List of proofs

Proof of Lemma 1. First, note that

$$\mathbf{E}\left[\|GX\|_1^2\right] = \mathbf{E}\Big[\Big(\sum\nolimits_{i=1}^{n}\|g_i\|_\infty\,|\hat{g}_i^{\mathrm{H}}X|\Big)^2\Big],$$

where g_i is the ith column of G and $\hat{g}_i = g_i/\|g_i\|_\infty$. Using the condition $\|G\|_{\infty,*} \le 1$ and Jensen's inequality, we find that

$$\mathbf{E}\left[\|GX\|_1^2\right] \le \sum\nolimits_{i=1}^{n}\|g_i\|_\infty\,\mathbf{E}\left[|\hat{g}_i^{\mathrm{H}}X|^2\right] \le \|\!|X|\!\|^2.$$

The other inequality follows by noting that for any $f \in \mathbb{C}^n$ with $\|f\|_\infty \le 1$, the matrix G with first row equal to f^{H} and all other rows zero satisfies the constraint $\|G\|_{\infty,*} \le 1$. That (26) is the dual norm of the ∞-norm follows from straightforward verification, or see [27, Proposition 7.2]. □

Proof of Lemma 2. The result holds in $n = 1$ dimensions. Suppose that the result holds in $n-1$ dimensions. We will show that it must also therefore hold in n dimensions and conclude, by induction, that the result holds in any dimension.

Let $\tilde{A}$ be the $(n-1)\times(n-1)$ principle submatrix of an $n \times n$ matrix A. For any vector $f \in \mathbb{C}^n$ we can write

$$\begin{aligned} f^{\mathrm{H}}Af &= \sum\nolimits_{i=1}^{n}|f_i|^2A_{ii} + 2\Re\left[\sum\nolimits_{i=1}^{n}\sum\nolimits_{j=1}^{i-1}\bar{f}_iA_{ij}f_j\right] \\ &= \tilde{f}^{\mathrm{H}}\tilde{A}\tilde{f} + |f_n|^2A_{nn} + 2\Re\left[\bar{f}_n\sum\nolimits_{j=1}^{n-1}A_{nj}\tilde{f}_j\right], \end{aligned}$$

where $\tilde{f} \in \mathbb{C}^{n-1}$ has entries equal to the first $n-1$ entries of f.

By the induction hypothesis, we can choose the first $n-1$ entries of f (i.e., $\tilde{f}$) so that the right-hand side of the last display is not less than

$$\sum\nolimits_{i=1}^{n-1}A_{ii} + |f_n|^2A_{nn} + 2\Re\left[\bar{f}_n\sum\nolimits_{j=1}^{n-1}A_{nj}\tilde{f}_j\right].$$

If, for this choice of $\tilde{f}$, $\sum_{j=1}^{n-1}A_{nj}\tilde{f}_j$ is nonzero, then choose f_n as

$$f_n = \frac{\sum_{j=1}^{n-1}A_{nj}\tilde{f}_j}{\left|\sum_{j=1}^{n-1}A_{nj}\tilde{f}_j\right|}.$$

Otherwise set $f_n = 1$. With the resulting choice of f_n,

$$|f_n|^2A_{nn} + 2\Re\left[\bar{f}_n\sum\nolimits_{j=1}^{n-1}A_{nj}\tilde{f}_j\right] \ge A_{nn}.$$

We have therefore shown that

$$\sup_{\|f\|_\infty \le 1} f^{\mathrm{H}}Af \ge \sum\nolimits_{i=1}^{n}A_{ii}.$$

□

Proof of Theorem 1. Let V_t^m be generated by (11). Let $Y_t^m = \Phi_t^m(V_t^m)$ and notice that

$$\begin{aligned} \mathcal{U}(V_t^m) &= \mathcal{U}(\mathcal{M}(Y_{t-1}^m)) \\ &\le R + \alpha\,\mathcal{U}(V_{t-1}^m) + \alpha\left(\mathcal{U}(Y_{t-1}^m) - \mathcal{U}(V_{t-1}^m)\right). \end{aligned}$$

Using the fact that $\mathcal{U}$ is twice differentiable with bounded second derivative, this last expression is bounded above by

$$\mathcal{U}(V_t^m) \le R + \alpha\,\mathcal{U}(V_{t-1}^m) + \alpha\nabla\mathcal{U}(V_{t-1}^m)\left(Y_{t-1}^m - V_{t-1}^m\right) + \frac{\alpha\sigma}{2}\|G\left(Y_{t-1}^m - V_{t-1}^m\right)\|_1^2.$$

Taking the expectation and using (30) yields

$$\mathbf{E}\left[\mathcal{U}(V_t^m)\right] \le R + \alpha\,\mathbf{E}\left[\mathcal{U}(V_{t-1}^m)\right] + \frac{\alpha\sigma}{2}\mathbf{E}\left[\|G\left(Y_{t-1}^m - V_{t-1}^m\right)\|_1^2\right].$$

An application of Lemma 1 reveals that

$$\mathbf{E}\left[\|G\left(Y_{t-1}^m - V_{t-1}^m\right)\|_1^2\right] \le |||Y_{t-1}^m - V_{t-1}^m|||^2.$$

As a consequence, noting (28), we arrive at the upper bound

$$\begin{aligned}\mathbf{E}\left[\mathcal{U}(V_t^m)\right] &\le R + \alpha\,\mathbf{E}\left[\mathcal{U}(V_{t-1}^m)\right] + \frac{\alpha\gamma^2\sigma}{2m}\mathbf{E}\left[\|V_{t-1}^m\|_1^2\right] \\ &\le R + \alpha\left(1 + \frac{\beta\gamma^2\sigma}{2m}\right)\mathbf{E}\left[\mathcal{U}(V_{t-1}^m)\right],\end{aligned}$$

from which we can conclude that

$$\mathbf{E}\left[\|V_t^m\|_1^2\right] \le \beta\mathbf{E}\left[\mathcal{U}(V_t^m)\right] \le \beta R\left[\frac{1-\alpha^t\left(1+\frac{\beta\gamma^2\sigma}{2m}\right)^t}{1-\alpha\left(1+\frac{\beta\gamma^2\sigma}{2m}\right)}\right] + \beta\alpha^t\left(1+\frac{\beta\gamma^2\sigma}{2m}\right)^t\mathcal{U}(V_0^m). \qquad \square$$

Proof of Theorem 2. We begin with a standard expansion of the scheme's error.

$$\begin{aligned}|||V_t^m - v_t||| &= \left|\left|\left|V_t^m - \mathcal{M}_0^t(v_0)\right|\right|\right| \\ &= \left|\left|\left|\sum\nolimits_{r=0}^{t-1}\mathcal{M}_{r+1}^t(V_{r+1}^m) - \mathcal{M}_r^t(V_r^m)\right|\right|\right|.\end{aligned}$$

Now notice that if we define $Y_r^m = \Phi_r^m(V_r^m)$, then $V_{r+1}^m = \mathcal{M}(Y_r^m)$ and the last equation becomes

$$|||V_t^m - v_t||| = \left|\left|\left|\sum\nolimits_{r=0}^{t-1}\mathcal{M}_r^t(Y_r) - \mathcal{M}_r^t(V_r^m)\right|\right|\right|.$$

The right-hand side of the last equation is bounded above by

$$\left|\left|\left|\sum\nolimits_{r=0}^{t-1}\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m]\right|\right|\right| + \sum\nolimits_{r=0}^{t-1}|||\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m)|||.$$

Considering the first term in the last display, note that, for any fixed $f \in \mathbb{C}^n$,

$$\begin{aligned}\mathbf{E}\Big[\big|f^{\mathrm{H}}\sum\nolimits_{r=0}^{t-1}\big(\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m]\big)\big|^2\Big] &= \sum\nolimits_{r=0}^{t-1}\mathbf{E}\Big[\big|f^{\mathrm{H}}\big(\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m]\big)\big|^2\Big] \\ &\quad + 2\sum\nolimits_{s=0}^{t-1}\sum\nolimits_{r=s+1}^{t-1}\Re\Big\{\mathbf{E}\Big[\big(f^{\mathrm{H}}(\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m])\big)\times\big(\overline{f^{\mathrm{H}}(\mathcal{M}_s^t(Y_s) - \mathbf{E}[\mathcal{M}_s^t(Y_s)\mid V_s^m])}\big)\Big]\Big\}.\end{aligned}$$

Letting $\mathcal{F}_r$ denote the σ-algebra generated by $\{V_s^m\}_{s=0}^r$ and $\{Y_r^m\}_{s=0}^{r-1}$, for $s < r$ we can write

$$\begin{aligned}&\mathbf{E}\Big[\big(f^{\mathrm{H}}(\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m])\big)\times\big(\overline{f^{\mathrm{H}}(\mathcal{M}_s^t(Y_s) - \mathbf{E}[\mathcal{M}_s^t(Y_s)\mid V_s^m])}\big)\Big] \\ &\qquad = \mathbf{E}\Big[\mathbf{E}\big[f^{\mathrm{H}}(\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m])\mid\mathcal{F}_r\big]\times\big(\overline{f^{\mathrm{H}}(\mathcal{M}_s^t(Y_s) - \mathbf{E}[\mathcal{M}_s^t(Y_s)\mid V_s^m])}\big)\Big].\end{aligned}$$

Because, conditioned on V_r^m, Y_r^m is independent of $\mathcal{F}_r$, the expression above vanishes exactly.

Supremizing over the choice of f, we have shown that

$$|||V_t^m - v_t||| \le \Big(\sum\nolimits_{r=0}^{t-1}|||\mathcal{M}_r^t(Y_r) - \mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m]|||^2\Big)^{1/2} + \sum\nolimits_{r=0}^{t-1}|||\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m)|||.$$

Expanding the term inside of the square root, we find that

$$\begin{aligned}|||V_t^m - v_t||| &\le \Big(\sum\nolimits_{r=0}^{t-1}\big(|||\mathcal{M}_r^t(Y_r) - \mathcal{M}_r^t(V_r^m)||| + |||\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m)|||\big)^2\Big)^{1/2} \\ &\qquad + \sum\nolimits_{r=0}^{t-1}|||\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m)||| \\ &\le \Big(\sum\nolimits_{r=0}^{t-1}|||\mathcal{M}_r^t(Y_r) - \mathcal{M}_r^t(V_r^m)|||^2\Big)^{1/2} + \Big(\sum\nolimits_{r=0}^{t-1}|||\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m)|||^2\Big)^{1/2} \\ &\qquad + \sum\nolimits_{r=0}^{t-1}|||\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m)|||,\end{aligned}$$

where, in the second inequality, we have used the triangle inequality for the ℓ^2-norm in $\mathbb{R}^t$. Noting that $\mathbf{E}\left[A(V_r^m)(Y_r - V_r^m)\mid V_r^m\right] = 0$ yields

$$\mathbf{E}[\mathcal{M}_r^t(Y_r)\mid V_r^m] - \mathcal{M}_r^t(V_r^m) = \mathbf{E}\big[\big(\mathcal{M}_r^t - A_r\big)(Y_r)\,\big|\,V_r^m\big] - \big(\mathcal{M}_r^t - A_r\big)(V_r^m).$$

As a consequence, applying our assumptions (31) and (32), we obtain the upper bound

$$|||V_t^m - v_t||| \le (L_1 + L_2)\Big(\sum\nolimits_{r=0}^{t-1}\alpha^{2(t-r)}|||\Phi_r^m(V_r^m) - V_r^m|||^2\Big)^{1/2} + L_2\sum\nolimits_{r=0}^{t-1}\alpha^{t-r}|||\Phi_r^m(V_r^m) - V_r^m|||^2.$$

Bounding the error from the random compressions, we arrive at the error bound

$$|||V_t^m - v_t||| \le \frac{\gamma(L_1 + L_2)}{\sqrt{m}}\Big(\sum\nolimits_{r=0}^{t-1}\alpha^{2(t-r)}\mathbf{E}\left[\|V_r^m\|_1^2\right]\Big)^{1/2} + \frac{\gamma^2 L_2}{m}\sum\nolimits_{r=0}^{t-1}\alpha^{t-r}\mathbf{E}\left[\|V_r^m\|_1^2\right]. \qquad \square$$

Proof of Corollary 1. We have already seen that when $\mathcal{M}(v) = Kv$ we can take $\alpha = \|K\|_1$ in the statement of Theorem 2 to verify conditions (31) and (32). We have also commented above that when K is nonnegative, the quantities $\mathbf{E}\left[\|V_r^m\|_1^2\right]$ can be bounded independently of n.

When $\mathcal{M}(v) = Kv/\|Kv\|_1$, bounding the size of the iterates is not an issue, but it becomes slightly more difficult to verify (31) and (32). That K is aperiodic and irreducible implies that the dominant left and right eigenvectors, v_L and v_R, of K are unique and have all positive entries. Because power iteration is invariant to scalar multiples of K we can assume that the dominant eigenvalue of K is 1. We will assume that v_L is normalized so that $\|v_L\|_\infty = 1$ and that v_R is normalized so that $v_L^{\mathrm{T}} v_R = 1$. Let D be the diagonal matrix with $D_{ii} = (v_L)_i$ (i.e., $D\mathbb{1} = v_L$). Our matrix K can be written $K = D^{-1}SD$ where S is an aperiodic, irreducible, column-stochastic matrix. Let

$$\widetilde{K} = K - v_R v_L^{\mathrm{T}} = D^{-1}SPD,$$

where we have defined the projection $P = I - Dv_R\mathbb{1}^{\mathrm{T}}$. Note that $\|P\|_1 \leq 2$ and that $PSP = SP$ so that for any positive integer r, $\tilde{K}^r = D^{-1}S^rPD$. Letting

$$C = \frac{1}{\min_j\{(v_L)_j\}} \geq 1$$

we find that, for any positive integer r,

$$\|\widetilde{K}^r\|_1 \leq \|D^{-1}\|_1\|D\|_1\|S^rP\|_1 \leq 2\,C \sup_{\substack{\|v\|_1=1\\ \mathbb{1}^{\mathrm{T}}v=0}} \|S^r v\|_1 \leq 2\,C\,\alpha^r$$

where

$$\alpha = \sup_{\substack{\|v\|_1=1\\ \mathbb{1}^{\mathrm{T}}v=0}} \|Sv\|_1$$

Aperiodicity and irreducibility of S implies that $\alpha < 1$. We also have that

$$\sup_{v_L^{\mathrm{T}}v=1} \|K^r v\|_1 \leq C \qquad \text{and} \qquad \inf_{\substack{v_L^{\mathrm{T}}v=1\\ v_j\geq 0}} \|K^r v\|_1 \geq 1.$$

Now let u and v be any two non-negative vectors normalized so that $v_L^{\mathrm{T}}u = v_L^{\mathrm{T}}v = 1$ and, for $\theta \in [0,1]$, define $w_\theta = (1-\theta)u + \theta v$. Note that w_θ also has non-negative entries and that $v_L^{\mathrm{T}}w_\theta = 1$. For any fixed $f \in \mathbb{R}^n$ with $\|f\|_\infty \leq 1$, define the function

$$\varphi_r(u,v;\theta) = \frac{f^{\mathrm{T}}K^r w_\theta}{\|K^r w_\theta\|_1} - \frac{f^{\mathrm{T}}K^r u}{\|K^r u\|_1}.$$

Our goal is to establish bounds on

$$\varphi_r(u,v;1) = \frac{f^{\mathrm{T}}K^r v}{\|K^r v\|_1} - \frac{f^{\mathrm{T}}K^r u}{\|K^r u\|_1}.$$

To that end note that

$$\frac{d}{d\theta}\varphi_r(u,v;\theta) = \frac{f^{\mathrm{T}}K^r(v-u)}{\|K^r w_\theta\|_1} - \frac{(f^{\mathrm{T}}K^r w_\theta)(\mathbb{1}^{\mathrm{T}}K^r(v-u))}{\|K^r w_\theta\|_1^2}$$

and

$$\frac{d^2}{d\theta^2}\varphi_r(u,v;\theta) = -2\frac{(f^{\mathrm{T}}K^r(v-u))(\mathbb{1}^{\mathrm{T}}K^r(v-u))}{\|K^r w_\theta\|_1^2} + 2\frac{(f^{\mathrm{T}}K^r w_\theta)(\mathbb{1}^{\mathrm{T}}K^r(v-u))^2}{\|K^r w_\theta\|_1^3}.$$

Observing that $K^r(v-u) = \tilde{K}^r(v-u)$, and applying our bounds we find that

$$\begin{aligned}
|\varphi_r(u,v;1)| &\leq \max_\theta \left|\frac{d}{d\theta}\varphi_r(u,v;\theta)\right| \\
&\leq |f^{\mathrm{T}}\tilde{K}^r(v-u)| + C|\mathbb{1}^{\mathrm{T}}\tilde{K}^r(v-u)| \\
&\leq 4\,C^2\,\alpha^r\,\|G(v-u)\|_1 \qquad (52)
\end{aligned}$$

where $G \in \mathbb{R}^{n\times n}$ is the matrix with first row equal to $f^{\mathrm{T}}\tilde{K}^r/\|2f^{\mathrm{T}}\tilde{K}^r\|_\infty$, second row equal to $\mathbb{1}^{\mathrm{T}}\tilde{K}^r/\|2\mathbb{1}^{\mathrm{T}}\tilde{K}^r\|_\infty$, and all other entries equal to 0.

Defining the matrix valued function

$$A_r(u) = \frac{1}{\|K^r u\|_1}\left[I - \frac{K^r u \mathbb{1}^{\mathrm{T}}}{\|K^r u\|_1}\right] K^r$$

we observe that

$$\frac{d}{d\theta}\varphi_r(u,v;0) = f^{\mathrm{T}} A_r(u)(v-u)$$

so that

$$\begin{aligned} |\varphi_r(u,v;1) - f^{\mathrm{T}}A_r(u)(v-u)| &\leq \frac{1}{2}\max_\theta\left|\frac{d^2}{d\theta^2}\varphi_r(u,v;\theta)\right| \\ &\leq |f^{\mathrm{T}}\tilde{K}^r(v-u)||\mathbb{1}^{\mathrm{T}}\tilde{K}^r(v-u)| + C|\mathbb{1}^{\mathrm{T}}\tilde{K}^r(v-u)|^2 \\ &\leq 16\, C^3\, \alpha^{2r}\,\|G(v-u)\|_1^2 \end{aligned} \tag{53}$$

Expressions (52) and (53) verify the stability conditions in the statement of Theorem 2 with L_1 and L_2 dependent only on C yielding the first term on the right-hand side of (33). The second term follows similarly when one observes that (31) implies

$$\sup_{v,\tilde{v}\in\mathcal{X}} \frac{\|\mathcal{M}_s^r(v) - \mathcal{M}_s^r(\tilde{v})\|_1}{\|v-\tilde{v}\|_1} \leq L_1\alpha^{r-s}.$$

□

Proof of Lemma 3. If $Y_t^m = \Phi_t^m(V_t^m)$, then

$$\begin{aligned} \mathbf{E}\left[|f^{\mathrm{H}}\Phi_t^m(V_t^m) - f^{\mathrm{H}}V_t^m|^2 \mid Y_{t-1}^m\right] &= \mathbf{E}\left[|f^{\mathrm{H}}\Phi_t^m\left(Y_{t-1}^m + \varepsilon b(Y_{t-1}^m)\right) - f^{\mathrm{H}}\left(Y_{t-1}^m + \varepsilon b(Y_{t-1}^m)\right)|^2 \mid Y_{t-1}^m\right] \\ &\leq \gamma_p \frac{\varepsilon}{m}\|b(Y_{t-1}^m)\|_1\|V_t^m\|_1 \end{aligned}$$

for some constant C. Our assumed bound on the growth of b along with (29) implies that

$$\mathbf{E}\left[\|b(Y_{t-1}^m)\|_1^2\right] \leq C'\left(1 + \mathbf{E}\left[\|V_{t-1}^m\|_1^2\right]\right)$$

for some constant C'. From these bounds it follows that for some constant $\tilde{\gamma}$,

$$|||\Phi_t^m(V_t^m) - V_t^m|||^2 \leq \tilde{\gamma}^2\frac{\varepsilon}{m}\sqrt{\mathbf{E}\left[\|V_t^m\|_1^2\right]}\sqrt{1 + \mathbf{E}\left[\|V_{t-1}^m\|_1^2\right]}.$$ □

Proof of Theorem 5. By exactly the same arguments used in the proof of Theorem 2 we arrive at the bound

$$\begin{aligned} |||V_t^m - v_t||| \leq (L_1 + L_2)\left(\sum_{r=0}^{t-1} e^{-2\beta(t-r)\varepsilon}|||\Phi_r^m(V_r^m) - V_r^m|||^2\right)^{1/2} \\ + L_2\sum_{r=0}^{t-1} e^{-\beta(t-r)\varepsilon}|||\Phi_r^m(V_r^m) - V_r^m|||^2. \end{aligned}$$

Bounding the error from the random compressions, we arrive at the error bound

$$\begin{aligned} |||V_t^m - v_t||| \leq \frac{\tilde{\gamma}(L_1+L_2)}{\sqrt{m}}\left(e^{-2\beta t\varepsilon}\mathbf{E}\left[\|V_0^m\|_1^2\right] + \varepsilon\sum_{r=1}^{t-1} e^{-2\beta(t-r)\varepsilon}\sqrt{\mathbf{E}\left[\|V_r^m\|_1^2\right]}\sqrt{1+\mathbf{E}\left[\|V_{r-1}^m\|_1^2\right]}\right)^{\frac{1}{2}} \\ + \frac{\tilde{\gamma}^2 L_2}{m}\sum_{r=1}^{t-1} e^{-\beta(t-r)\varepsilon}\sqrt{\mathbf{E}\left[\|V_r^m\|_1^2\right]}\sqrt{1+\mathbf{E}\left[\|V_{r-1}^m\|_1^2\right]}. \end{aligned}$$ □

Proof of Theorem 6. By an argument very similar to that in the proof of Theorem 2, we arrive at the bound

$$\begin{aligned} |||V_t^m - v_t||| \leq \left(\sum_{r=0}^{t-1}|||\mathcal{M}_{r+1}^t(V_r^m + \varepsilon b(Y_r^m)) - \mathcal{M}_{r+1}^t(V_r^m + \varepsilon b(V_r^m))|||^2\right)^{1/2} \\ + \left(\sum_{r=0}^{t-1}|||\mathbf{E}\left[\mathcal{M}_{r+1}^t(V_r^m + \varepsilon b(Y_r^m)) \mid V_r^m\right] - \mathcal{M}_r^t(V_r^m)|||^2\right)^{1/2} \\ + \sum_{r=0}^{t-1}|||\mathbf{E}\left[\mathcal{M}_{r+1}^t(V_r^m + \varepsilon b(Y_r^m)) \mid V_r^m\right] - \mathcal{M}_{r+1}^t(V_r^m + \varepsilon b(V_r^m))|||, \end{aligned}$$

which, also as in that proof, is bounded above by

$$|||V_t^m - v_t||| \leq (L_1 + L_2) \left(\varepsilon^2 \sum_{r=0}^{t-1} \alpha^{2(t-r-1)} |||b(Y_r^m) - b(V_r^m)|||^2\right)^{1/2} + L_2 \varepsilon^2 \sum_{r=0}^{t-1} \alpha^{t-r} |||b(Y_r^m) - b(V_r^m)|||^2.$$

From (37) and Lemma 1 we find that

$$|||b(Y_r^m) - b(V_r^m)||| \leq L_1 |||Y_r^m - V_r^m|||.$$

The rest of the argument proceeds exactly as in the proof of Theorem 2. □

Proof of Lemma 4. Observe that if $\tau_v^m > 0$, then condition

$$\sum_{j=\ell+1}^{n} |v_{\sigma_j}| \leq \frac{m-\ell}{m} \|v\|_1$$

holds for $\ell = 0$. Assume that

$$\sum_{j=\ell}^{n} |v_{\sigma_j}| \leq \frac{m-\ell+1}{m} \|v\|_1$$

for some $\ell \leq \tau_v^m$. From the definition of τ_v^m and the fact that $\ell \leq \tau_v^m$, we must also have that

$$\frac{1}{m-\ell} \sum_{j=\ell+1}^{n} |v_{\sigma_j}| < |v_{\sigma_{\ell+1}}|.$$

Combining the last two inequalities yields

$$\sum_{j=\ell+1}^{n} |v_{\sigma_j}| \leq \frac{m-\ell}{m} \|v\|_1.$$

□

Proof of Lemma 5. First we assume that, for all j, $|v_j + w_j| \leq \|v + w\|/m$. We will remove this assumption later. With this assumption in place, $N_j \in \{0, 1\}$ and the **while** loop in Algorithm 1 is inactive so that

$$f^{\mathrm{H}} \Phi_t(v+w) = \sum_{j=1}^{n} \bar{f}_j \frac{v_j + w_j}{|v_j + w_j|} \frac{\|v+w\|}{m} N_j,$$

$$\mathbf{E}\left[|f^{\mathrm{H}} \Phi_t(v+w) - f^{\mathrm{H}}(v+w)|^2\right] = \frac{\|v+w\|_1^2}{m^2} \mathbf{E}\left[\left|\sum_{j=1}^{n} \bar{f}_j \frac{v_j + w_j}{|v_j + w_j|} \left(N_j - \frac{m|v_j - w_j|}{\|v+w\|_1}\right)\right|^2\right].$$

The random variables in the sum are independent, so the last expression becomes

$$\mathbf{E}\left[|f^{\mathrm{H}} \Phi_t(v+w) - f^{\mathrm{H}}(v+w)|^2\right] = \frac{\|v+w\|_1^2}{m^2} \sum_{j=1}^{n} |f_j|^2 \mathbf{E}\left[\left|N_j - \frac{m|v_j - w_j|}{\|v+w\|_1}\right|^2\right] \leq \frac{\|v+w\|_1^2}{m^2} \sum_{j=1}^{n} \mathbf{var}\left[N_j\right].$$

Since $N_j \in \{0, 1\}$, the expression for the variance of N_j becomes

$$\mathbf{var}\left[N_j\right] = \mathbf{E}\left[N_j\right]\left(1 - \mathbf{E}\left[N_j\right]\right) = \frac{m|v_j + w_j|}{\|v+w\|_1} \left(1 - \frac{m|v_j + w_j|}{\|v+w\|_1}\right),$$

so that

$$\mathbf{E}\left[|f^{\mathrm{H}} \Phi_t(v+w) - f^{\mathrm{H}}(v+w)|^2\right] \leq \frac{\|v+w\|_1^2}{m^2} \left[m - \left(\frac{m}{\|v+w\|_1}\right)^2 \|v+w\|_2^2\right].$$

Because this scheme does not depend on the ordering of the entries of $v + w$ we can assume that the entries have been ordered so that $v_j = 0$ for $j > m$. In this case we can write

$$\|v+w\|_2^2 = \sum_{j=1}^{m} |v_j + w_j|^2 + \sum_{j=m+1}^{n} |w_j|^2 \geq \frac{1}{m} \left(\sum_{j=1}^{m} |v_j + w_j|\right)^2,$$

which then implies that

$$\mathbf{E}\left[|f^{\mathrm{H}}\Phi_t(v+w) - f^{\mathrm{H}}(v+w)|^2\right] \le \frac{\|v+w\|_1^2}{m}\left(1 - \frac{1}{\|v+w\|_1^2}\left(\|v+w\|_1 - \sum\nolimits_{j=m+1}^{n}|w_j|\right)^2\right)$$
$$\le \frac{2\|w\|_1\|v+w\|_1}{m}.$$

We now remove the assumption that $|v_j + w_j| \le \|v+w\|/m$. Let σ be a permutation of the indices of $v+w$ resulting in a vector $v_\sigma + w_\sigma$ with entries of nonincreasing magnitude. Since Algorithm 1 preserves the largest τ^m_{v+w} entries of $v+w$ and the remaining entries, $v_{\sigma_j} + w_{\sigma_j}$ for $j > \tau^m_{v+w}$, satisfy

$$|v_{\sigma_j} + w_{\sigma_j}| \le \frac{1}{m - \tau^m_{v+w}} \sum\nolimits_{k=\tau^m_{v+w}}^{n} |v_{\sigma_k} + w_{\sigma_k}|,$$

we can apply the sampling error bound just proved to find that

$$|\!|\!|\Phi_t(v+w) - v - w|\!|\!| \le \sqrt{2}\,\frac{\left(\sum_{j=\tau^m_{v+w}+1}^{n}|w_j|\right)^{\frac{1}{2}}\left(\sum_{j=\tau^m_{v+w}+1}^{n}|v_j+w_j|\right)^{\frac{1}{2}}}{\sqrt{m-\tau^m_{v+w}}}.$$

An application of Lemma 4 then yields (43).

In bounding the size of $\Phi^m_t(v+w)$ we will again assume that $\tau^m_{v+w} = 0$ and that the entries have been ordered so that $v_j = 0$ for $j > m$. The size of the resampled vector can be bounded by first noting that, since the N_j are independent and are in $\{0,1\}$,

$$\mathbf{E}\Big[\Big(\sum\nolimits_{j=1}^{n} N_j\Big)^2\Big] = \sum\nolimits_{j=1}^{n} \frac{m|v_j+w_j|}{\|v+w\|_1} + 2\sum\nolimits_{i=1}^{n}\sum\nolimits_{j=i+1}^{n} \frac{m|v_i+w_i|}{\|v+w\|_1}\frac{m|v_j+w_j|}{\|v+w\|_1}$$
$$= \sum\nolimits_{j=1}^{n}\left(\frac{m|v_j+w_j|}{\|v+w\|_1}\right)^2 + 2\sum\nolimits_{i=1}^{n}\sum\nolimits_{j=i+1}^{n} \frac{m|v_i+w_i|}{\|v+w\|_1}\frac{m|v_j+w_j|}{\|v+w\|_1}$$
$$+ \sum\nolimits_{j=1}^{n} \frac{m|v_j+w_j|}{\|v+w\|_1} - \left(\frac{m|v_j+w_j|}{\|v+w\|_1}\right)^2$$
$$= m^2 + \sum\nolimits_{j=1}^{n} \frac{m|v_j+w_j|}{\|v+w\|_1} - \left(\frac{m|v_j+w_j|}{\|v+w\|_1}\right)^2.$$

Breaking up the last sum in this expression, we find that

$$\sum\nolimits_{j=1}^{m} \frac{m|v_j+w_j|}{\|v+w\|_1} - \left(\frac{m|v_j+w_j|}{\|v+w\|_1}\right)^2 \le m\sum\nolimits_{j=1}^{m}\frac{|v_j+w_j|}{\|v+w\|_1} - m\left(\sum\nolimits_{j=1}^{m}\frac{|v_j+w_j|}{\|v+w\|_1}\right)^2$$
$$\le m\left(1 - \sum\nolimits_{j=1}\frac{|v_j+w_j|}{\|v+w\|_1}\right) \le \frac{m\|w\|_1}{\|v+w\|_1}$$

and that

$$\sum\nolimits_{j=m+1}^{n} \frac{m|w_j|}{\|v+w\|_1} - \left(\frac{m|w_j|}{\|v+w\|_1}\right)^2 \le \frac{m\|w\|_1}{\|v+w\|_1},$$

so that

$$\mathbf{E}\Big[\Big(\sum\nolimits_{j=1}^{n} N_j\Big)^2\Big] \le m^2 + 2\frac{m\|w\|_1}{\|v+w\|_1}.$$

It follows then that (at least when $\tau^m_{v+w} = 0$)

$$\mathbf{E}\left[\|\Phi^m_t(v+w)\|_1^2\right] \le \|v+w\|_1 + 2\frac{\|v+w\|_1\|w\|_1}{m}.$$

Writing the corresponding formula for $\tau^m_{v+w} > 0$ and applying Lemma 4 gives the bound in the statement of the lemma.

Finally we consider the probability of the event $\{\Phi_t^m(v+w)=0\}$. If $\tau_{v+w}^m=0$, then $N_j\in\{0,1\}$, so that $\mathbf{P}[N_j=0]=1-m|v_j+w_j|/\|v+w\|_1$, and, since the N_j are independent,

$$\mathbf{P}[N_j=0 \text{ for all } j]=\prod_{j=1}^{n}\left(1-\frac{m|v_j+w_j|}{\|v+w\|_1}\right)\leq\prod_{j\leq n,\ v_j\neq 0}\left(1-\frac{m|v_j+w_j|}{\|v+w\|_1}\right).$$

The first product in the last display is easily seen to be bounded above by e^{-m}. The second product is maximized subject to the constraint

$$\sum\nolimits_{j\leq n,\ v_j\neq 0}\left(1-\frac{m|v_j+w_j|}{\|v+w\|_1}\right)\leq\frac{m\|w\|_1}{\|v+w\|_1}$$

when the terms in the product are all equal, in which case we get

$$\mathbf{P}[N_j=0 \text{ for all } j]\leq\left(\frac{\|w\|_1}{\|v+w\|_1}\right)^m. \qquad \square$$