text{old}}$ . To prevent this from happening, we recorded a new noise realization provided that it satisfied either of the two conditions: (i)  $\sum_{i=1}^L |\Delta \hat{z}_i \hat{\kappa}_i| \geq A \cdot L |\max(\Delta \hat{z}) \max(\hat{\kappa})|$ , or (ii) a new realization has been accepted, but not recorded (since it failed to satisfy (i)) a number of  $B$  times. Here  $L$  is the number of simulated Fourier components of the noise and  $\hat{\kappa}$  is the amplitude characteristic of the front-end OBPF (which limits the amount of noise entering the regenerator). In the simulations reported in this work, we chose  $A = 0.3$  and  $B = 3$ . While imposition of these criteria ensured that the noise realizations recorded are “more distinct” than in our original procedure [12], it also made the process of collection of such “more distinct” realizations slower and even more sensitive to a “good” choice of the initial PDF  $\text{pdf}_0$  at the first iteration of the MMC.

To overcome these sensitivity and slowness, we made a second, and a key, modification to the procedure of [12]. Our idea was to use biased multicanonical sampling [23, 24]. Recall that a transition from the “old” to a “new” noise realization is accepted with probability given by (3). That is, the MMC biases noise realizations towards those regions where  $\text{pdf}_n(P)$  is small, and it is this bias that makes the algorithm visit each bin  $[P^{(k)}, P^{(k+1)}]$  with equal probability (as  $n \rightarrow \infty$ ) [13]. However, as we explained above, not each such visit results in an acceptance of the noise realization, and, given the modification described in the previous paragraph, even more rarely does it result in recording that realization. Therefore, we need to use additional bias to make the algorithm more frequently visit those bins where the number of recorded samples is small. This can be easily done by employing in (3) a product  $\text{pdf}_n(k) F_n(k)$  instead of  $\text{pdf}_n(k)$ , where  $F_n(k)$  can be taken simply proportional to the number of samples recorded in bin  $k$  from the first to the  $n$ th iterations of the MMC. However, then formula (4) for updating the PDF needs to be modified, and below we derive its modified form. Our derivation follows the lines of the original derivation by Berg [13] and of [25].

Let us redenote  $x_k \equiv P^{(k)}$ . From now on, subscripts  $k$  and  $(k + 1)$  will pertain to the bin number, and subscripts  $n$  and  $(n + 1)$  will pertain to the iteration number. We also denote  $\Delta \equiv x_{k+1} - x_k$  and assume that  $\Delta \ll 1$ . Since we are looking for PDFs that decay at an exponential rate, we introduce a function  $S_n(x)$  such that  $\text{pdf}_n(x) = \exp[-S_n(x)]$ . One key idea of Berg’s derivation is that the PDF is assumed to be a smooth function of  $x$ . Then we can write

$$\frac{\text{pdf}_n(x_{k+1})}{\text{pdf}_n(x_k)} = e^{-(S_n(x_{k+1}) - S_n(x_k))} \approx e^{-b_n(x_k)\Delta}, \quad (6)$$

where  $b_n(x) = dS_n(x)/dx$ . Our goal is to find  $b_{n+1}(x)$ . Then, taking an arbitrary value for  $\text{pdf}_{n+1}(x_1)$ , we can compute  $\text{pdf}_{n+1}(x_k)$  for all  $k$  from (6) and then normalize the PDF as in (5). As we will see, a nontrivial part of this problem will be to determine correct relative statistical weights of  $b_n$  and  $b_{n+1}$ , so as to account for the statistical information obtained not just at the  $(n + 1)$ th iteration, but also at all the preceding iterations.

To this end, we first obtain an unweighed approximation to the PDF at the  $(n+1)$ th iteration [23]:

$$\text{pdf}_{n+1,0}(x_k) = c_n \text{pdf}_n(x_k) H_{n+1}(x_k) F_n(x_k), \quad (7)$$

where  $c_n$  is a normalization constant. The only difference from earlier derivations [13, 25] is in the factor  $F_n(x_k)$ . Taking the logarithm of (7) and then differentiating the result with respect to  $x$ , we obtain, with the same accuracy as in (6):

$$b_{n+1,0}(x_k) = b_n(x_k) - \frac{1}{\Delta} \ln \frac{F_n(x_{k+1})}{F_n(x_k)} - \frac{1}{\Delta} \ln \frac{H_{n+1}(x_{k+1})}{H_{n+1}(x_k)}. \quad (8)$$

The statistical weight of  $b_{n+1,0}$  that we are trying to determine is inversely proportional to its variance,  $\sigma^2[b_{n+1,0}]$ , which occurs due to a *finite* number of random realizations during the  $(n+1)$ th iteration. (Here and below we omit the argument  $x_k$  or  $x_{k+1}$  when this does not lead to a confusion.) From (8) it is seen that the only contributor to this variance is the histogram  $H_{n+1}$ . Using the facts that: (i) for  $H_{n+1}^2 \gg \sigma^2[H_{n+1}]$ , one has  $\sigma^2[\ln H_{n+1}] \approx \sigma^2[H_{n+1}]/H_{n+1}^2$ , and (ii)  $\sigma^2[H_{n+1}] \approx \chi H_{n+1}$ , where  $\chi$  is some constant, from (8) we deduce that

$$\sigma^2[b_{n+1,0}(x_k)] = \frac{\chi}{\Delta^2} \left( \frac{1}{H_{n+1}(x_k)} + \frac{1}{H_{n+1}(x_{k+1})} \right). \quad (9)$$

We have also assumed that the fluctuations of  $H_{n+1}(x_k)$  and  $H_{n+1}(x_{k+1})$  are statistically independent. Equation (9) simply says that the more counts are in bin  $k$ , the more reliable is the information about the PDF in that bin.

Next, as we noted earlier, the final  $b_{n+1}$  at this iteration is a weighed average of  $b_{n+1,0}$  and  $b_n$ , where the latter has been computed at the previous iteration:

$$b_{n+1} = \alpha b_{n+1,0} + (1 - \alpha) b_n. \quad (10)$$

The factor  $\alpha$  is to be chosen so as to minimize the variance (and hence maximize the statistical weight) of  $b_{n+1}$ . Recall that the variance in question of  $b_{n+1,0}$  occurs due to a *finite* sample size during the  $(n+1)$ th iteration alone, and the variance of  $b_n$  has contributions due to the finite sample sizes of the  $n$ th and preceding iterations. Therefore, these variances are independent (contrary to what one may naively deduce from (8)), and hence from (10) one has:

$$\sigma^2[b_{n+1}] = \alpha^2 \sigma^2[b_{n+1,0}] + (1 - \alpha)^2 \sigma^2[b_n]. \quad (11)$$

The minimum of (11) is found from  $\partial \sigma^2[b_{n+1}]/\partial \alpha = 0$ , whence

$$\alpha_{\min} = \frac{\sigma^2[b_n]}{\sigma^2[b_{n+1,0}] + \sigma^2[b_n]}; \quad (12)$$

$$\sigma^2[b_{n+1}] \stackrel{\text{def}}{=} \min(\sigma^2[b_{n+1}]) = \frac{\sigma^2[b_{n+1,0}] \cdot \sigma^2[b_n]}{\sigma^2[b_{n+1,0}] + \sigma^2[b_n]}. \quad (13)$$

If we now define statistical weights  $g_n(x_k) = 1/\sigma^2[b_n(x_k)]$ ,  $g_{n+1,0}(x_k) = 1/\sigma^2[b_{n+1,0}(x_k)]$ , etc., then from (13) we obtain a recursion relation

$$g_{n+1} = g_n + g_{n+1,0}. \quad (14)$$

Equations (14) and (9) allow one to recursively compute the statistical weight  $g_{n+1}$  given a value for  $g_0$ ; since we assign no weight to the information found in the first iteration, we set

$g_0 = 0$ . Note that the value of  $\chi$  in (9) is unimportant, so for the sake of convenience one can set  $\chi = \Delta^2$ . Now, substituting (12) into (10), one has:

$$b_{n+1} = \frac{g_{n+1,0}}{g_n + g_{n+1,0}} b_{n+1,0} + \frac{g_n}{g_n + g_{n+1,0}} b_n. \quad (15)$$

Finally, we use (8) to eliminate  $b_{n+1,0}$  in (15). Substituting the result into (6), one obtains a recursion relation for  $\text{pdf}_{n+1}$  that generalizes (4):

$$\frac{\text{pdf}_{n+1}(x_{k+1})}{\text{pdf}_{n+1}(x_k)} = \frac{\text{pdf}_n(x_{k+1})}{\text{pdf}_n(x_k)} \cdot \left( \frac{F_n(x_{k+1})}{F_n(x_k)} \frac{H_{n+1}(x_{k+1})}{H_{n+1}(x_k)} \right)^{g_{n+1,0}/g_{n+1}}. \quad (16)$$

As a biasing function, we used a slightly more complex expression than just the vector containing numbers of recorded noisy samples per bin. Namely, once this number in the least populated bin has reached 25% of the targeted number,  $N_{\text{collected}}$ , of samples to be recorded, we implemented the following Matlab commands:

```
F_aux = N_recorded_samples.^3;   coef=polyfit(bins,F_aux,3);
F=max([ ones(size(F_aux)); min([F_aux; polyval(coef,bins)]) ]);
```

In the second line, the second argument of the max selects the smaller of  $F_{\text{aux}}$  and its smoothed replica, while the first argument ensures that  $F$  always remains positive.

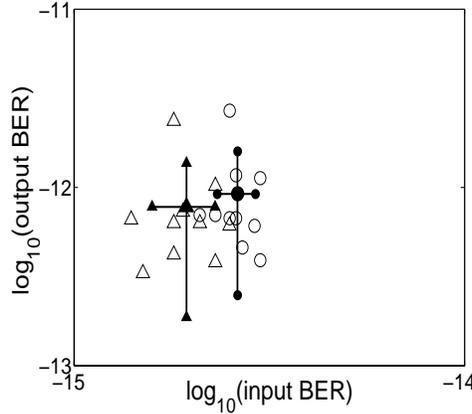


Fig. 5. BER computed without (circles) and with (triangles) acceleration of the sample collection via biasing, as described in the text. Large filled symbols show the mean values and the lines with small filled symbols at their ends show the corresponding standard deviations about these means.

In Fig. 5 we show the BER values computed with and without this biasing for the regenerator with  $D_{\text{av}} = -5$  ps/nm/km described in the main text. The FWHM of the front-OBPF and the input peak of the undistorted ONE were taken as 23 GHz and 350 mW, respectively. There are ten data points for each case (i.e., with and without biasing), which differ only by the seed of the random number generator. Figure 5 shows that a difference between the BERs computed with and without this biasing is within the statistical error of the computed BERs. Without the biasing, collection times of noisy samples for ONEs varied between 2.5 to 6.5 hours, while with

the biasing these times were reduced by factors between 2 and 4. Typically, a greater speed-up was achieved for a longer collection time. For ZEROs, collection times were considerably shorter (about 40 minutes), and little gain in collection speed was achieved by the biasing. Thus, the biasing procedure presented above removes the main drawback of the modified MMC proposed in [12]. Namely, with this biasing, the time of collecting a specified number of noisy samples becomes fairly insensitive to the choice of the initial PDF that needs to be assumed in the first stage of that modified MMC.

### **Acknowledgment**

This work was supported in part by NSF grants DMS-0507429 and DMS-0507540.