Review Article　　　　　　　　　　　　　　　　　**Open Access**

George C. Knee*, Joshua Combes, Christopher Ferrie, and Erik M. Gauger

Weak-value amplification: state of play

Abstract: Weak values arise in quantum theory when the result of a weak measurement is conditioned on a subsequent strong measurement. The majority of the trials are discarded, leaving only very few successful events. Intriguingly those can display a substantial signal amplification. This raises the question of whether weak values carry potential to improve the performance of quantum sensors, and indeed a number of impressive experimental results suggested this may be the case. By contrast, recent theoretical studies have found the opposite: using weak-values to obtain an amplification generally worsens metrological performance. This survey summarises the implications of those studies, which call for a reappraisal of weak values' utility and for further work to reconcile theory and experiment.

Weak measurements vs. weak values

A quantum weak measurement is a procedure whereby only a little bit of information about a quantum system is obtained; as a consequence, the system is only disturbed a little. This is in contrast to the usual strong measurements, which give a lot of information but inflict a large disturbance on the system.

*Corresponding Author: George C. Knee:** Department of Materials, University of Oxford, Oxford OX1 3PH, United Kingdom, E-mail: gk@physics.org

Joshua Combes: Center for Quantum Information and Control, University of New Mexico, Albuquerque, New Mexico, 87131-0001 USA, ARC Centre for Engineered Quantum Systems, School of Mathematics and Physics, The University of Queensland, St Lucia, QLD 4072, Australia

Christopher Ferrie: Center for Quantum Information and Control, University of New Mexico, Albuquerque, New Mexico, 87131-0001 USA

Erik M. Gauger: SUPA, Institute of Photonics and Quantum Sciences, Heriot Watt University, Edinburgh EH14 4AS, United Kingdom; Department of Materials, University of Oxford, Oxford OX1 3PH, United Kingdom

Imagine the needle on a poor-quality analogue voltmeter, which twitches or deflects in response to an electrical signal. In a weak measurement, the amount of deflection is only loosely correlated with the true voltage—because the needle also twitches about randomly. The expected value of the deflection, however, is precisely the true voltage.

Physically, a weak measurement can be implemented by introducing a weak interaction between the system of interest and an ancillary meter degree of freedom [1, 2]. Consider the following interaction Hamiltonian between system and meter:

$$H = gA \otimes P \,, \tag{1}$$

where g is a scalar quantity, A is the operator associated with the relevant system observable, and P is an operator effecting a shift of the meter variable. The dynamics induced by this Hamiltonian (we assume it is switched on and off again instantaneously) will build up correlations between the system and meter. The degree of correlations, or the 'measurement strength' can be controlled for example by changing g; with strong and weak limits being attained when $g \to \infty$ and $g \to 0$, respectively. Owing to these correlations, a subsequent measurement on the meter alone may reveal information about the system. The back-action imparted to the system can be understood by examining the effective measurement operators of the system (sometimes called POVM, or positive operator values measure elements [3]): strong measurement operators are projective, giving maximal information but also imparting the largest back-action. Intermediate strength measurements impart less back-action but give less information. In the limit of a vanishing interaction, the POVM elements become proportional to the identity matrix, and no measurement is performed at all. As long as a measurement of some finite strength is performed, however, the average deflection will reveal the true voltage with increasing precision.

Weak measurements have become part of the standard toolbox in the modern field of quantum control: they can be a useful method of stabilising a quantum computation, where information must be prevented from leaking into the environment [4]. If a weak measurement is repeated on a single system enough times to provide the same informa-

tion as a strong measurement, a comparable back action will be imparted.
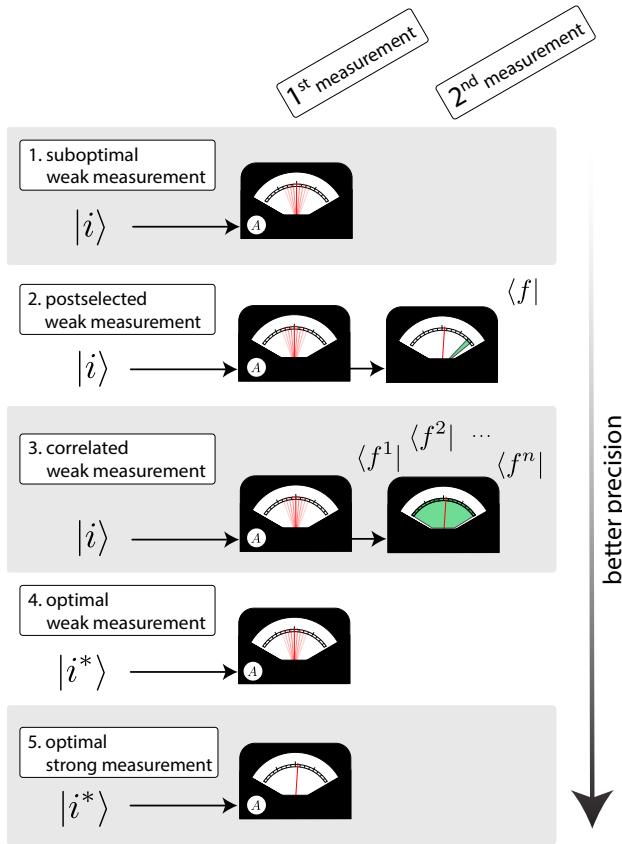
Figure 1: Various schemes for parameter estimation. The experimenter can choose the initial state $|i\rangle$, the amount of uncertainty in the measurement of observable A (depicted here by a fluctuating meter needle), and also whether to correlate the first measurement with a second one (which projects onto a final states $|f^i\rangle$). The green areas on the second meter depict a value that must occur for the experiment to be successful – otherwise it is rejected. It has recently been shown that a postselected weak measurement (2.) will not give more information than three simple alternatives, despite its association with large meter readings known as 'anomalous weak values'. It is better to either (3.) keep all of the data from both measurements, or (4.) prepare an optimal initial state $|i^{\star}\rangle$ and dispense with the second measurement altogether. A single strong measurement (5.) gives the most information of all. Scheme (1.) involves a suboptimal initial state and no conditioning, giving the worst performance of all.

The weak value is different to a weak measurement, but often defined in conjunction. A weak value, like an expected value, is a well-defined statistical quantity. It arises from applying standard quantum mechanics to a particular measurement protocol involving pre and postselection [5, 6]. The procedure to obtain a weak value is explained in detail in Ref. [7], but we sketch the idea here (see

Figure 1, case 2). First, the system of interest is prepared in a known initial state (e.g., a predetermined voltage) $|i\rangle$. It is measured weakly (for example, by using a poor quality voltmeter coupled to observable A) and then measured again, this time strongly (with a good quality meter). The first and second measurement are generally described by complementary observables. Finally, the data from the poor voltmeter are culled, a step known as postselection; only those instances where the second measurement reported a particular unlikely result are kept (for example the result corresponding to final state $|f\rangle$). The real part of the weak value

$$A_w = \frac{\langle f|A|i\rangle}{\langle f|i\rangle} \qquad (2)$$

is then (approximately) proportional to the expected value of the surviving data from the first measurement. When the reading on the second meter is a very rare one, the weak value can become quite large, and in particular larger in magnitude than the largest eigenvalue $\lambda^{\star} = \max |A|$ of the observable. Intuitively, the anomalously large deflection seems to indicate the potential to increase or amplify the precision of measuring devices—why live with a small signal when a large one can be arranged? Intuition, however, is notoriously unreliable.

The cost of amplification

The weak-value approximately determines the magnitude of needle deflection, and it can be larger than the expected value

$$\langle A \rangle = \langle i|A|i\rangle \leq \lambda^{\star} , \qquad (3)$$

which provides an upper bound on the deflection when no second measurement is performed, whilst in the postselected case $|A_w| > \lambda_{\star}$ is possible. This has motivated many experiments to convert tiny effects into larger ones using postselection, usually with the aim of estimating the coupling between two quantum degrees of freedom. The first such study was Onur Hosten and Paul Kwiat's 2008 experiment [8], which detected the spin-Hall effect of light (a coupling between the polarisation and transverse momentum of light at an interface between media with different refractive indices) with postselected weak measurements. As with many experiments with weak values, the polarisation of a beam of light played the role of the quantum system and the deflection of the beam replaced the twitching of the meter needle – other notable experiments include the measurement of a femtoradian mirror tilt [9] and other

small optical effects. However, the technique applies generally to different degrees of freedom in optical as well as in a range of other physical systems.

The recent experimental successes being reported lend weight to the longstanding question whether weak value amplification can deliver a fundamental advantage for parameter estimation, or whether it should merely be regarded as a convenient experimental tool in certain circumstances. In the last year, a number of researchers working independently have addressed the question of whether weak values do indeed unlock superior performance, by employing the rigorous framework of parameter estimation. Before describing the formalism of parameter estimation in order to understand those results, let us discuss the meaning of postselection and give two intuitive reasons (which can be made mathematically rigorous) why the amplification provided by weak values may not be a 'silver bullet' for precision measurements.

The first reason is that the data are necessarily extremely noisy; for the measurement to truly qualify as weak, the needle on the measuring device is continually wandering under quantum fluctuations. Any systematic deflection of the needle is, by design, hidden by the fundamental quantum uncertainty in any given run. Detecting the signal necessitates the use of a statistical approach. In any weak measurement, postselected or otherwise, a very large number of trials is vital for a significant conclusion to be reached. Only a strong measurement can provide a precise estimate after a single trial, when all sources of noise are eradicated.

The second reason is that the anomalously large deflections are very rare. The larger the 'amplification' that is desired, the more trials are required before an experiment succeeds. In optical experiments, a low success probability translates into a much reduced photon detection rate. Therefore, the effect leads to an *attenuation* as much as it leads to an amplification. This latter point is known as the problem of low postselection success probability.

It is important to distinguish two notions of 'postselection': On the one hand, it can be understood as a physical step – for example the inclusion of a polarising filter which only allows certain photons to reach the detector. The data from the first measurement are then only recorded *after* the second measurement has triggered the amplification. On the other hand, postselection may be understood as the rejection of certain events from a larger dataset – the amplification is then an artefact of data processing only. This distinction matters as the two cases lead to different ways of benchmarking the weak value technique, which can in turn give rise to slightly different conclusions.

Under the first definition, the weak value amplification approach and its standard benchmark strategy correspond to (slightly) different physical setups and protocols: one compares the use of postselection (case 2) against not using a second measurement at all (case 4 in Figure 1). This is the approach taken by Refs. [10–13].

Taking the second definition, the experimental setup remains identical (i.e. a second strong measurement is always performed) and postselection is understood as partitioning the recorded data into two sets, only one of which will be used for the parameter estimation process at hand. Then, one effectively compares case 2 with case 3 in Figure 1. This notion of postselection is employed by Refs. [15–18].

Parameter estimation

Gathering and interpreting data is the very essence of empirical science. For results to be meaningful, they should correspond accurately to those predicted by a theoretical model, and a statement about their uncertainty must be made. When writing up the results of an experiment, one states the *estimate* of the measured quantity along with an *uncertainty* – the voltage was 5.0V ±0.1V, for example. Both numbers are the output of statistical calculations. Consider estimating the voltage of a constant signal using a noisy voltmeter. The true voltage is denoted by x_{true}. Typically one looks at the voltmeter and records the (random) value x_i, the needle deflection in the i'th experiment. An *estimator* is a function of the data that produces an estimate for the true voltage. Often, the average of x over N trials is a good choice, $x_{\text{est}} = (1/N) \sum_i^N x_i$.

If the experimenter happens to be sitting at an angle to the voltmeter dial, she might consistently over- or underes-
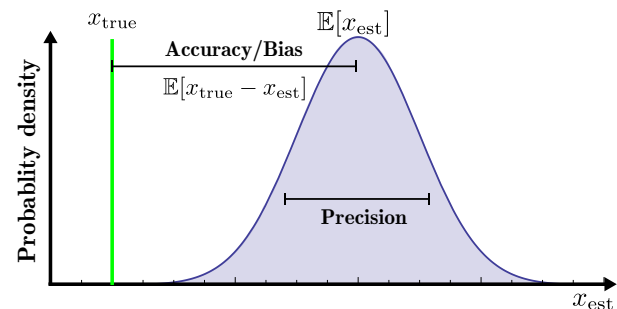
Figure 2: An unbiased estimator has the property that in an infinite number of trials its average value will equal the true value. The quality of an unbiased estimator is characterized by the variance of the estimator. The smaller the variance the more precise the estimator is.

timate the deflection. This will result in a biased estimate of the voltage. An unbiased estimator, on the other hand, is perfectly accurate on average $\mathbf{E}[x_{\text{true}} - x_{\text{est}}] = 0$, where the symbol $\mathbf{E}$ denotes the expected value. Aside from accuracy, the quality of the estimation procedure is characterized by the precision, defined as the variance of the estimator. See Figure 2 for a simple illustration.

In many instances it can be difficult to directly calculate the precision theoretically. A mathematical tool known as the Fisher information [19, 20] allows one to place an lower bound on the variance of an estimator. The Fisher information gives a single number F, providing a powerful link between theory and experiment. If one processes the data from the experiment with the best possible, or optimal estimator, the variance will be given by the inverse of the Fisher information. That means the experimenter will report e.g. $5.0\text{V} \pm (1/\sqrt{NF})\text{V}$. Clearly a higher F is better because it implies a lower uncertainty in the estimate. The magnitude of the Fisher information depends on exactly how the experiment is performed. It thus provides an excellent way of comparing different approaches to parameter estimation. The dependence on N means that, as long as the experiment is repeatable, the uncertainty can become arbitrarily small by increasing the number of trials.

An estimation inequality

Applying the above framework of parameter estimation to the case of weak value amplification, several recent theoretical studies have reached the conclusion that, given the same number of input resources, a weak-value strategy will generally not outperform the standard metrology strategy [12, 15–18, 21]. These results are essentially all captured in an inequality constraining the corrected Fisher information:

$$p(\checkmark)F_{\text{weak value}} \leq F_{\text{standard}} , \qquad (4)$$

where $p(\checkmark)$ is the postselection success probability. As discussed above, the standard Fisher information could be one of a number of alternatives (see Figure 1). For example keeping all of the data (rather than discarding most of it) will give higher Fisher information [15, 16, 22]; similarly, choosing an optimal initial state and not performing the second measurement at all will outperform a postselected weak measurement [12]. A strong measurement, if available, provides the highest precision of all [12, 15–18]. Furthermore, a simple estimator based upon the approximate weak value (rather than the true average of the

postselected weak measurement) is not unbiased: there is a systematic error in the estimate in any real experiment. Accuracy and precision are thus both worse when postselection is used.

It is possible, however, for $p(\checkmark)F_{\text{weak value}}$ to approach equality with F_{standard} in a restricted parameter regime. With weak-values, a simple estimator can match the accuracy of the standard strategy only when the measurement is infinitely weak, and the precision can match the standard strategy only when the postselection probability is zero. Neither of these is physically possible, although both can be approximated with arbitrary closeness: the fact that only a negligible amount of information is discarded through postselection has been confirmed in the laboratory [14]. It is therefore interesting to consider how the two approaches compare away from the ideal scenario. Below we list some of the ongoing investigations.

Estimation in non-ideal scenarios

The true limits to parameter estimation in the laboratory stem from a variety of sources other than the inherent quantum uncertainty. Are there scenarios where the above results do not apply, and where an advantage can be found?

Getting lucky

Sometimes $F_{\text{weak value}} \geq F_{\text{standard}}$ and postselection might temporarily provide more information, because the inequality in (4) only applies when the number of trials is large. However, the probability of a windfall persisting decreases exponentially as more trials are performed [16, 17].

Technical noise

Experiments exploiting different physics can exhibit varying robustness to noise and detector imperfections. This fact has motivated a long standing feeling that weak-value amplification can help to overcome such problems. However, we have shown that estimation inequalities analogous to (4) hold in the presence of a broad class of noise before, during, and after the weak measurement [12, 16, 17, 21]. Together these articles treat the most prevalent types of noise (including pixelation, detector jitter, and dephasing), and a similar approach can be used to analyse other imperfections. For instance, Jordan *et al.* have claimed that for specific noise models a weak-value-amplified tech-

nique gives higher Fisher information than conventional methods [13]. By engineering these noise models artificially, these predictions have been explored experimentally [14].

Suboptimal estimation strategies

Some researchers question the use of optimal estimators, and therefore question the relevance of the Fisher information to actual experiments. A recent manuscript imagines a type of noise that cannot be modelled statistically, and hence affects the uncertainty in the estimate in a more profound way [23]. Refs. [13, 24] argue that time-correlated noise, for example, renders the optimal estimator impractical due to computational complexity. This is not, by itself, a compelling argument for employing weak values, however: one must show that the entire weak value protocol, including the post processing, is not outperformed by a suitable benchmark strategy. Further, it has been shown that a computationally trivial estimator exists for time correlated noise and outperforms the weak-value-amplification technique [16]. If one has chosen a suboptimal initial state $|i\rangle \neq |i^{*}\rangle$, then performing an optimal postselection $\langle f^{*}|$ can lead to more (corrected) Fisher information than not performing the second measurement at all (case 1) [11] – otherwise, when $|i^{*}\rangle$ is chosen, the Fisher information of case 4 achieves the theoretical maximum and cannot be surpassed by any pre and postselected strategy, including an optimal one.

Imaginary weak values

Imaginary weak values occur when the postselected weak measurement is performed in a certain way: the amplification is then seen in Fourier space, rather than real space.

Whether or not imaginary weak values merit special consideration depends on what is meant by 'postselection'. Comparing cases 2 and 4 in Figure 1, i.e. protocols with different *physics*, there may be scope for altering the apparatus in a fashion which tailors it to better fit the available hardware. In a time-domain experiment, for example, a frequency analyser is sometimes preferred to a stopwatch [10, 13]. However, unless there is a severe mismatch between the quality of detection in the two variables, imaginary weak values will not provide a significant advantage [12]. On the other hand, when case 2 is compared to case 3 in Figure 1, then postselection is merely data rejection, which cannot improve estimation under the most general evolution allowed by quantum theory [17]. This lat-

ter analysis therefore covers experiments involving any defined quantities, including imaginary and even complex weak values.

The true cost of estimation

It is interesting to consider the different notions of the cost associated with an investigation. One must spend time and energy to perform the experiment; there is a financial cost accompanying the hardware; and a computational cost associated with data processing and estimation.

Weak value estimators have been conjectured to offer a computationally simpler alternative than standard techniques, despite the extra apparatus required [13]. Although this is an appealing idea, standard methods are often equally cheap to compute as the weak-value estimator whilst being more precise [16]. For example, for case 4 in Figure 1 the optimal estimator is simply proportional to the average measurement result. One should bear in mind that performing unbiased estimation with weak values requires more complicated postprocessing: one must account for non-linear effects not captured in the weak value, which is only an approximate quantity.

Another type of cost arises uniquely in optics experiments. Lasers can easily emit 10^{10} photons per nanosecond, making the creation-cost per photon almost negligible. The detection-cost of a photon, by contrast, is often effectively much higher, especially if the photodetector saturates very quickly [25]. A variant accounting philosophy, which weighs output resources more heavily than input resources, may give rise to a different conclusion to the one reached above [26, 27]. Interestingly, it is in exactly these special circumstances (those of large numbers of cheap input photons) that the weak-value phenomenon is known to have a classical explanation. By contrast, genuinely quantum-enhanced metrology typically exploits effects involving *single quanta* [28], and then the number of input resources becomes the limiting factor.

Experimental runs which fail the postselection test can be "recycled" for a further interaction with the unknown parameter [29], and this procedure repeated until they eventually pass. If the runs (typically thought of as photons) are reused in this way but are not *recounted* as a resource, the technique is superior to a one-interaction-per-photon method. This is of no great surprise, because the interaction with the unknown parameter is precisely when the information is imparted to the probe, and thus increasing the total number of interactions is worthwhile. Recycling is therefore a technique to increase Fisher information given a fixed laser input power, for example. It is

of course possible to recycle in the 'standard' strategies, by rerouting all photons (not just those that failed postselection).

Finally it is worth noting that experiments are frequently performed far from the ultimate performance limit. Some decisions are made in the laboratory merely for convenience, for example to cope with incidental factors. From this point of view, weak-value amplification may be judged 'handy' on a case-by-case basis.

Combination with other quantum effects

Another direction that has begun to be investigated is combining weak value approaches with more established quantum metrology methods (see e.g. [30–32]); for example with entanglement [33] or squeezing [11]. Both these effects are known to unlock a genuine quantum advantage for certain metrological tasks, but at this point it is not clear whether combining them with weak values can indeed lead to anything greater than the sum of its parts.

Conclusion

For a number of years, experimental successes involving the weak-value technique have lent credence to the intuition that an 'amplified' signal is always a good thing. By contrast, recent theoretical results cast doubt on whether weak value techniques can really offer a fundamental quantum advantage for metrology. However, the debate continues since the conclusion one reaches depends on various subtleties such as the the resource accounting philosophy, the choice of benchmark, and the notion of postselection employed. In any case, however, we here submit that, by itself, a large amplification factor is not sufficient for an advantage in parameter estimation.

References

[1] A. Barchielli, L. Lanz, and G. M. Prosperi, Il Nuovo Cimento B **72**, 79 (1982).

[2] C. M. Caves and G. J. Milburn, Phys. Rev. A **36**, 5543 (1987).

[3] M. A. Nielsen, I. L. Chuang, Quantum Computation and Quantum Inforamtion, 1st edition (Cambridge University Press, Cambridge, 2000)

[4] G. G. Gillett, R. B. Dalton, B. P. Lanyon, M. P. Almeida, M. Barbieri, G. J. Pryde, J. L. O'Brien, K. J. Resch, S. D. Bartlett, A. G. White, Phys. Rev. Lett. **104**, 080503 (2010)

[5] Y. Aharonov, D. Z. Albert, L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988)

[6] I. M. Duck, P. M. Stevenson, E. C. G. Sudarshan, Phys. Rev. D **40**, 2112 (1989)

[7] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Rev. Mod. Phys. **86**, 307 (2014)

[8] O. Hosten, P. Kwiat, Science **319**, 787 (2008)

[9] P. Dixon, D. J. Starling, A. N. Jordan, J. C. Howell, Phys. Rev. Lett. **102**, 173601 (2009)

[10] Y. Kedem, Phys. Rev. A **85**, 060102 (2012)

[11] S. Pang, T. A. Brun, Phys. Rev. Lett. **115**, 120401 (2015)

[12] G. C. Knee, E. M. Gauger, Phys. Rev. X **4**, 011032 (2014)

[13] A. N. Jordan, J. Martinez-Rincon, and J. C. Howell, Phys. Rev. X **4**, 011031 (2014)

[14] G. I. Viza, J. Martínez-Rincón, G. B. Alves, A. N. Jordan, J. C. Howell, Phys. Rev. A **92**, 032127 (2015)

[15] S. Tanaka, N. Yamamoto, Phys. Rev. A **88**, 042116 (2013)

[16] C. Ferrie, J. Combes, Phys. Rev. Lett. **112**, 040406 (2014)

[17] J. Combes, C. Ferrie, Z. Jiang, and C. M. Caves, Phys. Rev. A **89**, 052117 (2014)

[18] L. Zhang, A. Datta, I. A. Walmsley, Phys. Rev. Lett. **114**, 210801 (2015)

[19] Van Trees, Harry L. Detection, estimation, and modulation theory. John Wiley & Sons, 2004.

[20] L. Wasserman, All of Statistics: A Concise Course in Statistical Inference, (Springer Texts in Statistics. Springer, 2004)

[21] G. C. Knee, G. A. D. Briggs, S. C. Benjamin, E. M. Gauger, Phys. Rev. A **87**, 012115 (2013)

[22] G. Strübi, C. Bruder, Phys. Rev. Lett. **110**, 083605 (2013)

[23] J. Lee, I. Tsutsui, Quantum Stud.: Math. Found. **1**, 65 (2014)

[24] A. N. Jordan, J. Tollaksen, J. E. Troupe, J. Dressel, Y. Aharonov, Quantum Stud.: Math. Found. **2**, 5 (2015)

[25] L. Vaidman, arXiv:1402.0199v1 [quant-ph]

[26] J. Dressel, T. A. Brun, A. N. Korotkov, Phys. Rev. A **91**, 040301(R) (2014)

[27] J. Combes and C. Ferrie, Phys. Rev. A **92**, 022117 (2015)

[28] V. Giovannetti, S. Lloyd, L. Maccone, Science **306**, 1330 (2004)

[29] J. Dressel, K. Lyons, A. N. Jordan, T. M. Graham, and P. G. Kwiat, Phys. Rev. A **88**, 023821 (2013)

[30] M. Paris, Int. J. Quant. Inf. 7, 125 95 (2009)

[31] G. Tóth, I. Apellaniz, J. Phys. A: Math. Theor. **47**, 424006 (2014)

[32] R.-Demkowicz-Dobrzanski, M. Jarzyna, J. Kolodynski, Progress in Optics **60**, 345 (2015)

[33] S. Pang, J. Dressel, T. A. Brun, Phys. Rev. Lett. **113**, 030401 (2014)