Supplemental Material

Shengshi Pang[1,*] Justin Dressel[2,†] and Todd A. Brun[1‡]

[1] *Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA and*

[2] *Department of Electrical Engineering, University of California, Riverside, CA 92521, USA.*

This is a supplement for "Entanglement-assisted weak value amplification," wherein we provide the derivations for the technical results.

I. DERIVATION OF THE MAXIMUM POST-SELECTION PROBABILITY

To maximize $P_s \approx |\langle\Psi_f|\Psi_i\rangle|^2$ while keeping A_w and $|\Psi_i\rangle$ fixed, we note that the initial state can be decomposed into a piece parallel to $(\hat{A}-A_w)|\Psi_i\rangle$ and an orthogonal piece in the complementary subspace $\mathcal{V}^\perp$:

$$|\Psi_i\rangle = \frac{(\hat{A}-A_w)|\Psi_i\rangle\langle\Psi_i|(\hat{A}-A_w^*)|\Psi_i\rangle}{\langle\Psi_i|(\hat{A}-A_w^*)(\hat{A}-A_w)|\Psi_i\rangle} + \left(|\Psi_i\rangle - \frac{(\hat{A}-A_w)|\Psi_i\rangle\langle\Psi_i|(\hat{A}-A_w^*)|\Psi_i\rangle}{\langle\Psi_i|(\hat{A}-A_w^*)(\hat{A}-A_w)|\Psi_i\rangle}\right). \tag{1}$$

Since $|\Psi_f\rangle$ must also be in $\mathcal{V}^\perp$ by the definition of the weak value, it follows that the maximum P_s can be achieved for the post-selection state parallel to the component of $|\Psi_i\rangle$ in $\mathcal{V}^\perp$, i.e.,

$$|\Psi_f\rangle \propto |\Psi_i\rangle - \frac{(\hat{A}-A_w)|\Psi_i\rangle\langle\Psi_i|(\hat{A}-A_w^*)|\Psi_i\rangle}{\langle\Psi_i|(\hat{A}-A_w^*)(\hat{A}-A_w)|\Psi_i\rangle}. \tag{2}$$

After some calculation, it follows that

$$\max_{|\Psi_f\rangle\in\mathcal{V}^\perp} P_s = \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{\langle\Psi_i|\hat{A}^2|\Psi_i\rangle - 2\langle\Psi_i|\hat{A}|\Psi_i\rangle\mathrm{Re}A_w + |A_w|^2}, \tag{3}$$

where $\mathrm{Var}(\hat{A})_{|\Psi_i\rangle} = \langle\Psi_i|\hat{A}^2|\Psi_i\rangle - [\langle\Psi_i|\hat{A}|\Psi_i\rangle]^2$ is the variance of $\hat{A}$ in the state $|\Psi_i\rangle$.

For the purposes of weak value amplification, we usually require $|A_w|$ to be larger than any eigenvalue of $\hat{A}$, $|A_w| \gg |\Lambda|$. Therefore, this maximum P_s can be approximated as Eq. (9) in the main text.

II. DERIVATION OF THE OPTIMAL POST-SELECTION STATE

As noted in the previous section, the optimal post-selection state should be parallel to the component of $|\Psi_i\rangle$ in $\mathcal{V}^\perp$. The post-selection probability is then controlled by the variance $\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}$. This variance is maximized for a maximally entangled initial state $|\Psi_i\rangle = \frac{1}{\sqrt{2}}(|\lambda_{\max}\rangle^{\otimes n} + \mathrm{e}^{\mathrm{i}\theta}|\lambda_{\min}\rangle^{\otimes n})$. Hence, we can directly compute the optimal post-selected state to be

$$\begin{aligned}
|\Psi_f\rangle &\propto |\Psi_i\rangle - \frac{(\hat{A}-A_w)|\Psi_i\rangle\langle\Psi_i|(\hat{A}-A_w^*)|\Psi_i\rangle}{\langle\Psi_i|(\hat{A}-A_w^*)(\hat{A}-A_w)|\Psi_i\rangle} \\
&= \frac{1}{\sqrt{2}}(|\lambda_{\max}\rangle^{\otimes n} + \mathrm{e}^{\mathrm{i}\theta}|\lambda_{\min}\rangle^{\otimes n}) - \frac{1}{\sqrt{2}}((n\lambda_{\max}-A_w)|\lambda_{\max}\rangle^{\otimes n} \\
&\qquad + \mathrm{e}^{\mathrm{i}\theta}(n\lambda_{\min}-A_w)|\lambda_{\min}\rangle^{\otimes n})\frac{n\lambda_{\max}+n\lambda_{\min}-2A_w^*}{|n\lambda_{\max}-A_w|^2+|n\lambda_{\min}-A_w|^2} \\
&\propto (|n\lambda_{\min}-A_w|^2 - (n\lambda_{\max}-A_w)(n\lambda_{\min}-A_w^*))|\lambda_{\max}\rangle^{\otimes n} \\
&\qquad + \mathrm{e}^{\mathrm{i}\theta}(|n\lambda_{\max}-A_w|^2 - (n\lambda_{\min}-A_w)(n\lambda_{\max}-A_w^*))|\lambda_{\min}\rangle^{\otimes n}) \\
&\propto -(n\lambda_{\min}-A_w^*)|\lambda_{\max}\rangle^{\otimes n} + \mathrm{e}^{\mathrm{i}\theta}(n\lambda_{\max}-A_w^*)|\lambda_{\min}\rangle^{\otimes n}.
\end{aligned} \tag{4}$$

This is Eq. (12) in the main text.

III. QUANTUM FISHER INFORMATION

It is important to determine just how well the weak value amplification technique can estimate the small parameter g. There is some concern that the post-selection process will lead to a substantial reduction of the total obtainable information, since a large fraction of the potentially usable data is being thrown away (e.g., [1–3]). To assuage these concerns, we compare the maximum Fisher information about g that can be obtained without post-selection to the Fisher information that remains in the post-selected states used for weak value amplification.

We first recall a few general results from the study of quantum Fisher information. If one wishes to estimate a parameter g, then the minimum standard deviation of any unbiased estimator for g is given by the *quantum Cramér-Rao bound*: $I(g)^{-1/2}$. The function $I(g)$ is the *quantum Fisher information* [4]

$$I(g) = 4\frac{\mathrm{d}\langle\Phi_g|}{\mathrm{d}g}\frac{\mathrm{d}|\Phi_g\rangle}{\mathrm{d}g} - 4\left|\frac{\mathrm{d}\langle\Phi_g|}{\mathrm{d}g}|\Phi_g\rangle\right|^2, \tag{5}$$

which is determined by a quantum state $|\Phi_g\rangle$ that contains the information about g. If this state is prepared with some interaction Hamiltonian $|\Phi_g\rangle = \exp(-ig\hat{H})|\Phi\rangle$ then the Fisher information reduces to a simpler form [5]

$$I(g) = 4\mathrm{Var}(\hat{H})_{|\Phi\rangle}, \tag{6}$$

and is entirely determined by the variance of the Hamiltonian in the pre-interaction state $|\Phi\rangle$.

A. General Discussion

In the main text, the relevant Hamiltonian with a meter observable $\hat{F}$ is $\hat{H} = \hbar g\hat{A}\otimes\hat{F}\delta(t-t_0)$, where $\hat{A}$ is a sum of n ancilla observables $\hat{a}$ of dimension d. The joint state $|\Phi\rangle$ is also always prepared in a product state $|\Phi\rangle = |\Psi_i\rangle\otimes|\phi\rangle$ between the ancillas and the meter. If there is no post-selection then the quantum Fisher information is found to be

$$I(g) = 4\left[\langle\hat{A}^2\rangle_{|\Psi_i\rangle}\langle\hat{F}^2\rangle_{|\phi\rangle} - \left(\langle\hat{A}\rangle_{|\Psi_i\rangle}\langle\hat{F}\rangle_{|\phi\rangle}\right)^2\right]. \tag{7}$$

Now suppose we projectively measure the ancillas in order to make a post-selection. This measurement will produce d^n independent outcomes corresponding to some orthonormal basis $\{|\Psi_f^{(k)}\rangle\}_{k=1}^{d^n}$. In the linear response regime with $g \ll 1$, each of these outcomes prepares a particular meter state

$$|\phi'_k\rangle \propto \langle\Psi_f^{(k)}|\exp(-ig\hat{H})|\Psi_i\rangle|\phi\rangle \approx (\hat{1} - igA_w^{(k)}\hat{F})|\phi\rangle \tag{8}$$

with probability $P_s^{(k)} \approx |\langle\Psi_f^{(k)}|\Psi_i\rangle|^2$ that is governed by a different weak value

$$A_w^{(k)} = \frac{\langle\Psi_f^{(k)}|\hat{A}|\Psi_i\rangle}{\langle\Psi_f^{(k)}|\Psi_i\rangle}. \tag{9}$$

We can then compute the remaining Fisher information contained in each of the post-selected states $\sqrt{P_s^{(k)}}|\phi'_k\rangle$ using (5), which produces

$$I^{(k)}(g) \approx 4\,P_s^{(k)}|A_w^{(k)}|^2\left[\mathrm{Var}(\hat{F})_{|\phi\rangle} - \langle\hat{F}^2\rangle_{|\phi\rangle}\left(2g\mathrm{Im}A_w^{(k)}\langle\hat{F}\rangle_{|\phi\rangle} + |gA_w^{(k)}|^2\langle\hat{F}^2\rangle_{|\phi\rangle}\right)\right]. \tag{10}$$

Importantly, if we add the information from all d^n post-selections we obtain

$$\sum_{k=1}^{d^n} I^{(k)}(g) \approx 4\,\langle\hat{A}^2\rangle_{|\Psi_i\rangle}\,\mathrm{Var}(\hat{F})_{|\phi\rangle} - O(g). \tag{11}$$

With the condition $\langle\hat{F}\rangle_{|\phi\rangle} = 0$, this saturates the maximum in (7) up to small corrections, which indicates that the ancilla measurement does not lose information by itself. One can always examine all d^n ancilla outcomes to obtain the maximum information, as pointed out in [1–3].

Now let us focus on a particular post-selection $k = 1$, using an unbiased meter that satisfies $\langle \hat{F} \rangle_{|\phi\rangle} = 0$, as assumed in the main text. This produces the simplification

$$I^{(1)}(g) \approx 4\, P_s^{(1)} |A_w^{(1)}|^2 \left[1 - |g A_w^{(1)}|^2 \mathrm{Var}(\hat{F})\right]. \tag{12}$$

Now recall Eq. (15) of the main text, where we showed that if we fix $P_s^{(1)} \ll 1$ and picked a post-selection state that maximizes $A_w^{(1)}$ then we found

$$\max |A_w^{(1)}|^2 \approx \frac{1 - P_s^{(1)}}{P_s^{(1)}} \mathrm{Var}(\hat{A})_{|\Psi_i\rangle} \approx \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{P_s^{(1)}}. \tag{13}$$

For this strategically chosen post-selection with small $P_s^{(1)}$ and maximized $A_w^{(1)}$, it then follows that

$$I^{(1)}(g) \approx 4\, \mathrm{Var}(\hat{A})_{|\Psi_i\rangle} \left[1 - |g A_w^{(1)}|^2 \mathrm{Var}(\hat{F})\right] = I(g) \left[\frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{\langle \hat{A}^2 \rangle_{|\Psi_i\rangle}}\right] \left[1 - |g A_w^{(1)}|^2 \mathrm{Var}(\hat{F})\right], \tag{14}$$

which is Eq. (16) in the main text. That is, nearly *all* the Fisher information can be concentrated into a single (but rarely post-selected) meter state (see also [6]). The remaining information is distributed among the $(d^n - 1)$ remaining states, and could be retrieved in principle. The special post-selected meter state suffers an overall reduction factor of $\eta = \mathrm{Var}(\hat{A})/\langle \hat{A}^2 \rangle$, as well as a small loss $|g A_w^{(1)}|^2 \mathrm{Var}(\hat{F})$. However, most weak value amplification experiments operate in the linear response regime $g|A_w^{(1)}| \mathrm{Var}(\hat{F})^{\frac{1}{2}} \ll 1$ where this remaining loss is negligible. Moreover, the overall reduction factor η can even be set to unity by choosing ancilla observables that satisfy $\langle \hat{A} \rangle_{|\Psi_i\rangle} = 0$.

As carefully discussed in [2, 3], one cannot actually reach the optimal bound of (7) when making a post-selection. However, (14) shows that one can get remarkably close by carefully choosing which post-selection to make. It is quite surprising that one can even approximately saturate (7) while discarding the $(d^n - 1)$ much more probable outcomes. Rare post-selections can often be advantageous for independent reasons (e.g., to attenuate an optical beam down to a manageable post-selected beam power), so this property of weak value amplification makes it an attractive technique for estimating an extremely small parameter g that permits the linear response conditions [6].

B. Examples

To see how this works in more detail, let us examine the ancilla qubit post-selection examples used in the main text, where $g = \varphi/2$. For completeness, we will work through two examples. First, we consider a sub-optimal ancilla observable $\hat{a} = |1\rangle\langle 1|$. Second, we consider an optimal ancilla observable $\hat{a} = \hat{\sigma}_z$ to emphasize the practical difference.

1. Ancilla Projectors

A suboptimal choice of ancilla observable is the projector $\hat{a} = |1\rangle\langle 1|$ used in controlled qubit operations. From the optimal initial state given by Eq. (10) in the main text, we have $\langle \hat{A}^2 \rangle = n^2/2$ and $\langle \hat{A} \rangle = n/2$. Therefore, the maximum quantum Fisher information from (7) that we can expect for estimating φ is

$$I(\varphi) = \frac{n^2}{2}, \tag{15}$$

where the factor 1/2 in $g = \varphi/2$ has been taken into account, and the corresponding quantum Cramér-Rao bound is $\sqrt{2}/n$. This is the best (Heisenberg) scaling of the estimation precision that can be obtained by using n entangled ancillas with the given initial states.

Now, let us consider what happens when we make the optimal preparation and post-selections for weak value amplification. We expect from (14) that the maximum information which can be attained through post-selection will be reduced by a factor of

$$\eta = \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{\langle \hat{A}^2 \rangle_{|\Psi_i\rangle}} = \frac{1}{2}. \tag{16}$$

It is in this sense that the choice of $\hat{a}$ as a projector is suboptimal. We will see in the next section what happens with the optimal choice of $\hat{\sigma}_z$.

In the first case considered in the main text (i.e., increasing the post-selection probability with the weak value A_w fixed), the optimal post-selected state is

$$|\Psi_f\rangle \propto (A_w^*)|1\rangle^{\otimes n} + (n - A_w^*)|0\rangle^{\otimes n}. \tag{17}$$

Computing the post-selected meter state then produces

$$|\phi'\rangle_1 = \frac{\left[n - A_w[1-\cos(n\varphi/2)]\hat{1} - iA_w\sin(n\varphi/2)\hat{\sigma}_z\right]|\phi\rangle}{\left(n^2 + 2[|A_w|^2 - n\mathrm{Re}A_w][1-\cos(n\varphi/2)]\right)^{1/2}} \approx \left(\hat{1} - iA_w\frac{\varphi}{2}\hat{\sigma}_z\right)|\phi\rangle, \tag{18}$$

where we have used $\langle\phi|\hat{\sigma}_z|\phi\rangle = 0$, and then have made the small parameter approximation $n\varphi \ll 1$. This recovers the expected linear response result in (8). This state is post-selected with probability

$$p_1 = \frac{1}{2} - \cos(n\varphi/2)\frac{|A_w|^2 - n\mathrm{Re}A_w}{n^2 + 2[|A_w|^2 - n\mathrm{Re}A_w]} \approx \frac{n^2}{2n^2 + 4[|A_w|^2 - n\mathrm{Re}A_w]} \approx \frac{n^2}{4}|A_w|^{-2}, \tag{19}$$

where we have made the small parameter approximation $n\varphi \ll 1$, and then the large weak value assumption $n \ll |A_w|$.

Now computing the quantum Fisher information (5) with the post-selected meter state $\sqrt{p_1}\,|\phi'\rangle_1$ yields the simple expression

$$I_1(\varphi) \approx \frac{n^2}{4}\left(1 - \left|\frac{\varphi A_w}{2}\right|^2\right) \leq \frac{n^2}{4}, \tag{20}$$

in agreement with (14). The maximum achieves the best possible scaling of n^2 as in (15). Moreover, for the most frequently used linear response regime with $|A_w|\varphi \ll 1$, we achieve the expected maximum information of $\eta I(\varphi) = n^2/4$.

For the second case (i.e., increasing the weak value A_w with the post-selection probability fixed), we can obtain the results simply by rescaling $A_w \to \sqrt{n}A_w$ to produce $p_2 \propto n$, as shown in the main text. This produces,

$$|\phi'\rangle_2 \approx \left(\hat{1} - i\sqrt{n}A_w\frac{\varphi}{2}\hat{\sigma}_z\right)|\phi\rangle, \tag{21}$$

and

$$p_2 \approx \frac{n^2}{4}|\sqrt{n}A_w|^{-2} = \frac{n}{4}|A_w|^{-2}, \tag{22}$$

and yields the Fisher information

$$I_2(\varphi) \approx \frac{n^2}{4}\left(1 - n\left|\frac{\varphi A_w}{2}\right|^2\right) \leq \frac{n^2}{4}. \tag{23}$$

The increase of the amplification factor $|A_w|$ correspondingly decreases the remaining Fisher information, as expected from (20). However, since $n\varphi \ll 1$ and $\varphi|A_w| \ll 1$ in the linear response regime, this decrease is still small.

Alternatively, this second case can be computed explicitly as follows. For a fixed post-selection probability p, the post-selected state must be $|\Psi_f\rangle = \sqrt{p}|\Psi_i\rangle + \sqrt{1-p}|\Psi_i^{\perp}\rangle$, where the optimal $|\Psi_i^{\perp}\rangle$ is parallel to the component of $\hat{A}|\Psi_i\rangle$ in the complementary subspace orthogonal to $|\Psi_i\rangle$. Computing this yields

$$\begin{aligned}|\Psi_f\rangle &= \sqrt{p}|\Psi_i\rangle + \sqrt{1-p}\frac{\hat{A}|\Psi_i\rangle - |\Psi_i\rangle\langle\Psi_i|\hat{A}|\Psi_i\rangle}{\sqrt{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}} \\ &= \left(\sqrt{\frac{p}{2}} - \sqrt{\frac{1-p}{2}}\right)|0\rangle^{\otimes n} + \left(\sqrt{\frac{p}{2}} + \sqrt{\frac{1-p}{2}}\right)|1\rangle^{\otimes n}.\end{aligned} \tag{24}$$

Thus, computing the post-selected meter state yields

$$|\phi'\rangle_2 \propto \left(\left(\sqrt{\frac{p}{2}} - \sqrt{\frac{1-p}{2}}\right)\hat{1} + \left(\sqrt{\frac{p}{2}} + \sqrt{\frac{1-p}{2}}\right)\mathrm{e}^{-in\varphi\hat{\sigma}_z/2}\right)|\phi\rangle \approx \left(\hat{1} - i|A_w|\frac{\varphi}{2}\hat{\sigma}_z\right)|\phi\rangle, \tag{25}$$

where we have defined the effective weak value factor

$$|A_w| = \frac{n}{2}\left(1+\sqrt{\frac{1-p}{p}}\right) \approx \frac{n}{2}p^{-1/2}, \tag{26}$$

and have used the linear response approximations $n\varphi \ll 1$ and $\varphi|A_w| \ll 1$, as well as the small probability assumption $p \ll 1$. Computing the quantum Fisher information from (5) with the state $\sqrt{p}\,|\phi'\rangle_2$ then produces

$$I_2(\varphi) \approx p|A_w|^2\left(1-\left[\frac{\varphi|A_w|}{2}\right]^2\right) = \frac{n^2}{4}\left(1-\left[\frac{n\varphi}{4\sqrt{p}}\right]^2\right) \leq \frac{n^2}{4} \tag{27}$$

using the definition (26). This result precisely matches the form of (12). It is now clear that for quadratic scaling $p = n^2 p_0$ we recover (20) with the effective reference weak value $|A_w| = 1/(2\sqrt{p_0})$, while for linear scaling $p = np_0$ we recover (23).

2. Ancilla Z-operators

For contrast, an optimal choice of ancilla observable is $\hat{a} = \hat{\sigma}_z$, as used in the main text. From the optimal initial state given by Eq. (10) in the main text, we have $\langle \hat{A}^2\rangle = n^2$ and $\langle \hat{A}\rangle = 0$. Therefore, the maximum quantum Fisher information from (7) that we can expect for estimating φ is

$$I(\varphi) = n^2, \tag{28}$$

which is a factor of 2 larger than (15). The corresponding quantum Cramér-Rao bound is $1/n$. From (14), we expect that the reduction factor is

$$\eta = \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{\langle \hat{A}^2\rangle_{|\Psi_i\rangle}} = 1. \tag{29}$$

Thus, it is possible to saturate the optimal bound with this choice of $\hat{a}$.

In the first case considered in the main text (i.e., increasing the post-selection probability with the weak value A_w fixed), the optimal post-selected state is

$$|\Psi_f\rangle \propto (n + A_w^*)|1\rangle^{\otimes n} + (n - A_w^*)|0\rangle^{\otimes n}. \tag{30}$$

Computing the post-selected meter state then produces

$$|\phi'\rangle_1 = \frac{\left[n\cos(n\varphi/2)\hat{1} - iA_w\sin(n\varphi/2)\hat{\sigma}_z\right]|\phi\rangle}{\left(n^2\cos^2(n\varphi/2) + |A_w|^2\sin^2(n\varphi/2)\right)^{1/2}} \approx \left(\hat{1} - iA_w\frac{\varphi}{2}\hat{\sigma}_z\right)|\phi\rangle, \tag{31}$$

where we have used $\langle\phi|\hat{\sigma}_z|\phi\rangle = 0$, and then have made the small parameter approximation $n\varphi \ll 1$. This again recovers the expected linear response result in (8). This state is post-selected with probability

$$p_1 = \frac{n^2\cos^2(n\varphi/2) + |A_w|^2\sin^2(n\varphi/2)}{n^2 + [A_w|^2} \approx \frac{n^2}{n^2 + |A_w|^2} \approx n^2|A_w|^{-2}, \tag{32}$$

where we have made the small parameter approximation $n\varphi \ll 1$, and then the large weak value assumption $n \ll |A_w|$.

Now computing the quantum Fisher information (5) with the post-selected meter state $\sqrt{p_1}\,|\phi'\rangle_1$ yields the simple expression

$$I_1(\varphi) \approx n^2\left(1 - \left|\frac{\varphi A_w}{2}\right|^2\right) \leq n^2, \tag{33}$$

in agreement with (14). The maximum saturates the upper bound of n^2 in (28), as expected.

For the second case (i.e., increasing the weak value A_w with the post-selection probability fixed), we can again obtain the results simply by rescaling $A_w \to \sqrt{n}A_w$ to produce

$$|\phi'\rangle_2 \approx \left(\hat{1} - i\sqrt{n}A_w\frac{\varphi}{2}\hat{\sigma}_z\right)|\phi\rangle, \tag{34}$$

$$p_2 \approx n^2|\sqrt{n}A_w|^{-2} = n|A_w|^{-2}, \tag{35}$$

and the Fisher information

$$I_2(\varphi) \approx n^2 \left(1 - n\left|\frac{\varphi A_w}{2}\right|^2\right) \leq n^2. \tag{36}$$

Alternatively, computing the optimal post-selection state for a fixed post-selection probability p yields the same state as (24). Hence, computing the post-selected meter state yields

$$|\phi'\rangle_2 \propto \left(\left(\sqrt{\frac{p}{2}} - \sqrt{\frac{1-p}{2}}\right) e^{in\varphi\hat{\sigma}_z/2} + \left(\sqrt{\frac{p}{2}} + \sqrt{\frac{1-p}{2}}\right) \mathrm{e}^{-in\varphi\hat{\sigma}_z/2}\right)|\phi\rangle \approx \left(\hat{1} - i|A_w|\frac{\varphi}{2}\hat{\sigma}_z\right)|\phi\rangle, \tag{37}$$

where we have defined the effective weak value factor

$$|A_w| = n\sqrt{\frac{1-p}{p}} \approx np^{-1/2}, \tag{38}$$

in contrast to (26). Computing the quantum Fisher information from (5) with the state $\sqrt{p}\,|\phi'\rangle_2$ then produces

$$I_2(\varphi) \approx p|A_w|^2 \left(1 - \left[\frac{\varphi|A_w|}{2}\right]^2\right) = n^2\left(1 - \left[\frac{n\varphi}{\sqrt{p}}\right]^2\right) \leq n^2, \tag{39}$$

using the definition (38). As before, this result precisely matches the form of (12). It is now clear that for quadratic scaling $p = n^2 p_0$ we recover (33) with the effective reference weak value $|A_w| = 1/\sqrt{p_0}$, while for linear scaling $p = np_0$ we recover (36). Therefore, in both post-selected qubit examples considered in the main text we can nearly saturate the expected maximum of $I(\varphi) = n^2$ when the linear response conditions $n\varphi \ll 1$, $\varphi|A_w| \ll 1$, and the large weak value condition $n \ll |A_w|$ are met, despite the loss of data incurred by the post-selection.

---

* Electronic address: shengshp@usc.edu
† Electronic address: jdressel@ee.ucr.edu
‡ Electronic address: tbrun@usc.edu
[1] G. C. Knee, G. A. D. Briggs, S. C. Benjamin, and E. M. Gauger, Phys. Rev. A **87**, 012115 (2013).
[2] G. C. Knee and E. M. Gauger, Phys. Rev. X **4**, 011032, (2014).
[3] C. Ferrie and J. Combes, Phys. Rev. Lett. **112**, 040406 (2014).
[4] S. L. Braunstein, C. M. Caves and G. J. Milburn, Ann. Phys. **247**, 135 (1996).
[5] V. Giovannetti, S. Lloyd and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[6] A. N. Jordan, J. Martínez-Rincón, and J. C. Howell, Phys. Rev. X **4**, 011031 (2014).