Hidden-variable model for continuous-variable teleportation

C. M. Caves

2003 August 21

The material herein will be used in two papers, the first on a hidden-variable model for continuous-variable teleporation (by K. Wodkiewicz and CMC) and the second on teleportation fidelity as a probe of sub-Planck structure (by A. Scott and CMC).

1. The general scenario

The scenario in continuous-variable teleportation is the following. Alice has a mode with annihilation operator $a = (x_A + ip_A)/\sqrt{2}$, and Bob has a mode with annihilation operator $b = (x_B + ip_B)/\sqrt{2}$. We generally let $\alpha = (\alpha_1 + i\alpha_2)/\sqrt{2}$ and $\beta = (\beta_1 + i\beta_2)/\sqrt{2}$ denote corresponding c-number variables for these modes. These two modes are prepared in a joint quantum state ρ_{AB} that has Wigner function $W_{AB}(\alpha, \beta)$. Victor brings up to Alice a mode with annihilation operator $v = (x_V + ip_V)/\sqrt{2}$ [c-number variable $\nu = (\nu_1 + i\nu_2)/\sqrt{2}$]. Victor's mode is prepared in an input state $\rho = |\psi\rangle\langle\psi|$ that has Wigner function $W_\rho(\nu)$. The overall Wigner function is $W_\rho(\nu)W_{AB}(\alpha, \beta)$.

Alice measures

$$v + a^\dagger = \underbrace{\frac{1}{\sqrt{2}}(x_V + x_A)}_{=X} + i\underbrace{\frac{1}{\sqrt{2}}(p_V - p_A)}_{=P} \tag{1}$$

on modes v and a. We denote the c-number variable corresponding to $v + a^\dagger$ by

$$\xi = \nu + \alpha^* = \frac{1}{\sqrt{2}}(\nu_1 + \alpha_1) + \frac{i}{\sqrt{2}}(\nu_2 - \alpha_2)\ . \tag{2}$$

The probability density for getting result ξ is

$$\begin{aligned} p(\xi) &= \int d^2\nu\, d^2\alpha\, d^2\beta\, \delta(\nu + \alpha^* - \xi) W_\rho(\nu) W_{AB}(\alpha, \beta) \\ &= \int d^2\nu\, d^2\beta\, W_\rho(\nu) W_{AB}(\xi^* - \nu^*, \beta) \\ &= \int d^2\nu\, W_\rho(\nu) W_A(\xi^* - \nu^*)\ , \end{aligned} \tag{3}$$

where $W_A(\alpha)$ is the marginal Wigner function for mode a, and the state of Bob's mode after the measurement, conditioned on result ξ, has Wigner function

$$\begin{aligned} W'(\beta|\,\xi) &= \frac{1}{p(\xi)} \int d^2\nu\, d^2\alpha\, \delta(\nu + \alpha^* - \xi) W_\rho(\nu) W_{AB}(\alpha, \beta) \\ &= \frac{1}{p(\xi)} \int d^2\nu\, W_\rho(\nu) W_{AB}(\xi^* - \nu^*, \beta)\ . \end{aligned} \tag{4}$$

Alice now sends the measurement result ξ to Bob, who displaces his mode by this amount, i.e., x_B is displaced by $\sqrt{2}X$ and p_B is displaced by $\sqrt{2}P$, giving an output state $\rho_{\rm out}(\xi)$ with Wigner function

$$W_{\rm out}(\beta|\,\xi) = W'(\beta - \xi|\,\xi) = \frac{1}{p(\xi)} \int d^2\nu\, W_\rho(\nu) W_{AB}(\xi^* - \nu^*, \beta - \xi)\ . \tag{5}$$

This state averages to

$$\overline{\rho}_{\rm out} = \int d^2\xi\, p(\xi)\rho_{\rm out}(\xi)\ , \tag{6}$$

which has Wigner function

$$\begin{aligned} W_{\overline{\rho}_{\rm out}}(\beta) &= \int d^2\xi\, p(\xi) W_{\rm out}(\beta|\,\xi) \\ &= \int d^2\nu\, W_\rho(\nu) \int d^2\xi\, W_{AB}(\xi^* - \nu^*, \beta - \xi) \\ &= \int d^2\nu\, W_\rho(\nu) \int d^2\alpha\, W_{AB}(\alpha, \beta - \nu - \alpha^*) \\ &= \int d^2\nu\, P(\beta - \nu) W_\rho(\nu) \\ &= \int d^2\nu\, P(\nu) W_\rho(\beta - \nu)\ , \end{aligned} \tag{7}$$

where

$$P(\nu) = \int d^2\alpha\, W_{AB}(\alpha, \nu - \alpha^*) = \int d^2\alpha\, d^2\beta\, \delta(\beta + \alpha^* - \nu) W_{AB}(\alpha, \beta) \tag{8}$$

is the probability to obtain result ν in a measurement of

$$b + a^\dagger = \frac{1}{\sqrt{2}}(x_B + x_A) + \frac{i}{\sqrt{2}}(p_B - p_A) \tag{9}$$

on modes a and b.

Since $W_\rho(\beta - \nu)$ is the Wigner function for the displaced state $D(v,\nu)\rho D^\dagger(v,\nu)$, i.e.,

$$W_{D(v,\nu)\rho D^\dagger(v,\nu)}(\beta) = W_\rho(\beta - \nu)\ , \tag{10}$$

we can write the average output state (6) as

$$\overline{\rho}_{\rm out} = \int d^2\nu\, P(\nu) D(b,\nu)\rho D^\dagger(b,\nu)\ , \tag{11}$$

where we now regard the initial state ρ as a state of Bob's mode. We now have two different ensemble decompositions for $\overline{\rho}_{\rm out}$: Eq. (6) gives $\overline{\rho}_{\rm out}$ in terms of an average over the teleported states $\rho_{\rm out}(\xi)$, whereas Eq. (11) gives $\overline{\rho}_{\rm out}$ in terms an average over displaced input states.

One can see more directly the relation between the two decompositions by defining new variables

$$\begin{aligned} \nu &\equiv \beta + \alpha^* = \frac{1}{\sqrt{2}}(\beta_1 + \alpha_1) + \frac{i}{\sqrt{2}}(\beta_2 - \alpha_2)\ , \\ \mu &\equiv \frac{1}{2}(\beta - \alpha^*) = \frac{1}{2\sqrt{2}}(\beta_1 - \alpha_1) + \frac{i}{2\sqrt{2}}(\beta_2 + \alpha_2)\ , \end{aligned} \tag{12}$$

and defining the normalized conditional quasidistribution

$$R(\mu|\,\nu) \equiv \frac{W_{AB}(\alpha,\beta)}{P(\nu)}\ . \tag{13}$$

This allows us to rewrite Eq. (5) in the form

$$
\begin{aligned}
W_{\text{out}}(\beta|\xi) &= \frac{1}{p(\xi)}\int d^2\nu\, W_\rho(\nu)P(\beta-\nu)R\big((\beta+\nu)/2-\xi|\beta-\nu\big) \\
&= \frac{1}{p(\xi)}\int d^2\nu\, W_\rho(\beta-\nu)P(\nu)R(\beta-\xi-\nu/2|\,\nu)\;,
\end{aligned}
\tag{14}
$$

giving

$$
p(\xi) = \int d^2\beta\, W_{\text{out}}(\beta|\xi)p(\xi) = \int d^2\beta\, d^2\nu\, W_\rho(\beta-\nu)P(\nu)R(\beta-\xi-\nu/2|\,\nu) \tag{15}
$$

and

$$
\begin{aligned}
W_{\overline{\rho}_{\text{out}}}(\beta) &= \int d^2\xi\, p(\xi)W_{\text{out}}(\beta|\,\xi) \\
&= \int d^2\nu\, W_\rho(\beta-\nu)P(\nu)\int d^2\xi\, R(\beta-\xi-\nu/2|\,\nu) \\
&= \int d^2\nu\, W_\rho(\beta-\nu)P(\nu)\;.
\end{aligned}
\tag{16}
$$

The decomposition (11) turns out to be the more useful decomposition even though the states in the decomposition are not the teleported states. It is easy to see that if the initial state $|\psi\rangle$ is displaced, the average output state $\overline{\rho}_{\text{out}}$ is displaced by the same amount. If $P(\nu)$ is even under parity, i.e., $P(-\nu) = P(\nu)$, the average output state can be written as

$$
\overline{\rho}_{\text{out}} = \int d^2\nu\, P(\nu)D(b,-\nu)\rho D^\dagger(b,-\nu) = \int d^2\nu\, P(\nu)D^\dagger(b,\nu)\rho D(b,\nu)\;. \tag{17}
$$

If $P(\nu)$ is a function only of $|\nu|$, a rotation of the input state leads to the same rotation of $\overline{\rho}_{\text{out}}$.

For outcome ξ, the fidelity of the output state with the input state is

$$
F_\xi = \langle\psi|\rho_{\text{out}}(\xi)|\psi\rangle\;. \tag{18}
$$

Thus the average fidelity is given by

$$
\overline{F} = \int d^2\xi\, p(\xi)F_\xi = \langle\psi|\,\overline{\rho}_{\text{out}}|\psi\rangle = \int d^2\nu\, P(\nu)|\langle\psi|D(b,\nu)|\psi\rangle|^2\;. \tag{19}
$$

The symmetrically ordered characteristic function for $\overline{\rho}_{\text{out}}$ is

$$
\begin{aligned}
\Phi_{\overline{\rho}_{\text{out}}}(\mu) &= \operatorname{tr}\big(\overline{\rho}_{\text{out}}D(b,\mu)\big) \\
&= \int d^2\nu\, P(\nu)\operatorname{tr}\big(\rho\, \underbrace{D^\dagger(b,\nu)D(b,\mu)D(b,\nu)}_{=D(\nu,\mu)D(b,\mu)}\big) \\
&= \Phi_\rho(\mu)\int d^2\nu\, P(\nu)D(\nu,\mu) \\
&= \pi\tilde{P}(\mu)\Phi_\rho(\mu)\;,
\end{aligned}
\tag{20}
$$

where $\tilde{P}(\mu)$ is the Fourier transform of $P(\nu)$. This result is the Fourier transform of the corresponding Wigner function relations in Eq. (7).

This leaves us with a variety of forms for the average fidelity:

$$\begin{aligned}
\overline{F} &= \int d^2\nu\, P(\nu)|\Phi_\rho(\nu)|^2 \\
&= \pi \int d^2\beta\, W_{\overline{\rho}_{\rm out}}(\beta) W_\rho(\beta) = \pi \int d^2\beta\, d^2\nu\, P(\beta-\nu) W_\rho(\beta) W_\rho(\nu) \\
&= \int \frac{d^2\mu}{\pi}\, \Phi^*_{\overline{\rho}_{\rm out}}(\mu)\Phi_\rho(\mu) = \int d^2\mu\, \tilde{P}(\mu)|\Phi_\rho(\mu)|^2 \\
&= \pi \int d^2\beta\, d^2\nu\, \tilde{P}(\beta-\nu) W_\rho(\beta) W_\rho(\nu)\ .
\end{aligned} \tag{21}$$

The first form is just a rewrite of Eq. (19). The second line comes from rewriting $\overline{F} = \langle\psi|\overline{\rho}_{\rm out}|\psi\rangle$ as an overlap of the input and output Wigner functions, $W_{\overline{\rho}_{\rm out}}(\beta)$ and $W_\rho(\beta)$, and then using Eq. (7) for $W_{\overline{\rho}_{\rm out}}(\beta)$. Similarly, the third line comes from rewriting $\overline{F} = \langle\psi|\overline{\rho}_{\rm out}|\psi\rangle$ as an overlap of the input and output characteristic functions, $\Phi_{\overline{\rho}_{\rm out}}(\mu)$ and $\Phi_\rho(\mu)$, and then using Eq. (20) for $\Phi_{\overline{\rho}_{\rm out}}(\mu)$. The second and third lines are related to one another by a Fourier transform of the quantities in the integrand, as are the first and last lines.

2. Squeezed-state teleportation

We do a good job of teleporting when the distribution $P(\nu)$ is narrow, i.e., when x_A and x_B are tightly anti-correlated, and p_A and p_B are tightly correlated. Introducing modes

$$\begin{aligned}
c &= \frac{1}{\sqrt{2}}(a+b) = \frac{1}{\sqrt{2}}(x_C + ip_C) \qquad & x_C &= \frac{1}{\sqrt{2}}(x_A+x_B) \\
& & p_C &= \frac{1}{\sqrt{2}}(p_A+p_B) \\
d &= \frac{1}{\sqrt{2}}(a-b) = \frac{1}{\sqrt{2}}(x_D + ip_D) \qquad & x_D &= \frac{1}{\sqrt{2}}(x_A-x_B) \\
& & p_D &= \frac{1}{\sqrt{2}}(p_A-p_B)
\end{aligned}\ , \tag{22}$$

with c-number variables $\gamma = (\alpha+\beta)/\sqrt{2} = (\gamma_1+i\gamma_2)/\sqrt{2}$ and $\delta = (\alpha-\beta)/\sqrt{2} = (\delta_1+i\delta_2)/\sqrt{2}$, we see that we want the variances of x_C and p_D to be small. Thus a natural choice is to use squeezed vacuum states for modes c and d, with the variances of the quadrature components given by

$$\begin{aligned}
(\Delta x_C)^2 &= \frac{1}{2}\,e^{-2r} \qquad & (\Delta p_C)^2 &= \frac{1}{2}\,e^{2r} \\
(\Delta x_D)^2 &= \frac{1}{2}\,e^{2r} \qquad & (\Delta p_D)^2 &= \frac{1}{2}\,e^{-2r}
\end{aligned}\ . \tag{23}$$

This state is a two-mode squeezed vacuum state for modes a and b. The Wigner function is given by

$$
\begin{aligned}
W_{CD}(\gamma,\delta) &= \frac{4}{\pi^2}\exp\Big(-\frac{\gamma_1^2}{e^{-2r}}-\frac{\gamma_2^2}{e^{2r}}-\frac{\delta_1^2}{e^{2r}}-\frac{\delta_2^2}{e^{-2r}}\Big) \\
&= \frac{4}{\pi^2}\exp\Big(-2(|\gamma|^2+|\delta|^2)\cosh 2r-[\gamma^2+\gamma^{*2}-\delta^2-\delta^{*2}]\sinh 2r\Big) \\
&= \frac{4}{\pi^2}\exp\Big(-2(|\alpha|^2+|\beta|^2)\cosh 2r-2(\alpha\beta+\alpha^*\beta^*)\sinh 2r\Big) = W_{AB}(\alpha,\beta)\ .
\end{aligned}
\tag{24}
$$

Since $P(\nu)$ is the probability to obtain result ν in a measurement of

$$
b+a^\dagger = \frac{1}{\sqrt{2}}(x_B+x_A)+\frac{i}{\sqrt{2}}(p_B-p_A) = x_C - ip_D\ , \tag{25}
$$

the Wigner function immediately implies that

$$
P(\nu) = \frac{e^{2r}}{\pi}\,e^{-e^{2r}|\nu|^2} = \frac{2}{\pi t}\,e^{-2|\nu|^2/t} \quad\Longrightarrow\quad \tilde{P}(\mu) = \frac{1}{\pi}\,e^{-e^{-2r}|\mu|^2} = \frac{1}{\pi}\,e^{s|\mu|^2/2}\ , \tag{26}
$$

where $s=-2e^{-2r}$. We define $t\equiv -s = 2e^{-2r}$ for convenience since the only interesting values of s are negative. Using the previous definitions (12), we have

$$
\nu = \beta+\alpha^* = \gamma_1 - i\delta_2 \qquad \mu = \frac{1}{2}(\beta-\alpha^*) = \frac{1}{2}(-\delta_1+i\gamma_2)\ , \tag{27}
$$

and we can immediately write

$$
W_{AB}(\alpha,\beta) = \frac{4}{\pi^2}\exp\big(e^{2r}|\nu|^2-4e^{-2r}|\mu|^2\big) = \underbrace{\frac{e^{2r}}{\pi}e^{e^{2r}|\nu|^2}}_{=P(\nu)}\ \underbrace{\frac{4e^{-2r}}{\pi}e^{-4e^{-2r}|\mu|^2}}_{=R(\mu|\,\nu)=R(\mu)}\ . \tag{28}
$$

We can now specialize the important results of the preceding section to the case of squeezed-state teleportation. The average output state (11) at Bob's end becomes

$$
\overline{\rho}_{\rm out} = e^{2r}\int\frac{d^2\nu}{\pi}\,e^{-e^{2r}|\nu|^2}D(b,\nu)\rho D^\dagger(b,\nu) = \frac{2}{t}\int\frac{d^2\nu}{\pi}\,e^{-2|\nu|^2/t}D(b,\nu)\rho D^\dagger(b,\nu)\ . \tag{29}
$$

Recall that if the input state is displaced, $\overline{\rho}_{\rm out}$ is displaced by the same amount. Because the Gaussian is a function only of $|\nu|$, we have that a rotation of the input state leads to the same rotation of $\overline{\rho}_{\rm out}$ and that

$$
\overline{\rho}_{\rm out} = \frac{2}{t}\int\frac{d^2\nu}{\pi}\,e^{-2|\nu|^2/t}D^\dagger(b,\nu)\rho D(b,\nu)\ . \tag{30}
$$

The symmetrically ordered characteristic function for $\overline{\rho}_{\rm out}$ is the s-ordered characteristic function for ρ,

$$
\Phi_{\overline{\rho}_{\rm out}}(\mu) = \pi\tilde{P}(\mu)\Phi_\rho(\mu) = e^{s|\mu|^2/2}\Phi_\rho(\mu) = \Phi_\rho^{(s)}(\mu)\ , \tag{31}
$$

and the Wigner function for $\overline{\rho}_{\rm out}$ is the s-ordered quasidistribution for ρ,

$$W_{\overline{\rho}_{\rm out}}(\beta) = W_\rho^{(s)}(\beta) = \int d^2\nu\, P(\beta-\nu) W_\rho(\nu) = \frac{2}{t}\int \frac{d^2\nu}{\pi}\, W_\rho(\nu) e^{-2|\beta-\nu|^2/t}\;. \tag{32}$$

The various forms (21) for the average fidelity become

$$\begin{aligned}
\overline{F}_\rho(t) &= \frac{2}{t}\int \frac{d^2\nu}{\pi}\, e^{-2|\nu|^2/t} |\langle\psi|D(b,\nu)|\psi\rangle|^2 \\
&= \frac{2}{t}\int \frac{d^2\nu}{\pi}\, e^{-2|\nu|^2/t} |\Phi_\rho(\nu)|^2 \\
&= \pi \int d^2\beta\, W_\rho^{(s)}(\beta) W_\rho(\beta) = \frac{2}{t}\int d^2\beta\, d^2\nu\, e^{-2|\beta-\nu|^2/t} W_\rho(\beta) W_\rho(\nu) \\
&= \int \frac{d^2\mu}{\pi}\, \Phi_\rho^{(s)*}(\mu)\Phi_\rho(\mu) = \int \frac{d^2\mu}{\pi}\, e^{-t|\mu|^2/2} |\Phi_\rho(\mu)|^2 \\
&= \int d^2\beta\, d^2\nu\; e^{-t|\beta-\nu|^2/2} W_\rho(\beta) W_\rho(\nu)\;.
\end{aligned} \tag{33}$$

The second and fourth forms (and the third and fifth) show us that

$$\overline{F}_\rho(t) = \frac{2}{t}\overline{F}_\rho(4/t)\;. \tag{34}$$

We can relate these forms to Zurek's work on sub-Planck structures [W. H. Zurek, *Nature* **412**, 712 (2001)]. For a given input state ρ, the characteristic function $\Phi_\rho(\nu)$ has two important scales: (i) a small scale ℓ over which, in some phase-space direction(s), it plunges from 1 at $\nu = 0$ to close to zero and (ii) a large scale L over which it remains nonnegligible. These scales satisfy $\ell L \sim 1$. The Wigner function being the Fourier transform of the characteristic function, these scales appear inversely in the Wigner function: (i) $1/L \sim \ell$ is the scale of the finest structure in the Wigner function, and (ii) $1/\ell \sim L$ is the scale over which the Wigner function is nonnegligible.

One way to display these scales is to consider the fidelity between the input state and the output state (30), which is obtained from the input state by Gaussian phase-space displacements of characteristic size $\sqrt{t/2} = e^{-r}$. The second form in Eq. (33) shows that to get high fidelity between input and output, the scale of the displacements should satisfy $\sqrt{t/2} = e^{-r} \lesssim \ell$; when this is true, the fourth form in Eq. (33) re-assures us that the fidelity is near one, since $\sqrt{2/t} = e^r \gtrsim L$. Thus the second and fourth forms express the reciprocal relation between ℓ and L. These same conclusions can also be read off the Wigner-function forms of the fidelity. The third form in Eq. (33) tells us that the fidelity is close to one as long as the scale of the phase-displacements satisfies $\sqrt{t/2} = e^{-r} \lesssim \ell$, whereas the last form in Eq. (33) assures us of high fidelity as long as $\sqrt{2/t} = e^r \gtrsim L$.

These conclusions have an immediate interpretation in terms of high-fidelity teleportation. To get good fidelity in the teleportation process, the two-mode squeezed state used as the entanglement resource must have a small scale $e^{-r} = \sqrt{t/2}$ somewhat smaller than the scale ℓ of the finest phase-space structure in the state to be teleported, which also means that $e^r = \sqrt{2/t}$ is somewhat larger than the scale L over which the Wigner function is nonnegligible. To put it succintly, good fidelity requires the squeezing to be large enough that the smallest sub-Planck structures are teleported faithfully.

3. A closer look at the teleported state in the high-fidelity limit

For the case of squeezed-state teleportation, we can get a better idea of what is going on by taking a closer look at the probability (15) that Alice gets result ξ in her measurement,

$$p(\xi) = \int d^2\beta\, d^2\nu\, W_\rho(\beta - \nu) P(\nu) R(\beta - \xi - \nu/2) \tag{35}$$

and at the Wigner function (14) of the teleported state,

$$W_{\rm out}(\beta|\xi) = \frac{1}{p(\xi)} \int d^2\nu\, W_\rho(\beta - \nu) P(\nu) R(\beta - \xi - \nu/2) \;, \tag{36}$$

As discussed above, high-fidelity teleportation is achieved when $P(\nu)$ is narrow enough that e^{-r} is small compared to the smallest-scale structure in the Wigner function, which also means that e^r is large compared to the size of the region over which the Wigner function is nonnegligible. In the limit of high-fidelity teleportation, where $P(\nu)$ is narrow and $R(\mu)$ is broad, we can get a good approximation by setting $\nu = 0$ and $\beta = \langle v \rangle$ in the broad Gaussian, which leaves us with

$$p(\xi) = R(\langle v \rangle - \xi) = \frac{4e^{-2r}}{\pi} e^{-4e^{-2r}|\xi - \langle v \rangle|^2} \;, \tag{37}$$

$$W_{\rm out}(\beta|\,\xi) = \int d^2\nu\, P(\nu) W_\rho(\beta - \nu) \;. \tag{38}$$

Thus, in the limit of high-fidelity teleportation, the output state $\rho_{\rm out}(\xi)$ is independent of the measurement result ξ and is the same as $\overline{\rho}_{\rm out}$.

The high-fidelity limit can be described by a very simple model. The mean values of the measurement results, X and P, are given by $\langle X \rangle = \langle x_V \rangle/\sqrt{2}$ and $\langle P \rangle = \langle p_V \rangle/\sqrt{2}$, i.e, $\langle \xi \rangle = \langle v \rangle$; the corresponding variances are

$$\begin{aligned}
\langle (\Delta X)^2 \rangle &= \frac{1}{2}\Big(\langle (\Delta x_V)^2 \rangle + \langle (\Delta x_A)^2 \rangle\Big) \simeq \frac{1}{2}\langle (\Delta x_A)^2 \rangle = \frac{1}{8}(e^{2r} + e^{-2r}) \simeq \frac{1}{8}e^{2r} \;, \\
\langle (\Delta P)^2 \rangle &= \frac{1}{2}\Big(\langle (\Delta p_V)^2 \rangle + \langle (\Delta p_A)^2 \rangle\Big) \simeq \frac{1}{2}\langle (\Delta p_A)^2 \rangle = \frac{1}{8}(e^{2r} + e^{-2r}) \simeq \frac{1}{8}e^{2r} \;.
\end{aligned} \tag{39}$$

These means and variances are described by the probability distribution $p(\xi) = R(\xi - \langle v \rangle)$. With the measurement results X and P in hand, we have that

$$\begin{aligned}
x_B &= -x_A + \sqrt{2}x_C = x_V - \sqrt{2}X + \sqrt{2}x_C \;, \\
p_B &= p_A - \sqrt{2}p_D = p_V - \sqrt{2}P - \sqrt{2}p_D \;.
\end{aligned} \tag{40}$$

Displacing x_B by $\sqrt{2}X$ and p_B by $\sqrt{2}P$ gives

$$\begin{aligned}
x_B &= x_V + \sqrt{2}x_C \;, \\
p_B &= p_V - \sqrt{2}p_D \;.
\end{aligned} \tag{41}$$

Since the variances of $\sqrt{2}x_C$ and $\sqrt{2}p_D$ are both equal to e^{-2r}, the convolution (38) describes the output probability corresponding to the transformation (41).

We can estimate how much information must be transmitted to make continuous-variable teleportation work with high fidelity. The possible values of X and P are distributed over a range given roughly by the uncertainties in Eq. (39): $\Delta X = \Delta P = e^r/2\sqrt{2}$. The values of X and P transmitted to Bob must allow him to perform displacements with accuracy somewhat better than the uncertainties in x_C and p_D: $\Delta x_C = \Delta p_D = e^{-r}/\sqrt{2}$. Thus the required number of bits in the transmission of X or P must be roughly $\log(\Delta X/\Delta x_C)$, giving a total amount of information of order $2\log(\Delta X/\Delta x_C) = 2\log(e^{2r}/2) \simeq 4r/\ln 2$. The squeeze parameter must be large enough to teleport the smallest phase-space structures faithfully; i.e., e^{-r} must be smaller than the smallest-scale structure in the Wigner function of ρ. Calling this smallest scale $L^{-1} = \ell \sim e^{-r}$, we have that Alice must transmit roughly $2\log(1/2\ell^2) \simeq 4\log(1/\ell) = 4\log L$ bits in order to transmit X and P with sufficient accuracy for high-fidelity teleportation.

Since L^2 is approximately the Hilbert-space dimension needed to represent ρ, the number of bits needed has the standard form $2\log L^2$. Indeed, this gives us an independent, phase-space way to interpret why we need 2 bits of classical information per qubit of quantum information. The teleportation process must be able to distinguish $(L/\ell)^2$ phase-space regions to transmit all of the sub-Planck structure, and this means transmitting $\log(L/\ell)^2 = \log(L^2)^2 = 2\log L^2$ bits of classical information.

4. Examples

All coherent states give the same average fidelity as the vacuum state, for which $\Phi_{|0\rangle\langle 0|}(\mu) = \langle 0|D(b,\mu)|0\rangle = e^{-|\mu|^2/2}$, so the coherent-state fidelity is

$$\overline{F}_{\rm coh}(t) = \int \frac{d^2\mu}{\pi}\, e^{-(1+t/2)|\mu|^2} = \frac{1}{1+t/2}\,. \tag{42}$$

All squeezed states with the same squeeze parameter u give the same average fidelity, so we can calculate the fidelity for the squeezed vacuum state $S(u,0)|0\rangle$, for which the characteristic function is $\Phi_{S(u,0)|0\rangle\langle 0|S^\dagger(u,0)}(\mu) = e^{-(e^{2u}\mu_1^2+e^{-2u}\mu_2^2)/4}$, so the squeezed-state fidelity is

$$\begin{aligned}\overline{F}_{\rm sq}(t) &= \int \frac{d^2\mu}{\pi}\, e^{-t|\mu|^2/2} e^{-(e^{2u}\mu_1^2+e^{-2u}\mu_2^2)/2} \\ &= \int \frac{d\mu_1\, d\mu_2}{2\pi}\, e^{-(e^{2u}+t/2)\mu_1^2/2} e^{-(e^{-2u}+t/2)\mu_2^2)/2} \\ &= \frac{1}{\sqrt{(e^{2u}+t/2)(e^{-2u}+t/2)}} \\ &= \frac{1}{\sqrt{1+t\cosh 2u + t^2/4}}\,.\end{aligned} \tag{43}$$

The average fidelity for an input number state, calculated in Section 6, is given by

$$\overline{F}_{|n\rangle\langle n|}(t) = \frac{(1-t/2)^n}{(1+t/2)^{n+1}}\, P_n\!\left(\frac{1+t^2/4}{1-t^2/4}\right)\,, \tag{44}$$

where $P_n(x)$ is a Legendre polynomial.

5. Hidden-variable model

Gaussian states are the only states that have nonnegative Wigner functions. For Gaussian input states and for the two-mode squeezed vacuum state shared by Alice and Bob, the Wigner-function description provides a hidden-variable model for teleportation: the hidden variables are the quadrature components of all the modes, and the Wigner function is a probability distribution for the hidden variables.

Non-Gaussian input states have Wigner functions that take on negative values. To get the hidden-variable model to work, one imagines that Alice takes Victor's input state ρ and kicks it randomly in phase space, the probability distribution of the random kicks being described by a Gaussian. The resulting state is

$$\rho' = \frac{2}{t}\int \frac{d^2\nu}{\pi}\, e^{-2|\nu|^2/t} D(v,\nu)\rho D^\dagger(v,\nu)\ . \tag{45}$$

The Wigner function of ρ' is the s-ordered quasidistribution for ρ, i.e., $W_{\rho'}(\nu) = W_\rho^{(s)}(\nu)$. The strength of the kicks, t, is chosen to be the minimum value that yields a nonnegative Wigner function $W_{\rho'}(\nu)$. For all non-Gaussian states, this minimum kicking strength is $s=-1$, meaning that the new Wigner function is the Husimi Q function of the original state: $W_{\rho'}(\nu) = W_\rho^{(-1)}(\nu) = Q(\nu) = \langle\nu|\rho|\nu\rangle/\pi$. The primed state having a nonnegative Wigner function, it, along with the two-mode squeezed state used for teleportation, can be accommodated within the hidden-variable model. To get a limit on the fidelity within the hidden-variable model, imagine that ρ' is teleported with perfect fidelity. Then the overall fidelity of the average output state with the input state ρ is given by the overlap of the Wigner and Husimi Q functions of ρ, i.e., by the $s=-1$ fidelity (33) (notice that higher kicking strengths would lead to higher values of t and thus to smaller fidelities, so we want to use the minimum kicking strengths that yields a positive Wigner function).

These considerations, together with wanting to know the maximum teleportation fidelity for a given amount of squeezing, motivate the following problem: Find the maximum value of the average fidelity $\overline{F}_\rho(t)$ over input pure states $\rho = |\psi\rangle\langle\psi|$, especially for $s=-1$. Since the average fidelity is the overlap of the Wigner function and the s-ordered quasidistribution for ρ, the task can be restated as finding the pure state that maximizes this overlap. Within the hidden-variable model, the maximum fidelity for $s=-1$ becomes a gold standard for quantum teleportation: If a non-Gaussian input state is teleported with fidelity exceeding this maximum, it is guaranteed that the teleportation cannot be described within a hidden-variable model based on the quadrature components.

A direct approach to finding the maximum is to vary $\overline{F} - 2\lambda(\langle\psi|\psi\rangle - 1)$, giving

$$0 = 2\langle\delta\psi|\overline{\rho}_{\rm out}|\psi\rangle - 2\lambda\langle\delta\psi|\psi\rangle + \text{h.c.}\ , \tag{46}$$

which implies that

$$\overline{\rho}_{\rm out}|\psi\rangle = \lambda|\psi\rangle = \overline{F}|\psi\rangle\ . \tag{47}$$

Thus the condition for a fidelity extremum is that the input state be an eigenstate of the average output state.

It is easy to see that if $|\psi\rangle$ satisfies the condition (47), then $D(b,\beta)|\psi\rangle$ and $e^{-i\theta b^\dagger b}|\psi\rangle$ also satisfy it. In particular, the vacuum state satisfies Eq. (47), since $\overline{\rho}_{\rm out}$ is diagonal in the number basis, so all coherent states also satisfy it. I have long thought that the coherent states give the maximum average fidelity, $\overline{F}_{\rm coh}(t) = (1 + t/2)^{-1}$, but I hadn't made much progress in showing this till 2003 March 7, when I constructed the (trivial) proof given at the end of this section.

The reason the eigenvalue equation (47) doesn't provide a solution to the problem is that it is nonlinear in $|\psi\rangle$, having come from minimizing a quantity that is quartic, instead of quadratic, in the input state $|\psi\rangle$. As a consequence, the eigenvalue equation has many more solutions than just an orthonormal set of states. In particular, we can easily see that all the number states satisfy Eq. (47), because in this case, $\overline{\rho}_{\rm out}$ is invariant under rotations and thus has number states as eigenstates.

Trivial bounds on the average fidelity, coming from using

$$|\Phi_\rho(\mu)|^2 \le 1 \qquad \text{and} \qquad 1 = \mathrm{tr}(\rho^2) = \int \frac{d^2\mu}{\pi}\,|\Phi_\rho(\mu)|^2\;, \tag{48}$$

are

$$\overline{F}(t) = \int \frac{d^2\mu}{\pi}\,e^{-t|\mu|^2/2}|\Phi_\rho(\mu)|^2 \le \begin{cases} \displaystyle\int \frac{d^2\mu}{\pi}\,e^{-t|\mu|^2/2} = \frac{2}{t} \\ \displaystyle\int \frac{d^2\mu}{\pi}\,|\Phi_\rho(\mu)|^2 = 1 \end{cases} . \tag{49}$$

We get more useful information from the first two derivatives,

$$\frac{d\overline{F}}{dt} = -\frac{1}{2}\int \frac{d^2\mu}{\pi}\,|\mu|^2 e^{-t|\mu|^2/2}|\Phi_\rho(\mu)|^2 < 0\;, \tag{50}$$

$$\frac{d^2\overline{F}}{dt^2} = \frac{1}{4}\int \frac{d^2\mu}{\pi}\,|\mu|^4 e^{-t|\mu|^2/2}|\Phi_\rho(\mu)|^2 > 0\;, \tag{51}$$

which together imply that $\overline{F}(t)$ is a strictly decreasing, strictly concave function of t.

Defining

$$I(x) \equiv \int \frac{d^2\mu}{\pi}\,e^{-x|\mu|^2}|\Phi_\rho(\mu)|^2\;, \tag{52}$$

we have

$$\overline{F}(t) = I(t/2) = (2/t)I(2/t)\;, \tag{53}$$

which leads us to Eq. (34). We can also write

$$\frac{\overline{F}(t)}{\overline{F}_{\rm coh}(t)} = (1+t/2)\overline{F}(t) = \overline{F}(t) + \overline{F}(4/t)\;. \tag{54}$$

In addition, we have

$$\frac{d\overline{F}}{dt} = -\frac{1}{2}\int d^2\beta\,d^2\nu\,|\beta-\nu|^2 e^{-t|\beta-\nu|^2/2}W_\rho(\beta)W_\rho(\nu)\;, \tag{55}$$

which implies

$$
\begin{aligned}
\left.\frac{d\overline{F}}{dt}\right|_{t=0} &= -\frac{1}{2}\int d^2\beta\, d^2\nu\, |\beta-\nu|^2 W_\rho(\beta)W_\rho(\nu) \\
&= -\frac{1}{2}\int d^2\beta\, d^2\nu\, \big(|\beta|^2+|\nu|^2-\beta\nu^*-\beta^*\nu\big)W_\rho(\beta)W_\rho(\nu) \\
&= -\frac{1}{2}\big(\langle\psi|(b^\dagger b+bb^\dagger)|\psi\rangle - 2|\langle\psi|b|\psi\rangle|^2\big) \\
&= -\frac{1}{2}\big(\langle\psi|(\Delta x_B)^2|\psi\rangle + \langle\psi|(\Delta p_B)^2|\psi\rangle\big) \\
&= -\frac{1}{2}\big((\Delta x_B-\Delta p_B)^2 + 2\Delta x_B\Delta p_B\big) \\
&\le -\Delta x_B\Delta p_B \\
&\le -\frac{1}{2}\ ,
\end{aligned}
\tag{56}
$$

with equality if and only if $|\psi\rangle$ is a coherent state. Thus coherent states have the smallest initial rate of decrease of average fidelity as t increases from zero. Of course, the same information is contained in Eq. (50) specialized to $t=0$,

$$
\left.\frac{d\overline{F}}{dt}\right|_{t=0} = -\frac{1}{2}\int \frac{d^2\mu}{\pi}\, |\mu|^2|\Phi_\rho(\mu)|^2\ , \tag{57}
$$

except that here this information is encoded in the large-scale structure of the characteristic function and, hence, in the small-scale structure of the Wigner function. Fourier transform yields an expression in terms of the Wigner function:

$$
\begin{aligned}
\left.\frac{d\overline{F}}{dt}\right|_{t=0} &= -\frac{1}{2}\int \frac{d^2\mu}{\pi}\, |\mu|^2|\Phi_\rho(\mu)|^2 \\
&= -\frac{\pi}{2}\int d^2\nu\, \left|\frac{\partial W_\rho}{\partial\nu^*}\right|^2 \\
&= -\frac{\pi}{4}\int d^2\nu\, \left|\frac{\partial W_\rho}{\partial\nu_1}+i\frac{\partial W_\rho}{\partial\nu_2}\right|^2 \\
&= -\frac{\pi}{4}\int d^2\nu\, \left[\left(\frac{\partial W_\rho}{\partial\nu_1}\right)^2+\left(\frac{\partial W_\rho}{\partial\nu_2}\right)^2\right] \\
&= -\frac{\pi}{8}\int d\nu_1\, d\nu_2\, \left[\left(\frac{\partial W_\rho}{\partial\nu_1}\right)^2+\left(\frac{\partial W_\rho}{\partial\nu_2}\right)^2\right] \\
&= -\frac{\pi}{8}\int d\nu_1\, d\nu_2\, |\nabla W_\rho|^2
\end{aligned}
\tag{58}
$$

This derivative provides a sensible measure of the small-scale structure in the Wigner function. If we approximate $\overline{F}(t)\simeq 1+t(d\overline{F}/dt)_{t=0}$ and ask when this approximation goes to zero, we get a critical value of t given by

$$t_0 = -\frac{1}{(d\overline{F}/dt)_{t=0}} = \frac{1}{\big|(d\overline{F}/dt)_{t=0}\big|}\;; \tag{59}$$

the appropriate measure of the size of the small-scale structure is

$$\begin{aligned}
\ell \equiv \sqrt{t_0/2} &= \frac{1}{\sqrt{2\big|(d\overline{F}/dt)_{t=0}\big|}} \\
&= \frac{1}{\sqrt{\langle\psi|(\Delta x_B)^2|\psi\rangle + \langle\psi|(\Delta p_B)^2|\psi\rangle}} \\
&= \left(\frac{\pi}{4}\int d\nu_1\, d\nu_2\, |\nabla W_\rho|^2\right)^{-1/2}\;,
\end{aligned} \tag{60}$$

which is exactly what we might have guessed for a good measure of small-scale structure without any consideration of teleportation fidelity. For coherent states, we get $\ell = 1$, and for squeezed states with squeeze parameter u, we get $\ell = 1/\sqrt{\cosh 2u}$.

We can get improved bounds on average fidelity by returning to the forms

$$\overline{F}(t) = \frac{2}{t}\int \frac{d^2\mu}{\pi}\, e^{-2|\mu|^2/t}|\Phi_\rho(\mu)|^2 = \int \frac{d^2\mu}{\pi}\, e^{-t|\mu|^2/2}|\Phi_\rho(\mu)|^2\;. \tag{61}$$

Given the constraints (48) on $|\Phi_\rho(\mu)|^2$, the average fidelity is bounded above by the case where $|\Phi_\rho(\mu)|^2$ is as tightly confined about $\mu = 0$ as possible, i.e., $|\Phi_\rho(\mu)|^2$ vanishes outside a circle of radius $|\nu| = 1$ and is equal to 1 inside the circle. This will not give a tight upper bound because no pure state has this characteristic function. Thus we get

$$\begin{aligned}
\overline{F}(t) &\le \frac{2}{t}\int_{|\mu|\le 1} \frac{d^2\mu}{\pi}\, e^{-2|\mu|^2/t} = \frac{2}{t}\int_0^1 dx\, e^{-2x/t} = 1 - e^{-2/t}\;, \\
\overline{F}(t) &\le \int_{|\mu|\le 1} \frac{d^2\mu}{\pi}\, e^{-t|\mu|^2/2} = \int_0^1 dx\, e^{-tx/2} = \frac{2}{t}(1 - e^{-t/2})\;.
\end{aligned} \tag{62}$$

We can also use the Schwarz inequality in the following way:

$$\begin{aligned}
\overline{F}(t) &= \left(\frac{2}{t}\int \frac{d^2\mu}{\pi}\, e^{-2|\mu|^2/t}|\Phi_\rho(\mu)|^2\right)^{1/2}\left(\int \frac{d^2\mu}{\pi}\, e^{-t|\mu|^2/2}|\Phi_\rho(\mu)|^2\right)^{1/2} \\
&\ge \sqrt{\frac{2}{t}}\int \frac{d^2\mu}{\pi}\, e^{-|\mu|^2(1/t+t/4)}|\Phi_\rho(\mu)|^2 \\
&= \sqrt{\frac{2}{t}}\, I\!\left(\frac{1}{t} + \frac{t}{4}\right)\;.
\end{aligned} \tag{63}$$

This implies two inequalities:

$$\begin{aligned}
\overline{F}(t) &\ge \sqrt{\frac{2}{t}}\,\overline{F}\!\left(\frac{2}{t} + \frac{t}{2}\right)\;, \\
\overline{F}(t) &\ge \sqrt{\frac{2}{t}}\,\frac{2}{2/t + t/2}\,\overline{F}\!\left(\frac{4}{2/t + t/2}\right)\;.
\end{aligned}$$

The most useful results of all these manipulations are the various forms for $d\overline{F}/dt|_{t=0}$ and the relation of this derivative to the scale of sub-Planck structure. All the bounds turn out to be irrelevant now, because I am able to show that the maximum average fidelity is given by the coherent-state fidelity for all values of t. To do so, return to the last expression for the average fidelity in Eq. (33),

$$\overline{F}_\rho(t) = \int d^2\beta\, d^2\nu\, e^{-t|\beta-\nu|^2/2} W_\rho(\beta) W_\rho(\nu) \ . \tag{64}$$

Notice that this fidelity can be thought of as the average value of $e^{-t|\beta-\nu|^2/2}$ with respect to a product state $\rho\otimes\rho$ of two modes, b and v, the joint Wigner function for the two modes being $W_{BV}(\beta,\nu) = W_\rho(\beta)W_\rho(\nu)$. Introducing modes $c = (b+v)/\sqrt{2}$ and $d = (b-v)/\sqrt{2}$, with corresponding c-number variables $\gamma = (\beta+\nu)/\sqrt{2}$ and $\delta = (\beta-\nu)/\sqrt{2}$, we can re-write the average fidelity (64) as

$$\overline{F}_\rho(t) = \int d^2\gamma\, d^2\delta\, e^{-t|\delta|^2} W_{CD}(\gamma,\delta) \ , \tag{65}$$

where $W_{CD}(\gamma,\delta) = W_{BV}(\beta,\nu)$.

What we see is that the average fidelity is the expectation value of the mode-d operator A_t whose symmetrically ordered associated function is $e^{-t|\delta|^2}$. Letting $t = 1/(\bar{n}+1/2)$, we see that A_t is given by $\bar{n}+1/2$ times the density operator for a thermal state of mode d whose mean number of photons is

$$\bar{n} = \frac{1}{t}\left(1-\frac{t}{2}\right) \quad \Longrightarrow \quad 1+\bar{n} = \frac{1}{t}\left(1+\frac{t}{2}\right) \ , \tag{66}$$

i.e.,

$$\begin{aligned} A_t &= \frac{\bar{n}+1/2}{1+\bar{n}}\left(\frac{\bar{n}}{1+\bar{n}}\right)^{d^\dagger d} \\ &= \frac{1}{1+t/2}\left(\frac{1-t/2}{1+t/2}\right)^{d^\dagger d} \\ &= \frac{1}{1+t/2}\left(\frac{1-t/2}{1+t/2}\right)^{(b-v)^\dagger(b-v)/2} \ . \end{aligned} \tag{67}$$

Thus we can write the average fidelity as

$$\overline{F}_\rho(t) = \mathrm{tr}\big(\rho\otimes\rho A_t\big) \tag{68}$$

For $t=0$, $A_t = 1$, and we recover the result that the average fidelity is 1 regardless of the input state. Using

$$\left.\frac{dA_t}{dt}\right|_{t=0} = -\left(d^\dagger d + \frac{1}{2}\right) = -\frac{1}{2}\big(d^\dagger d + dd^\dagger\big) = -\frac{1}{4}\big(b^\dagger b + bb^\dagger + v^\dagger v + vv^\dagger - 2b^\dagger v - 2v^\dagger b\big) \ , \tag{69}$$

we can re-derive Eq. (56), i.e.,

$$\left.\frac{d\overline{F}}{dt}\right|_{t=0} = \mathrm{tr}\left(\rho\otimes\rho\left.\frac{dA_t}{dt}\right|_{t=0}\right) = -\frac{1}{2}\big(\langle b^\dagger b + bb^\dagger\rangle - 2|\langle b\rangle|^2\big) \ , \tag{70}$$

since b and v are in the same state ρ.

Generally, we can use Eq. (67) to bound the average fidelity by the largest eigenvalue of A_t:

$$\overline{F}_\rho(t) \le \begin{pmatrix} \text{largest} \\ \text{eigenvalue} \\ \text{of } A_t \end{pmatrix} = \frac{1}{1+t/2} \ . \tag{71}$$

The reason this is the largest eigenvalue is that the factor in large parentheses in the expression (67) for A_t has magnitude ≤ 1, which means that the largest eigenvalue, corresponding to the vacuum state for mode d, is $(1+t/2)^{-1}$. Since coherent states saturate the upper bound, we can write

$$\overline{F}_{\max}(t) = \overline{F}_{\text{coh}}(t) = \frac{1}{1+t/2} \ . \tag{72}$$

Notice that the bound on the expectation value of A_t holds for all joint states of modes b and v, with arbitrary Wigner functions $W_{BV}(\beta,\nu)$, not just the pure product copy states that are relevant to the average fidelity. In general, we can say that the bound is saturated if and only if mode d is in vacuum. Considering just pure product states $|\Psi\rangle = |\psi_B\rangle \otimes |\psi_B\rangle$, however, we can immediately show that only copy coherent states achieve the fidelity bound. Saturating the fidelity bound requires that mode d be in vacuum, i.e., that

$$0 = d|\Psi\rangle = \frac{1}{\sqrt{2}}(b-v)|\Psi\rangle \ . \tag{73}$$

This condition becomes $b|\psi_B\rangle \otimes |\psi_V\rangle = |\psi_B\rangle \otimes v|\psi_V\rangle$, which requires that $b|\psi_B\rangle = |\psi_B\rangle\langle\psi_V|v|\psi_V\rangle$ and $v|\psi_V\rangle = |\psi_V\rangle\langle\psi_B|b|\psi_B\rangle$, i.e., that $|\psi_B\rangle = |\psi_V\rangle$ be a coherent state. Thus the fidelity bound, which requires that the state of B and V be a product copy state, is saturated only by coherent states.

Our conclusion is that the gold standard for quantum teleportation is teleporting a non-Gaussian state with fidelity exceeding $\overline{F}_{\max}(1) = 2/3$.

Another way of thinking is to adopt the point of view advanced for qubit teleportation by Toner and Bacon. We can make an exact hidden-variable model for any continuous-variable teleportation if Alice is allowed to transmit the vacuum-strength ($t=1$) kick to Bob, who removes it from his mode along with the measured values X and P. The typical phase-space size of the kick is $1/\sqrt{2}$, and it must be removed to the same accuracy, i.e., $e^{-r}/\sqrt{2}$, as for the measured values. Thus the number of bits required to transmit the two quadrature components of the kick is roughly $2\log(e^r) \sim 2\log(1/\ell) = \log L^2$, i.e., half the number of bits required for sufficiently accurate transmission of X and P.

Another application of the technique used to show the fidelity bound (71) is to show that the largest fidelity for teleporting coherent states using a separable state for A and B is $1/2$. To see this, first assume that the state of modes A and B is a pure separable state, i.e., a pure product state. Then use $|\Phi_{\text{coh}}(\nu)|^2 = e^{-|\nu|^2}$ to write the average fidelity for teleporting a coherent state as

$$\begin{aligned} \overline{F} &= \int d^2\nu\, P(\nu)\, e^{-|\nu|^2} \\ &= \int d^2\nu\, d^2\alpha\, e^{-|\nu|^2} W_A(\alpha) W_B(\nu-\alpha^*) \\ &= \int d^2\alpha\, d^2\beta\, e^{-|\alpha-\beta|^2} W_A(-\alpha^*) W_B(\beta) \ . \end{aligned} \tag{74}$$

The quasidistribution $W_A(-\alpha^*)$ is the Wigner function for the time-reversed, parity-inverted state of mode A. We can now use the general bound to obtain $\overline{F} \le 1/2$ ($t = 2$), with equality holding if and only if mode A is in a coherent state $|\alpha\rangle$ and mode B is in the time-reversed, parity-inverted coherent state $|-\alpha^*\rangle$. Now suppose the state of modes A and B is a separable state, thus having a product-pure-state ensemble decomposition. The fidelity is now the average over the product-pure-state ensemble, which shows that the fidelity is bounded above by 1/2, with equality if and only if the state is a mixture of product states of the form $|\alpha\rangle \otimes |-\alpha^*\rangle$.* Notice that the correlations in this kind of state are a classical version of the quantum correlations in the squeezed state (24).

6. Number-state analysis

The diffusion superoperator that takes the input to the average output is given by

$$\mathcal{D} \equiv \frac{2}{t} \int \frac{d^2\nu}{\pi}\, e^{-2|\nu|^2/t} D(b,\nu) \odot D^\dagger(b,\nu)\;, \tag{75}$$

i.e., $\overline{\rho}_{\rm out} = \mathcal{D}(\rho)$. It is useful to find the number-state representation of $\mathcal{D}$:

$$\begin{aligned} \mathcal{D}_{lj,mk} &= \big\langle l \big| \mathcal{D}(|j\rangle\langle k|) \big| m \big\rangle \\ &= \frac{2}{t} \int \frac{d^2\nu}{\pi}\, e^{-2|\nu|^2/t} \langle l|D(b,\nu)|j\rangle\langle k|D^\dagger(b,\nu)|m\rangle \\ &= \frac{2}{t} \int \frac{d^2\nu}{\pi}\, e^{-2|\nu|^2/t} \langle l|D(b,\nu)|j\rangle\langle m|D(b,\nu)|k\rangle^*\;. \end{aligned} \tag{76}$$

Using

$$\langle m|D(b,\nu)|n\rangle = \begin{cases} \sqrt{\dfrac{n!}{m!}}\, e^{-|\nu|^2/2} \nu^{m-n} L_n^{(m-n)}\big(|\nu|^2\big)\;, & m \ge n \\ \sqrt{\dfrac{m!}{n!}}\, e^{-|\nu|^2/2} (-\nu^*)^{n-m} L_m^{(n-m)}\big(|\nu|^2\big)\;, & m \le n \end{cases}\;, \tag{77}$$

we see that the phase integral in $d^2\mu$ makes $\mathcal{D}_{lj,mk}$ vanish unless $l - j = m - k$. This allows us

* A mixed state that has $\overline{F} = 1/2$ has other other ensemble decompositions, of course. In such a decomposition, the entangled states can give fidelity $> 1/2$, and product states give fidelities $\le 1/2$. In any separable decomposition, which has *only* product states, all of these product states must be of the form $|\alpha\rangle \otimes |-\alpha^*\rangle$.

to write

$$
\begin{aligned}
\mathcal{D}_{lj,mk} &= \delta_{l-j,m-k}\frac{2}{t}\int\frac{d^2\nu}{\pi}\,e^{-|\nu|^2(1+2/t)} \\
&\qquad\qquad\times\begin{cases}\sqrt{\dfrac{j!\,k!}{l!\,m!}}|\nu|^{2(m-k)}L_j^{(m-k)}\big(|\nu|^2\big)L_k^{(m-k)}\big(|\nu|^2\big)\;, & m\ge k\\ \sqrt{\dfrac{l!\,m!}{j!\,k!}}|\nu|^{2(k-m)}L_l^{(k-m)}\big(|\nu|^2\big)L_m^{(k-m)}\big(|\nu|^2\big)\;, & m<k\end{cases} \\
&= \delta_{l-j,m-k}\frac{2}{t}\int dx\,e^{-x(1+2/t)}\begin{cases}\sqrt{\dfrac{j!\,k!}{l!\,m!}}x^{m-k}L_j^{(m-k)}(x)L_k^{(m-k)}(x)\;, & m\ge k\\ \sqrt{\dfrac{l!\,m!}{j!\,k!}}x^{k-m}L_l^{(k-m)}(x)L_m^{(k-m)}(x)\;, & m<k\end{cases}\;.
\end{aligned}
\tag{78}
$$

Now we use G&R 7.414.4 to evaluate the integral in terms of a hypergeometric function:

$$
\begin{aligned}
\mathcal{D}_{lj,mk} &= \delta_{l-j,m-k}\frac{(j+m)!}{\sqrt{l!\,m!\,j!\,k!}} \\
&\qquad\times\begin{cases}\dfrac{(2/t)^{j+k+1}}{(1+2/t)^{j+m+1}}F\big(-k,-j;-j-m;1-t^2/4\big)\;, & m\ge k\\ \dfrac{(2/t)^{l+m+1}}{(1+2/t)^{j+m+1}}F\big(-m,-l;-j-m;1-t^2/4\big)\;, & m<k\end{cases} \\
&= \delta_{l-j,m-k}\frac{(j+m)!}{\sqrt{l!\,m!\,j!\,k!}} \\
&\qquad\times\begin{cases}\dfrac{(t/2)^{m-k}}{(1+t/2)^{j+m+1}}F\big(-k,-j;-j-m;1-t^2/4\big)\;, & m\ge k\\ \dfrac{(t/2)^{k-m}}{(1+t/2)^{j+m+1}}F\big(-m,-l;-j-m;1-t^2/4\big)\;, & m<k\end{cases}\;.
\end{aligned}
\tag{79}
$$

As Andrew Scott pointed out to me, the property $F(a,b;c;z)=(1-z)^{c-a-b}F(c-a,c-b;c;z)=(1-z)^{c-a-b}F(c-b,c-a;c;z)$ shows that the two expressions involving the hypergeometric function are the same when $l-j=m-k$, so we can use either one, i.e.,

$$
\mathcal{D}_{lj,mk}=\delta_{l-j,m-k}\frac{(j+m)!}{\sqrt{l!\,m!\,j!\,k!}}\frac{(t/2)^{m-k}}{(1+t/2)^{j+m+1}}F\big(-k,-j;-j-m;1-t^2/4\big)\;. \tag{80}
$$

We retain the two expressions in what follows, however, because it makes the subsequent conversion to Jacobi polynomials easier.

An important special case is

$$
\begin{aligned}
\mathcal{D}_{ln,mn} &= \big\langle l\big|\mathcal{D}(|n\rangle\langle n|)\big|m\big\rangle \\
&= \frac{2}{t}\int\frac{d^2\nu}{\pi}\,e^{-2|\nu|^2/t}\langle l|D(b,\nu)|n\rangle\langle m|D(b,\nu)|n\rangle^* \\
&= \delta_{lm}\frac{(n+m)!}{n!\,m!}\begin{cases}\dfrac{(t/2)^{m-n}}{(1+t/2)^{n+m+1}}F\big(-n,-n;-n-m;1-t^2/4\big)\ , & m\ge n,\\ \dfrac{(t/2)^{n-m}}{(1+t/2)^{n+m+1}}F\big(-m,-m;-n-m;1-t^2/4\big)\ , & m<n,\end{cases} \\
&= \delta_{lm}\frac{(n+m)!}{n!\,m!}\frac{(t/2)^{|m-n|}}{(1+t/2)^{n+m+1}}F\big(-N,-N;-n-m;1-t^2/4\big)\ ,
\end{aligned}
\tag{81}
$$

where $N\equiv\min(n,m)$. We can write this in terms of the Jacobi polynomials, $P_n^{(\alpha,\beta)}(x)$, by using A&S 22.5.43 with $n\to N$, $\alpha\to 0$, $\beta\to|m-n|$,

$$
x\to-\frac{1+t^2/4}{1-t^2/4}\ ,\quad\text{i.e.,}\quad g(x)=\frac{2}{1-x}\to 1-t^2/4\ ,
$$

which gives the other parameters in A&S 22.5.43 as $a=-n\to-N$, $b=-n\to-N$, $c=-2n-\alpha-\beta\to-n-m$, and

$$
d=\binom{2n+\alpha+\beta}{n}\left(\frac{x-1}{2}\right)^n\to\frac{(n+m)!}{n!\,m!}\left(-\frac{1}{1-t^2/4}\right)^N\ .
$$

The result is

$$
\begin{aligned}
\frac{(n+m)!}{n!\,m!}F\big(-N,-N;-n-m;1-t^2/4\big) &= (-1)^N\big(1-t^2/4\big)^N P_N^{(0,|m-n|)}\bigg(-\frac{1+t^2/4}{1-t^2/4}\bigg) \\
&= \big(1-t^2/4\big)^N P_N^{(|m-n|,0)}\bigg(\frac{1+t^2/4}{1-t^2/4}\bigg)\ ,
\end{aligned}
\tag{82}
$$

where we use $(-1)^nP_n^{(\alpha,\beta)}(-x)=P_n^{(\beta,\alpha)}(x)$ (A&S 22.4.1) in the last line. This allows us to rewrite Eq. (81) as

$$
\begin{aligned}
\mathcal{D}_{ln,mn} &= \delta_{lm}\frac{(t/2)^{|m-n|}}{(1+t/2)^{n+m+1}}\big(1-t^2/4\big)^N P_N^{(|m-n|,0)}\bigg(\frac{1+t^2/4}{1-t^2/4}\bigg) \\
&= \delta_{lm}\frac{(t/2)^{|m-n|}(1-t/2)^N}{(1+t/2)^{M+1}}\,P_N^{(|m-n|,0)}\bigg(\frac{1+t^2/4}{1-t^2/4}\bigg)\ ,
\end{aligned}
\tag{83}
$$

where $M\equiv\max(n,m)$.

This particular case allows us to find an explicit expression for the average output state when the input is a number state $|n\rangle$,

$$
\overline{\rho}_{\rm out}=\mathcal{D}\big(|n\rangle\langle n|\big)=\sum_{l,m}\mathcal{D}_{ln,mn}|l\rangle\langle m|=\sum_m\mathcal{D}_{mn,mn}|m\rangle\langle m|\ .
\tag{84}
$$

This confirms the earlier assertion that for an input number state, $\overline{\rho}_{\rm out}$ is diagonal in the number-state representation, and it gives us an explicit formula for the average fidelity for input number states:

$$\overline{F}_{|n\rangle\langle n|}(t)=\mathcal{D}_{nn,nn}=\frac{(1-t/2)^n}{(1+t/2)^{n+1}}\,P_n\!\left(\frac{1+t^2/4}{1-t^2/4}\right)\;, \tag{85}$$

where we use $P_n^{(0,0)}(x)=P_n(x)$ (A&S 22.5.35) to reduce the Jacobi polynomial to a Legendre polynomial. It is easy to check that this expression satisfies the general relation (34).

For $s=-1$, the average fidelity is

$$\overline{F}_{|n\rangle\langle n|}(1)=\frac{2}{3^{n+1}}P_n(5/3)\;. \tag{86}$$

The case $s=-2$ is the only one for which we have to think a bit, since the argument of the Legendre polynomial blows up. Since P_n is a polynomial of rank n, the singularity in the argument is compensated by the factor in front, and the only term in $P_n(x)$ that survives is the term of highest rank, $a_n x^n$, where

$$a_n=d_n c_0=\frac{1}{2^n}\binom{2n}{n}=\frac{1}{2^n}\frac{(2n)!}{(n!)^2}$$

(A&S 22.3.8). The average fidelity for $s=-2$ becomes

$$\overline{F}_{|n\rangle\langle n|}(2)=\frac{2^n}{2^{2n+1}}a_n=\frac{1}{2^{2n+1}}\frac{(2n)!}{(n!)^2}\;. \tag{87}$$

This being enough rummaging around in the A&S chapter on orthogonal polynomials, I close this section.

7. Mixed-state teleportation and entanglement fidelity

An appropriate measure for assessing the fidelity with which a mixed state ρ is teleported is the *entanglement fidelity*, which is the fidelity for teleporting Victor's half of a purification of ρ, thus transferring the entanglement to Bob (entanglement swapping). It is quite easy to see how to generalize all of our results to entanglement fidelity. Victor's mode is now entangled with a mode U that has annihilation operator u (c-number variable μ); the joint state of U and V is a pure state $\rho_{UV}=|\psi_{UV}\rangle\langle\psi_{UV}|$, which purifies Victor's state ρ, i.e., $\mathrm{tr}_U(\rho_{UV})=\rho$. Any such purification can be written as

$$|\psi_{UV}\rangle=1\otimes\sqrt{\rho}|\phi_{UV}\rangle=\sum_{n=0}^{\infty}|n\rangle\otimes\sqrt{\rho}|n\rangle\;, \tag{88}$$

where $|\phi_{UV}\rangle=\sum_{n=0}^{\infty}|n\rangle\otimes|n\rangle$ and $|n\rangle$ denotes arbitrary orthonormal bases in U and V.

The results of Section 1 generalize to the following:

$$p(\xi)=\int d^2\mu\,d^2\nu\,W_{\rho_{UV}}(\mu,\nu)W_A(\xi^*-\nu^*)\;, \tag{89}$$

$$W_{\text{out}}(\mu,\beta|\,\xi)=\frac{1}{p(\xi)}\int d^2\nu\,W_{\rho_{UV}}(\mu,\nu)W_{AB}(\xi^*-\nu^*,\beta-\xi)\;, \tag{90}$$

$$\overline{\rho}_{\text{out}}=\int d^2\nu\,P(\nu)1\otimes D(b,\nu)\rho_{UV}1\otimes D^\dagger(b,\nu)\;, \tag{91}$$

$$W_{\overline{\rho}_{\text{out}}}(\mu,\beta)=\int d^2\nu\,P(\nu)W_{\rho_{UV}}(\mu,\beta-\nu)\;, \tag{92}$$

$$\Phi_{\overline{\rho}_{\text{out}}}(\nu,\alpha)=\pi\tilde{P}(\alpha)\Phi_{\rho_{UV}}(\nu,\alpha)\;. \tag{93}$$

The average entanglement fidelity is

$$\begin{aligned}
\overline{F}_{\text{ent}}&=\int d^2\xi\,p(\xi)\langle\psi_{UV}|\rho_{\text{out}}(\xi)|\psi_{UV}\rangle\\
&=\langle\psi_{UV}|\,\overline{\rho}_{\text{out}}|\psi_{UV}\rangle\\
&=\int d^2\nu\,P(\nu)|\langle\psi_{UV}|1\otimes D(b,\nu)|\psi_{UV}\rangle|^2\\
&=\int d^2\nu\,P(\nu)\big|\text{tr}\big(\rho D(b,\nu)\big)\big|^2\\
&=\int d^2\nu\,P(\nu)|\Phi_\rho(\nu)|^2\\
&=\pi\int d^2\beta\,d^2\nu\,\tilde{P}(\beta-\nu)W_\rho(\beta)W_\rho(\nu)\;;
\end{aligned} \tag{94}$$

i.e., the average entanglement fidelity is given by the first and last forms of the average fidelity (21). In contrast, the analogue of the second form in Eq. (21) is

$$\begin{aligned}
\overline{F}_{\text{ent}}&=\langle\psi_{UV}|\,\overline{\rho}_{\text{out}}|\psi_{UV}\rangle\\
&=\pi^2\int d^2\mu\,d^2\beta\,W_{\overline{\rho}_{\text{out}}}(\mu,\beta)W_{\rho_{UV}}(\mu,\beta)\\
&=\pi\int d^2\mu\left(\pi\int d^2\beta\,d^2\nu\,P(\beta-\nu)W_{\rho_{UV}}(\mu,\beta)W_{\rho_{UV}}(\mu,\nu)\right)\;,
\end{aligned} \tag{95}$$

and the analogue of the third form is

$$\overline{F}_{\text{ent}}=\langle\psi_{UV}|\,\overline{\rho}_{\text{out}}|\psi_{UV}\rangle=\int\frac{d^2\nu\,d^2\alpha}{\pi^2}\,\Phi^*_{\overline{\rho}_{\text{out}}}(\nu,\alpha)\Phi_{\rho_{UV}}(\nu,\alpha)=\int\frac{d^2\nu}{\pi}\,d^2\alpha\,\tilde{P}(\alpha)|\Phi_{\rho_{UV}}(\nu,\alpha)|^2\;. \tag{96}$$

The meaning of these differences is the following. The characteristic function $\Phi_\rho(\nu)$ has fine structure on a small scale ℓ and coarse structure on a large scale L. These are inverted in the Wigner function: $W_\rho(\beta)$ has coarse structure on the scale $1/\ell$ and small-scale structure on the scale $1/L$. The difference with mixed states is that there is no necessary connection between the small and large scales; ℓ and L, instead of being related by $\ell L\sim 1$ as for pure states, satisfy $\ell^2\lesssim\ell L\lesssim 1$. Thus the small-scale structure $1/L$ in the Wigner function can be at any scale from ℓ all the way up to $1/\ell$, where there is no fine structure at all.

What the bottom two expressions for the entanglement fidelity in Eq. (94) mean is that to get high entanglement fidelity, $P(\nu)$ must be narrower than the fine scale ℓ in the characteristic function, meaning that $\tilde{P}(\alpha)$ is broader than the coarse scale $1/\ell$ in the Wigner function. This is just as it is for pure states, and it is sufficient for high-fidelity teleportation of entanglement. That the expressions for the entanglement fidelity in Eqs. (95) and (96) are different than for pure states means that the appropriate coarse structure in the characteristic function and the corresponding fine structure in the Wigner function are only captured in the functions for a joint purified state. In other words, the potential Wigner-function fine structure associated with a purification of ρ—structure that must be teleported to achieve high entanglement fidelity—is expressed in the quantity

$$\pi \int d^2\mu\, W_{\rho_{UV}}(\mu,\beta) W_{\rho_{UV}}(\mu,\nu) = \pi \int d^2\mu\, W_{\rho_{UV}}(\mu^*,\beta) W_{\rho_{UV}}(\mu^*,\nu) \tag{97}$$

or, equivalently, as course structure, in the quantity

$$\int \frac{d^2\nu}{\pi} |\Phi_{\rho_{UV}}(\nu,\alpha)|^2 = \int \frac{d^2\nu}{\pi} |\Phi_{\rho_{UV}}(-\nu^*,\alpha)|^2 \,. \tag{98}$$

The reason for the second forms in these expressions becomes clear as we manipulate them below.

Since the entanglement fidelity is independent of which purification is used, we can put these two quantities solely in terms of ρ. To do so, assume that the states $|n\rangle$ denote number states for U and V. The characteristic-function expression is preferable because it is a function of only one complex variable, so we deal with it first:

$$\begin{aligned}
\Phi_{\rho_{UV}}(-\nu^*,\alpha) &= \mathrm{tr}\big(\rho_{UV} D(u,-\nu^*) D(b,\alpha)\big) \\
&= \langle\psi_{UV}|D(u,-\nu^*) D(b,\alpha)|\psi_{UV}\rangle \\
&= \langle\phi_{UV}|(1\otimes\sqrt{\rho})\big(D(u,-\nu^*)\otimes D(b,\alpha)\big)(1\otimes\sqrt{\rho})|\phi_{UV}\rangle \\
&= \sum_{n,m} \langle n|D(u,-\nu^*)|m\rangle\langle n|\sqrt{\rho}D(b,\alpha)\sqrt{\rho}|m\rangle \\
&= \sum_{n,m} \langle m|D^T(u,-\nu^*)|n\rangle\langle n|\sqrt{\rho}D(b,\alpha)\sqrt{\rho}|m\rangle \\
&= \mathrm{tr}\big(D^T(b,-\nu^*)\sqrt{\rho}D(b,\alpha)\sqrt{\rho}\big) \\
&= \mathrm{tr}\big(D(b,\nu)\sqrt{\rho}D(b,\alpha)\sqrt{\rho}\big) \,.
\end{aligned} \tag{99}$$

The transposition and complex conjugation are taken with respect to the number basis; in the final step, we use

$$D^T(b,-\nu^*) = \big([D(b,-\nu^*)]^\dagger\big)^* = D^*(b,\nu^*) = D(b,\nu) \,, \tag{100}$$

remembering that the creation and annihilation operators are real in the number representation. The coarse structure in the characteristic function

$$\Phi_{\rho_{UV}}(-\nu^*,\alpha) = \mathrm{tr}\big(D(b,\nu)\sqrt{\rho}D(b,\alpha)\sqrt{\rho}\big) \tag{101}$$

captures the fine structure in the joint Wigner function for the particular purification based on the number representation. In its coarse structure, the corresponding integral (98),

$$\int \frac{d^2\nu}{\pi}\,|\Phi_{\rho_{UV}}(-\nu^*,\alpha)|^2 = \int \frac{d^2\nu}{\pi}\,\big|\mathrm{tr}\big(D(b,\nu)\sqrt{\rho}D(b,\alpha)\sqrt{\rho}\big)\big|^2\;, \tag{102}$$

captures directly the potential fine-scale Wigner-function structure in any purification. It is this potential fine structure that must be teleported to achieve high fidelity.

Turning to the Wigner-function expression, we start with

$$W_{\rho_{UV}}(\mu,\beta) = \mathrm{tr}\big(\rho_{UV}\tilde{D}^{(0)}(u,\mu)\tilde{D}^{(0)}(b,\beta)\big)\;, \tag{103}$$

where

$$\tilde{D}^{(0)}(b,\beta) = \int \frac{d^2\alpha}{\pi}\,D(b,\alpha)D(\alpha,\beta) = 2D(b,\beta)PD^\dagger(b,\beta) = 2PD(b,-2\beta) \tag{104}$$

is the Fourier transform of the displacement operator (P is the parity operator). Now proceeding just as for the characteristic function, we have

$$\begin{aligned}
W_{\rho_{UV}}(\mu^*,\beta) &= \mathrm{tr}\big(\rho_{UV}\tilde{D}^{(0)}(u,\mu^*)\tilde{D}^{(0)}(b,\beta)\big)\\
&= \langle\psi_{UV}|\tilde{D}^{(0)}(u,\mu^*)\tilde{D}^{(0)}(b,\beta)|\psi_{UV}\rangle\\
&= \langle\phi_{UV}|(1\otimes\sqrt{\rho})\big(\tilde{D}^{(0)}(u,\mu^*)\tilde{D}^{(0)}(b,\beta)\big)(1\otimes\sqrt{\rho})|\phi_{UV}\rangle\\
&= \sum_{n,m}\langle n|\tilde{D}^{(0)}(u,\mu^*)|m\rangle\langle n|\sqrt{\rho}\tilde{D}^{(0)}(b,\beta)\sqrt{\rho}|m\rangle\\
&= \sum_{n,m}\langle m|\tilde{D}^{(0)T}(u,\mu^*)|n\rangle\langle n|\sqrt{\rho}\tilde{D}^{(0)}(b,\beta)\sqrt{\rho}|m\rangle\\
&= \mathrm{tr}\big(\tilde{D}^{(0)T}(b,\mu^*)\sqrt{\rho}\tilde{D}^{(0)}(b,\beta)\sqrt{\rho}\big)\\
&= \mathrm{tr}\big(\tilde{D}^{(0)}(b,\mu)\sqrt{\rho}\tilde{D}^{(0)}(b,\beta)\sqrt{\rho}\big)\;,
\end{aligned} \tag{105}$$

where in the last step, we use

$$\tilde{D}^{(0)T}(b,\mu^*) = \big([\tilde{D}^{(0)}(b,\mu^*)]^\dagger\big)^* = \tilde{D}^{(0)*}(b,\mu^*) = \tilde{D}^{(0)}(b,\mu)\;. \tag{106}$$

The fine structure in the two-variable function

$$\begin{aligned}
\pi\int d^2\mu\,W_{\rho_{UV}}(\mu^*,\beta)W_{\rho_{UV}}(\mu^*,\nu) = \pi\int d^2\mu\,&\mathrm{tr}\big(\tilde{D}^{(0)}(b,\mu)\sqrt{\rho}\tilde{D}^{(0)}(b,\beta)\sqrt{\rho}\big)\\
&\times\mathrm{tr}\big(\tilde{D}^{(0)}(b,\mu)\sqrt{\rho}\tilde{D}^{(0)}(b,\nu)\sqrt{\rho}\big)
\end{aligned} \tag{107}$$

captures directly the potential fine-scale Wigner-function structure in any purification of ρ.

8. Required classical communication and sub-Planck structure

We can get a more quantitative idea of the classical communication required for continuous-variable teleportation—and its relation to the scale of sub-Planck structure—by considering the effect of having Alice send her information through a noisy Gaussian channel. Formally, we mean that the probability that Bob receives complex variable η, given that Alice sends ξ, is a Gaussian

$$p(\eta|\,\xi) = \frac{1}{\pi\sigma_1^2} e^{-|\eta-\xi|^2/\sigma_1^2} \; . \tag{108}$$

The state of Bob's mode, given that he receives η, becomes

$$\begin{aligned}
W'(\beta|\,\eta) &= \int d^2\xi \, W'(\beta|\,\xi) p(\xi|\,\eta) \\
&= \int d^2\xi \, W'(\beta|\,\xi) \frac{p(\eta|\,\xi)p(\xi)}{p(\eta)} \\
&= \frac{1}{p(\eta)} \int d^2\nu \, W_\rho(\nu) \int d^2\xi \, p(\eta|\,\xi) W_{AB}(\xi^* - \nu^*, \beta) \\
&= \frac{1}{p(\eta)} \int d^2\nu \, W_\rho(\nu) \underbrace{\int \frac{d^2\xi}{\pi\sigma_1^2} \, e^{-|\xi|^2/\sigma_1^2} W_{AB}(\eta^* - \nu^* + \xi^*, \beta)}_{= \hat{W}_{AB}(\eta^* - \nu^*, \beta)} \; .
\end{aligned} \tag{109}$$

The effect of the noisy channel is thus just the same as changing the initial state of modes A and B to a state with Wigner function

$$\hat{W}_{AB}(\alpha, \beta) = \int \frac{d^2\xi}{\pi\sigma_1^2} \, e^{-|\xi|^2/\sigma_1^2} W_{AB}(\alpha + \xi^*, \beta) = \int \frac{d^2\xi}{\pi\sigma_1^2} \, e^{-|\xi|^2/\sigma_1^2} P(\nu + \xi) R(\mu - \xi/2|\,\nu + \xi) \; , \tag{110}$$

i.e., a convolution of the original Wigner function with a Gaussian. In the last expression in Eq. (110), we use the previous definitions (12) for ν and μ. The probability for Bob to receive η becomes

$$p(\eta) = \int d^2\beta \, W'(\beta|\,\eta) p(\eta) = \int d^2\nu \, W_\rho(\nu) \hat{W}_A(\eta^* - \nu^*) \; . \tag{111}$$

The entire effect of the new Wigner function (110) can be described by saying that in calculating expectation values involving α and β, α is replaced by $\alpha - \xi^*$, where α and ξ are independent complex random variables distributed according to $W_{AB}(\alpha, \beta)$ and $e^{-|\xi|^2/\sigma_1^2}/\pi\sigma_1^2$, respectively. Translated to the variables $\gamma = (\alpha + \beta)/\sqrt{2}$ and $\delta = (\alpha - \beta)\sqrt{2}$, this becomes

$$\begin{aligned}
\gamma_1 &\to \gamma_1 - \xi_1/\sqrt{2} \; , \\
\gamma_2 &\to \gamma_2 + \xi_2/\sqrt{2} \; , \\
\delta_1 &\to \delta_1 - \xi_1/\sqrt{2} \; , \\
\delta_2 &\to \delta_2 + \xi_2/\sqrt{2} \; ,
\end{aligned} \tag{112}$$

where ξ_1 and ξ_2 are uncorrelated, zero-mean Gaussian random variables, both having variance σ_1^2.

Now specialize to squeezed-state teleportation. We get the following correlation matrix:

$$\begin{aligned}
\langle\gamma_1^2\rangle &= \langle\delta_2^2\rangle = \frac{1}{2}(e^{-2r}+\sigma_1^2)\ ,\\
\langle\gamma_2^2\rangle &= \langle\delta_1^2\rangle = \frac{1}{2}(e^{2r}+\sigma_1^2)\ ,\\
\langle\gamma_1\delta_1\rangle &= \langle\gamma_2\delta_2\rangle = \frac{1}{2}\sigma_1^2\ ,\\
\langle\gamma_1\gamma_2\rangle &= \langle\gamma_1\delta_2\rangle = \langle\gamma_2\delta_1\rangle = \langle\delta_1\delta_2\rangle = 0\ .
\end{aligned} \tag{113}$$

This means that we can update all our expressions involving the average fidelity by replacing $P(\nu)$ by

$$\hat{P}(\nu) = \int \frac{d^2\xi}{\pi\sigma_1^2}\, e^{-|\xi|^2/\sigma_1^2} P(\nu+\xi) = \frac{1}{\pi(e^{-2r}+\sigma_1^2)}\, e^{-|\nu|^2/(e^{-2r}+\sigma_1^2)} = \frac{2}{\pi\hat{t}}\, e^{-2|\nu|^2/\hat{t}}\ , \tag{114}$$

i.e., by replacing t everywhere by $\hat{t} = 2(e^{-2r}+\sigma_1^2) = t + 2\sigma_1^2$. Squeezed-state teleportation also allows us to simplify Eq. (110) to

$$\hat{W}_{AB}(\alpha,\beta) = \int \frac{d^2\xi}{\pi\sigma_1^2}\, e^{-|\xi|^2/\sigma_1^2} P(\nu+\xi) R(\mu-\xi/2)\ . \tag{115}$$

Nothing up till now has required any assumptions about the amount of squeezing. We're now going to work in the high-fidelity limit. In this limit we can write the average fidelity as

$$\overline{F}(\hat{t}) = \overline{F}(0) + \hat{t}\,\frac{d\overline{F}}{dt}\bigg|_{t=0} = 1 - \frac{e^{-2r}+\sigma_1^2}{t_0/2} = 1 - \frac{e^{-2r}+\sigma_1^2}{\ell^2}\ . \tag{116}$$

The high-fidelity limit requires that e^{-2r} and σ_1^2 both be small compared to the small-scale structure in the Wigner function; i.e., both are small compared to $t_0/2 = \ell^2$, as one can see from Eq. (116). In this limit, we can set $\xi = 0$ in the broad Gaussian in Eq. (115), obtaining

$$\hat{W}_{AB}(\alpha,\beta) = \underbrace{R(\mu)}_{=\hat{R}(\mu)} \underbrace{\int \frac{d^2\xi}{\pi\sigma_1^2}\, e^{-|\xi|^2/\sigma_1^2} P(\nu+\xi)}_{=\hat{P}(\nu)}\ . \tag{117}$$

In the high-fidelity limit, the probability for Bob to receive η becomes

$$p(\eta) = \int d^2\beta\, d^2\nu\, W_\rho(\beta-\nu)\hat{P}(\nu)\hat{R}(\beta-\eta-\nu/2) = \hat{R}(\langle v\rangle-\eta) = \frac{4e^{-2r}}{\pi} e^{-4e^{-2r}|\,\eta-\langle v\rangle|^2}\ , \tag{118}$$

where in the latter two expressions we make the same approximations as in Eq. (37).

The information Alice sends to Bob can be quantified by the mutual information between the complex variables η and ξ,

$$I_1 = -\int d^2\eta\, d^2\xi\, p(\eta|\,\xi)p(\xi)\log\!\left(\frac{p(\eta)}{p(\eta|\,\xi)}\right) = \log\!\left(\frac{e^{2r}/4}{\sigma_1^2}\right)\ , \tag{119}$$

which relates the variance of the noisy channel to the transmitted information,

$$\sigma_1^2 = \frac{1}{4}\, e^{2r} 2^{-I_1} = \frac{1}{4}\, 2^{2r/\ln 2 - I_1} \; . \tag{120}$$

Putting all this together, we get a final expression for the average fidelity:

$$\overline{F}(\hat{t}) = 1 - \frac{e^{-2r} + 2^{2r/\ln 2 - I_1}/4}{\ell^2} = 1 - \frac{e^{-2r}}{\ell^2} - \frac{1}{4}\, 2^{2r/\ln 2 + \log(1/\ell^2) - I_1} \; . \tag{121}$$

This expression tells us that to achieve good fidelity, we need to have sufficient squeezing that e^{-r} is somewhat smaller than ℓ and that I_1 is somewhat larger than $2r/\ln 2 + \log(1/\ell^2)$, which is somewhat larger than $\log(1/\ell^4)$.

Now let's consider another scenario, this time of the Bacon-Toner type. Suppose Alice gives Victor's state a phase-space kick τ, distributed according to

$$p(\tau) = \frac{2}{\pi} e^{-2|\tau|^2} \; . \tag{122}$$

This kicking strength, averaged over τ, turns the original Wigner function into the (positive) Husimi Q function. The state that Alice teleports is $D(v,\tau)\rho D^\dagger(v,\tau)$, with Wigner function $W_{D(v,\tau)\rho D^\dagger(v,\tau)}(\nu) = W_\rho(\nu - \tau)$. Alice communicates her measurement result ξ to Bob, after which the state of mode B is

$$W''(\beta|\,\xi,\tau) = \frac{1}{p(\xi)} \int d^2\nu\, W_\rho(\nu - \tau) W_{AB}(\xi^* - \nu^*, \beta) \; . \tag{123}$$

Now suppose Alice also communicates the kick strength τ over a noisy Gaussian channel where the probability that Bob receives kick strength κ, given that Alice sends τ, is

$$p(\kappa|\tau) = \frac{1}{\pi\sigma_2^2} e^{-|\kappa|^2/\sigma_2^2} \; . \tag{124}$$

Bob displaces mode B by $-\kappa$, leaving a Wigner function for mode B, after averaging over τ and κ, given by

$$\begin{aligned} W'(\beta|\,\xi) &= \int d^2\tau\, d^2\kappa\, p(\kappa|\tau) p(\tau) W''(\beta + \kappa|\,\xi,\tau) \\ &= \frac{1}{p(\xi)} \int d^2\tau\, d^2\kappa\, d^2\nu\, p(\kappa|\tau) p(\tau) W_\rho(\nu - \tau) W_{AB}(\xi^* - \nu^*, \beta + \kappa) \\ &= \frac{1}{p(\xi)} \int d^2\nu\, W_\rho(\nu) \underbrace{\int d^2\tau\, d^2\kappa\, p(\kappa|\tau) p(\tau) W_{AB}(\xi^* - \nu^* - \tau^*, \beta + \kappa)}_{= \check{W}_{AB}(\xi^* - \nu^*, \beta)} \; . \end{aligned} \tag{125}$$

The effect of the kicks is thus just the same as changing the initial state of modes A and B to a state with Wigner function

$$\begin{aligned} \check{W}_{AB}(\alpha,\beta) &= \int d^2\tau\, d^2\kappa\, p(\kappa|\tau) p(\tau) W_{AB}(\alpha - \tau^*, \beta + \kappa) \\ &= \int d^2\tau\, p(\tau) \int \frac{d^2\chi}{\pi\sigma_2^2} e^{-|\chi|^2/\sigma_2^2} W_{AB}(\alpha - \tau^*, \beta + \tau + \chi) \\ &= \int d^2\tau\, p(\tau) \int \frac{d^2\chi}{\pi\sigma_2^2}\, e^{-|\chi|^2/\sigma_2^2} P(\nu + \chi) R(\mu + \tau + \chi/2|\nu + \chi) \; . \end{aligned} \tag{126}$$

By integrating over μ, we get

$$\check{P}(\nu)=\int d^2\mu\,\check{W}_{AB}(\alpha,\beta)=\int d^2\tau\,p(\tau)\int\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\chi|^2/\sigma_2^2}P(\nu+\chi)=\int\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\chi|^2/\sigma_2^2}P(\nu+\chi)\;, \tag{127}$$

which replaces $P(\nu)$ in all calculations involving the average fidelity.

We could go on to analyze this scenario in the case of high-fidelity squeezed-state teleportation, but what we really want to do is to combine the two scenarios. We see that the effect of the two scenarios is the same as replacing the initial state of modes A and B by a state with Wigner function

$$\begin{aligned}\hat{W}_{AB}(\alpha,\beta)&=\int\frac{d^2\xi}{\pi\sigma_1^2}\,e^{-|\xi|^2/\sigma_1^2}\check{W}_{AB}(\alpha+\xi^*,\beta)\\&=\int d^2\tau\,p(\tau)\int\frac{d^2\xi}{\pi\sigma_1^2}\,\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\xi|^2/\sigma_1^2}\,e^{-|\chi|^2/\sigma_2^2}W_{AB}(\alpha+\xi^*-\tau^*,\beta+\tau+\chi)\\&=\int d^2\tau\,p(\tau)\\&\qquad\times\int\frac{d^2\xi}{\pi\sigma_1^2}\,\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\xi|^2/\sigma_1^2}\,e^{-|\chi|^2/\sigma_2^2}P(\nu+\xi+\chi)R(\mu+\tau-\xi/2+\chi/2|\nu+\xi+\chi)\;.\end{aligned} \tag{128}$$

By integrating over μ, we get

$$\hat{P}(\nu)=\int d^2\mu\,\hat{W}_{AB}(\alpha,\beta)=\int\frac{d^2\xi}{\pi\sigma_1^2}\,\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\xi|^2/\sigma_1^2}\,e^{-|\chi|^2/\sigma_2^2}P(\nu+\xi+\chi)\;, \tag{129}$$

which replaces $P(\nu)$ in all calculations involving the average fidelity.

Specializing to squeezed-state teleportation, we get

$$\hat{P}(\nu)=\frac{1}{\pi(e^{-2r}+\sigma_1^2+\sigma_2^2)}\,e^{-|\nu|^2/(e^{-2r}+\sigma_1^2+\sigma_2^2)}=\frac{2}{\pi\hat{t}}\,e^{-2|\nu|^2/\hat{t}}\;; \tag{130}$$

i.e., we can update all our results on average fidelity to this scenario by replacing t everywhere by $\hat{t}=2(e^{-2r}+\sigma_1^2+\sigma_2^2)=t+2\sigma_1^2+2\sigma_2^2$. Squeezed-state teleportation also allows us to simplify Eq. (128) to

$$\hat{W}(\alpha,\beta)=\int d^2\tau\,p(\tau)\int\frac{d^2\xi}{\pi\sigma_1^2}\,\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\xi|^2/\sigma_1^2}\,e^{-|\chi|^2/\sigma_2^2}P(\nu+\xi+\chi)R(\mu+\tau-\xi/2+\chi/2)\;. \tag{131}$$

In the high-fidelity limit, we can write the average fidelity as

$$\overline{F}(\hat{t})=\overline{F}(0)+\hat{t}\,\frac{d\overline{F}}{dt}\bigg|_{t=0}=1-\frac{e^{-2r}+\sigma_1^2+\sigma_2^2}{t_0/2}=1-\frac{e^{-2r}+\sigma_1^2+\sigma_2^2}{\ell^2}\;. \tag{132}$$

In the high-fidelity limit, we can also make the same approximations as before [these approximations are not so good for $p(\tau)$], obtaining

$$\hat{W}_{AB}(\alpha,\beta)=\underbrace{R(\mu)}_{=\hat{R}(\mu)}\underbrace{\int\frac{d^2\xi}{\pi\sigma_1^2}\,\frac{d^2\chi}{\pi\sigma_2^2}\,e^{-|\xi|^2/\sigma_1^2}\,e^{-|\chi|^2/\sigma_2^2}P(\nu+\xi+\chi)}_{=\hat{P}(\nu)}\;. \tag{133}$$

This allows us to conclude that in the high-fidelity limit the probability for Bob to receive η is given by Eq. (118), implying that the mutual information between η and ξ is given by Eq. (119).

We also need the mutual information between τ and κ. The unconditioned probability for κ is given by

$$p(\kappa) = \int d^2\tau\, p(\kappa|\tau)p(\tau) = \frac{1}{\pi(1/2+\sigma_2^2)} e^{-|\kappa|^2/(1/2+\sigma_2^2)} \; . \tag{134}$$

The mutual information between τ and κ is

$$I_2 = -\int d^2\kappa\, d^2\tau\, p(\kappa|\tau)p(\tau)\log\left(\frac{p(\kappa)}{p(\kappa|\tau)}\right) = \log\left(\frac{1/2+\sigma_2^2}{\sigma_2^2}\right) \; , \tag{135}$$

which in the high-fidelity limit, we approximate as $I_2 = \log(1/2\sigma_2^2)$, giving

$$\sigma_2^2 = 2^{-I_2-1} \; . \tag{136}$$

This allows us to put the fidelity in the final form

$$\overline{F}(\hat{t}) = 1 - \frac{e^{-2r} + 2^{2r/\ln 2 - I_1}/4 + 2^{-I_2-1}}{\ell^2} = 1 - \frac{e^{-2r}}{\ell^2} - \frac{1}{4}\, 2^{2r/\ln 2 + \log(1/\ell^2) - I_1} - 2^{\log(1/\ell^2) - I_2 - 1} \; . \tag{137}$$

To achieve high-fidelity in the Bacon-Toner scenario, I_2 must be somewhat larger than $\log(1/\ell^2)$, whereas I_1 must be somewhat larger than $\log(1/\ell^4)$, meaning that this scenario, which operates on the shared randomness of the hidden variables and the additional communication regarding the phase-space kicks, requires roughly 50% more classical communication than quantum teleportation, as already explained heuristically in Section 5.