Corrigendum to
"Lieb's concavity theorem, matrix geometric means, and semidefinite optimization"

Hamza Fawzi[*] James Saunderson[†]

April 13, 2020

We correct a mistake in [FS17] with the semidefinite representation of the matrix geometric mean. We explicitly describe the changes that need to be made to fix the error. We also show that the SDP representations of related functions derived in [FS17] and [FSP19], such as Lieb's function and approximations to the quantum relative entropy, were not affected by this mistake.

Notations We first recall some notation from [FS17]. The t-weighted matrix geometric mean of $A, B \in \mathbf{H}^n_{++}$ (Hermitian positive deinite matrices) is $A\#_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$. When $t = 1/2$ we write $A\#_{1/2}B = A\#B$. The matrix geometric mean is jointly operator concave in both its arguments (A, B) and we let $\text{hyp}_t := \{(A, B, T) : A\#_t B \succeq T\}$.

1 Mistake and fixes

The mistake In Lemma 4 of [FS17] it is claimed that

$$A\#B \succeq T \quad \iff \quad \begin{bmatrix} A & T \\ T & B \end{bmatrix} \succeq 0. \tag{1}$$

The direction $\Leftarrow$ is correct and can be proved using the Schur complement lemma, and monotonicity of the matrix square root. The direction $\Rightarrow$ however is incorrect. For example when $B = I_n$ (identity matrix), then $A\#B = A^{1/2}$ and so $A\#B \succeq T \Leftrightarrow A^{1/2} \succeq T$. On the other hand, we have, by the Schur complement lemma $\left[\begin{smallmatrix} A & T \\ T & I \end{smallmatrix}\right] \succeq 0$ if and only $A \succeq T^2$. To find a counterexample to $\Rightarrow$ in (1) it thus suffices to find A, T such that $A^{1/2} \succeq T$ and yet $A \not\succeq T^2$. Such matrices can be found since it is known that the matrix square function is not monotone.

The fix Fortunately however, one can fix this problem by introducing an additional variable. We have:

$$A\#B \succeq T \quad \iff \quad \exists W \in \mathbf{H}^n : \begin{bmatrix} A & W \\ W & B \end{bmatrix} \succeq 0 \text{ and } W \succeq T. \tag{2}$$

[*]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK. Email: h.fawzi@damtp.cam.ac.uk

[†]Department of Electrical and Computer Systems Engineering, Monash University, VIC 3800, Australia. Email: james.saunderson@monash.edu

Proof of Equation (2). The direction $\Leftarrow$ is similar as before: if $\left[\begin{smallmatrix} A & W \\ W & B \end{smallmatrix}\right] \succeq 0$ then $A\#B \succeq W$ and so using $W \succeq T$ we get $A\#B \succeq T$. For the direction $\Rightarrow$, it suffices to take $W = A\#B$. One can verify that the block matrix $\left[\begin{smallmatrix} A & A\#B \\ A\#B & B \end{smallmatrix}\right]$ is positive semidefinite since

$$\begin{bmatrix} A & A\#B \\ A\#B & B \end{bmatrix} = \begin{bmatrix} A^{1/2} \\ A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2} \end{bmatrix} \begin{bmatrix} A^{1/2} \\ A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2} \end{bmatrix}^* \succeq 0.$$

□

Other changes The wrong expression for $\mathrm{hyp}_{1/2}$ was subsequently used in Lemma 5 and Proposition 1. We have corrected these statements in the new arXiv version of the paper, at the URL https://arxiv.org/abs/1512.03401 (V3). The main change is the SDP description given in Proposition 1 which should include an additional LMI. The corrected statement is

$$\mathrm{hyp}_{p/2^\ell} = \left\{ (A,B,T) : \exists Z_1,\ldots,Z_{\ell-1}, Z_\ell \in \mathbf{H}^n \text{ s.t. } \begin{bmatrix} A\#_{m_i}B & Z_i \\ Z_i & Z_{i-1} \end{bmatrix} \succeq 0 \text{ for } i = 2,3,\ldots,\ell, \right. \\ \left. \begin{bmatrix} A & Z_1 \\ Z_1 & B \end{bmatrix} \succeq 0,\ Z_\ell \succeq T \right\}. \quad (3)$$

2 Comment

In the paper [FS17] we used the SDP representation of the matrix geometric mean to give SDP representations for other matrix functions, in particular Lieb's function $F_t(A,B) = \mathrm{tr}[K^* A^{1-t} K B^t]$. It turns out that, even though our representation of the matrix geometric mean was incorrect, the SDP representation for F_t *is* correct. This is because the SDP representation of F_t only relied on specific properties that are satisfied by our LMI construction. Namely, the SDP representation works with any set $C_t \subset \mathbf{H}^n_{++} \times \mathbf{H}^n_{++} \times \mathbf{H}^n$ that satisfies the following properties:

$$\begin{aligned} &(i) \quad (A,B,T) \in C_t \quad \Rightarrow \quad A\#_t B \succeq T \\ &(ii) \quad \forall A,B \in \mathbf{H}^n_{++}, \ \ (A,B,A\#_t B) \in C_t. \end{aligned} \quad (4)$$

Remark 1. *Given a set C_t satisfying* (4) *one can construct an SDP representation of the hypograph* hyp_t *at the price of a single additional LMI:*

3. $\{(A,B,T) : A\#_t B \succeq T\} = \{(A,B,T) : \exists W$ *s.t.* $(A,B,W) \in C_t$ *and* $W \succeq T\}$.

This is essentially the change in Equation (2) *vs.* (1).

Lieb's function Our representation of Lieb's function $F_t(A,B) = \mathrm{tr}[K^* A^{1-t} K B^t]$ uses the fact (Equation (7) in [FS17]) that

$$\mathrm{tr}[K^* A^{1-t} K B^t] \geq t \quad \iff \quad \exists T \text{ s.t. } \begin{cases} (A \otimes I, I \otimes \bar{B}, T) \in \mathrm{hyp}_t \\ \mathrm{vec}(K)^* T \,\mathrm{vec}(K) \geq t. \end{cases}$$

We claim that a similar equivalence is true if we replace hyp_t by C_t. Namely we have:

$$\mathrm{tr}[K^* A^{1-t} K B^t] \geq t \quad \iff \quad \exists T \text{ s.t. } \begin{cases} (A \otimes I, I \otimes \bar{B}, T) \in C_t \\ \mathrm{vec}(K)^* T \,\mathrm{vec}(K) \geq t. \end{cases}$$

For the direction $\Rightarrow$ it suffices to take $T = A^{1-t} \otimes \bar{B}^t$. For the reverse direction we use the fact that $(A \otimes I, I \otimes \bar{B}, T) \in C_t \Rightarrow A^{1-t} \otimes \bar{B}^t \succeq T$. This shows that the SDP representation of Lieb's function is correct.

Other functions In [FS17] we also derive SDP representations for the functions $S_t(A,B) = \frac{1}{t}\mathrm{tr}[A - A^{1-t}B^t]$ and $\Upsilon_t(X) = \mathrm{tr}[(K^*A^tK)^{1/t}]$. These rely only on F_t (and not directly on the matrix geometric mean) and so do not require changes.

Implications for the paper "Semidefinite approximations of the matrix logarithm" In the subsequent paper [FSP19] the SDP representation of the matrix geometric mean was used to give an SDP approximation of the quantum relative entropy. We show that, just like with Lieb's function, the set of LMI constraints given in [FSP19] are valid. Let r_m be the rational function from [FSP19, Equation (9)] and let $P_{r_m}(X,Y) = Y^{1/2}r_m(Y^{-1/2}XY^{-1/2})Y^{1/2}$ be its matrix perspective. Consider the following cone defined in [FSP19, Equation (15)], which serves as an approximation of the operator relative entropy cone:

$$K_{m,k} = \left\{(X,Y,T) : P_{r_m}(X\#_{2^{-k}}Y, X) \succeq -T/2^k\right\}.$$

This cone can be written equivalently as

$$K_{m,k} = \left\{(X,Y,T) : \exists Z \text{ s.t. } P_{r_m}(Z,X) \succeq -T/2^k \text{ and } (X,Y,Z) \in \mathrm{hyp}_{2^{-k}}\right\}.$$

We claim that this representation still holds if we replace $\mathrm{hyp}_{2^{-k}}$ by $C_{2^{-k}}$. The argument is similar as above. For the inclusion $\subseteq$ we let $Z = A\#_{2^{-k}}B$. The reverse inclusion holds because $C_{2^{-k}} \subseteq \mathrm{hyp}_{2^{-k}}$. The SDP representations in the paper [FSP19] of the operator relative entropy (and other derived functions such as the Umegaki relative entropy) are thus not affected.

Acknowledgments HF would like to thank Omar Fawzi who pointed out the mistake.

References

[FS17] Hamza Fawzi and James Saunderson. Lieb's concavity theorem, matrix geometric means, and semidefinite optimization. *Linear Algebra and its Applications*, 513:240–263, 2017.

[FSP19] Hamza Fawzi, James Saunderson, and Pablo A. Parrilo. Semidefinite approximations of the matrix logarithm. *Foundations of Computational Mathematics*, 19:259–296, 2019. Package `cvxquad`: https://github.com/hfawzi/cvxquad.