16.5

The answer to this question is yes - if we come up with a clever enough choice for g, we can actually turn this into a factor 3/4 approximation.

To start, I'll outline the properties we will need this function g to satisfy, and then I'll demonstrate one particular g that works. First, notice that with the algorithm from class we can write

$$E[\text{value(output)}] = \sum_C w_C \left(1 - \prod_{i:x_i \in S_C^+} (1-y_i^*) \prod_{i:x_i \in S_C^-} y_i^*\right)$$

where I am using the notation from the book to mean that S_C^+ is the set of variables in the clause that are unnegated and S_C^- are the ones that are negated. Since the only way for a clause to be unsatisfied is if all its literals are set to False (also recall that y_i^* represents the probability that x_i is set to True), the term we are multiplying by the weight is just the probability that this clause is satisfied. Therefore, we are looking for a $g : [0,1] \to [0,1]$ such that

$$1 - \prod_{S_C^+} (1 - g(y_i^*)) \prod_{S_C^-} g(y_i^*) \geq \frac{3}{4} z_C^*, \tag{1}$$

as this would imply that (using the $g(y_i^*)$ as our probabilities this time):

$$\begin{aligned} E[\text{value(output)}] &= \sum_C w_C \left(1 - \prod_{S_C^+} (1-g(y_i^*)) \prod_{S_C^-} g(y_i^*)\right) \\ &\geq \frac{3}{4} \sum_C w_C z_C^* \\ &= \frac{3}{4} \cdot OPT_{LP} \\ &\geq \frac{3}{4} \cdot OPT \end{aligned}$$

and we would have a factor 3/4 approximation algorithm. If we plug our constraints from the linear program into Equation (1), we find that we need a g such that

$$1 - \prod_{S_C^+} (1 - g(y_i^*)) \prod_{S_C^-} g(y_i^*) \geq \frac{3}{4} \min\left(1, \sum_{S_C^+} y_i + \sum_{S_C^-} (1 - y_i)\right). \tag{2}$$

The idea for g is to emulate the behavior of the 3/4-approximation we examined in class and to find a compromise between having probabilities of 1/2 and having probabilities of y_i^*. This motivates my choice of g, which is defined as follows:

$$g(y) = \begin{cases} \frac{3}{4}y + \frac{1}{4} & \text{if } 0 \leq y \leq \frac{1}{3} \\ \frac{1}{2} & \text{if } \frac{1}{3} \leq y \leq \frac{2}{3} \\ \frac{3}{4}y & \text{if } \frac{2}{3} \leq y \leq 1 \end{cases}$$

In the same way that we did in class, we can rename and reorder the variables of our particular clause of size k to assume that its variables are $x_1, \ldots, x_k$ and that none of the variables are negated. Note that one reason this is okay is that g satisfies the property that

$$1 - g(1-y) = g(y)$$

so in particular the probabilities will work out to be the same. Now, we simplify Equation (2) to see that we need to prove that g is such that

$$1 - \prod(1 - g(y_i^*)) \geq \frac{3}{4}\min\left(1, \sum y_i^*\right).$$

To make things a little easier, we can rewrite this as $G \geq \frac{3}{4}M$. First, we need to make a few simplifying assumptions about the y_i^*. Before we make these assumptions note that any changes we make to the y_i^* that do not increase G or decrease M are fine, since if we prove that $G' \geq \frac{3}{4}M$ for some $G' \leq G$ this still implies that $G \geq \frac{3}{4}M$, and similarly for an $M' \geq M$.

First, we can assume that no y_i^* fall in our intermediate interval - namely $[\frac{1}{3}, \frac{2}{3})$. Note that if we did have such a y_i^*, we would be able to change it to $\frac{2}{3}$ without affecting our left side G, and getting a new M' such that $M' \geq M$.

We can also assume that all the y_i^* in our lower interval $[0, \frac{1}{3})$ are equal. This is because of the arithmetic mean/geometry mean inequality, which says that for n numbers $x_1, \ldots, x_n$

$$\frac{x_1 + \ldots + x_n}{2} \geq (x_1 \cdot \ldots \cdot x_n)^{1/2}$$

where equality holds if and only if $x_1 = \ldots x_n$. So assuming that $y_1^* = \ldots = y_n^*$ where they are all equal to their arithmetic mean doesn't affect M and produces a G' such that $G' \leq G$, since the product is maximized and therefore G is minimized.

Finally, we assume that if there is some y_i in the lower interval $[0, \frac{1}{3})$ then there isn't a y_j in the higher interval $(\frac{2}{3}, 1]$. This works because we can just increase y_i by ϵ and decrease y_j by ϵ without affecting M and producing a G' such that $G' \leq G$.

Now that we have made all our assumptions, we can break our proof down into 4 cases:

- Case 1: We assume that $k = 1$. Then $G = 1 - (1 - g(y_1^*)) = 1 - g(1 - y_1^*) = g(y_1^*)$, and by the nature of g we know this is greater than or equal to $\frac{3}{4}y_1^*$. So we know that $G \geq \frac{3}{4}M$.
- Case 2: We assume that $k \geq 2$, and that all the y_i^* fall in the lower interval. By our second assumption above, this means that they are all equal to some value $y \leq \frac{1}{3}$. This further

implies that $ky \geq M$. Then we have that

$$\begin{aligned}
G &= 1 - (1 - g(y))^k \\
&= 1 - \left(1 - \left(\frac{3}{4}y + \frac{1}{4}\right)\right)^k \\
&= 1 - \left(\frac{3}{4}\right)^k (1 - y)^k \\
&\geq 1 - \left(\frac{3}{4}\right)^k (1 - \frac{M}{k})^k \qquad \text{because } ky \geq M \\
&\geq 1 - \left(\frac{3}{4}\right)^k e^{-M} \qquad \text{because } (1 - \frac{M}{k})^k \leq e^{-M} \\
&\geq 1 - \frac{9}{16}e^{-M} \qquad \text{because } k \geq 2 \\
&= (1 - \frac{9}{16}e^{-M} - \frac{3}{4}M) + \frac{3}{4}M \\
&\geq \min\left(\frac{7}{16}, \frac{1}{4} - \frac{9}{16} \cdot \frac{1}{e}\right) + \frac{3}{4}M \qquad \text{by the concavity of } 1 - \frac{9}{16}e^{-M} - \frac{3}{4}M \\
&\geq \frac{3}{4}M
\end{aligned}$$

- Case 3: We assume that $k \geq 2$, and that all but one of the y_i^* are in the lower interval (and again equal to their arithmetic mean y). This means, by our first and third assumptions above, that there is one $y_j^* = \frac{2}{3}$. This means that $(k-1)y + \frac{2}{3} \geq M$, so we have that

$$\begin{aligned}
G &= 1 - g\left(\frac{2}{3}\right)(1-g(y))^{k-1} \\
&= 1 - \frac{1}{2}\left(1 - \left(\frac{3}{4}y + \frac{1}{4}\right)\right)^{k-1} \\
&= 1 - \frac{1}{2}\left(\frac{3}{4} - \frac{3}{4}y\right)^{k-1} \\
&\geq 1 - \frac{1}{2}\left(\frac{3}{4}\right)^{k-1}(1-y)^{k-1} \\
&= 1 - \frac{1}{2}\left(\frac{3}{4}\right)^{k-1}\left(1 - \frac{M - \frac{2}{3}}{k-1}\right)^{k-1} \\
&\geq 1 - \frac{1}{2}\left(\frac{3}{4}\right)^{k-1} e^{-M+\frac{2}{3}} \\
&\geq 1 - \frac{3}{8}e^{-M+\frac{2}{3}} \qquad \text{because } k \geq 2 \\
&= \left(1 - \frac{3}{8}e^{-M+\frac{2}{3}} - \frac{3}{4}M\right) + \frac{3}{4}M \\
&\geq \min\left(1 - \frac{3}{8}e^{\frac{2}{3}}, \frac{1}{4} - \frac{3}{8}e^{-\frac{1}{3}}\right) + \frac{3}{4}M \\
&\geq \frac{3}{4}M
\end{aligned}$$

- Case 4: Finally, we assume that at least two of the y_i^* are in the higher interval, so $y_i^* \geq \frac{2}{3}$ for at least two values of i. Then clearly $M = 1$, so we need to show $G \geq \frac{3}{4}$. We can assume that y_1^* and y_2^* are the ones in the higher interval (there could be more)

$$\begin{aligned}
G &= 1 - \prod(1 - g(y_i^*)) \\
&\geq 1 - g(y_1^*)g(y_2^*)\prod_{i\geq 3}(1 - g(y_i^*)) \\
&\geq 1 - \frac{1}{4}\prod_{i\geq 3}(1 - g(y_i^*)) \\
&\geq \frac{3}{4}
\end{aligned}$$

where this last part follows because a product of things of the form $1-x$ can never exceed 1. So we have that $G \geq \frac{3}{4}M$ and we are done.